

SHARP TOTAL VARIATION RESULTS FOR MAXIMAL FUNCTIONS

João P. G. Ramos

Rheinische Friedrich-Wilhelms-Universität Bonn, Mathematisches Institut
Endenicher Allee 60, D-53115 Bonn, Germany; jpgramos@math.uni-bonn.de

Abstract. In this article, we prove some total variation inequalities for maximal functions. Our results deal with two possible generalizations of the results contained in Aldaz and Pérez Lázaro’s work [1], one of whose considers a variable truncation of the maximal function, and the other one interpolates the centered and the uncentered maximal functions. In both contexts, we find sharp constants for the desired inequalities, which can be viewed as progress towards the conjecture that the best constant for the variation inequality in the centered context is one. We also provide counterexamples showing that our methods do not apply outside the stated parameter ranges.

1. Introduction

An object of major interest in harmonic analysis is the Hardy–Littlewood maximal function, which can be defined as

$$Mf(x) = \sup_{t \in \mathbf{R}_+} \frac{1}{2t} \int_{x-t}^{x+t} |f(s)| \, ds.$$

Alternatively, one can also define its *uncentered* version as

$$\tilde{M}f(x) = \sup_{x \in I} \frac{1}{|I|} \int_I |f(s)| \, ds.$$

The most classical result about these maximal functions is perhaps the Hardy–Littlewood–Wiener theorem, which states that both M and \tilde{M} map $L^p(\mathbf{R})$ into itself for $1 < p \leq \infty$, and that in the case $p = 1$ they satisfy a weak type inequality:

$$|\{x \in \mathbf{R} : Mf(x) > \lambda\}| \leq \frac{C}{\lambda} \|f\|_1,$$

where $C = \frac{11+\sqrt{61}}{12}$ is the best constant possible found by Melas [12] for M . The same inequality also holds in the case of \tilde{M} above, but this time with $C = 2$ being the best constant, as shown by F. Riesz [13].

In the remarkable paper [6], Kinnunen proves, using functional analytic techniques and the aforementioned theorem, that, in fact, M maps the Sobolev spaces $W^{1,p}(\mathbf{R})$ into themselves, for $1 < p \leq \infty$. Kinnunen also proves that this result holds if we replace the standard maximal function by its uncentered version. This opened a new field of studies, and several other properties of this and other related maximal functions were studied. We mention, for example, [4, 3, 5, 7, 9].

Since the Hardy–Littlewood maximal function fails to be in L^1 for every nontrivial function f and the tools from functional analysis used are not available either in the

<https://doi.org/10.5186/aasfm.2019.4409>

2010 Mathematics Subject Classification: Primary 26A45, 42B25, 46E35.

Key words: Hardy–Littlewood maximal function, functions of bounded variation, sharp estimates.

case $p = 1$, an important question was whether a bound of the form $\|(Mf)'\|_1 \leq C\|f'\|_1$ could hold for every $f \in W^{1,1}$.

In the uncentered case, Tanaka [15] provided us with a positive answer to this question. Explicitly, Tanaka proved that, whenever $f \in W^{1,1}(\mathbf{R})$, then $\tilde{M}f$ is weakly differentiable, and it satisfies that $\|(\tilde{M}f)'\|_1 \leq 2\|f'\|_1$. Here, $W^{1,1}(\mathbf{R})$ stands for the Sobolev space $\{f: \mathbf{R} \rightarrow \mathbf{R}: \|f\|_1 + \|f'\|_1 < +\infty\}$.

Some years later, Aldaz and Pérez Lázaro [1] improved Tanaka's result, showing that, whenever $f \in BV(\mathbf{R})$, then the maximal function $\tilde{M}f$ is in fact *absolutely continuous*, and $\mathcal{V}(\tilde{M}f) = \|(\tilde{M}f)'\|_1 \leq \mathcal{V}(f)$, with $C = 1$ being sharp, where we take the *total variation of a function* to be $\mathcal{V}(f) := \sup_{\{x_1 < \dots < x_N\} = \mathcal{P}} \sum_{i=1}^{N-1} |f(x_{i+1}) - f(x_i)|$, and consequently the space of *bounded variation functions* as the space of functions $f: \mathbf{R} \rightarrow \mathbf{R}: \exists g; f = g$ a.e. and $\mathcal{V}(g) < +\infty$. In this direction, Bober, Carneiro, Hughes and Pierce [2] studied the discrete version of this problem, obtaining similar results.

In the centered case, many questions remain unsolved. Surprisingly, it turned out to be *harder* than the uncentered one, due to the contrast in smoothness of Mf and $\tilde{M}f$. In [8], Kurka showed the endpoint question to be true, that is, that $\mathcal{V}(Mf) \leq C\mathcal{V}(f)$, with $C = 240,004$. Unfortunately, his method does not give the best constant possible, with the standing conjecture being that $C = 1$ is the sharp constant.

In [16], Temur studied the discrete version of this problem, proving that for every $f \in BV(\mathbf{Z})$ we have $\mathcal{V}(Mf) \leq C'\mathcal{V}(f)$, where $C' > 10^6$ is an absolute constant. The standing conjecture is again that $C' = 1$ in this case, which was in part backed up by Madrid's optimal results [11]: If $f \in \ell^1(\mathbf{Z})$, then $Mf \in BV(\mathbf{Z})$, and $\mathcal{V}(Mf) \leq 2\|f\|_1$, with 2 being sharp in this inequality.

Our main theorems deal with—as far as the author knows—the first attempt to prove sharp bounded variation results for classical Hardy–Littlewood maximal functions. Indeed, we may see the classical, uncentered Hardy–Littlewood maximal function as

$$\tilde{M}f(x) = \sup_{x \in I} \frac{1}{|I|} \int_I |f(s)| \, ds = \sup_{(y,t): |x-y| \leq t} \frac{1}{2t} \int_{y-t}^{y+t} |f(s)| \, ds.$$

Notice that this supremum is *not* necessarily attained for every function f and at every point $x \in \mathbf{R}$, but this shall not be a problem for us in the most diverse cases, as we will see throughout the text. This way, we may look at this operator as a particular case of the wider class of *nontangential* maximal operators

$$M^\alpha f(x) = \sup_{|x-y| \leq \alpha t} \frac{1}{2t} \int_{y-t}^{y+t} |f(s)| \, ds.$$

Indeed, from this new definition, we get directly that

$$\begin{cases} M^\alpha f = Mf, & \text{if } \alpha = 0, \\ M^\alpha f = \tilde{M}f, & \text{if } \alpha = 1. \end{cases}$$

As in the uncentered case, we can still define ‘truncated’ versions of these operators, by imposing that $t \leq R$. These operators are far from being a novelty: several references consider those all around mathematics, among those the classical [14, Chapter 2], and the more recent, yet related to our work, [3]. An easy argument

(see Section 5.1 below) proves that, if $\alpha < \beta$, then

$$\mathcal{V}(M^\beta f) \leq \mathcal{V}(M^\alpha f).$$

This implies already, by the main Theorem in [8], that there exists a constant $A \geq 0$ such that $\mathcal{V}(M^\alpha f) \leq A\mathcal{V}(f)$, for all $\alpha > 0$. In the intention of sharpening this result, our first result reads, then, as follows:

Theorem 1. *Fix any $f \in BV(\mathbf{R})$. For every $\alpha \in [\frac{1}{3}, +\infty)$, we have that*

$$(1) \quad \mathcal{V}(M^\alpha f) \leq \mathcal{V}(f).$$

There exists an extremizer f for the inequality (1). If $\alpha > \frac{1}{3}$, then any positive extremizer f to inequality (1) satisfies:

- $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow +\infty} f(x)$.
- There is x_0 such that f is non-decreasing on $(-\infty, x_0)$ and non-increasing on $(x_0, +\infty)$.

Finally, for every $\alpha \geq 0$ and $f \in W^{1,1}(\mathbf{R})$, $M^\alpha f \in W_{\text{loc}}^{1,1}(\mathbf{R})$.

Notice that stating that a function $g \in W_{\text{loc}}^{1,1}(\mathbf{R})$ is the same as asking it to be locally absolutely continuous. Our ideas to prove this theorem and theorem 3 are heavily inspired by the ones in [1]. Our aim will always be to prove that, when $f \in BV(\mathbf{R})$, then the maximal function $M^\alpha f$ is well-behaved on the detachment set

$$E_\alpha = \{x \in \mathbf{R} : M^\alpha f(x) > f(x)\}.$$

Namely, we seek to obtain that the maximal function does not have any local maxima in the set where it disconnects from the original function. Such an idea, together with the concept of the detachment set E_α , are also far from being new, having already appeared at [1, 4, 3, 15], and recently at [10]. More specific details of this can be found in the next section.

In general, our main ideas are contained in Lemma 2, where we prove that the region in the upper half plane that is taken into account for the supremum that defines

$$M_{\equiv R}^1 f = \sup_{x \in I : |I| \leq 2R} \int_I |f(s)| \, ds,$$

where we define

$$\int_I g(s) \, ds := \frac{1}{|I|} \int_I g(s) \, ds,$$

is actually a (rotated) *square*, and not a triangle—as a first glance might impress on someone—, and in the comparison of $M^\alpha f$ and $M_{\equiv R}^1$ over a small interval, in order to establish the maximal attachment property.

We may ask ourselves if, for instance, we could go lower than $1/3$ with this method. Our next result, however, shows that this is the optimal bound for this technique:

Theorem 2. *Let $\alpha < \frac{1}{3}$. Then there exists $f \in BV(\mathbf{R})$ such that $f \geq 0$, $f(x) = \limsup_{y \rightarrow x} f(y)$ and a point $x_\alpha \in \mathbf{R}$ such that x_α is a local maximum of $M^\alpha f$, but $M^\alpha f(x_\alpha) > f(x_\alpha)$.*

We could, alternatively, use other normalizations on f more suitable to each $M^\alpha f$. See the next section for further definitions and motivations for such normalizations.

We can inquire ourselves whether we can generalize the results from Aldaz and Pérez Lázaro in yet another direction, though. With this in mind, we notice that Kurka [8] mentions in his paper that his techniques allow one to prove that some

Lipschitz truncations of the center maximal function, that is, maximal functions of the form

$$M_N^0 f(x) = \sup_{t \leq N(x)} \frac{1}{2t} \int_{x-t}^{x+t} |f(s)| \, ds,$$

are bounded from $BV(\mathbf{R})$ to $BV(\mathbf{R})$ —with some possibly big constant—if $\text{Lip}(N) \leq 1$. Inspired by it, we define the N -truncated uncentered maximal function as

$$M_N^1 f(x) = \sup_{|x-y| \leq t \leq N(x)} \int_{y-t}^{y+t} |f(s)| \, ds.$$

The next result deals then with an analogous of Kurka’s result in the case of the centered maximal functions. In fact, we achieve even more in this case, as we have also the explicit sharp constants for that. In details, the result reads as follows:

Theorem 3. *Let $N: \mathbf{R} \rightarrow \mathbf{R}_+$ be a measurable function. If $\text{Lip}(N) \leq \frac{1}{2}$, we have that, for all $f \in BV(\mathbf{R})$,*

$$\mathcal{V}(M_N^1 f) \leq \mathcal{V}(f).$$

Moreover, the result is sharp, in the sense that there are non-constant functions f such that $\mathcal{V}(f) = \mathcal{V}(M_N^1 f)$.

Again, we are also going to use a careful maxima analysis in this case. Actually, we are going to do it both in theorems 1 and 3 for the non-endpoint cases $\alpha > \frac{1}{3}$ and $\text{Lip}(N) < \frac{1}{2}$, while the endpoints are treated with a limiting argument.

In the same way, one may ask whether we can ask our Lipschitz constant to be greater than $\frac{1}{2}$ in this result. Regarding this question, we prove in section 4.3 the following negative answer:

Theorem 4. *Let $c > \frac{1}{2}$ and*

$$f(x) = \begin{cases} 1, & \text{if } x \in (-1, 0); \\ 0, & \text{otherwise.} \end{cases}$$

Then there is a function $N: \mathbf{R} \rightarrow \mathbf{R}_{\geq 0}$ such that $\text{Lip}(N) = c$ and

$$\mathcal{V}(M_N^1 f) = +\infty.$$

Acknowledgements. The author would like to thank Christoph Thiele, for the remarks that led him to the full range $\alpha \geq \frac{1}{3}$ at Theorem 1, as well as to the proof that this is sharp for this technique, and Olli Saari, for enlightening discussions about the counterexamples in the proof of Theorem 4 and their construction. He would also like to thank Emanuel Carneiro and Mateus Sousa for helpful comments and discussions, many of which took place during the author’s visit to the International Centre for Theoretical Physics in Trieste, to which the author is grateful for its hospitality, and Diogo Oliveira e Silva, for his thorough inspection and numerous comments on the preliminary versions of this paper. The author would like to thank also the anonymous referees, whose corrections and ideas, among which the use of a new normalization, have simplified and cleaned a lot this manuscript. Finally, the author acknowledges financial support from the Hausdorff Center of Mathematics and the DAAD.

2. Basic definitions and properties

Throughout the paper, I and J will usually denote open intervals, and $l(I)$, $l(J)$, $r(I)$, $r(J)$ their left and right endpoints, respectively. We also denote, for $f \in BV(\mathbf{R})$, the *one-sided limits* $f(a+)$ and $f(a-)$ to be

$$f(a+) = \lim_{x \searrow a} f(x) \quad \text{and} \quad f(a-) = \lim_{x \nearrow a} f(x).$$

We also define, for a general function $N: \mathbf{R} \rightarrow \mathbf{R}$, its *Lipschitz constant* as

$$\text{Lip}(N) := \sup_{x \neq y \in \mathbf{R}} \frac{|N(x) - N(y)|}{|x - y|}.$$

By considering the arguments and techniques contained in the lemmata from [1], we may consider sometimes a function in $BV(\mathbf{R})$ endowed with the normalization $f(x) = \limsup_{y \rightarrow x} f(y)$, $\forall x \in \mathbf{R}$. At some other times, however, we might need to work with a normalization a little more friendly to the maximal functions involved. Let, then, for a fixed $\alpha \in (0, 1]$,

$$\mathcal{N}_\alpha f(x) = \limsup_{(y,t) \rightarrow (x,0): |y-x| \leq \alpha t} \frac{1}{2t} \int_{y-t}^{y+t} |f(s)| \, ds.$$

This coincides, by definition, with f almost everywhere, as bounded variation functions are continuous almost everywhere. Moreover, this normalization can be stated, in a pointwise context, as

$$\mathcal{N}_\alpha f(x) = \frac{(1 + \alpha) \limsup_{y \rightarrow x} |f(y)| + (1 - \alpha) \liminf_{y \rightarrow x} |f(y)|}{2}.$$

With this normalization, we see that, for any $f \in BV(\mathbf{R})$,

$$M^\alpha f(x) \geq \mathcal{N}_\alpha f(x), \quad \text{for each } x \in \mathbf{R}.$$

This normalization, however, is *not* friendly to boundary points: the sets $\{M^\alpha f > f\}$ might not be open when we adopt it, as the example of $f = \chi_{(0, \frac{1-\alpha}{4}]} + \frac{1}{2} \chi_{(\frac{1-\alpha}{4}, \frac{1-\alpha}{2}]} + \chi_{(\frac{1-\alpha}{2}, 1]}$ endowed with $\mathcal{N}_\alpha f$ shows. This function has the property that $M^\alpha f(\frac{1-\alpha}{2}) > \mathcal{N}_\alpha f(\frac{1-\alpha}{2})$, but $M^\alpha f = f$ at $(\frac{1-\alpha}{2}, 1)$.

Consider then $\mathcal{N}_\alpha f$, and notice that the situations as in the example above can only happen if $\mathcal{N}_\alpha f$ is *discontinuous* at a point x . We then let

$$(2) \quad \tilde{\mathcal{N}}_\alpha f(x) = \begin{cases} \mathcal{N}_\alpha f(x), & \text{if } M^\alpha f(x) > \limsup_{y \rightarrow x} f(y); \\ M^\alpha f(x), & \text{if } \limsup_{y \rightarrow x} f(x) \geq M^\alpha f(x) \geq \mathcal{N}_\alpha f(x). \end{cases}$$

Of course, we are only changing the points in which $\liminf_{y \rightarrow x} f(y) < \mathcal{N}_\alpha f < \limsup_{y \rightarrow x} f(y)$, and thus this normalization does not increase the variation, i.e., $\mathcal{V}(\tilde{\mathcal{N}}_\alpha f) \leq \mathcal{V}(f)$. Again, by adapting the lemmata in [1] to this context, one checks that we may assume, without loss of generality, that our function has this normalization. We will, for shortness, say we are using $NORM(\alpha)$ whenever we use this normalization. Notice that $NORM(1)$ is the normalization used by Aldaz and Pérez Lázaro.

We mention also a couple of words about the maxima analysis performed throughout the paper. In the paper [1], the authors developed an ingenious way to prove the sharp bounded variation result for the uncentered maximal function. Namely,

they proved that, whenever $f \in BV(\mathbf{R})$, then the maximal function $\tilde{M}f$ is actually continuous, and the (open) set

$$E = \{\tilde{M}f > f\} = \bigcup_j I_j$$

satisfies that, in each of the intervals I_j , $\tilde{M}f$ has no local maxima. More specifically, they observed that every local maximum x_0 of $\tilde{M}f$ satisfies that $\tilde{M}f(x_0) = f(x_0)$. In our case, we are going to need the general version of this property, as the statement with local maxima of $M^\alpha f(x_0)$ may not hold. It is much more of an informal principle than a property itself, but we shall state it nonetheless, for the sake of stressing its impact on our methods.

Property 1. We say that an operator \mathcal{O} defined on the class of bounded variation functions has a good attachment at local maxima if, for every $f \in BV(\mathbf{R})$ and local maximum x_0 of $\mathcal{O}f$ over an interval (a, b) , with $\mathcal{O}f(x_0) > \max(\mathcal{O}f(a), \mathcal{O}f(b))$, then either $\mathcal{O}f(x_0) = |f(x_0)|$ or there exists an interval $(a, b) \supset I$ such that $\mathcal{O}f$ is constant on I and there is $y \in I$ such that $\mathcal{O}f(y) = |f(y)|$.

The intuition behind this principle is that, for such operators, one usually has that $\mathcal{V}(\mathcal{O}f) \leq \mathcal{V}(f)$, as skimming through the proofs in [1] suggests. This is, as one should expect, the main tool to prove Theorems 1 and 3.

3. Proof of Theorems 1 and 2

In what follows, let $f \in BV(\mathbf{R})$ have either $NORM(1)$ or $NORM(\alpha)$, where the specified normalization used will be stated in each context.

3.1. Analysis of maxima for M^α , $\alpha > \frac{1}{3}$. Here, we prove some major facts that will facilitate our work. Let then $[a, b]$ be an interval, and suppose that $M^\alpha f$ has a *strict* local maximum at $x_0 \in (a, b)$. That is, we suppose that $M^\alpha f(x_0)$ is maximal over $[a, b]$, with $M^\alpha f(x_0) > \max\{M^\alpha f(a), M^\alpha f(b)\}$. Suppose also that $M^\alpha f(x_0) = u(y, t)$, for some $(y, t) \in \{(z, s); |z - x_0| \leq \alpha s\}$, where we define the function $u: \mathbf{R} \times \mathbf{R}_+ \rightarrow \mathbf{R}_+$ as

$$u(y, t) = \frac{1}{2t} \int_{y-t}^{y+t} |f(s)| ds.$$

Such an assumption is possible, as we would otherwise have that either

- a sequence $(y, t) \rightarrow (x_0, 0)$ such that $\frac{1}{2t} \int_{y-t}^{y+t} |f(s)| ds \rightarrow M^\alpha f(x_0)$, which implies $|f(x_0)| = M^\alpha f(x_0)$ by the normalization;
- a sequence (y, t) with $t \rightarrow \infty$ such that $\frac{1}{2t} \int_{y-t}^{y+t} |f(s)| ds \rightarrow M^\alpha f(x_0)$, which implies that either $M^\alpha f(a)$ or $M^\alpha f(b)$ is bigger than or equal to $M^\alpha f(x_0)$, a contradiction.

As $M^\alpha f(x_0) = u(y, t)$, we have that $M^\alpha f(x_0) = M^\alpha f(y)$. Moreover, we claim that

$$[y - \alpha t, y + \alpha t] \subset (a, b).$$

If this did not hold, then $[y - \alpha t, y + \alpha t] \ni$ either a or b . Let us suppose, without loss of generality, that $a \in [y - \alpha t, y + \alpha t]$. But then

$$a \geq y - \alpha t \implies |a - y| \leq \alpha t \implies M^\alpha f(a) \geq M^\alpha f(y) \geq M^\alpha f(x_0),$$

a contradiction to our assumption of strictness of the maximum. This implies that, as for any $z \in [y - \alpha t, y + \alpha t] \implies |z - y| \leq \alpha t$, the maximal function $M^\alpha f$ is

constant over the interval $[y - \alpha t, y + \alpha t]$. Moreover, we have that the supremum of

$$u(z, s), \text{ for } (z, s) \in \bigcup_{z' \in [y - \alpha t, y + \alpha t]} \{(z'', s'') : |z'' - z'| \leq \alpha s''\} =: C(y, \alpha, t),$$

is attained for $(z, s) = (y, t)$.

By standard techniques, we shall assume $f \geq 0$ from now on. Our next step is then to find a subinterval I of $[y - \alpha t, y + \alpha t]$ and a $R = R(y, \alpha, t)$ such that, over this interval I , it holds that

$$M_{\equiv R}^1 f \equiv M^\alpha f.$$

Here, $M_{\equiv R}^1$ stands for the operator $\sup_{x \in I, |I| \leq 2R} \int_I |f(s)| ds$. For that, we need to investigate a few properties of the restricted maximal function $M_{\equiv R}^1 f$. This is done via the following:

Lemma 1. (Boundary Projection Lemma) *Let $(y, t) \in \mathbf{R} \times \mathbf{R}_+$. Let us denote*

$$\frac{1}{2t} \int_{y-t}^{y+t} f(s) ds = u(y, t).$$

If $(y, t) \in \{(z, s); 0 < |z - x| \leq s\}$, then

$$u(y, t) \leq \max \left\{ u \left(\frac{x + y - t}{2}, \frac{x - y + t}{2} \right), u \left(\frac{x + y + t}{2}, \frac{y - x + t}{2} \right) \right\}.$$

Proof. The proof is simple: in case $|x - y| = t$, then the inequality is trivial, so we assume $|x - y| < t$. We then just have to write

$$\begin{aligned} u(y, t) &= \frac{1}{2t} \int_{y-t}^{y+t} f(s) ds = \frac{1}{2t} \int_{y-t}^x f(s) ds + \frac{1}{2t} \int_x^{y+t} f(s) ds \\ &= \frac{x - y + t}{2t} \frac{1}{x - y + t} \int_{y-t}^x f(s) ds + \frac{y - x + t}{2t} \frac{1}{y - x + t} \int_x^{y+t} f(s) ds \\ &= \frac{x - y + t}{2t} u \left(\frac{x + y - t}{2}, \frac{x - y + t}{2} \right) + \frac{y - x + t}{2t} u \left(\frac{x + y + t}{2}, \frac{y - x + t}{2} \right) \\ &\leq \max \left\{ u \left(\frac{x + y - t}{2}, \frac{x - y + t}{2} \right), u \left(\frac{x + y + t}{2}, \frac{y - x + t}{2} \right) \right\}. \end{aligned}$$

□

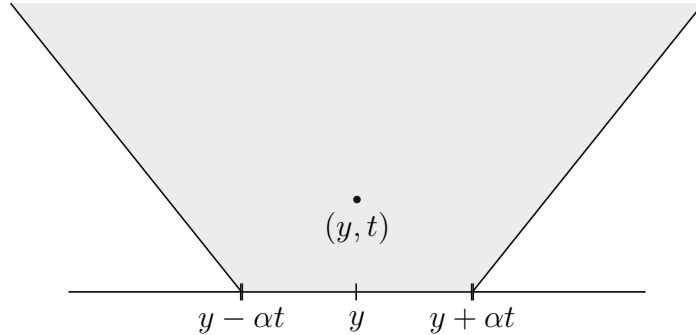


Figure 1. The region $C(y, \alpha, t)$.

Let $M_{r,A} f(x) = \sup_{0 \leq t \leq 2A} \frac{1}{t} \int_x^{x+t} |f(s)| ds$, and define $M_{l,A} f$ in a similar way, there the subindexes “ r ” and “ l ” represent, respectively, “right” and “left”. These operators are present out of the context of sharp regularity estimates for maximal functions, just like in [13]. In the realm of regularity of maximal function, though,

the first to introduce this notion was Tanaka [15]. As a corollary, we may obtain the following:

Corollary 1. For every $f \in L^1_{\text{loc}}(\mathbf{R})$, it holds that

$$\sup_{|z-x|+|t-R|\leq R} u(z, t) \leq \max\{M_{r,R}f(x), M_{l,R}f(x)\}.$$

From this last corollary, we are able to establish the following important—and, as far as the author knows, new—lemma:

Lemma 2. For every $f \in L^1_{\text{loc}}(\mathbf{R})$, we have also that

$$M^1_{\equiv R}f(x) = \sup_{|z-x|+|t-R|\leq R} u(z, t).$$

Proof. From Corollary 1, we have that

$$\begin{aligned} M^1_{\equiv R}f(x) &:= \sup_{|x-y|\leq t\leq R} u(y, t) \leq \sup_{|z-x|+|t-R|\leq R} u(z, t) \\ &\leq \max\{M_{r,R}f(x), M_{l,R}f(x)\} \leq M^1_{\equiv R}f(x). \end{aligned}$$

That is exactly what we wanted to prove. \square

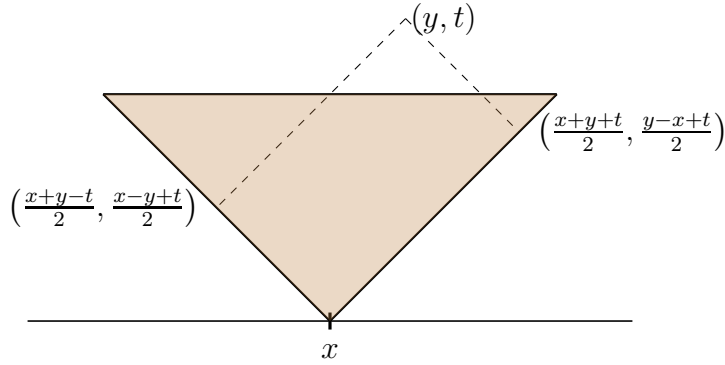


Figure 2. Illustration of Lemma 1: the points $(\frac{x+y-t}{2}, \frac{x-y+t}{2})$ and $(\frac{x+y+t}{2}, \frac{y-x+t}{2})$ are the projections of (y, t) over the lines $t = y - x$ and $t = y + x$, respectively.

Let R be then selected such that $\frac{t}{2} < R$ and $R(1 - \alpha) < \alpha t$. For $\alpha > \frac{1}{3}$ this is possible. This condition is *exactly* the condition so that the region

$$\{(z, t') : |z - y| + |t' - R| \leq R\} \subset C(y, \alpha, t).$$

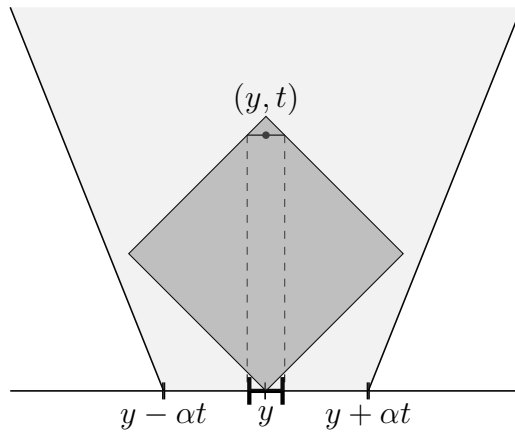


Figure 3. In the figure, the **dark gray** area represents the region that our Lemma gives, for some $\frac{1}{2}t < R < \frac{\alpha}{1-\alpha}t$, and the **black** interval is one in which $M^\alpha f = M^1_{\equiv R}f \equiv M^\alpha f(y)$.

Now we are able to end the proof: if I is a sufficiently small interval around y , then, by continuity, it must hold true that the regions

$$\{(z, t') : |z - y'| + |t' - R| \leq R\} \subset C(y, \alpha, t),$$

for all $y' \in I$. This is our desired interval for which $M^\alpha f \equiv M_{\equiv R}^1 f$. But we already know that, from [1, Lemma 3.6], $M_{\equiv R}^1 f$ satisfies a *stronger* property of control of maxima. Indeed, in order to fit it into the context of Aldaz and Pérez Lázaro, we note that, by adopting $NORM(1)$, f becomes automatically upper semicontinuous, and also $f \leq M_{\equiv R}^1 f$ everywhere. In particular, we know that, if $M_{\equiv R}^1 f$ is constant in an interval, then it must be *equal* to the function f at *every* point of that interval. But this is exactly our case, as we have already noticed that $M^\alpha f$ is constant on $[y - \alpha t, y + \alpha t]$, and therefore also on I . This implies, in particular, that

$$M^\alpha f(y) = M_{\equiv R}^1 f(y) = f(y),$$

which concludes our analysis of local maxima.

3.2. Proof of $\mathcal{V}(M^\alpha f) \leq \mathcal{V}(f)$, for $\alpha \geq \frac{1}{3}$. We remark, before beginning, that this strategy, from now on, is essentially the same as the one contained in [1]. We will, therefore, assume that $f \geq 0$ throughout.

First, we say that a function $g: I \rightarrow \mathbf{R}$ is *V-shaped* if there exists a point $c \in I$ such that

$$g|_{(I, c)} \text{ is non-increasing and } g|_{(c, r(I))} \text{ is non-decreasing.}$$

We then present two different proofs of this inequality, the first using an approximation and the second working directly with general *BV* functions.

First proof. For this, we will suppose that f has $NORM(1)$ as normalization. One can easily check then that $M^\alpha f \in C(\mathbf{R})$ for f a Lipschitz function. In fact, it is not difficult to show also that $M^\alpha f$ is continuous at x if f is continuous at x . Moreover, we may prove an additional property about it that will help us later:

Lemma 3. (Reduction to the Lipschitz case) *Suppose we have that*

$$\mathcal{V}(M^\alpha f) \leq \mathcal{V}(f), \quad \forall f \in BV(\mathbf{R}) \cap \text{Lip}(\mathbf{R}).$$

Then the same inequality holds for all Bounded Variation functions, that is,

$$\mathcal{V}(M^\alpha f) \leq \mathcal{V}(f), \quad \forall f \in BV(\mathbf{R}).$$

Proof. Let $\varphi \in \mathcal{S}(\mathbf{R})$ be a smooth, nonnegative function such that $\int_{\mathbf{R}} \varphi(t) dt = 1$, $\text{supp}(\varphi) \subset [-1, 1]$, φ is even and non-increasing on $[0, 1]$. Call $\varphi_\varepsilon(x) = \frac{1}{\varepsilon} \varphi(\frac{x}{\varepsilon})$. We define then $f_\varepsilon(x) = f * \varphi_\varepsilon(x)$. Notice that these functions are all Lipschitz (in fact, smooth) functions. Moreover, by standard theorems on approximate identities, we have that $f_\varepsilon(x) \rightarrow f(x)$ almost everywhere. Therefore, assuming the theorem to hold for Lipschitz functions, we have:

$$\begin{aligned} \mathcal{V}(M^\alpha f_\varepsilon) &\leq \mathcal{V}(f_\varepsilon) = \sup_{x_1 < \dots < x_N} \sum_{i=1}^{N-1} |f_\varepsilon(x_{i+1}) - f_\varepsilon(x_i)| \\ &\leq \int_{\mathbf{R}} \varphi_\varepsilon(t) \sup_{x_1 < \dots < x_N} \left(\sum_{i=1}^{N-1} |f(x_{i+1} - t) - f(x_i - t)| \right) dt \leq \mathcal{V}(f). \end{aligned}$$

Thus, it suffices to prove that

$$(3) \quad \limsup_{y \rightarrow x} M^\alpha f(y) \geq \limsup_{\varepsilon \rightarrow 0} M^\alpha f_\varepsilon(x) \geq \liminf_{\varepsilon \rightarrow 0} M^\alpha f_\varepsilon(x) \geq \liminf_{y \rightarrow x} M^\alpha f(y), \quad \forall x \in \mathbf{R},$$

as then

$$(4) \quad \mathcal{V}(M^\alpha f) = \mathcal{V}(\liminf_{\varepsilon \rightarrow 0} M^\alpha f_\varepsilon) = \mathcal{V}(\limsup_{\varepsilon \rightarrow 0} M^\alpha f_\varepsilon) = \mathcal{V}(\lim_{j \rightarrow \infty} M^\alpha f_{\varepsilon_j}) \leq \mathcal{V}(f),$$

which follows from the following

Lemma 4. *Let g_ε, g be bounded functions such that*

$$(5) \quad \limsup_{y \rightarrow x} g(y) \geq \limsup_{\varepsilon \rightarrow 0} g_\varepsilon(x) \geq \liminf_{\varepsilon \rightarrow 0} g_\varepsilon(x) \geq \liminf_{y \rightarrow x} g(y),$$

for all $x \in \mathbf{R}$. Assume that each g_ε is continuous, $\forall \varepsilon > 0$, and that g is continuous up to a countable set, in which the lateral limits $g(x-), g(x+)$ still exist, and it holds that $g(x) \in [\min\{g(x-), g(x+)\}, \max\{g(x-), g(x+)\}]$. Then

$$\mathcal{V}(g) = \mathcal{V}(\limsup_{\varepsilon \rightarrow 0} g_\varepsilon) = \mathcal{V}(\liminf_{\varepsilon \rightarrow 0} g_\varepsilon).$$

Proof of Lemma 4. Let $g_1 = \limsup_{\varepsilon \rightarrow 0} g_\varepsilon$, $g_2 = \liminf_{\varepsilon \rightarrow 0} g_\varepsilon$. We first prove that $\mathcal{V}(g)$ is less than both $\mathcal{V}(g_1)$, $\mathcal{V}(g_2)$. For that, fix a finite partition of the real line $\mathcal{P} = \{x_1 < \dots < x_N\}$. In order to estimate the variation $\mathcal{V}_{\mathcal{P}}(g)$, we need to divide into two cases: (i) if g is continuous at every x_i , then we let the partition remain as it is; (ii) if g is *not* continuous at a certain x_i , we then pick two points x'_i, x''_i such that g is continuous at *both* of them, $x_{i-1} < x'_i < x_i < x''_i < x_{i+1}$ and $g(x_i)$ lies between $g(x'_i)$ and $g(x''_i)$. The assumptions on g show that this is *always* possible. Add these new points to the partition \mathcal{P} and call the new one \mathcal{P}' . By the way we picked the points x'_i, x''_i , we see that the existence of the points x_i in the new partition is superfluous, and therefore we might think of \mathcal{P}' as consisting of the points of \mathcal{P} where g is continuous and the x'_i, x''_i . By (5), we see that the limit $\lim_{\varepsilon \rightarrow 0} g_\varepsilon$ in fact *exists* for each point of this \mathcal{P}' and is equal to g . From that,

$$\mathcal{V}_{\mathcal{P}}(g) \leq \mathcal{V}_{\mathcal{P}'}(g) = \mathcal{V}_{\mathcal{P}'}(g_1) = \mathcal{V}_{\mathcal{P}'}(g_2).$$

The desired inequality then follows by taking supremum over all finite partitions of the real line. The reverse inequality consists then of applying the exact same strategy to g_0, g_1 , by noticing that, from (5), they satisfy the same assumptions as g . \square

The reader might notice one still has to prove that $M^\alpha f := g$ satisfies the assumptions in Lemma 4. Indeed, one straightforward way to do so is to use the result in subsection 5.1 to conclude that $M^\alpha f$, as a BV -function, must fulfill all the properties above. A proof without resorting to this result is however also possible, but we omit it for shortness.

Let us suppose, for the sake of a contradiction, that either the first or the third inequalities in 3 are not fulfilled. Therefore, we focus on the first inequality: suppose that there exists a real number x_0 , a sequence $\varepsilon_k \rightarrow 0$ and a positive real number $\eta > 0$ such that

$$M^\alpha f_{\varepsilon_k}(x_0) > (1 + 2\eta) \limsup_{y \rightarrow x_0} M^\alpha f(y).$$

By definition, there exists a sequence (y_k, r_k) with $|y_k - x_0| \leq \alpha r_k$ and

$$\int_{y_k - r_k}^{y_k + r_k} f_{\varepsilon_k}(s) ds > (1 + \eta) \limsup_{y \rightarrow x_0} M^\alpha f(y).$$

Case 1. Suppose $r_k \rightarrow 0$. By the way we normalized f , there is an interval $I \ni x_0$ such that $f(y) \leq (1 + \eta/4)f(x_0), \forall y \in I$. But then, by the support properties of φ and for k sufficiently large, we would have that $(1 + \eta/2)f(x_0) \geq M^\alpha f_{\varepsilon_k}(x_0)$, which is a contradiction, as $\limsup_{y \rightarrow x_0} M^\alpha f(y) \geq f(x_0)$.

Case 2. Let then $\inf_k r_k > 0$. Then, by Fubini's theorem and manipulations,

$$\begin{aligned} \int_{y_k-r_k}^{y_k+r_k} f_{\varepsilon_k}(s) ds &= \int_{y_k-r_k}^{y_k+r_k} \left(\int_{-\varepsilon_k}^{\varepsilon_k} \varphi_{\varepsilon_k}(t) f(s-t) dt \right) ds \\ &= \int_{-\varepsilon_k}^{\varepsilon_k} \varphi_{\varepsilon_k}(t) \left(\int_{y_k-r_k}^{y_k+r_k} f(s-t) ds \right) dt \leq \frac{r_k + \varepsilon_k}{r_k} M^\alpha f(x_0). \end{aligned}$$

This implies $r_k \leq \frac{\varepsilon_k}{\eta} \rightarrow 0$, which is another contradiction.

For the third inequality, we divide it once again: if $M^\alpha f(x_0) = u(y, t)$ for some $(y, t) \neq (x_0, 0)$, then, by L^1 convergence of approximate identities, one easily gets that $\liminf_{\varepsilon \rightarrow 0} M^\alpha f_\varepsilon(x_0) \geq M^\alpha f(x_0)$. If not, pick (y, t) such that $M^\alpha f(x_0) \leq u(y, t) + \frac{\delta}{2}$. Use then the L^1 convergence of approximate identities in the interval $(y-t, y+t)$. The reverse inequality, and therefore the lemma, is proved, as $M^\alpha f(x) \geq \liminf_{y \rightarrow x} M^\alpha f(y)$. \square

Our main claim is then the following:

Lemma 5. *Let $f \in \text{Lip}(\mathbf{R}) \cap \text{BV}(\mathbf{R})$. Then, over every interval of the set*

$$E_\alpha = \{x \in \mathbf{R} : M^\alpha f(x) > f(x)\} = \bigcup_{j \in \mathbf{Z}} I_j^\alpha,$$

it holds that $M^\alpha f$ is either monotone or V shaped in I_j^α .

Proof. The proof goes roughly as the first paragraph of the proof of Lemma 3.9 in [1]: let $I_j^\alpha = (l(I_j^\alpha), r(I_j^\alpha)) =: (l_j, r_j)$, and suppose that $M^\alpha f$ is *not* V shaped there. Therefore, there would be a maximal point $x_0 \in I_j^\alpha$ and an interval $J \subset I_j^\alpha$ such that $M^\alpha f$ has a *strict* local maximum at x_0 over J . Then, by the maxima analysis we performed, we see that we have reached a contradiction from this fact alone, as $J \subset E_\alpha$. We omit further details, as they can be found, as already mentioned, at [1, Lemma 3.9]. \square

We also need the following

Lemma 6. *If $f \in \text{BV}(\mathbf{R}) \cap \text{Lip}(\mathbf{R})$, then, for every (maximal) open interval $I_j^\alpha \subset E_\alpha$, we have that*

$$M^\alpha f(l(I_j^\alpha)) = f(l(I_j^\alpha)),$$

and an analogous identity holds for $r(I_j^\alpha)$.

The proof of this Lemma is straightforward, and we therefore skip it. To finalize the proof in this case for $\alpha > \frac{1}{3}$, we just notice that we can, in fact, bound the variation of $M^\alpha f$ *inside* every interval I_j^α . In fact, we have directly from the last claim that, in case $M^\alpha f$ is V shaped on I_j^α , then there exists $c_j \in I_j^\alpha$ such that $M^\alpha f$ is non-increasing on (l_j, c_j) and non-decreasing on (c_j, r_j) . We then calculate:

$$\begin{aligned} \mathcal{V}_{I_j^\alpha}(M^\alpha f) &= |M^\alpha f(l(I_j^\alpha)) - M^\alpha f(c_j)| + |M^\alpha f(r(I_j^\alpha)) - M^\alpha f(c_j)| \\ &\leq |f(l(I_j^\alpha)) - f(c_j)| + |f(r(I_j^\alpha)) - f(c_j)| \leq V_{I_j^\alpha}(f). \end{aligned}$$

The way to formally end the proof is the following: let $\mathcal{P} = \{x_1 < \dots < x_N\}$, and let $A := \{j \in \mathbf{N} : \exists x_i \in \mathcal{P} \cap I_j^\alpha\}$. Clearly, the index set A is finite. Moreover, there are at most two $j \in \mathbf{N}$ such that I_j^α is not a bounded interval. With this in mind, we refine the partition \mathcal{P} by adding to it the following points:

- If $j \in A$ and $M^\alpha f$ is *monotone* over I_j^α , then add l_j, r_j to the partition;
- If $j \in A$ and $M^\alpha f$ is *V shaped* over I_j^α , then add l_j, r_j and the point c_j to the partition.

Notice that this covers only the case of I_j^α being bounded. For the case of unbounded intervals, one might proceed in a similar fashion, by adding directly a ‘‘sufficiently large’’ point in each interval instead of the (missing) endpoint. Notice this strategy allows us to deal with unbounded intervals over which the maximal function is either monotone or V shaped: indeed, if $\lim_{x \rightarrow -\infty} f(x) = L$, $\lim_{x \rightarrow +\infty} f(x) = M$, and we suppose $L \geq M$ (without loss of generality), then:

- (A) if there is an interval $I_{j_0}^\alpha = (-\infty, r_{j_0})$, then it is easy to prove that $\lim_{x \rightarrow -\infty} M^\alpha f(x) = L$. Therefore, by ‘choosing a point’ x' sufficiently large, we see that $|M^\alpha f(x') - f(x')|$ has to be small, and the argument follows;
- (B) if there is an interval $I_{j_1}^\alpha = (l_{j_1}, +\infty)$, then a standard argument proves that $\lim_{x \rightarrow +\infty} M^\alpha f(x) = \frac{1+\alpha}{2}L + \frac{1-\alpha}{2}M$. But, also by a standard argument, one proves that $M^\alpha f(x) \geq \frac{1+\alpha}{2}L + \frac{1-\alpha}{2}M$ for each $x \in \mathbf{R}$. This shows that $M^\alpha f$ cannot be ‘‘strictly’’ V shaped over $I_{j_1}^\alpha$. As we know from Lemma 6 that $M^\alpha f(l_{j_1}) = f(l_{j_1})$ and that $M^\alpha f$ has to be *non-increasing* over $I_{j_1}^\alpha$, choosing a ‘sufficiently large’ point only helps us, as

$$\mathcal{V}_{I_{j_1}^\alpha}(M^\alpha f) = f(l_{j_1}) - \left(\frac{1+\alpha}{2}L + \frac{1-\alpha}{2}M \right) \leq f(l_{j_1}) - M \leq \mathcal{V}_{I_{j_1}^\alpha}(f).$$

Call the new partition obtained by this procedure \mathcal{P}' . By the calculation above and the fact that, if $f \in \text{Lip}(\mathbf{R}) \implies M^\alpha f \geq f$ everywhere, and in particular $M^\alpha f = f$ at $\mathbf{R} \setminus E_\alpha$, one obtains that

$$\mathcal{V}_{\mathcal{P}}(M^\alpha f) \leq \mathcal{V}_{\mathcal{P}'}(M^\alpha f) \leq \mathcal{V}(f).$$

By taking a supremum over all partitions, we finish the result for $\alpha > \frac{1}{3}$. On the other hand, it is straight from the definition that

$$\beta \leq \alpha \implies \frac{\beta}{\alpha} M^\alpha f \leq M^\beta f \leq M^\alpha f.$$

This implies that, for a partition \mathcal{P} as above,

$$\sum_{i=1}^{N-1} |M^{\frac{1}{3}} f(x_{i+1}) - M^{\frac{1}{3}} f(x_i)| \leq \lim_{\alpha \searrow \frac{1}{3}} \sum_{i=1}^{N-1} |M^\alpha f(x_{i+1}) - M^\alpha f(x_i)| \leq \mathcal{V}(f).$$

The theorem follows, again, as before.

Second proof. For this part, we assume that f has $NORM(\alpha)$ normalization. The argument here is morally the same, with just a couple of minor modifications. Therefore, this section might seem a little bit superfluous now, even though its reason of being is going to be shown while we characterize the extremizers.

Claim 1. *Let $E_\alpha = \{x \in \mathbf{R}: M^\alpha f(x) > f(x)\}$. This set is open for any $f \in BV(\mathbf{R})$ normalized with $NORM(\alpha)$ and therefore can be decomposed as*

$$E_\alpha = \bigcup_{j \in \mathbf{Z}} I_j^\alpha,$$

where each I_j^α is an interval. Furthermore, the restriction of $M^\alpha f$ to each of those intervals is either a monotone function or a V shaped function with a minimum at $c_j \in I_j^\alpha$. Moreover, $M^\alpha f(c_j) < \min\{M^\alpha f(l(I_j^\alpha)), M^\alpha f(r(I_j^\alpha))\}$.

Proof of the claim. The claim seems quite sophisticated, but its proof is simple, once one has done the maxima analysis we have done. The fact that E_α is open is

easy to see. In fact, let $x_0 \in E_\alpha$. By the lower semicontinuity of $M^\alpha f$ at x_0 and the fact that we normalized f with $NORM(\alpha)$,

$$\liminf_{z \rightarrow x_0} M^\alpha f(z) \geq M^\alpha f(x_0) > \limsup_{z \rightarrow x_0} f(z).$$

This shows that, for z close to x_0 , the strict inequality should still hold, as desired.

The second part follows in the same fashion as the proof of Lemma 5, and we therefore omit it. \square

To finish the proof of the fact that $\mathcal{V}_{I_j^\alpha}(M^\alpha f) \leq \mathcal{V}_{I_j^\alpha}(f)$ also in this case we just need one more lemma:

Lemma 7. *For every (maximal) open interval $I_j^\alpha \subset E_\alpha$ we have that*

$$M^\alpha f(l(I_j^\alpha)) = f(l(I_j^\alpha)),$$

and an analogous identity holds for $r(I_j^\alpha)$.

This is, just like Lemma 6, direct from the definition and the maximality of the intervals I_j^α . The conclusion in this case uses Lemma 7 in a direct fashion, combined with the strategy for the first proof: namely, the estimate

$$\begin{aligned} \mathcal{V}_{I_j^\alpha}(M^\alpha f) &\leq |M^\alpha f(l(I_j^\alpha)) - M^\alpha f(c_j)| + |M^\alpha f(r(I_j^\alpha)) - M^\alpha f(c_j)| \\ &\leq |f(l(I_j^\alpha)) - f(c_j)| + |f(r(I_j^\alpha)) - f(c_j)| \leq V_{I_j^\alpha}(f) \end{aligned}$$

still holds, by Lemma 7 and by the fact that $c_j \in I_j^\alpha$. This finishes finally the second proof of Theorem 1.

3.3. Absolute continuity on the detachment set. We prove briefly the fact that, for $f \in W^{1,1}(\mathbf{R})$, then we have that $M^\alpha f \in W_{\text{loc}}^{1,1}(\mathbf{R})$ for any $1 > \alpha > 0$, as the case $\alpha = 0$ has been dealt with by Kurka [8], in Corollary 1.4.

Indeed, let

$$E_{\alpha,k} = \{x \in E_\alpha : M^\alpha f(x) = \sup_{(y,t): |y-x| \leq \alpha t, t \geq \frac{1}{2k}} \frac{1}{2t} \int_{y-t}^{y+t} |f(s)| ds\}.$$

Then we see that $E_\alpha = \bigcup_{k \geq 1} E_{\alpha,k}$. Moreover, for $x, y \in E_{\alpha,k}$, let then (y_1, t_1) have this property for x . Suppose also, without loss of generality, that $y \geq x$ and $M^\alpha f(x) > M^\alpha f(y)$. By assuming that $y > y_1 + \alpha t_1$ —as otherwise $M^\alpha f(x) \leq M^\alpha f(y)$ —, we have that

$$\begin{aligned} M^\alpha f(x) - M^\alpha f(y) &\leq \frac{1}{2t_1} \int_{y_1-t_1}^{y_1+t_1} |f(s)| ds - u \left(\frac{y + \alpha y_1 - \alpha t_1}{1 + \alpha}, \frac{y - y_1 + t_1}{1 + \alpha} \right) \\ &\leq \frac{\frac{2}{1+\alpha}(y - y_1) - \frac{2\alpha}{1+\alpha}t_1}{2t_1 \cdot \frac{2}{1+\alpha}(y - y_1 + t_1)} \int_{y_1-t_1}^{y_1+t_1} |f(s)| ds \\ &\leq \frac{\frac{2}{1+\alpha}|y - x|}{\frac{2}{1+\alpha}(y - y_1 + t_1)} \|f\|_\infty \leq \frac{|x - y|}{(1 + \alpha)t_1} \|f\|_\infty \leq \frac{2}{1 + \alpha} k |x - y| \|f\|_\infty. \end{aligned}$$

This shows that $M^\alpha f$ is Lipschitz continuous with constant $\leq \frac{2}{1+\alpha} k \|f\|_\infty$ on each $E_{\alpha,k}$. The proof of the asserted fact, however, follows from this, by using the well-known Banach–Zarecki lemma:

Lemma 8. (Banach–Zarecki) *A function $g: I \rightarrow \mathbf{R}$ is absolutely continuous if and only if the following conditions hold simultaneously:*

- (A) g is continuous;
- (B) g is of bounded variation;

(C) $g(S)$ has measure zero for every set $S \subset I$ with $|S| = 0$.

In fact, let S be then a null-measure set on the real line and $f \in W^{1,1}(\mathbf{R})$ —which implies that $M^\alpha f \in C(\mathbf{R})$ and, by the comments in subsection 5.1, $\forall \alpha > 0$ and $f \in W^{1,1}(\mathbf{R})$, $M^\alpha f \in BV(\mathbf{R})$ —, and let us invoke [1, Lemma 3.1]:

Lemma 9. *Let $f: I \rightarrow \mathbf{R}$ be a continuous function. Let also $E \subset \{x \in I: |\overline{D}f(x)| := |\limsup_{h \rightarrow 0} \frac{f(x+h)-f(x)}{h}| \leq k\}$. Then*

$$m^*(f(E)) \leq km^*(E),$$

where $m(S) = |S|$ stands for the Lebesgue measure of S .

It is easy to see that the maximal functions $M^\alpha f$ are, in fact, *continuous* on the open set E_α . Thus, we may use Lemmas 8 and 9 in each of the connected components of E_α :

$$|M^\alpha f(S \cap I_j^\alpha)| \leq \sum_{k \geq 1} |M^\alpha f(S \cap E_{\alpha,k} \cap I_j^\alpha)| = 0,$$

where we used that $M^\alpha f$ is Lipschitz over each $E_{\alpha,k}$. But this implies that

$$|M^\alpha f(S)| \leq |M^\alpha f(S \cap E_\alpha^c)| + \sum_{j \in \mathbf{Z}} |M^\alpha f(S \cap I_j^\alpha)| = |f(S \cap E_\alpha^c)| = 0,$$

by Lemma 8 and the fact that $f \in W_{\text{loc}}^{1,1}(\mathbf{R})$. This finishes this part of the analysis.

3.4. Sharpness of the inequality and extremizers. In this part, we prove that the best constant in such inequalities is indeed 1, and characterize the extremizers for such. Namely, we mention promptly that the inequality *must* be sharp, as $f = \chi_{(-1,0)}$ realizes equality. It is easy to see that, to do so, we may assume that f still has $NORM(\alpha)$ normalization.

Claim 2. *Let $f \in BV(\mathbf{R})$ normalized as before satisfy $\mathcal{V}(f) = \mathcal{V}(M^\alpha f)$. If we decompose $E_\alpha = \cup_j I_j^\alpha$, where each of the I_j^α is open and maximal, then*

$$\mathcal{V}_{I_j^\alpha}(f) = \mathcal{V}_{I_j^\alpha}(M^\alpha f).$$

Proof. Let \mathcal{P}, \mathcal{Q} be two finite partitions of \mathbf{R} such that

$$(6) \quad \begin{cases} \mathcal{V}(M^\alpha f) \leq \mathcal{V}_{\mathcal{P}}(M^\alpha f) + \frac{\varepsilon}{20}, \\ \mathcal{V}(f) \leq \mathcal{V}_{\mathcal{Q}}(f) + \frac{\varepsilon}{20}. \end{cases}$$

Now let the mutual refinement of those be $\mathcal{S} = \mathcal{P} \cup \mathcal{Q}$. We consider the intersection $\mathcal{S} \cap E_\alpha$: if the finite set $A := \{j: I_j^\alpha \cap \mathcal{S} \neq \emptyset\}$ satisfies that

$$(7) \quad \sum_{j \in A} \mathcal{V}_{I_j^\alpha}(f) \geq \sum_{j \in \mathbf{N}: I_j^\alpha \neq \emptyset} \mathcal{V}_{I_j^\alpha}(f) - \frac{\varepsilon}{20},$$

then keep the partition as it is before advancing. If not, then add to \mathcal{S} finitely many points, all of them contained in intervals of the form $\overline{I_j^\alpha}$, such that inequality 7 holds. Call this new partition \mathcal{S} again, as it still satisfies the inequalities 6.

We finally add some other points to the partition \mathcal{S} : If $j \notin A$, do not add any points from the interval. If $j \in A$, then do the following:

(A) As $f = M^\alpha f$ on the boundary of an interval I_j^α , we add to the collection both endpoints $r(I_j^\alpha), l(I_j^\alpha)$.

- (B) If $M^\alpha f$ is V shaped over the interval I_j^α , then there is a point c_j such that $M^\alpha f$ is non-increasing on (l_j, c_j) and non-decreasing on (c_j, r_j) . Add such a point to our partition.
- (C) If $\mathcal{V}_{I_j^\alpha}(f) > \mathcal{V}_{\{x_i \in \mathcal{S}: x_i \in I_j^\alpha\}}(f) + \frac{\varepsilon}{2^{20|j|}}$, then add finitely many points to the partition to make the reverse inequality hold (here, $\mathcal{V}_{\{x_i \in \mathcal{S}: x_i \in A\}}(g)$ stands for the variation along the finite partition composed solely by elements in the set A).

It is easy to see that, if we denote by \mathcal{S}' the partition obtained by the prescribed procedure above, then, as $\mathcal{V}(f) = \mathcal{V}(M^\alpha f)$ and $f = M^\alpha f$ on $\mathbf{R} \setminus E_\alpha$,

$$|\mathcal{V}_{\mathcal{S}' \cap E_\alpha}(f) - \mathcal{V}_{\mathcal{S}' \cap E_\alpha}(M^\alpha f)| \leq 2\varepsilon,$$

which then implies that, by the considerations above,

$$\begin{aligned} \sum_{j \in \mathbf{Z}} \mathcal{V}_{I_j^\alpha}(f) - \frac{\varepsilon}{4} &\leq \sum_{j \in A} \mathcal{V}_{I_j^\alpha}(f) \leq \sum_{j \in A} \mathcal{V}_{\{x_i \in \mathcal{S}': x_i \in I_j^\alpha\}}(f) + \varepsilon \\ (8) \qquad \qquad \qquad &\leq \sum_{j \in A} \mathcal{V}_{\{x_i \in \mathcal{S}': x_i \in I_j^\alpha\}}(M^\alpha f) + 3\varepsilon \leq \sum_{j \in \mathbf{Z}} \mathcal{V}_{I_j^\alpha}(M^\alpha f) + 3\varepsilon. \end{aligned}$$

As ε was arbitrary, comparing the first and last terms above and looking back to our proof that in each of the I_j^α the variation of f controls that of the maximal function, we conclude that, for each $j \in \mathbf{Z}$,

$$(9) \qquad \qquad \qquad \mathcal{V}_{I_j^\alpha}(f) = \mathcal{V}_{I_j^\alpha}(M^\alpha f).$$

This finishes the proof of this claim. \square

Claim 3. *Let f, I_j^α as above. Then f and $M^\alpha f$ are monotone in the closure $\overline{I_j^\alpha}$.*

Proof. Suppose first that $M^\alpha f$ is *not* monotone there. Then it must be V shaped on $\overline{I_j^\alpha}$, and then, by Claim 2, we see that the only possibility for that to happen is if $M^\alpha f(c_j) = f(c_j)$, $c_j \in I_j^\alpha$. This is clearly not possible by the definition of I_j^α , and we reach a contradiction.

Suppose now that f is *not* monotone over $\overline{I_j^\alpha}$. As $\mathcal{V}_{I_j^\alpha}(f) = \mathcal{V}_{I_j^\alpha}(M^\alpha f)$ by Claim 2, and $\mathcal{V}_{I_j^\alpha}(M^\alpha f) = |f(r_j) - f(l_j)|$, then it is easy to see that, no matter what configuration of non-monotonicity we have, it yields a contradiction with the equality for the variations over the interval I_j^α . We skip the details, for they are routinely verified. \square

Remark 1. Note that this last claim proves also that, if I_j^α is *bounded*, f is *non-decreasing* over it and l_j is its left endpoint, then $f(l_j-) \leq f(l_j+)$, as otherwise we would arrive at a contradiction with the fact that $\mathcal{V}_{I_j^\alpha}(f) = \mathcal{V}_{I_j^\alpha}(M^\alpha f)$. An analogous statement holds for the right endpoint, and analogous conclusions if f is *non-increasing* instead of non-decreasing over the interval.

Next, we suppose without loss of generality that the function f is non-decreasing on $\overline{I_j^\alpha}$, as the other case is completely analogous.

Claim 4. *Such an f is, in fact, non-decreasing on $(-\infty, r(I_j^\alpha)]$.*

Proof. Our proof of this fact will go by contradiction: First, let $a_j = \inf\{t \in \mathbf{R}; f \text{ is non-decreasing in } [t, r(I_j^\alpha)]\}$, and define $b_j < a_j$ such that the minimum of f in $[b_j, r_j]$ happens *inside* (b_j, r_j) . Of course, such a minimum need not happen at a point, but it surely does happen at a *lateral limit* of a point.

Subclaim 1. $M^\alpha f(a_j) = f(a_j)$ and $f(a_j-) = f(a_j+)$.

Proof. If $M^\alpha f(a_j) > f(a_j)$, then there exists a maximal open interval $E_\alpha \supset J'_j \ni a_j$, and, as we proved before, f must be monotone in such an interval. By the definition of a_j , we see that, at least on $(a_j, r(I_j^\alpha)) \cap \overline{J'_j} =: K_j$, the function f has to be non-decreasing. If f is non-constant in K_j , then, by maximality and Claim 3, we see that f is non-decreasing in J'_j , which contradicts the choice of a_j . Then f has to be constant in K_j and, therefore, non-increasing in J'_j . We wish to show that this cannot happen, so that we conclude the desired equality. We seek to contradict Lemma 7, in the sense that we wish to prove that, actually, $M^\alpha f(r(J'_j)) > f(r(J'_j))$. Before we start doing so, we notice that, as f is non-decreasing (and non-constant) on $(a_j, r(I_j^\alpha)]$, and non-increasing on $\overline{J'_j}$, then $r(J'_j) \in (l(J'_j), r(I_j^\alpha))$ and f attains a minimum over $[l(J'_j), r(I_j^\alpha)]$ at $r(J'_j)$.

We consider two cases: if $|r(J'_j) - l(J'_j)| \leq |r(J'_j) - r(I_j^\alpha)|$, then

$$M^\alpha f(r(J'_j)) \geq \int_{l(J'_j)}^{2r(J'_j) - l(I_j^\alpha)} f > f(r(J'_j)),$$

where the strict inequality comes from the facts that (i) $[l(J'_j), 2r(J'_j) - l(I_j^\alpha)] \subset [l(J'_j), r(I_j^\alpha)]$; (ii) f has a local minimum at $r(J'_j)$ in $[l(J'_j), r(I_j^\alpha)]$; (iii) f is not constant over the whole interval J'_j (as otherwise it would yield a contradiction to the definition of a_j). If, on the other hand, $|r(J'_j) - l(J'_j)| > |r(J'_j) - r(I_j^\alpha)|$, we may consider, for $\delta > 0$ sufficiently small,

$$M^\alpha f(r(J'_j)) \geq \int_{2r(J'_j) - r(I_j^\alpha) - \delta}^{r(I_j^\alpha) + \delta} f > f(r(J'_j)),$$

as (i) $[2r(J'_j) - r(I_j^\alpha) - \delta, r(I_j^\alpha)] \subset [l(J'_j), r(I_j^\alpha)]$; (ii) f has a local minimum at $r(J'_j)$ in $[l(J'_j), r(I_j^\alpha)]$; (iii) if f is constant on $(r(J'_j), r(I_j^\alpha))$, then, for $r(I_j^\alpha) + \delta > p > r(I_j^\alpha)$, it holds that $f(p) > f(r(J'_j))$, as long as we choose $\delta > 0$ to be small enough (this holds because f is normalized). As we covered the only two possible cases, we arrive at a contradiction, namely, that $M^\alpha f(r(J'_j)) > f(r(J'_j))$. This finishes the first part of our subclaim.

Now for the second equality: if it were not true, then a_j would be, again, one of the endpoints of a maximal interval $J_j \subset E_\alpha$. If a_j is the left-endpoint, then it means that $f(a_j-) > f(a_j+)$. But this is a contradiction, as f then must be non-decreasing on J_j , and therefore we would again have that $\mathcal{V}_{J_j}(f) > \mathcal{V}_{J_j}(M^\alpha f)$. Therefore, a_j is the right endpoint, and also $f(a_j-) < f(a_j+)$. At the present moment an analysis as in Remark 1 is already available, and thus we conclude that f shall be non-decreasing on J_j , which is again a contradiction to the definition of J_j . \square

We must prove yet another fact that will help us:

Subclaim 2. *Let*

$$\mathcal{D} = \{x \in (b_j, r_j) : \min(f(x-), f(x+)) \text{ attains the minimum in } (b_j, r_j)\}.$$

Then there exists $d \in \mathcal{D}$ such that $f(d-) = f(d+)$ and $M^\alpha f(d) = f(d)$.

Proof. If $a_j \in \mathcal{D}$, then our assertion is proved by Subclaim 1. If not, then $\mathcal{D} \subset (b_j, a_j)$. In this case, pick any point d_0 in this intersection.

Case 1. $f(d_0+) = f(d_0-)$. In this case, there is nothing left undone if $f(d_0) = M^\alpha f(d_0)$. Otherwise, we would have that $M^\alpha f(d_0) > f(d_0)$, and then there would be an interval $E_\alpha \supset J_0 \ni d_0$. By the fact that *all* the points in \mathcal{D} must lie in (b_j, a_j) , and that f is monotone on J_0 , we see automatically that either $f(b_j) \leq f(d_0)$, a

contradiction, or the right endpoint of J_0 satisfies $f(r(J_0)) \leq f(d_0)$. By the definition of d_0 , this inequality has to be an equality, and also f must be continuous at $r(J_0)$, by the argument of Remark 1. As an endpoint of a maximal interval $J_0 \subset E_\alpha$, we have then $M^\alpha f(r(J_0)) = f(r(J_0))$.

Case 2. $f(d_{0+}) > f(d_{0-})$. It is easy to see that, in this case, there is an open interval $J \subset E_\alpha$ such that either $J \ni d_0$ or d_0 is its right endpoint. In either case, we see that f must be non-decreasing over this interval J , and let again l_0 be its left endpoint. As we know, $l_0 \in \mathcal{D}$ again, $l_0 \in (b_j, r_j)$ and, by Remark 1, we must have that $f(l_{0-}) = f(l_{0+})$. Of course, by being the endpoint we have automatically again that $M^\alpha f(l_0) = f(l_0)$. This concludes again this case, and therefore the proof of the subclaim. \square

The concluding argument for the proof of the Claim 4 goes as follows: let d be the point from Subclaim 2. Then we must have that

$$f(d) = M^\alpha f(d) \geq Mf(d) \geq \int_{d-\delta}^{d+\delta} f.$$

For small δ , it is easy to get a contradiction from that. Indeed, by the properties of the interval $(b_j, r_j]$ one can ensure that it is only needed to analyze $\delta \leq |d - b_j|$. The details are omitted.

This contradiction came from the fact that we supposed that $a_j > -\infty$, and our claim is established. \square

Now we finish the proof: If $M^\alpha f \leq f$ always, we get to the case of a *superharmonic function*, i.e., a function which satisfies $\int_{x-r}^{x+r} f(s) ds \leq f(x)$ for all $r > 0$. That is going to be handled in a while. If not, then we analyze the detachment set:

- (A) If all intervals in the detachment set are of *one single type*, that is, either all *non-increasing* or all *non-decreasing*, our function must then admit a point x_0 such that f is either non-decreasing on $(-\infty, x_0]$ (resp. non-decreasing on $[x_0, +\infty)$), and $f = M^\alpha f$ on $(x_0, +\infty)$ (resp. on $(-\infty, x_0)$).
- (B) If there is at least one interval of each type, then we must have an interval $[R, S]$ such that
 - f is non-decreasing on $(-\infty, R]$;
 - f is non-increasing on $[S, +\infty)$;
 - $f = M^\alpha f$ on (R, S) .

The analysis is then easily completed for every one of the cases above: If $f = M^\alpha f$ over an interval, then, as $M^\alpha f \geq Mf$, we conclude that f must be superharmonic there, where by ‘‘locally subharmonic’’ we mean a function that satisfies $f(x) \geq \int_{x-r}^{x+r} f(s) ds$ for all $0 \leq r \ll_x 1$. As superharmonic in one dimension coincides with concave, and concave functions have *at most* one global maximum, then the first case above gives that f is either monotone or has exactly one point x_1 such that it is exactly non-decreasing until a point x_1 , non-increasing after. The case of monotone functions is easily ruled out, as if $\lim_{x \rightarrow \infty} f = L$, $\lim_{x \rightarrow -\infty} f = M \implies \mathcal{V}(f) = |M - L|$, $\mathcal{V}(M^\alpha f) \leq \frac{|M-L|}{2}$. The second case is treated in the exact same fashion, and the result is the same: in the end, the only possible extremizers for this problem are functions f such that there is a point x_1 such that f is non-decreasing on $(-\infty, x_1)$, and f is non-increasing on $(x_1, +\infty)$. The theorem is then complete.

3.5. Proof of Theorem 2. We start our discussion by pointing out that the measure $d\mu = \delta_0 + \delta_1$ satisfies our Theorem.

Proposition 1. *Let $0 \leq \alpha < \frac{1}{3}$. Then*

$$+\infty = M^\alpha \mu(0) > M^\alpha \mu\left(\frac{1}{3}\right) < M^\alpha \mu\left(\frac{1}{2}\right) > M^\alpha \mu\left(\frac{2}{3}\right).$$

That is, $M^\alpha \mu$ has a nontrivial local maximum.

Proof. By the symmetries of our measure, $M^\alpha \mu\left(\frac{1}{3}\right) = M^\alpha \mu\left(\frac{2}{3}\right)$. A simple calculation then shows that $M^\alpha f\left(\frac{1}{3}\right) = \frac{3(\alpha+1)}{2}$, if $\alpha < \frac{1}{3}$. As $M^\alpha \mu\left(\frac{1}{2}\right) \geq M\mu\left(\frac{1}{2}\right) = 2 > \frac{3\alpha+3}{2} \iff \alpha < \frac{1}{3}$, we are done with the proof of this proposition. \square

Before proving our Theorem, we mention that our choice of $\frac{1}{3}, \frac{1}{2}, \frac{2}{3}$ was not random: $\frac{1}{2}$ is actually a *local maximum* of $M^\alpha \mu$, while $\frac{1}{3}, \frac{2}{3}$ are *local minima*.

Proof of Theorem 2. Let $f_n(x) = n(\chi_{[0, \frac{1}{n}]} + \chi_{[1-\frac{1}{n}, 1]})$. It is easy to see that $\int g f_n dx \rightarrow \int g d\mu(x)$, for each $g \in L^\infty(\mathbf{R})$ that is continuous on $[0, t_0) \cup (t_1, 1]$, for some $t_0 < t_1$.

We prove that $M^\alpha f_n(x) \rightarrow M^\alpha \mu(x)$, $\forall x \in [0, 1]$. This is clearly enough to conclude our Theorem, as then, if we fix $\alpha < \frac{1}{3}$, there will be $n(\alpha) > 0$ such that, for $N \geq n(\alpha)$,

$$0 = f_N\left(\frac{1}{3}\right) < M^\alpha f_N\left(\frac{1}{3}\right) < M^\alpha f_N\left(\frac{1}{2}\right) > M^\alpha f_N\left(\frac{2}{3}\right) > f_N\left(\frac{2}{3}\right) = 0.$$

To prove convergence, we argue in two steps.

The first step is to prove that $\liminf_{n \rightarrow +\infty} M^\alpha f_n(x) \geq M^\alpha \mu(x)$. It clearly holds for $x \in \{0, 1\}$. For $x \in (0, 1)$, we see that

$$M^\alpha f_n(x) = \sup_{|x-y| \leq \alpha t \leq 3\alpha} \frac{1}{2t} \int_{y-t}^{y+t} f_n(s) ds.$$

But then

$$\begin{aligned} M^\alpha \mu(x) &= \sup_{|x-y| \leq \alpha t \leq 3\alpha} \frac{1}{2t} \int_{y-t}^{y+t} d\mu(s) \\ &= \sup_{|x-y| \leq \alpha t \leq 3\alpha; t \geq \delta(x) > 0} \lim_{n \rightarrow \infty} \frac{1}{2t} \int_{y-t}^{y+t} f_n(s) ds \leq \liminf_{n \rightarrow \infty} M^\alpha f_n(x), \end{aligned}$$

where $\delta(x) > 0$ is a multiple of the minimum of the distances of x to either 1 or 0. This completes this part.

The second step is to establish that, for every $\varepsilon > 0$, $(1 + \varepsilon)M^\alpha \mu(x) \geq \limsup_{N \rightarrow \infty} M^\alpha f_N(x)$. This readily implies the result.

To do so, notice that, as $1 > x > 0$, then for N sufficiently large, the average that realizes the supremum on the definition of M^α has a positive radius bounded below and above in N . Specifically, we have that

$$M^\alpha f_N(x) = \int_{y_N - t_N}^{y_N + t_N} f_N(s) ds, \quad \Delta(x) \geq t_N \geq \delta(x) > 0.$$

This shows also that $\{y_N\}$ and $\{t_N\}$ must be bounded sequences. Therefore, using compactness,

$$\begin{aligned} \limsup_{N \rightarrow \infty} M^\alpha f_N(x) &= \limsup_{N \rightarrow \infty} \int_{y_N - t_N}^{y_N + t_N} f_N(s) \, ds = \lim_{k \rightarrow \infty} \int_{y_{N_k} - t_{N_k}}^{y_{N_k} + t_{N_k}} f_{N_k}(s) \, ds \\ &\leq (1 + \eta) \frac{1}{2t} \limsup_{N \rightarrow \infty} \int_{y - (1 + \varepsilon/2)t}^{y + (1 + \varepsilon/2)t} f_N(s) \, ds \\ &= (1 + \eta)(1 + \varepsilon/2) \int_{y - (1 + \varepsilon/2)t}^{y + (1 + \varepsilon/2)t} d\mu(s) \leq (1 + \varepsilon) M^\alpha \mu(x), \end{aligned}$$

where we assume that the sequence $\{n_k\}$ is suitably chosen so that the convergence requirements all hold. If we make N sufficiently large, and take η depending on ε such that $(1 + \eta)(1 + \varepsilon/2) < 1 + \varepsilon$, we are done with the second part. \square

4. Proof of Theorems 3 and 4

The idea for this proof is basically the same as before: analyze local maxima in the detachment set in this Lipschitz case, proving that the maximal function is either V shaped or monotone in its composing intervals, *if* the Lipschitz constant into consideration is less than $\frac{1}{2}$. The endpoint case is done by approximation, and we comment on how to do it later. By the end, we sketch on how to build the mentioned counterexamples.

4.1. Analysis of maxima of $M_N^1 f$ for $\text{Lip}(N) < \frac{1}{2}$. We assume, first of all, that f has $NORM(1)$ normalization. Let (a, b) be an interval on the real line, such that there exists a point x_0 , maximum of $M_N^1 f$ over (a, b) , with the property that

$$M_N^1 f(x_0) > \max\{M_N^1 f(a), M_N^1 f(b)\}.$$

Therefore, we wish to prove that, for some point in (a, b) , $M_N^1 f = f$. We begin with the general strategy: let us suppose that this is not the case. Then there must be an average $u(y, t) = \frac{1}{2t} \int_{y-t}^{y+t} |f(s)| \, ds$ with $N(x_0) \geq t > 0$, $|x_0 - y| \leq t$ and $M_N^1 f(x_0) = u(y, t)$.

Now we want to find a neighbourhood of x_0 such that there is $R = R(x_0) > 0$ such that, for all $x \in I$, $M_{\equiv R}^1 f(x) = M_N^1 f(x_0)$. By Lemma 1, we can suppose that either $y = x_0 - t$ or $y = x_0 + t$. Without loss of generality, let us assume that $y = x_0 - t$.

Case (a). $t < N(x_0)$. This is the easiest case, and we rule it out with a simple observation: let I be an interval for which x_0 is an endpoint and such that, for all $x \in \bar{I}$, $N(x) > t$. Assume, without loss of generality, that x_0 is the right endpoint $r(I)$. We claim then that, for $x \in I$, $M_{\equiv t + \varepsilon}^1 f(x) = M_N^1 f(x_0)$, if ε is sufficiently small. Indeed, if ε is sufficiently small, then $M_{\equiv t + \varepsilon}^1 f(x) \leq M_N^1 f(x) (\leq M_N^1 f(x_0))$ for every $x \in I$. But then we see also that $(x_0 - t, t)$ belongs to the region $\{(z, s) : |x - z| \leq s \leq N(x)\}$, as then $|(x_0 - t) - x| = x + t - x_0 \leq t < t + \varepsilon < N(x)$. This shows that

$$M_N^1 f(x_0) \leq \inf_{x \in I} M_{\equiv t + \varepsilon}^1 f(x) \leq \sup_{x \in I} M_{\equiv t + \varepsilon}^1 f(x) \leq M_N^1 f(x_0).$$

As before, we finish this case with [1, Lemma 3.6], as then it guarantees us that $M_{\equiv t + \varepsilon}^1 f(x) = f(x)$ for every point in this interval I .

Case (b). $t = N(x_0)$. In this case, we have to use Lemma 2. Namely, we wish to include the point $(x_0 - N(x_0), N(x_0))$ in the region

$$\{(z, s): |z - x| + |s - N(x)| \leq N(x)\},$$

for $x_0 - \delta < x < x_0$, δ sufficiently small.

Let then $\varepsilon > 0$ and x close to x_0 be such that $N(x) \geq N(x_0) - \varepsilon$. We have already a comparison of the form

$$M_N^1 f(x) \geq M_{\equiv N(x_0) - \varepsilon}^1 f(x).$$

We want to conclude that there is an interval I such that $M_{\equiv N(x_0) - \varepsilon}^1 f$ is constant on I . We want then the point $(x_0 - N(x_0), N(x_0))$ to lie on the set

$$\{(z, s): |z - x| + |s - N(x_0) + \varepsilon| \leq N(x_0) - \varepsilon\}.$$

But this is equivalent to

$$x - x_0 + N(x_0) + \varepsilon \leq N(x_0) - \varepsilon \iff |x - x_0| \geq 2\varepsilon.$$

So, we can only afford to do this if x is somewhat not too close to x_0 . But, as $\text{Lip}(N) < \frac{1}{2}$ in this case, we see that

$$\begin{aligned} |N(x) - N(x_0)| \leq \text{Lip}(N)|x - x_0| &\implies N(x) \geq N(x_0) - \text{Lip}(N)|x - x_0| \geq N(x_0) - \varepsilon \\ &\iff |x - x_0| \leq \frac{1}{\text{Lip}(N)}\varepsilon. \end{aligned}$$

Therefore, we conclude that, on the non-trivial set

$$\left\{ x \in \mathbf{R}: \frac{1}{\text{Lip}(N)}\varepsilon \geq |x - x_0| \geq 2\varepsilon \right\},$$

it holds that $M_N^1 f(x_0) \geq M_N^1 f(x) \geq M_{\equiv N(x_0) - \varepsilon}^1 f(x) \geq M_N^1 f(x_0) \geq M_N^1 f(x)$. By [1, Lemma 3.6], $M_{\equiv N(x_0) - \varepsilon}^1 f(x) = M_N^1 f(x) = f(x)$. This concludes then that, whenever there is a “strict” local maximum (with respect to the endpoints) of $M_N^1 f$ over an interval (a, b) , then there is a $x \in (a, b)$ such that $M_N^1 f(a) = f(a)$, as the finishing argument here is then the same as the one used in Theorem 1, and we therefore omit it.

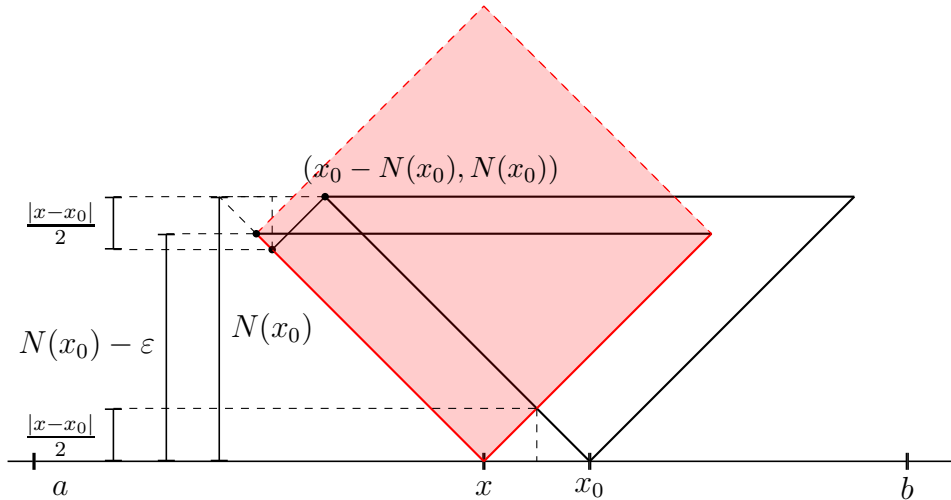


Figure 4. Illustration of proof of case (b).

4.2. The critical case $\text{Lip}(N) = \frac{1}{2}$. The argument is pretty simple: we build explicitly a suitable sequence of approximations of N such that they all have

Lipschitz constants less than $\frac{1}{2}$. By our already proved results, this will give us the result also in this case.

Explicitly, let N be such that $\text{Lip}(N) = \frac{1}{2}$ and $f \in BV(\mathbf{R})$. Let then $\mathcal{P} = \{x_1 < \dots < x_M\}$ be any partition of the real line. Let $J \gg 1$ be a large integer, and divide the interval $[x_1, x_M]$ into J equal parts, that we call (a_j, a_{j+1}) , where $j = 1, \dots, J$. Define also the numbers

$$\Delta_j = \frac{N(a_{j+1}) - N(a_j)}{a_{j+1} - a_j}.$$

We know, by hypothesis, that $\Delta_j \in [-1/2, 1/2]$. Let then $\tilde{\Delta}_j = \Delta_j - \frac{1}{J^2}$, and define the function

$$\tilde{N}(x) = \begin{cases} N(x_1), & \text{if } x \leq x_1, \\ N(x_1) + \tilde{\Delta}_1(x - x_1), & \text{if } x \in (a_1, a_2], \\ \tilde{N}(a_j) + \tilde{\Delta}_j(x - a_j), & \text{if } x \in (a_j, a_{j+1}], \\ \tilde{N}(a_{J+1}), & \text{if } x \geq x_M. \end{cases}$$

It is obvious that this function is continuous and Lipschitz with constant at most $\frac{1}{2} - \frac{1}{J^2}$. If $x \in (a_j, a_{j+1}]$, then

$$\begin{aligned} |\tilde{N}(x) - N(x)| &\leq |\tilde{N}(x) - \tilde{N}(a_j)| + |\tilde{N}(a_j) - N(x)| \\ &\leq \frac{|x_1 - x_M|}{2J} + |N(a_j) - N(x)| + |\tilde{N}(a_j) - N(a_j)| \\ \text{by the definition of } \tilde{N} &\leq \frac{|x_1 - x_M|}{J} + |\tilde{N}(a_{j-1}) - N(a_{j-1})| + \frac{|a_j - a_{j-1}|}{J^2} \\ \text{by an inductive argument} &\leq \frac{2|x_1 - x_M|}{J} + |\tilde{N}(x_1) - N(x_1)| = \frac{2|x_1 - x_M|}{J}. \end{aligned}$$

We now choose J such that the right hand side above is less than $\delta > 0$, which is going to be chosen as follows: for the same partition \mathcal{P} , we let $\delta > 0$ be such that

$$|\tilde{N}(x_i) - N(x_i)| < \delta \implies |M_N^1 f(x_j) - M_{\tilde{N}}^1 f(x_j)| < \frac{\varepsilon}{2M}.$$

This can, by continuity, always be accomplished. This implies that, using the previous case,

$$\mathcal{V}_{\mathcal{P}}(M_N^1 f) \leq \mathcal{V}_{\mathcal{P}}(M_{\tilde{N}}^1 f) + \varepsilon \leq \mathcal{V}(M_{\tilde{N}}^1 f) + \varepsilon \leq \mathcal{V}(f) + \varepsilon.$$

Taking the supremum over all possible partitions and then taking $\varepsilon \rightarrow 0$ finishes also this case, and thus the proof of Theorem 3.

4.3. Counterexample for $\text{Lip}(N) > \frac{1}{2}$. Finally, we build examples of functions with $\text{Lip}(N) > \frac{1}{2}$ and $f \in BV(\mathbf{R})$ such that

$$\mathcal{V}(M_N f) = +\infty.$$

Fix then $\beta > \frac{1}{2}$ and let a function N with $\text{Lip}(N) = \beta$ be defined as follows:

- (A) First, let $x_0 = \frac{2}{2\beta+1}$. Let then $N(0) = 1$, $N(x_0) = \frac{x_0}{2}$ and extend it linearly in $(0, x_0)$.
- (B) Let x'_K be the solution to the equation $\beta x - \beta x_{K-1} + \frac{x_{K-1}}{2} = \frac{x+1}{2} \iff x'_K = x_{K-1} + \frac{1}{\beta - \frac{1}{2}}$.
- (C) At last, take $x_K = x'_K + \frac{1}{2\beta+1}$, and define for all $K \geq 1$ $N(x_K) = \frac{x_K}{2}$, $N(x'_K) = \frac{x'_K+1}{2}$, extending it linearly on (x_{K-1}, x'_K) and (x'_K, x_K) .

As $\{x'_K\}_{K \geq 1}$ is an arithmetic progression, we see that

$$\sum_{K \geq 1} \frac{1}{x'_K} = +\infty.$$

Moreover, define $f(x) = \chi_{(-1,0)}(x)$. We will show that, for this N , we have that

$$\mathcal{V}(M_N^1 f) = +\infty.$$

In fact, it is not difficult to see that:

- (A) $M_N^1 f(x_K) = 0, \forall K \geq 0$. This is due to the fact that the maximal intervals $(y-t, y+t)$ that satisfy $|x_K - y| \leq t \leq N(x_K)$ are still contained in $[0, +\infty)$, which is of course disjoint from $(-1, 0)$.
- (B) $M_N^1 f(x'_K) \geq \frac{1}{x'_K + 1}$. This follows from

$$M_N^1 f(x'_K) \geq \frac{1}{2N(x'_K)} \int_{-1}^{x'_K} f(t) dt = \frac{1}{x'_K + 1}.$$

This shows that

$$\mathcal{V}(M_N^1 f) \geq \sum_{K=1}^{\infty} |M_N^1 f(x'_K) - M_N^1 f(x_K)| = \sum_{K=1}^{\infty} \frac{1}{x'_K + 1} = +\infty.$$

This construction therefore proves Theorem 4.

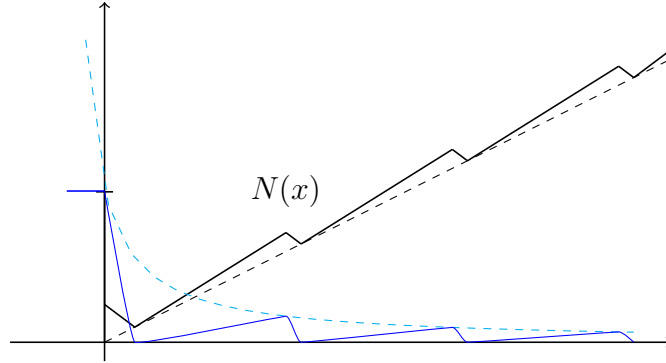


Figure 5. A counterexample in the case of $\text{Lip}(N) = \frac{3}{4}$. The dashed lines are the graphs of $\frac{x}{2}$ and $\frac{1}{1+x}$, and the non-dashed ones the graphs of $M_N^1 f$ and N in this case.

5. Comments and remarks

5.1. Monotonicity of maximal BV -norms. Theorem 1 proves that, if we define

$$B(\alpha) := \sup_{f \in BV(\mathbf{R}) : \mathcal{V}(f) \neq 0} \frac{\mathcal{V}(M^\alpha f)}{\mathcal{V}(f)},$$

then $B(\alpha) = 1$ for all $\alpha \in [\frac{1}{3}, 1]$. We can, however, with the same technique, show that $B(\alpha)$ is non-increasing in $\alpha > 0$, and also that $B(\alpha) \equiv 1 \forall \alpha \in [\frac{1}{3}, +\infty)$. Indeed, we show that, for $f \in BV(\mathbf{R})$ endowed with $NORM(1)$ normalization and $\beta > \alpha$, then $\mathcal{V}(M^\alpha f) \geq \mathcal{V}(M^\beta f)$. This allows one to conclude, without glancing at Theorem 1, that $\mathcal{V}(M^\alpha f) \leq C \cdot \mathcal{V}(f)$, for $C = 240.004$ and all $\alpha \geq 0$, as a consequence of Kurka's [8] result. The argument uses the maximal attachment property and is *independent* of the proof of Theorem 1—which allows us, for instance, to make use of this fact in the proof of Theorem 1, as indicated in the subsection 3.2: we first

assume f to be positive, without loss of generality. Let, as usual, (a, b) be an interval where $M^\beta f$ has a local maximum *inside* it, at, say, x_0 , and

$$M^\beta f(x_0) > \max(M^\beta f(a), M^\beta f(b)).$$

Then, as we have that $M^\beta f \geq M^\alpha f$ *everywhere*, we have two options:

- If $M^\beta f(x_0) = f(x_0)$, we do not have absolutely anything to do, as then also $M^\alpha f(x_0) = M^\beta f(x_0)$.
- If $M^\beta f(x_0) = u(y, t)$, for $t > 0$, we have—as in subsection 3.1—that $(y - \beta t, y + \beta t) \subset (a, b)$. But it is then obvious that

$$M^\alpha f(y) \geq u(y, t) = M^\beta f(x_0) \geq M^\beta f(y) \geq M^\alpha f(y).$$

Therefore, we have obtained a form of the maximal attachment property, and therefore we can apply the standard techniques that have been used through the paper to this case, and it is going to yield our result.

This shows directly that $B(\alpha) \leq 1, \forall \alpha \geq 1$, but taking $f(x) = \chi_{(0,1)}$ as we did several times shows that actually $B(\alpha) = 1$ in this range.

5.2. Nontangential maximal functions and classical results. Here, we investigated mostly the regularity aspect of our family M^α of nontangential maximal functions, and looked for the sharp constants in such bounded variation inequalities. One can, however, still ask about the most classical aspect studied by Melas [12]: Let C_α be the least constant such that we have the following inequality:

$$|\{x \in \mathbf{R}: M^\alpha f(x) > \lambda\}| \leq \frac{C_\alpha}{\lambda} \|f\|_1.$$

By [12], we have that, for when $\alpha = 0$, then $C_0 = \frac{11+\sqrt{61}}{12}$, and the classical argument of Riesz [13] that $C_1 = 2$. Therefore, $\frac{11+\sqrt{61}}{12} \leq C_\alpha \leq 2, \forall \alpha \in (0, 1)$. Nevertheless, the exact values of those constants is, as long as the author knows, still unknown.

5.3. Bounded variation results for mixed Lipschitz and nontangential maximal functions. In Theorems 3 and 4, we proved that, for the *uncentered* Lipschitz maximal function M_N , we have sharp bounded variation results for $\text{Lip}(N) \leq \frac{1}{2}$, and, if $\text{Lip}(N) > \frac{1}{2}$, we cannot even assure *any* sort of bounded variation result.

We can ask yet another question: if we define the *nontangential Lipschitz maximal function*

$$M_N^\alpha f(x) = \sup_{|x-y| \leq \alpha t \leq \alpha N(x)} \frac{1}{2t} \int_{y-t}^{y+t} |f(s)| ds,$$

then what should be the best constant $L(\alpha)$ such that, for $\text{Lip}(N) \leq L(\alpha)$, then we have some sort of bounded variation result like $\mathcal{V}(M_N^\alpha f) \leq A\mathcal{V}(f)$, and, for each $\beta > L(\alpha)$, there exists a function N_β and a function $f_\beta \in BV(\mathbf{R})$ such that $\text{Lip}(N_\beta) = \beta$ and $\mathcal{V}(M_{N_\beta} f_\beta) = +\infty$? Regarding this question, we cannot state any kind of sharp constant bounded variation result, but the following is still attainable: it is possible to show that the first two lemmas of Kurka [8] are adaptable in this context *if* we suppose that

$$\text{Lip}(N) \leq \frac{1}{\alpha + 1},$$

and then we obtain our results, with a constant that is even independent of $\alpha \in (0, 1)$. On the other hand, our example used above in the proof of Theorem 4 is easily adaptable as well, and therefore one might prove the following Theorem:

Theorem 5. *Let $\alpha \in [0, 1]$ and N be a Lipschitz function with $Lip(N) \leq \frac{1}{\alpha+1}$. Then, for every $f \in BV(\mathbf{R})$, we have that*

$$\mathcal{V}(M_N^\alpha f) \leq C\mathcal{V}(f),$$

where C is independent of N, f, α . Moreover, for all $\beta > \frac{1}{\alpha+1}$, there is a function N_β and

$$f(x) = \begin{cases} 1, & \text{if } x \in (-1, 0); \\ 0, & \text{otherwise,} \end{cases}$$

with $Lip(N_\beta) = \beta$ and $\mathcal{V}(M_{N_\beta}^\alpha f) = +\infty$.

References

- [1] ALDAZ, J. M., and J. PÉREZ LÁZARO: Functions of bounded variation, the derivative of the one dimensional maximal function, and applications to inequalities. - Trans. Amer. Math. Soc. 359:5, 2007, 2443–2461.
- [2] BOBER, J., E. CARNEIRO, K. HUGHES, and L. PIERCE: On a discrete version of Tanaka's theorem for maximal functions. - Proc. Amer. Math. Soc. 140, 2012, 1669–1680.
- [3] CARNEIRO, E., R. FINDER, and M. SOUSA: On the variation of maximal operators of convolution type II. - Rev. Mat. Iberoam. (to appear).
- [4] CARNEIRO, E., and B. F. SVAITER: On the variation of maximal operators of convolution type. - J. Funct. Anal. 265, 2013, 837–865.
- [5] HAJLASZ, P., and J. ONNINEN: On boundedness of maximal functions in Sobolev spaces. - Ann. Acad. Sci. Fenn. Math. 29:1, 2004, 167–176.
- [6] KINNUNEN, J.: The Hardy–Littlewood maximal function of a Sobolev function. - Israel J. Math. 100, 1997, 117–124.
- [7] KINNUNEN, J., and P. LINDQVIST: The derivative of the maximal function. - J. Reine Angew. Math. 503, 1998, 161–167.
- [8] KURKA, O.: On the variation of the Hardy–Littlewood maximal function. - Ann. Acad. Sci. Fenn. Math. 40, 2015, 109–133.
- [9] LUIRO, H.: Continuity of the maximal operator in Sobolev spaces. - Proc. Amer. Math. Soc. 135:1, 2007, 243–251.
- [10] LUIRO, H.: The variation of the maximal function of a radial function. - Preprint, arXiv:1702.00669.
- [11] MADRID, J.: Sharp inequalities for the variation of the discrete maximal function. - Preprint, arXiv:1512.04112.
- [12] MELAS, A.: The best constant for the centered Hardy–Littlewood maximal inequality. - Ann. of Math. (2) 157, 2003, 647–688.
- [13] RIESZ, F.: Sur un théorème de maximum de Mm. Hardy et Littlewood. - J. London Math. Soc. 7:1, 1932, 10–13.
- [14] STEIN, E.: Harmonic analysis: Real variable methods, orthogonality, and oscillatory integrals. - Princeton Univ. Press, Princeton, NJ, 1993.
- [15] TANAKA, H.: A remark on the derivative of the one-dimensional Hardy–Littlewood maximal function. - Bull. Aust. Math. Soc. 65:2, 2002, 253–258.
- [16] TEMUR, F.: On regularity of the discrete Hardy–Littlewood maximal function. - Preprint, arXiv:1303.3993.

Received 16 March 2017 • First revised received 30 January 2018

Second revised received 12 June 2018 • Accepted 23 July 2018