

# ON THE CONTINUITY OF DISCRETE MAXIMAL OPERATORS IN SOBOLEV SPACES

Hannes Luiro and Juho Nuutinen

University of Jyväskylä, Department of Mathematics and Statistics  
P.O. Box 35, FI-40014 Jyväskylä, Finland; hannes.s.luiro@jyu.fi

University of Jyväskylä, Department of Mathematics and Statistics  
P.O. Box 35, FI-40014 Jyväskylä, Finland; juho.nuutinen@jyu.fi

**Abstract.** We investigate the continuity of discrete maximal operators in Sobolev space  $W^{1,p}(\mathbf{R}^n)$ . A counterexample is given as well as it is shown that the continuity follows under certain sufficient assumptions. Especially, our research verifies that for the continuity in Sobolev spaces the role of the partition of the unity used in the construction of the maximal operator is very delicate.

## 1. Introduction

In the case of the classical Hardy–Littlewood maximal operator  $\mathcal{M}$  it is well known, by [5], [7] and [4], that it is bounded in the first order Sobolev spaces, when  $p > 1$ . Observe that in general bounded nonlinear operators do not need to be continuous, as is the case for some natural maximal operators, see e.g. [10]. By the main result of [9], the classical Hardy–Littlewood maximal operator  $\mathcal{M}$  is continuous in  $W^{1,p}(\mathbf{R}^n)$ , when  $p > 1$ . In the case of metric measure spaces, as an example in [2] shows, there is an unexpected obstruction concerning the regularity of  $\mathcal{M}$ . Indeed, it might even happen that  $\mathcal{M}f$  of a Lipschitz continuous function  $f$  may fail to be continuous. For this reason, in the metric setting a so-called discrete maximal function is often considered.

The discrete maximal function, as in [1], [6] and [8], of a Sobolev function  $f$  is constructed in a metric measure space  $X = (X, d, \mu)$  in terms of a covering and a partition of unity. First fix a scale  $r > 0$  and choose a family of balls  $\{B(x_j, r)\}$  that cover the space so that the dilated balls  $B(x_j, 6r)$  are of bounded overlap, i.e.

$$\sum_{j=1}^{\infty} \chi_{B(x_j, 6r)} \leq N < \infty,$$

where the constant  $N$  is independent of  $r$ . Then a partition of unity subordinate to the cover  $\{B(x_j, r)\}$  of  $X$  is constructed by choosing functions  $\psi_j^r$ ,  $j = 1, 2, \dots$ , on  $X$  such that  $0 \leq \psi_j^r \leq 1$ ,  $\psi_j^r = 0$  on  $X \setminus B(x_j, 6r)$  and  $\psi_j^r \geq C$  on  $B(x_j, 3r)$ . Functions  $\psi_j^r$  are also assumed to be Lipschitz continuous with constant  $C/r$ , where  $C$  is independent of  $r$ , and  $\sum_{j=1}^{\infty} \psi_j^r = 1$  on  $X$ . We define the approximation of  $f$  at

---

doi:10.5186/aasfm.2014.3916

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

Key words: Discrete maximal operator, Sobolev spaces, continuity.

The authors are supported by the Academy of Finland, grants no. 259069 (H.L.) and 135561 (J.N.).

the scale of  $3r$  by setting

$$(1.1) \quad T_r f(x) = \sum_{j=1}^{\infty} \psi_j^r(x) f_{B(x_j, 3r)}$$

for every  $x$ . Here  $f_{B(x_j, 3r)}$  denotes the integral average of  $f$  in  $B(x_j, 3r)$  with respect to  $\mu$ . The function  $T_r f$  is also called a *discrete convolution* of  $f$ . For some applications of the partition of unity and the discrete convolution to the analysis in metric measure spaces, we refer to [3] and [11]. See also pages 290–292 of [12]. The *discrete maximal function* (corresponding to the above given partition of unity) is then defined by

$$(1.2) \quad M^* f(x) = \sup_{r \in \mathbf{Q}_+} T_r |f|(x),$$

where  $\mathbf{Q}_+$  denotes the positive rationals.

Observe that for each scale there are many possible choices for the covering, but we simply take one of those. Note that the boundedness of  $M^*$  in Sobolev spaces does not depend on the chosen coverings and partitions of unity (see e.g. [6]). The main point of this paper is to show that the same is not true in the case of continuity even in the Euclidean space. For this reason, to avoid some inessential technical difficulties, we consider throughout this work the discrete maximal function of  $f$  on a discrete set of scales  $r = 2^{-k}$ , defined by

$$(1.3) \quad M^* f(x) = \sup_{k \in \mathbf{N}} T_k |f|(x)$$

for all  $x$ , where  $T_k |f| := T_{2^{-k}} |f|$ .

We will show in Section 2 that, kind of surprisingly, it might happen that the discrete maximal operator, as defined in (1.3), may fail to be continuous in  $W^{1,p}(\mathbf{R})$ . In the light of this example it is obvious that the same may happen for the standard discrete maximal operator (1.2). In Section 3 we show that under assumption

$$(1.4) \quad D_i T_k |f|(x) \rightarrow D_i |f|(x) \quad \text{a.e. as } k \rightarrow \infty, \quad 1 \leq i \leq n,$$

for any  $f \in W_{loc}^{1,1}(\mathbf{R}^n)$ , we get the continuity of  $M^*$  in  $W^{1,p}(\mathbf{R}^n)$ <sup>1</sup>. A similar kind of an assumption, related to the continuity of the mapping  $r \rightarrow D_i T_r |f|(x)$  would yield the continuity in the standard case (1.2), as is briefly noted in Remark 3.4. Finally, in Section 4 we give a one-dimensional example of a partition of unity for which assumption (1.4) holds and thus the corresponding discrete maximal operator is continuous in  $W^{1,p}(\mathbf{R})$ , for  $p > 1$ . We also give an example indicating both that the assumption (1.4) is definitely not necessary for the continuity and also that our counterexample in Section 2 can not be remarkably simplified.

*Acknowledgements.* We would like to thank Juha Lehrbäck, Heli Tuominen, Juha Kinnunen and Daniel Aalto for useful discussions and comments on the manuscript.

## 2. Counterexample

**2.1. Partition of unity.** For a fixed  $k \in \mathbf{N}$ , we define intervals  $I_j^k := [j 2^{-k}, (j+1) 2^{-k}]$ , where  $j \in \mathbf{Z}$ . We also denote

$$I_j^k =: B(x_j, r_k),$$

<sup>1</sup>Let  $D_i f$ ,  $1 \leq i \leq n$  and  $Df$  denote the standard weak partial derivatives and weak gradient (respectively) of  $f$ .

where  $x_j = 2^{-k} (j + \frac{1}{2})$  and  $r_k = 2^{-k-1}$ . Thus  $\mathbf{R}$  is covered by a countable union of balls  $B(x_j, r_k)$  such that the dilated balls  $B(x_j, 6r_k)$ ,  $j \in \mathbf{Z}$ , are of bounded overlap. Now let us define functions  $\psi_j^k$  so that

$$\psi_j^k = \frac{1}{4} \quad \text{in } I_{j-1}^k \cup I_j^k \cup I_{j+1}^k$$

and

$$\psi_j^k = 0 \quad \text{outside } I_{j-2}^k \cup I_{j-1}^k \cup I_j^k \cup I_{j+1}^k \cup I_{j+2}^k.$$

Here the delicate point is how we pass from  $\frac{1}{4}$  to 0 on intervals  $I_{j-2}^k$  and  $I_{j+2}^k$ . For that purpose we define functions  $w_k: [0, 2^{-k}] \rightarrow [0, 2^{k-1}]$  in the following way. Suppose  $l \in \{0, 1, \dots, 2^{9k} - 1\}$  and define

$$(2.1) \quad w_k(t) = \begin{cases} 2^{k-1}, & \text{if } t \in [l 2^{-10k}, (l+1) 2^{-10k}], \text{ } l \text{ is even, and} \\ 0, & \text{otherwise.} \end{cases}$$

Then define

$$\psi_j^k(x) = \begin{cases} \int_0^{x-(j-2)2^{-k}} w_k(t) dt, & \text{if } x \in I_{j-2}^k, \text{ and} \\ \frac{1}{4} - \int_0^{x-(j+2)2^{-k}} w_k(t) dt, & \text{if } x \in I_{j+2}^k. \end{cases}$$

So here we divide intervals  $I_{j-2}^k$  and  $I_{j+2}^k$  into  $2^{9k}$  subintervals of the form  $[(j-2) 2^{-k} + l 2^{-10k}, (j-2) 2^{-k} + (l+1) 2^{-10k}]$  and  $[(j+2) 2^{-k} + l 2^{-10k}, (j+2) 2^{-k} + (l+1) 2^{-10k}]$ . Notice that when  $k$  is large,  $\psi_j^k$  approximates a straight line very precisely on intervals  $I_{j-2}^k$  and  $I_{j+2}^k$ , but the approximation is kind of serrated and slightly antisymmetric, which is essential for this example to work.

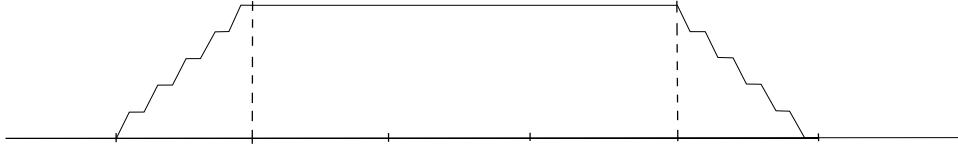


Figure 1. A sketch of the above partition of unity.

The discrete convolution and the discrete maximal function are defined as in (1.1) and (1.3) in the introduction.

### 2.2. Discontinuity.

**Lemma 2.1.** *Let the partition of unity be defined as above and let  $f(x) = ax + b$ ,  $a < 0$ . Then each discrete convolution of  $f$  is strictly below  $f$  except in points that are of the form  $i 2^{-k} + l 2^{-10k}$  for some  $i \in \mathbf{Z}$ ,  $k \in \mathbf{N}$  and  $l \in \mathbf{N}$ , where  $l$  is even. Moreover,*

$$(2.2) \quad T_k f(x) \geq f(x) - |a| 2^{-10k}$$

for every  $x \in \mathbf{R}$ .

*Proof.* Fix  $k$ , let  $x \in \mathbf{R}$  and take  $i \in \mathbf{Z}$  such that  $x \in I_i^k$ . Then

$$T_k f(x) = \sum_{j \in \mathbf{Z}} \psi_j^k(x) f_{B(x_j, 3r_k)} = \sum_{j=i-2}^{i+2} \psi_j^k(x) f_{B(x_j, 3r_k)}.$$

The definition of functions  $\psi_j^k$  and the fact that  $f_{B(x_j, 3r_k)} = ax_j + b$  for each scale  $3r_k$  gives

$$\begin{aligned} T_k f(x) &= \left( \frac{1}{4} - \int_0^{x-i2^{-k}} w_k(t) dt \right) f_{B(x_{i-2}, 3r_k)} + \frac{1}{4} f_{B(x_{i-1}, 3r_k)} \\ &\quad + \frac{1}{4} f_{B(x_i, 3r_k)} + \frac{1}{4} f_{B(x_{i+1}, 3r_k)} + \left( \int_0^{x-i2^{-k}} w_k(t) dt \right) f_{B(x_{i+2}, 3r_k)} \\ &= a \frac{x_{i-2} + x_{i-1} + x_i + x_{i+1}}{4} + b + a \left( \int_0^{x-i2^{-k}} w_k(t) dt \right) (x_{i+2} - x_{i-2}) \\ &= ai2^{-k} + b + a2^{-k+2} \int_0^{x-i2^{-k}} w_k(t) dt. \end{aligned}$$

Suppose then that

$$x = i2^{-k} + l2^{-10k} + c, \text{ where } l \in \{0, 1, \dots, 2^{9k} - 1\} \text{ and } c \in [0, 2^{-10k}].$$

Since

$$f(x) = ax + b = ai2^{-k} + b + a(l2^{-10k} + c),$$

we get from above that

$$(2.3) \quad T_k f(x) - f(x) = a \left( 2^{-k+2} \int_0^{x-i2^{-k}} w_k(t) dt - l2^{-10k} - c \right).$$

By the definition of  $w_k(t)$  in (2.1), it is easy to compute that

$$\int_0^{x-i2^{-k}} w_k(t) dt = \begin{cases} \frac{l}{2} 2^{k-1} 2^{-10k} + c2^{k-1}, & \text{when } l \text{ is even,} \\ \frac{l+1}{2} 2^{k-1} 2^{-10k}, & \text{when } l \text{ is odd.} \end{cases}$$

Substitution of this formula to (2.3) then gives us that

$$(2.4) \quad T_k f(x) - f(x) = \begin{cases} ac, & \text{when } l \text{ is even,} \\ a(2^{-10k} - c), & \text{when } l \text{ is odd.} \end{cases}$$

Our claim is an immediate consequence of this formula.  $\square$

**Lemma 2.2.** *Let the partition of unity be defined as previously and let  $f(x) = ax + b$ ,  $a < 0$ . Then the derivative of  $T_k f$  is either 0 or  $2a$  outside the set  $\{i2^{-k} + l2^{-10k} : i \in \mathbf{Z}, l \in \{0, 1, \dots, 2^{9k} - 1\}\}$ .*

*Proof.* Let  $x = i2^{-k} + l2^{-10k} + c$ , where  $l \in \{0, 1, \dots, 2^{9k} - 1\}$  and  $c \in [0, 2^{-10k}]$ . In the proof of the previous lemma we computed in (2.4) that

$$T_k f(x) = ax + b + \begin{cases} ac, & \text{when } l \text{ is even,} \\ a(2^{-10k} - c), & \text{when } l \text{ is odd.} \end{cases}$$

When  $0 < c < 2^{-10k}$ , one can straightforwardly differentiate  $T_k f$  using this formula to obtain

$$DT_k f(x) = \begin{cases} 2a & \text{if } l \text{ is even,} \\ 0 & \text{if } l \text{ is odd.} \end{cases}$$

This implies the claim.  $\square$

**Lemma 2.3.** *Let the partition of unity be defined as previously, let  $f(x) = ax + b$ ,  $a < 0$ , and define  $f^\epsilon(x) = f(x) + \epsilon(x^2 - 10x)$ , for  $\epsilon > 0$ . Then*

$$M^* f^\epsilon(x) > f^\epsilon(x) \quad \text{for all } x \in \mathbf{R}.$$

*Proof.* Let  $x_0 \in \mathbf{R}$  and let  $L^{x_0}(x) = a'(x) + b'$  denote the tangent line of  $f^\epsilon$  at point  $x_0$ . Then it holds e.g. by Taylor's formula that

$$f^\epsilon(x) = L^{x_0}(x) + \epsilon(x - x_0)^2 \quad \text{for all } x \in \mathbf{R}.$$

Denote  $g(x) = (x - x_0)^2$ . Using the linearity of the discrete convolution we then get that

$$(2.5) \quad T_k f^\epsilon(x_0) = T_k(L^{x_0})(x_0) + \epsilon T_k(g)(x_0).$$

Furthermore, it follows from Lemma 2.1 that

$$(2.6) \quad T_k(L^{x_0})(x_0) \geq L^{x_0}(x_0) - |(L^{x_0})'|2^{-10k} = f^\epsilon(x_0) - |(L^{x_0})'|2^{-10k}.$$

On the other hand, an elementary calculation shows that

$$(2.7) \quad T_k(g)(x_0) \geq c(2^{-k})^2$$

for any  $k \in \mathbf{N}$  and  $c > 0$  independent of  $k$  and  $x_0$ . Combining (2.5), (2.6) and (2.7) gives that

$$(2.8) \quad T_k f^\epsilon(x_0) \geq f^\epsilon(x_0) + c' \epsilon 2^{-2k}$$

if  $k$  is sufficiently big. Since  $M^* f^\epsilon(x_0) \geq T_k |f^\epsilon|(x_0) \geq T_k f^\epsilon(x_0)$ , our proof is complete.  $\square$

We will also need the following elementary fact.

**Proposition 2.4.** *Let  $f_k: \mathbf{R} \rightarrow \mathbf{R}$ ,  $1 \leq k \leq k_0$ , be a collection of continuous functions, which are differentiable up to a finite set and let  $F(x) = \max_{1 \leq k \leq k_0} f_k(x)$ . Then  $F$  is continuous and a.e. differentiable such that a.e.  $DF(x) = Df_{k(x)}$  for some  $1 \leq k(x) \leq k_0$ .*

The following Theorem verifies the discontinuity of the discrete maximal operator in  $W^{1,p}(\mathbf{R})$  for the above given partition of unity.

**Theorem 2.5.** *Let the partition of unity be defined as above, let  $f(x) = \max\{0, 10 - |x|\}$  and  $f^\epsilon(x) = f(x) + \chi_{[0,10]}(x)\epsilon(x^2 - 10x)$ . Then*

$$D M^* f^\epsilon \not\rightarrow D M^* f \quad \text{in } L^p \quad \text{when } \epsilon \rightarrow 0.$$

*Proof.* Denote  $I = [4, 5]$ ,  $g(x) = -x + 10$  and  $g^\epsilon(x) = g(x) + \epsilon(x^2 - 10x)$ . Since  $r_k \leq \frac{1}{2}$  for all  $k \in \mathbf{N}$ , it follows that  $M^* f = M^* g$  and  $M^* f^\epsilon = M^* g^\epsilon$  on  $I$ . Therefore we may consider  $f$  and  $f^\epsilon$  as if they were  $g$  and  $g^\epsilon$  (respectively), to obtain that previous lemmas of this section are in our use.

First of all, Lemma 2.1 implies that  $M^* f = f$  on  $I$ , and so we have  $D M^* f = -1$  on  $I$ . Moreover, by Lemma 2.2 we obtain that for every  $k \in \mathbf{N}$  the derivative of  $T_k f$  is equal to 0 or  $-2$  a.e. in  $I$ .

Consider then the derivative of  $T_k f^\epsilon$ . By the linearity of the discrete convolution and the derivative it follows that

$$(2.9) \quad DT_k f^\epsilon = DT_k f + \epsilon DT_k(x^2 - 10x).$$

One can easily check that

$$(2.10) \quad |DT_k((x^2 - 10x))| < C \quad \text{a.e.}$$

for some constant  $C > 0$ , independent of  $k$ . Since  $DT_k(f)$  is equal to 0 or  $-2$  on  $I$ , it follows from (2.9) and (2.10) that also the derivative of  $DT_k(f^\epsilon)$  is a.e. close to 0 or  $-2$ . Especially, there is  $\epsilon_0 > 0$  s.t. for any  $\epsilon < \epsilon_0$  it holds that

$$(2.11) \quad |DT_k(f^\epsilon) - (-1)| \geq \frac{1}{2} \text{ a.e. in } I.$$

Moreover, by using Lemma 2.3 and the continuity of functions  $M^*f^\epsilon$  and  $f^\epsilon$ , we know that there exists  $\lambda > 0$  such that  $M^*f^\epsilon > f^\epsilon + \lambda$  on  $I$ . This implies that there exists  $k_0 = k_0(\epsilon) \in \mathbf{N}$  such that for every  $x \in I$

$$(2.12) \quad M^*f^\epsilon(x) = \max_{1 \leq k \leq k_0} T_k f^\epsilon(x).$$

Then it follows by Proposition 2.4 that for a.e.  $x \in I$

$$(2.13) \quad DM^*f^\epsilon(x) = DT_{k(x)}f^\epsilon(x).$$

for some  $1 \leq k(x) \leq k_0$ . It is easy to see that combining this fact with (2.11) completes the proof.  $\square$

**Remark 2.6.** While the above theorem verifies the discontinuity of  $M^*$  (for the above given partition of unity) in Sobolev spaces, it also provides an example where  $Df^\epsilon \rightarrow Df$  pointwise uniformly but still  $|DM^*f^\epsilon(x) - DM^*f(x)| \rightarrow 1$ , as  $\epsilon \rightarrow 0$ , outside a countable set  $E$ . This can be seen by carefully revising the proof of the above theorem.

### 3. Positive results

In this section we establish the continuity of discrete maximal operators in  $W^{1,p}(\mathbf{R}^n)$ ,  $1 < p < \infty$ , in the case where derivatives of the discrete convolutions behave nicely in small scales. More precisely, we assume, throughout this section that

$$(3.1) \quad D_i T_k f(x) \rightarrow D_i f(x) \text{ a.e. as } k \rightarrow \infty$$

for all  $1 \leq i \leq n$  and  $f \in W_{loc}^{1,1}(\mathbf{R}^n)$ . We will show that this additional assumption yields the continuity of  $M^*$  in  $W^{1,p}(\mathbf{R}^n)$ , when  $1 < p < \infty$ . The proof follows essentially the same lines as in [9].

Let us begin with some notation: The Lebesgue measure of  $A \subset \mathbf{R}^n$  is denoted by  $|A|$ . For  $A \subset \mathbf{R}^n$  and  $\delta > 0$ , let  $|x - A| = \inf_{y \in A} |x - y|$  and

$$A_{(\delta)} = \{x \in \mathbf{R}^n : |x - A| \leq \delta\}.$$

Let us first define the sets  $\mathcal{R}f(x)$  of 'best scales' by

$$(3.2) \quad \mathcal{R}f(x) = \{2^{-k} : M^*f(x) = T_k|f|(x)\}.$$

This definition was introduced in the case of the usual Hardy–Littlewood maximal function in [9]. Here we also make a convention  $T_\infty f = |f|$  and  $2^{-\infty} = 0$ , thus if  $M^*f(x) = |f(x)|$ , then  $0 \in \mathcal{R}f(x)$ . Since almost every point is a Lebesgue point, one can easily check that  $\mathcal{R}f(x)$  is non-empty and closed set for a.e.  $x$ .

The following Lemma is proven in [9] for the usual Hardy–Littlewood maximal operator and exactly the same proof applies also in the case of the discrete maximal operator.

**Lemma 3.1.** *Let  $1 \leq p < \infty$  and suppose  $f_j, f \in L^p(\mathbf{R}^n)$  s.t.  $\|f_j - f\| \rightarrow 0$  when  $j \rightarrow \infty$ . Then for all  $R > 0$  and  $\lambda > 0$  it holds that*

$$|\{x \in B(0, R) : \mathcal{R}f_j(x) \not\subset \mathcal{R}f(x)_{(\lambda)}\}| \rightarrow 0 \text{ if } j \rightarrow \infty. \quad \square$$

Let us then define for  $A, B \subset \mathbf{R}^n$  a distance function  $\Delta$  by

$$\Delta(A, B) := \inf\{\delta > 0 : A \subset B_{(\delta)}\}.$$

With this new notation Lemma 3.1 says that

$$|\{x \in B(0, R) : \Delta(\mathcal{R}f_j(x), \mathcal{R}f(x)) > \lambda\}| \rightarrow 0 \text{ when } j \rightarrow \infty.$$

That is to say  $\Delta(\mathcal{R}f_j(x), \mathcal{R}f(x)) \rightarrow 0$  in measure, as  $j \rightarrow \infty$ . It is well known that in this case there exists a subsequence  $f_{j_i}$  so that

$$\Delta(\mathcal{R}f_{j_i}(x), \mathcal{R}f(x)) \rightarrow 0, \text{ for a.e. } x \in B(0, R), \text{ as } i \rightarrow \infty.$$

This will be exploited in the proof of the continuity.

The following theorem provides a formula for the derivative of the discrete maximal operator, corresponding to [9, Theorem 3.1]. Since the set  $\{2^{-k} : k \in \mathbf{Z}\} \cup \{0\}$  has only one density point, the result is more evident. However, unlike in the case of Hardy–Littlewood maximal operator, we had to take into account the fact that the operators  $T_k$  do not commute with translations.

**Theorem 3.2.** *Let  $p > 1$  and  $f \in W^{1,p}(\mathbf{R}^n)$ . Then*

$$(3.3) \quad D_i M^* f(x) = D_i T_k |f|(x) \text{ for every } 2^{-k} \in \mathcal{R}f(x).$$

*Proof.* We may assume that  $f$  is non-negative. Let  $h_j > 0$  and  $h_j \rightarrow 0$  as  $j \rightarrow \infty$ . Since we know that  $M^* f \in W_{\text{loc}}^{1,1}(\mathbf{R}^n)$  as well as  $T_k f \in W_{\text{loc}}^{1,1}(\mathbf{R}^n)$  for every  $k \in \mathbf{N}$  (see [6, Lemma 3.3 and Theorem 3.6]), it follows for a.e.  $x \in \mathbf{R}^n$  that

$$\begin{aligned} D_i M^* f(x) &= \lim_{j \rightarrow \infty} \frac{1}{h_j} (M^* f(x + h_j e_i) - M^* f(x)) \\ &= \lim_{j \rightarrow \infty} \frac{1}{h_j} (M^* f(x) - M^* f(x - h_j e_i)) \end{aligned}$$

and

$$\begin{aligned} D_i T_k f(x) &= \lim_{j \rightarrow \infty} \frac{T_k f(x + h_j e_i) - T_k f(x)}{h_j} \\ &= \lim_{j \rightarrow \infty} \frac{T_k f(x) - T_k f(x - h_j e_i)}{h_j} \end{aligned}$$

for every  $k \in \mathbf{N}$ .

Let then  $k \in \mathcal{R}f(x)$ , whence it holds a.e. that

$$\frac{M^* f(x + h_j) - M^* f(x)}{h_j} \geq \frac{T_k f(x + h_j e_i) - T_k f(x)}{h_j} \xrightarrow{j \rightarrow \infty} D_i T_k f(x).$$

On the other hand,

$$\frac{M^* f(x) - M^* f(x - h_j e_i)}{h_j} \leq \frac{T_k f(x) - T_k f(x - h_j e_i)}{h_j} \xrightarrow{j \rightarrow \infty} D_i T_k f(x).$$

Clearly above inequalities imply the claim. □

**Theorem 3.3.** *If assumption (3.1) is valid, then  $M^* : W^{1,p}(\mathbf{R}^n) \mapsto W^{1,p}(\mathbf{R}^n)$  is continuous for all  $1 < p < \infty$ .*

*Proof.* Let  $f_j \rightarrow f$  in  $W^{1,p}(\mathbf{R}^n)$  when  $j \rightarrow \infty$ . We may assume that functions  $f_j$  and  $f$  are non-negative. Because we know the continuity of  $M^*$  in  $L^p(\mathbf{R}^n)$ , it suffices to prove that  $\|D_i M^* f_j - D_i M^* f\|_p \rightarrow 0$  for all  $i$ ,  $1 \leq i \leq n$ . Let us assume, on the contrary, that this does not hold. Thus by extracting a subsequence, if needed, we have

$$(3.4) \quad \|D_i M^* f_j - D_i M^* f\|_p > \lambda > 0 \text{ for every } j \in \mathbf{N}.$$

Let then  $R > 0$ . Again, by a possible choice of a subsequence, we may assume that

$$(3.5) \quad \Delta(\mathcal{R}f_j(x), \mathcal{R}f(x)) \rightarrow 0, \text{ for a.e. } x \in B(0, R), \text{ as } j \rightarrow \infty$$

and (by boundedness of  $M^*$  in  $L^p$ )

$$(3.6) \quad M^*(|D(f_j - f)|)(x) \rightarrow 0, \text{ for a.e. } x \in B(0, R), \text{ as } j \rightarrow \infty.$$

Then, let us choose for almost every  $x$  a pair  $2^{-k_j(x)} := 2^{-k_j} \in \mathcal{R}f_j(x)$  and  $2^{-k'_j(x)} := 2^{-k'_j} \in \mathcal{R}f(x)$  so that they minimize the distance between the sets  $\mathcal{R}f_j(x)$  and  $\mathcal{R}f(x)$ . Now (3.5) implies that  $|2^{-k_j} - 2^{-k'_j}| \rightarrow 0$  as  $j \rightarrow \infty$ . Moreover, it follows from Theorem 3.2 and the linearity of operators  $T_k$  that for almost every  $x \in B(0, R)$

$$\begin{aligned} |D_i M^* f_j(x) - D_i M^* f(x)| &= |D_i T_{k_j} f_j(x) - D_i T_{k'_j} f(x)| \\ &\leq |D_i T_{k_j} f_j(x) - D_i T_{k_j} f(x)| + |D_i T_{k_j} f(x) - D_i T_{k'_j} f(x)| \\ &= |T_{k_j}(D_i f_j) - T_{k_j}(D_i f)(x)| + |D_i T_{k_j} f(x) - D_i T_{k'_j} f(x)| \\ &\leq C(M^*(|D(f_j - f)|)(x)) + |D_i T_{k_j} f(x) - D_i T_{k'_j} f(x)| \\ &=: s_j^1(x) + s_j^2(x). \end{aligned}$$

Now (3.6) implies that  $s_j^1 \rightarrow 0$  a.e. as  $j \rightarrow \infty$ . On the other hand, our assumption (3.1) guarantees that for almost every  $x$  the function

$$(3.7) \quad 2^{-k} \mapsto D_i T_k f(x), \quad 0 \mapsto D_i f(x)$$

is uniformly continuous on  $\{2^{-k} : k \in \mathbf{N}\} \cup \{0\}$ . Since  $|2^{-k_j} - 2^{-k'_j}| \rightarrow 0$ , it follows that also  $s_j^2 \rightarrow 0$  a.e. in  $B(0, R)$  as  $j \rightarrow \infty$ . Here  $R > 0$  was arbitrarily chosen and thus we have shown that  $D_i M^* f_j$  converges to  $D_i M^* f$  pointwise a.e.

The final component of the proof is the following uniform pointwise estimate; Using the fact  $|DM^* f| \leq CM^*(|Df|)$  (see e.g. [1]), we get

$$\begin{aligned} |D_i M^* f(x) - D_i M^* f_j(x)| &\leq |D_i M^* f(x)| + |D_i M^*(f_j)(x)| \\ &\leq C(M^*(|Df|)(x) + M^*(|Df_j|)(x)) \\ &= C(M^*(|Df|)(x) + M^*(|Df_j| - |Df| + |Df|)(x)) \\ &\leq 2CM^*(|Df|)(x) + CM^*(|Df_j| - |Df|) =: F(x) + F_j(x). \end{aligned}$$

Here  $F \in L^p(\mathbf{R}^n)$  and  $F_j \rightarrow 0$  in  $L^p(\mathbf{R}^n)$ . An elementary modification of the Lebesgue dominated convergence Theorem combined with the pointwise a.e. convergence implies the desired contradiction with (3.4).  $\square$

**Remark 3.4.** For the continuity of the discrete maximal operator in the case of (1.2), where suprema is considered with respect to the everywhere dense set  $\mathbf{Q}_+$ , the above proof would apply if it is assumed that for almost every  $x \in \mathbf{R}^n$  the function

$$(3.8) \quad r \rightarrow D_i T_r f(x)$$



is uniformly continuous on  $\mathbf{Q}_+$ .

#### 4. Two examples and important comments

What can we say about the sharpness of Theorem 3.3? Do we have any example on the partition of unity related to the coverings  $\{I_j^k\}$  such that assumption (3.1) holds and the corresponding  $M^*$  is continuous? Or does there exist much easier examples of discontinuity than it was given in section 2? In this section these questions are discussed.

**4.1. Standard partition of unity in  $\mathbf{R}$ .** Consider the case, where the most simple partition of unity related to the coverings  $\{I_j^k\}$ ,  $I_j^k$  as in Section 2, is used. Indeed, let  $\psi_j^k$  be such that  $\psi_j^k$  is  $\frac{1}{2}$  on  $I_j^k$  and tends linearly to zero on intervals  $I_{j-1}^k$  and  $I_{j+1}^k$ . Let us call this as the *standard* partition of unity on  $\mathbf{R}$ . We show below that this partition of unity satisfies the assumption (3.1) given in Section 3, and therefore the corresponding discrete maximal operator is continuous in  $W^{1,p}(\mathbf{R})$ , when  $1 < p < \infty$ .

**Proposition 4.1.** *Condition (3.1) holds for the above defined standard partition of unity on  $\mathbf{R}$ .*

*Proof.* Suppose that  $f \in W_{loc}^{1,1}(\mathbf{R})$ , denote the center points of intervals  $I_j^k$  by  $x_j^k$  and fix  $k \in \mathbf{N}$ . Recall first that  $f$  is differentiable a.e. and thus it suffices to show the claim at points  $x \in \mathbf{R}$  such that

$$(4.1) \quad f(y) = f(x) + (y - x)Df(x) + \epsilon_x(y - x),$$

where  $\frac{\epsilon_x(h)}{|h|} \rightarrow 0$  as  $h \rightarrow 0$ . Then for every  $k$  denote by  $I_{j'}^k$  the unique interval so that  $x \in I_{j'}^k$ . By the definition of the discrete convolutions  $T_k$  we have

$$\begin{aligned} DT_k f(x) &= \sum_{j \in \mathbf{Z}} f_{B(x_j^k, 3r_k)} D\psi_j^k(x) \\ &= f_{B(x_{j'-1}^k, 3r_k)} D\psi_{j'-1}^k(x) + f_{B(x_{j'}^k, 3r_k)} D\psi_{j'}^k(x) + f_{B(x_{j'+1}^k, 3r_k)} D\psi_{j'+1}^k(x) \\ &= f_{B(x_{j'-1}^k, 3r_k)} \left(-\frac{1}{4r_k}\right) + f_{B(x_{j'+1}^k, 3r_k)} \frac{1}{4r_k} \\ &= \frac{1}{4r_k} \left( f_{B(x_{j'+1}^k, 3r_k)} - f_{B(x_{j'-1}^k, 3r_k)} \right). \end{aligned}$$

Substituting the formula (4.1) into the above average integrals the above equation is reduced by elementary calculations to

$$DT_k f(x) = Df(x) + \frac{1}{4r_k} \left( \frac{1}{6r_k} \left( \int_{B(x_{j'+1}^k, 3r_k)} \epsilon_x(y - x) dy + \int_{B(x_{j'-1}^k, 3r_k)} \epsilon_x(y - x) dy \right) \right).$$

Finally, it is easy to check that  $\frac{\epsilon_x(h)}{|h|} \rightarrow 0$  implies that the error term above tends to zero as  $k \rightarrow \infty$ . This completes the proof.  $\square$

At this moment we do not know any concrete partitions of unity in dimensions  $n \geq 2$  for which the assumption (3.1) holds. The complete understanding of this question seems to require its own survey, also in the 1-dimensional case. Thus, we would like to pose the following question:

**Question 4.2.** For general dimension  $n$ , characterize (in some natural way) those partitions of unity for which (3.1) holds or, alternatively, for which

$$\|DT_k f - Df\|_p \rightarrow 0 \text{ as } k \rightarrow \infty.$$

**4.2. Are there more simple counterexamples?** There is evidently a big 'gap' between the assumption (3.1) guaranteeing the continuity and the complexity of the partition of unity in our counterexample in Section 2. There are examples related to the coverings  $\{I_j^k\}$  of Section 2 for which (3.1) is not valid and which are simpler than there. Perhaps the simplest such example is as follows. Suppose that the above defined standard partition of unity is modified so that  $\psi_j^k$  remains the same on the interval  $I_j^k$ , but on the interval  $I_{j+1}^k$  we let  $\psi_j^k(x) = 0$  on the right-hand side of  $I_{j+1}^k$  i.e. on  $[x_{j+1}^k, x_{j+1}^k + r_k]$  and let  $\psi_j^k$  tend linearly from  $\frac{1}{2}$  to 0 on  $[x_{j+1}^k - r_k, x_{j+1}^k]$ . Moreover, on  $I_{j-1}^k$  we define  $\psi_j^k(x) = \frac{1}{2} - \psi_j^k(x + 4r_k)$ . While it is easy to show that in this case (3.1) fails, it turns out that  $M^*$  will still be continuous in  $W^{1,p}(\mathbf{R})$ ,  $1 < p < \infty$ . In the following we sketch a proof for this.

**Theorem 4.3.** *If the partition of the unity is as above then (3.1) fails but  $M^*$  is continuous in  $W^{1,p}(\mathbf{R})$ .*

*Sketch of the proof.* First we remark that Lemma 3.1 and Theorem 3.2 do not use assumption (3.1) and so are applicable here. Assumption (3.1) comes into play in the proof of Theorem 3.3, only because we needed the continuity condition (3.7). In addition to the choices of subsequences in the proof of Theorem 3.3, let us also choose a subsequence of  $(f_j)$  such that for a.e.  $x \in \mathbf{R}$  it holds that

$$(4.2) \quad M^*(Df_j - Df)(x) \rightarrow 0 \text{ as } j \rightarrow \infty.$$

Suppose that  $2^{-k_j(x)} \in \mathcal{R}f_j(x)$  and  $2^{-k'_j(x)} \in \mathcal{R}f(x)$  are chosen as in the proof of Theorem 3.3. When revising the proof, one observes that we had to assume (3.1) only to get from the fact

$$|2^{-k_j(x)} - 2^{-k'_j(x)}| \rightarrow 0 \text{ as } j \rightarrow \infty$$

that

$$(4.3) \quad |D_i T_{k_j} f(x) - D_i T_{k'_j} f(x)| \rightarrow 0 \text{ as } j \rightarrow \infty.$$

It is also clear that this convergence may only fail in the case where  $2^{-k_j(x)}$  and  $2^{-k'_j(x)}$  both tend to zero as  $j \rightarrow \infty$  (we have to choose a subsequence if needed). In particular, in this possible bad set  $M^*f(x) = f(x)$  a.e. The crucial point is that now, if  $|Df(x)| \neq 0$ , then the 'antisymmetry' of our partition of unity guarantees that a.e. in this set actually also  $M^*f_j(x) = f_j(x)$  if  $j$  is big enough. To show this, consider first the case, where  $Df(x) > 0$  (and  $f$  is differentiable at  $x$ ). In this case, it is easy to check that the antisymmetry of our partition of unity implies  $T_k f(x) > f(x)$  for  $k$  large enough. Especially,  $M^*f(x) > f(x)$ , which contradicts the assumption  $M^*f(x) = f(x)$ . Thus, it suffices to check the case  $Df(x) < 0$ . In this case, since each discrete convolution weights the values on the right-hand side of  $x$ , one can use the assumption (4.2) and straightforwardly compute that there exists  $j(x)$  such that  $M^*f_j(x) = f_j(x)$  for any  $j \geq j(x)$ .

Summing up, it follows that in this possible bad set it holds a.e. that

$$(4.4) \quad |D_i M^* f(x) - D_i M^* f_j(x)| = |D_i f_j(x) - D_i f(x)|$$

for any sufficiently big  $j$ . Since we assumed (or may assume) that  $D_i f_j \rightarrow D_i f$  a.e., this case is not a problem anymore.

After that, we only have to check the case  $Df(x) = 0$ . However, it is again easy to see that in this case the convergence  $D_i T_k f(x) \rightarrow D_i f(x) = 0$  as  $k \rightarrow \infty$  is valid (contrary to case  $Df(x) \neq 0$ ) implying finally the desired convergence (4.3).  $\square$

The above proof shows that for the continuity of the discrete maximal operator, assumption (3.1) is far from sharp, but can evidently be replaced by a suitable assumption about the 'antisymmetry' of the partition of unity. In the light of the above example one might also see the challenges which have to be faced when trying to construct a discontinuous discrete maximal operator in  $W^{1,p}(\mathbf{R}^n)$ . Namely, avoiding (3.1) roughly saying means that for any linear function  $L \neq 0$  it holds that  $DT_k L \not\rightarrow DL$  as  $k \rightarrow \infty$  (say, in  $W^{1,p}(\mathbf{R}^n)$ ). Furthermore, if this holds in a manner that is 'too strict', we can not avoid the same phenomenon as in the proof of the previous example and thus it follows that in the only possible 'bad' set,

$$(4.5) \quad \{x : M^* f(x) = |f(x)|, |Df(x)| \neq 0\},$$

it typically happens that also  $M^* f_j(x) = |f_j(x)|$  for suitably large  $j$  if  $f_j \rightarrow f$  in  $W^{1,p}(\mathbf{R}^n)$ .

## References

- [1] AALTO, D., and J. KINNUNEN: The discrete maximal operator in metric spaces. - J. Anal. Math. 111, 2010, 369–390.
- [2] BUCKLEY, S. M.: Is the maximal function of a Lipschitz function continuous? - Ann. Acad. Sci. Fenn. Math. 24, 1999, 519–528.
- [3] COIFMAN, R. R., and G. WEISS: Analyse harmonique non-commutative sur certains espaces homogènes. - Lecture Notes in Math. 242, Springer-Verlag, 1971.
- [4] HAJLÁSZ, P., and J. ONNINEN: On boundedness of maximal functions in Sobolev spaces. - Ann. Acad. Sci. Fenn. Math. 29, 2004, 167–176.
- [5] KINNUNEN, J.: The Hardy–Littlewood maximal function of a Sobolev function. - Israel J. Math. 100, 1997, 117–124.
- [6] KINNUNEN, J., and V. LATVALA: Lebesgue points for Sobolev functions on metric spaces. - Rev. Mat. Iberoamericana 18:3, 2002, 685–700.
- [7] KINNUNEN, J., and P. LINDQVIST: The derivative of the maximal function. - J. Reine Angew. Math. 503, 1998, 161–167.
- [8] KINNUNEN, J., and H. TUOMINEN: Pointwise behaviour of  $M^{1,1}$  Sobolev functions. - Math. Z. 257:3, 2007, 613–630.
- [9] LUIRO, H.: Continuity of the Hardy–Littlewood maximal operator in Sobolev spaces. - Proc. Amer. Math. Soc. 135:1, 2007, 243–251.
- [10] LUIRO, H.: Regularity properties of maximal operators. - Rep. Univ. Jyväskylä Dept. Math. Stat. 114, 2008.
- [11] MACÍAS, R. A., and C. SEGOVIA: A decomposition into atoms of distributions on spaces of homogeneous type. - Adv. Math. 33, 1979, 271–309.
- [12] SEMMES, S.: Finding curves on general spaces through quantitative topology with applications to Sobolev and Poincaré inequalities. - Selecta Math. (N.S.) 2, 1996, 155–295.