Fine properties of maximal operators

Cristian González Riquelme
Advisor: Emanuel Carneiro
Instituto de Matemática Pura e Aplicada

Thesis submitted in fulfillment of the requirements for the degree of Doctor in Philosophy in Mathematics to the Postgraduate Program in Mathematics at Instituto Nacional de Matemática Pura e Aplicada.

$\otimes$

A mi mamá y mi papá.

Abstract

In this Ph.D. thesis we investigate fine properties of maximal operators in continuous and discrete settings. We obtain both quantitative and qualitative results regarding the oscillatory behavior of such operators. More concretely, we deal with the following topics: (i) boundedness and continuity for maximal operators in Sobolev and bounded variation spaces; (ii) sharp inequalities for maximal operators on graphs.

Acknowledgements

This thesis is the result of a journey that started many years ago. Consequently, I have to acknowledge many people.

First, I start by acknowledging the members of my thesis committee that include: Emanuel Carneiro, Milton Jara, Carlos Perez, Juha Kinnunen, Felipe Linares and Pedro Tradacete. Their comments and questions about this thesis were very helpful.

Secondly, I would like to acknowledge all the professors at IMPA; their very instructive courses were essential for my mathematical education. Also, I would like to thank all the administrative staff from IMPA; they were very helpful whenever needed, making my life at IMPA easier. Also, I am thankful to ICTP for being a very welcoming place in my stay during my PhD. I recognize, also, the financial support of both CAPES and the STEP program.

I would not be presenting this thesis if I had not been at IMPA in the first place, for this, I am thankful to both Milton Jara and Claudio Landim. They invited me to come to IMPA and this thesis is one of the consequences of this decision.

Many of the results contained in this thesis would not have been possible without the help of my collaborators: David Beltran, Dariusz Kosz, José Madrid and Julian Weigt. Also, I would like to acknowledge Andrés Chirre, Mateus Sousa and Oscar Quesada for their willingness to help me in any mathematical doubt that I had.

My time at IMPA would have not been the same without the company of many friends, this includes (but it is not restricted to): Oscar Riaño, Gonzalo Jimenez, Enrique Chavez, Jesús Ramos, Diego Navarro, Clarisse Neto, Leonardo Dinamarca, Juan Paucar, Alex Agudelo, Miguel Ibieta and Jennifer Loria. In particular, I am very grateful to Zoraida Fernandez and Enzo Aljovin for all the lunches and coffees that we shared.

Also, I very much appreciate the people that I met during my stay at ICTP and that made my stay there more enjoyable. In particular, I would like to thank Asem Abdelraouf, Giulia Gugiatti, Andrea Olivo, Alejandra Rincón and Zakarias Sjöström for their friendly welcome at ICTP.

My mathematical journey started a long time ago, when preparing for the mathematical olympiads. I want to express my acknowledgments to some friends that I made there and with whom, to this day, I enjoy talking about math as much as 10 years ago: Felipe Arbulú, Fernando Figueroa, Sebastián Pávez and Renato Velozo. Some people that I met later in Beaucheff that have been also good company are Matías Alvarado, Felipe Alarcón, Christopher Cabezas and José Manuel Palacios. I would like to thank Matías, in particular, for his support during the pandemic.

A special place in my acknowledgements is reserved to my friends Daniel Contreras, Hernán Iriarte and Roberto Villaflor. I lived with them for many years, and we spent many endless and fruitful discussions about many topics, including mathematics. I would like to express my gratitude to Daniel for always being willing to help me in any administrative matter; my time at IMPA certainly would have been much more difficult without his advice.

Also, I am very grateful to Angélica Garrido for her support and all her help during my PhD. Also, for making with her presence these years much more enjoyable.

This thesis certainly would not have been possible without the constant encouragement, help and advice of Emanuel Carneiro, my advisor. He introduced me to the topic that we treat in this thesis and guided me from the beginning to the end. His perspective on mathematics certainly has helped to shape mine. Moreover, I would like to thank him for never losing hope in me.

Finally, I would like to express my deepest gratitude to my mom and my dad, Elena and Héctor, for all their support and love. Also, I would like to thank my brother, Héctor, for the many good moments that we spent together.

Contents

Introduction

This thesis is inserted in the field of harmonic analysis. We investigate some fine properties of maximal operators, which are central objects in this area. Our main aim is to provide a deeper understanding of the oscillatory behavior of such objects. More concretely, we deal with the following topics: (i) boundedness and continuity of maximal operators in Sobolev and bounded variation spaces; (ii) sharp inequalities for maximal operators on graphs. The progress we made in these topics resulted in the following research articles (presented in chronological order), on which this document is based:

- [CGR21] Gradient bounds for radial maximal functions (with E. Carneiro), *Ann. Fenn. Math.*, 46(1), 495-521, 2021.
- [GR20] Sobolev regularity of polar fractional maximal functions, *Nonlinear Anal.*, 198: 111889, 2020.
- [GRM21] Sharp inequalities for maximal operators on finite graphs (with J. Madrid), *J. Geom. Anal.*, 31(10); 9708-9744, 2021.
- [CGRM20] Sunrise strategy for the continuity of maximal operators, (with E. Carneiro and J. Madrid), preprint, to appear in *J. Anal. Math.*, 2020.
- [GRK21] BV continuity for the uncentered Hardy–Littlewood maximal operator (with D. Kosz), *J. Funct. Anal.*, 281(2):109037, 2021.
- [GRM22] Sharp inequalities for maximal operators on finite graphs, II (with J. Madrid), J. Math. Anal. Appl., 506(2): 125647, 2022.
- [BGRMW21] Continuity of the gradient of the fractional maximal operator on $W^{1,1}(\mathbb{R}^d)$ (with D. Beltran, J. Madrid and J. Weigt), preprint, to appear in *Math. Res. Lett.*, 2021.
- [GR21b] On the continuity of maximal operators of convolution type at the derivative level, preprint, to appear in *Israel J. Math.*, 2021.
- [GR21a] Continuity for the one-dimensional centered Hardy-Littlewood maximal operator at the derivative level, preprint, 2021.

0.1 A historical summary

Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$. We define the Hardy-Littlewood maximal operator for $x \in \mathbb{R}^d$ as

$$Mf(x) = \sup_{r>0} \frac{\int_{B(x,r)} |f|}{|B(x,r)|} =: \sup_{r>0} \fint_{B(x,r)} |f|, \tag{1}$$

where $|X|$ is the Lebesgue measure of the measurable set $X \subset \mathbb{R}^d$. We define $\widetilde{M}$ as its uncentered version. These operators are bounded from $L^p(\mathbb{R}^d)$ to itself when $p > 1$ (see [Ste70, Chapter III, Theorem 2]). This boundedness provides a control over the size of the maximal functions and is very useful for applications, for instance, the Lebesgue differentiation theorem and for the a.e. pointwise convergence of solutions for the PDEs to the initial/boundary data.

Understanding the oscillatory behavior of this operator is another interesting theme of study. That is, how much can Mf oscillate given the oscillation of the original f? This question, in several cases, is more complicated than the one related to the size of the maximal function, given that M combines the smoothing effect of taking averages with the sometimes rough process of taking the supremum. This interplay suggests, in some cases, that it is plausible to expect certain smoothing properties of M. However, actually establishing such results is generally a non-trivial task.

0.1.1 Boundedness and continuity of maximal operators in Sobolev spaces and bounded variation spaces

Sobolev spaces are a natural framework to consider the oscillation of a function. Kinnunen, in his seminal work [Kin97], proved that the map $M : W^{1,p}(\mathbb{R}^d) \to W^{1,p}(\mathbb{R}^d)$ is bounded for $p > 1$. This theorem provides then the first result concerning some form of variation of a maximal function and its proof can be adapted to more general maximal operators. Motivated by this result, several interesting contributions to the regularity theory of maximal operators have been made over the past decades. These contributions were partially summarized in the survey [Car].

The endpoint case $p = 1$ of Kinnunen's result certainly does not hold, since $Mf \notin L^1(\mathbb{R}^d)$ whenever $f \in L^1(\mathbb{R}^d)\setminus\{0\}$. However, since we are interested in the behavior at the derivative level of these operators, the analogue of Kinnunen's result at the endpoint $p = 1$ would be given by the following conjecture.

$W^{1,1}$-conjecture: The map $f \mapsto \left|\nabla Mf\right|$ from $W^{1,1}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$ is bounded.

This was formally proposed for the first time by Hajlasz and Onninen in [HO04] and is an important open problem that drives our research program. Note that the conjecture also involves proving that if $f \in W^{1,1}(\mathbb{R}^d)$ then Mf is weakly differentiable. It has been solved just in dimension one by Kurka in the centered case [Kur15] and by Tanaka in the uncentered case [Tan02]. The uncentered result was later refined by Aldaz and Pérez-Lázaro in [APL07], where a sharp version of this boundedness was obtained. Moreover, they also proved that the map $f \mapsto \widetilde{M}f$ is bounded from $BV(\mathbb{R})$ to itself. Here, the space of functions of bounded variation $BV(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R}; \operatorname{Var} f < \infty\}$ is endowed with the norm $\|f\|_{BV} := |f(-\infty)| + \operatorname{Var} f$, where $\operatorname{Var} f$ denotes the total variation of a real-valued function. A version of the $W^{1,1}$-conjecture for uncentered cubes was very recently obtained by Weigt [Wei21b].

The continuity for this type of operators has also been extensively studied. We notice that since these objects are not necessarily sublinear, any boundedness result would not imply directly the continuity of the corresponding map. In this regard, Luiro [Lui07] proved that both M and $\widetilde{M}$ are continuous from $W^{1,p}(\mathbb{R}^d)$ to itself, when $p > 1$. This solved a question attributed to T. Iwaniec [HO04, Question 3]. The methods developed in Luiro's work can be adapted to several other maximal operators in the range $p > 1$. The endpoint case $p = 1$ is significantly more involved. For the uncentered Hardy-Littlewood maximal operator, the continuity of the map

$$f \mapsto \left(\widetilde{M}f\right)' \tag{2}$$

from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$ was proved by Carneiro, Madrid and Pierce in [CMP17]. In this thesis we address a number of problems in this research theme, concerning boundedness and continuity of maximal operators. In the work [GR21a] we established the continuity for the centered Hardy-Littlewood maximal operator, solving a question posed by Carneiro, Madrid and Pierce in [CMP17, Question A] and establishing, in the one-dimensional case, the endpoint version of [HO04, Question 3] at the derivative level.

Theorem A. *We have that the map*

$$f \mapsto (Mf)'$$

is continuous from $W^{1,1}(\mathbb{R})$ *to* $L^1(\mathbb{R})$.

We now turn our attention to the space of functions of bounded variation. The functions belonging to this space lack some of the regularity properties of $W^{1,1}(\mathbb{R})$ that were relevant in [CMP17], making the implementation of some of the tools presented there unsuitable. In [GRK21] we proved the following result, solving a question originally posed in [CMP17, Question B]:

Theorem B. *The map* $f \mapsto \widetilde{M}f$ *is continuous from* $BV(\mathbb{R})$ *to itself.*

Considering the progress made in this thesis, we summarize the situation of the *endpoint continuity program* (originally proposed in [CMP17, Table 1]) in the table below. The word YES in a box means that the continuity of the corresponding map has been proved, whereas the word NO means that it has been shown that it fails. We notice that after this work the only open problem in this program is to determine if the map $f \mapsto Mf$ is continuous from $BV(\mathbb{R})$ to itself, marked with OPEN in the table below.

Table 1: Endpoint continuity program

———	$W^{1,1}-$continuity; continuous setting	$BV-$continuity; continuous setting	$W^{1,1}-$continuity; discrete setting	$BV-$continuity; discrete setting
Centered classical maximal operator	YES: Theorem A	OPEN	YES[2]	YES[4]
Uncentered classical maximal operator	YES[1]	YES: Theorem B	YES[2]	YES[1]
Centered fractional maximal operator	YES[5]	NO[1]	YES[3]	NO[1]
Uncentered fractional maximal operator	YES[4]	NO[1]	YES[3]	NO[1]

[1] Result previously obtained in [CMP17].
[2] Result previously obtained in [CH12, Theorem 1].
[3] Result previously obtained in [CM17, Theorem 3].
[4] Result previously obtained in [Mad19].
[5] Result previously obtained in [BM20].
[6] Result previously obtained in [GRK21]

Another important result regarding the boundedness of these maps was due to Luiro [Lui18]. He proved that the map

$$f \mapsto \left|\nabla \widetilde{M} f\right|, \tag{3}$$

is bounded from $W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$, where $W^{1,1}_{\text{rad}}(\mathbb{R}^d) \subset W^{1,1}(\mathbb{R}^d)$ is the subspace of radial functions. Our next result, obtained in [CGRM20], establishes the continuity of this map. This provides the first continuity result at the endpoint for such operators in higher dimensions.

Theorem C. *The map* $f \mapsto \left|\nabla \widetilde{M} f\right|$ *is continuous from* $W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ *to* $L^1(\mathbb{R}^d)$.

One crucial step in the proof of this result is the construction of suitable higher dimensional analogues of lateral versions of the Hardy-Littlewood maximal operator. This construction allows us to establish the analogous continuity results for several other maximal operators of interest.

Maximal operators associated to smooth kernels

Another point of view in the topic presented above was introduced by Carneiro and Svaiter [CS13], who extended some of the previous results for maximal operators associated to smooth kernels. Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}_{\geq 0}$ be a radially non increasing function with $\int_{\mathbb{R}^d} \phi(x)\,\mathrm{d}x = 1$. We define, as usual, $\phi_t(x) := \frac{1}{t^d}\phi(\frac{x}{t})$. Then, given an initial datum $u_0 : \mathbb{R}^d \to \mathbb{R}$, we define the extension $u : \mathbb{R}^d \times (0,\infty) \to \mathbb{R}$ as

$$u(x,t) = |u_0| * \phi_t(x).$$

The maximal operator associated to the kernel ϕ is defined as

$$u^*(x) = \sup_{t>0} u(x,t), \tag{4}$$

where we omit the dependence to ϕ as it is clear from the context. This notion recovers the classical Hardy-Littlewood maximal operator, by choosing $\phi = \frac{\chi_{B(0,1)}}{|B(0,1)|}$. The following kernels are of major relevance for our purposes:

$$\varphi_1(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}}\,\frac{1}{(|x|^2+1)^{(d+1)/2}} \qquad \text{(Poisson kernel)} \tag{5}$$

$$\varphi_2(x) = \frac{1}{(4\pi)^{d/2}}\,e^{-|x|^2} \qquad \text{(Heat kernel).} \tag{6}$$

In [CS13], the authors proved that the map

$$u_0 \mapsto (u^*)' \tag{7}$$

is bounded from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$ for $\phi \in \{\varphi_1, \varphi_2\}$. A property that plays a major role in the proof of this result is the so called *subharmonicity property*. That is, u^* is subharmonic in the set $\{x \in \mathbb{R}; u^*(x) > u(x)\}$. The proof of such result is based on relations between these kernels and the underlying partial differential equation. This property holds in higher dimensions, and is used in [CS13] to prove that the L^2-norm of the gradient of u^* is not greater than the L^2-norm of the gradient of the original u. Some extensions of these results were later obtained by Carneiro, Finder and Svaiter in [CFS18] and by Bortz, Egert and Saari [BES19]. Another related point of view was proposed by Pérez, Picon, Saari and Sousa in [PPSS18], where a version of the $W^{1,1}$-conjecture for smooth kernels in the context of Hardy-Sobolev spaces was settled.

In this context, in [CGR21] we proved the first result in higher dimensions for centered maximal operators at the endpoint:

Theorem D. *Let $\phi \in \{\varphi_1, \varphi_2\}$. The map $u_0 \mapsto |(u^*)'|$ is bounded from $W^{1,1}_{\mathrm{rad}}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$.*

This result is based on a *comparison criterion*, where Luiro's boundedness result [Lui18, Theorem 3.11] plays a major role, along with the *subharmonicity property* of the kernels ϕ aforementioned. Also, in [GR21b] we obtained the analogue of the main result of [CMP17] for these operators. This was the first result in this direction for a centered maximal operator.

Theorem E. *Let $\phi \in \{\varphi_1, \varphi_2\}$. Then the map*

$$u_0 \mapsto (u^*)'$$

is continuous from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$.

The fractional Hardy-Littlewood maximal operator

Other kinds of operators of interest are the so-called fractional maximal operators. These have in general a smoother behavior since, in this case, the radii at which the suprema are attained are larger. For every $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ and $0 < \beta < d$ we define its centered Hardy-Littlewood fractional maximal function as

$$M_\beta f(x) := \sup_{r>0} r^\beta \fint_{B(x,r)} |f|. \tag{8}$$

We call $\widetilde{M}_\beta$ its uncentered version. For $q := \frac{dp}{d-\beta p}$, Kinnunen and Saksman in [KS03] proved that the map $M_\beta : W^{1,p}(\mathbb{R}^d) \to W^{1,q}(\mathbb{R}^d)$ is bounded when $p > 1$. This extends Kinnunen's result about the classical Hardy-Littlewood maximal operator. Recently, after partial progress made in [CM17, LM19, BM20], it was proved by Weigt [Wei21a] that the map

$$f \mapsto |\nabla M_\beta f|,$$

is bounded from $W^{1,1}(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$, when $0 < \beta < d$. This establishes the fractional version of the $W^{1,1}$-conjecture. He also proved the analogous result for $\widetilde{M}_\beta$. In this context, in [BGRMW21] we established the continuity of those maps.

Theorem F. *Let $0 < \beta < d$ and $q = \frac{d}{d-\beta}$. The map $f \mapsto |\nabla M_\beta f|$ is continuous from $W^{1,1}(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$. The same holds for $\widetilde{M}_\beta$.*

This theorem establishes the first continuity result in full generality at the endpoint $p = 1$. Previous radial versions of this result were obtained in [BM19, BM20].

Maximal operators on the sphere

Part of this theory was extended to the context of the sphere $\mathbb{S}^d$. Due to the geometric differences between the Euclidean and spherical spaces, several tools used in the classical setting are not immediately available for the $\mathbb{S}^d$ case. The first results in this direction were established in [CFS18], where the authors established the L^2-norm reduction of the gradient for some maximal operators of convolution type, when acting on the sphere. This result was based on the aforementioned *subharmonicity property*, this time in the spherical context.

We call a function $f : \mathbb{S}^d \to \mathbb{R}$ polar if it is invariant by rotations acting on $\mathbb{S}^d$ that fix the north pole. Here $\widetilde{\mathcal{M}}$ is the uncentered Hardy-Littlewood maximal operator defined over the sphere. In this context, in [CGR21] we established the following, where:

Theorem G. *If $f \in W^{1,1}(\mathbb{S}^d)$ is a polar function, then $\widetilde{M}f$ is weakly differentiable and*

$$\|\nabla \widetilde{\mathcal{M}} f\|_1 \lesssim_d \|\nabla f\|_1.$$

This establishes the spherical version of Luiro's result for radial functions. Some of the major difficulties that needed to be overcome arose from the geometric arguments in Luiro's original work [Lui18]. There, the invariance by dilations of the Euclidean space plays a subtle but important role. Novel geometric estimates were needed to surpass this difficulty. Also, an adaptation of the aforementioned *comparison criterion* was needed in order to establish the analogous result for some maximal operators of convolution type on $\mathbb{S}^d$. Later, in [GR20], by going further in this direction we established the analogue of the result of Luiro and Madrid [LM19, Theorem 1.1] on $\mathbb{S}^d$.

0.1.2 Sharp constants for maximal operators on finite graphs

Discrete analogues in harmonic analysis have been an active topic of research over the last decades. For instance, in [MSW02], boundedness for discrete spherical maximal functions were obtained, and in [Mel03], a discretization approach was used in order to compute the norm $\|M\|_{L^1\to L^{1,\infty}(\mathbb{R})}$. In this thesis, we are also interested in sharp constants for the centered Hardy-Littlewood maximal operator on finite graphs, both in the case of the p-norm and the p-variation of this maximal operator.

We define the Hardy-Littlewood maximal function of f along G at the point $v \in V$ by

$$M_G f(v) := \max_{r\geq 0} \frac{1}{|B(v,r)|} \sum_{m\in B(v,r)} |f(m)|,$$

where $B(v,r) = \{m \in V; d_G(v,m) \leq r\}$, where d_G is the metric induced by the edges of G (that is, the distance between two vertices is the number of edges in a shortest path connecting them) and $|X|$ is the quantity of elements of a set X. We define

$$\|M_G\|_p = \sup_{\substack{f:V\to\mathbb{R}_+ \\ f\neq 0}} \left(\frac{\sum_{v\in V} M_G f(v)^p}{\sum_{v\in V} |f(v)|^p}\right)^{\frac{1}{p}}.$$

One of our major goals is to try to understand how these constants behave and compute them when possible. In [ST16], Soria and Tradacete proved that for every graph $G_n = (V,E)$ of n vertices and every $p \in (0,1]$, we have

$$\|M_{K_n}\|_p \leq \|M_{G_n}\|_p \leq \|M_{S_n}\|_p, \tag{9}$$

where K_n is the complete graph and S_n is the star graph (where all the vertices are only connected to a central one). This, combined with the fact that K_n and S_n are model graphs of particular interest by their combinatorial and geometric properties, suggests that these graphs are a natural place to start developing our theory. For the range $0 < p \leq 1$ in [ST16] the authors computed $\|M_{S_n}\|_p$ and $\|M_{K_n}\|_p$. One of the purposes of our works [GRM21] and [GRM22] was to extend the understanding of this problem for $p > 1$. This case dramatically differs from the case $p \leq 1$. For instance, the concavity of the function $x \mapsto x^p$ for $p \leq 1$ was used in [ST16] to prove that $\|M_G\|_p$ is attained by some *Dirac's delta*, and this is not the case for general G when $p > 1$. It was proved by Soria and Tradacete (see [ST16, Proposition

3.4]) that

$$\left(1 + \frac{n-1}{2^p}\right) \leq \|M_{S_n}\|_p^p \leq \left(\frac{n+5}{2}\right). \tag{10}$$

They also presented similar bounds for $\|M_{K_n}\|_p^p$. We notice that both lower bounds go to 1 when $p \to \infty$. In [GRM21] both $\|M_{S_n}\|_2$ and $\|M_{K_n}\|_2$ were exactly computed, and extremizers were provided. In [GRM22], we observed that for $p > 1$ the extremizers for $\|M_{K_n}\|_p$ only take two values, and the extremizers for $\|M_{S_n}\|_p$ take at most three values. This last result can be improved for $p \in (1,2)$, where in fact the extremizers have a similar profile as in the case $p = 2$. Also, we obtained the following asymptotic result:

Theorem H. *For $n \geq 25$ we have that*

$$\lim_{p\to\infty} \|M_{S_n}\|_p^p = \frac{1+\sqrt{n}}{2}.$$

This, in particular, improves qualitatively the aforementioned estimate (10).

Best constants for the p-variation of maximal functions

For a finite graph $G = (V, E)$ we write $(v_1, v_2) := e \in E$ if the edge e connects v_1 with v_2. For a function $g : V \to \mathbb{R}_+$ we write

$$\operatorname{Var}_p g = \left(\sum_{(v_1,v_2)=e\in E} |g(v_1) - g(v_2)|^p\right)^{\frac{1}{p}},$$

and we define

$$\mathbf{C}_{G,p} = \sup_{\substack{f:V\to\mathbb{R} \\ \operatorname{Var} f\neq 0}} \frac{\operatorname{Var}_p M_G f}{\operatorname{Var}_p f}.$$

Motivated by the aforementioned results for the p-norm and the works about the p-variation in the Euclidean setting, one can ask the following.

Question. *Given a finite graph G and $p > 1$, what is the value of $\mathbf{C}_{G,p}$?*

Liu and Xue proposed some conjectures related to this question in [LX20]. They conjectured that $\mathbf{C}_{S_n,p} = 1 - \frac{1}{n}$ for $0 < p \leq 1$ and $\mathbf{C}_{K_n,p} = 1 - \frac{1}{n}$ for $0 < p < \infty$. In both cases this value is attained for some appropriate *Dirac's delta*. They also proved both conjectures for $n \leq 3$. In our work [GRM21], we proved both conjectures for a large range of p and general n. One of our results is the following.

Theorem I. *For every $p \in \left(\frac{\log 4}{\log 6}, \infty\right)$ we have $\mathbf{C}_{K_n,p} = 1 - \frac{1}{n}$.*

The most involved case of this result occurs when $p < 1$, due in particular to the concavity of the map $x \mapsto x^p$ in that range. Therefore, a more refined method is needed in that case. The strategy developed in [GRM21] is based on an inductive procedure, where the geometric properties of K_n play a major role. We also settled the aforementioned question for the star graph S_n in the range $p \in [\frac{1}{2}, 1]$. These results (and their proofs) show that the geometric properties of the graphs play a significant role in these kinds of questions.

0.2 Organization

Over the next nine chapters we elaborate on the brief outline presented in this Introduction. This thesis can be broadly divided into three parts, as described below.

In the first two chapters we present our progress regarding boundedness problems for maximal operators:

- In Chapter 1 we discuss developments made by the author (in colaboration with E. Carneiro) in problems regarding radial versions of (7). When ϕ is the Poisson or the heat kernel, we prove that the map (7) is bounded when restricted to radial functions. This is, we prove Theorem D. This chapter is based on the paper [CGR21].

- In Chapter 2 we discuss versions of our problems when in the sphere setting. We discuss the endpoint Sobolev boundedness when acting on $\mathbb{S}^d$ of the uncentered Hardy-Littlewood maximal operator, the fractional Hardy-Littlewood maximal operator and maximal operators of convolution type, when restricted to polar functions. In particular, we prove Theorem G. This chapter is based on the papers [CGR21] and [GR20].

In the next five chapters we present our progress regarding continuity problems for maximal operators:

- In Chapter 3 we discuss developments made by the author (in colaboration with D. Beltran, J. Madrid and J. Weigt) on the continuity at the derivative level of the fractional maximal operator in higher dimensions. In particular, we prove Theorem F. This chapter is based on the paper [BGRMW21].

- In Chapter 4 we discuss developments made by the author (in colaboration with E. Carneiro and J. Madrid) in problems regarding the continuity of (3) in the radial setting. In particular, we prove Theorem C. This chapter is based on the paper [CGRM20].

- In Chapter 5 we discuss developments made by the author (in colaboration with D. Kosz) regarding the continuity of the uncentered Hardy-Littlewood maximal operator from $BV(\mathbb{R})$ to itself. In particular, we prove Theorem B. This chapter is based on the paper [GRK21].

- In Chapter 6 we discuss developments made by the author regarding the continuity at the derivative level for maximal operators of convolution type. In particular, we prove Theorem E. This chapter is based on the paper [GR21b].

- In Chapter 7 we discuss developments made by the author regarding the continuity at the derivative level for the centered Hardy-Littlewood maximal operator. In particular, we prove Theorem A. This chapter is based on the paper [GR21a].

In the last two chapters we present our progress regarding sharp inequalities for maximal operators acting on finite graphs:

- In Chapters 8 and 9 we discuss developments made by the author (in colaboration with J. Madrid) about sharp constants regarding the size and the variation for maximal functions on finite graphs. In particular, we prove Theorems H and I. These chapters are based on the papers [GRM21] and [GRM22].

A word on notation

In what follows we write $A \lesssim_d B$ if $A \leq CB$ for a certain constant $C > 0$ that may depend on the dimension d. We say that $A \simeq_d B$ if $A \lesssim_d B$ and $B \lesssim_d A$. If there are other parameters of dependence, they will also be indicated. The surface area of the sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ is denoted by ω_d. The characteristic function of a generic set H is denoted by χ_H.

Chapter 1

Gradient bounds for radial maximal functions

1.1 Introduction

In this chapter we investigate the higher dimensional $W^{1,1}$-problem for certain centered maximal operators of convolution type associated to partial differential equations, *in the case of radial data*, establishing a result analogous to that of Luiro [Lui18]. Here, the Poisson kernel and heat kernel are given by (5) and (6), respectively. The maximal function u^* is defined as in (4). In this context we obtain the following:

Theorem 1.1.1. *Let ϕ be given by the Poisson kernel or the heat kernel. If $u_0 \in W^{1,1}(\mathbb{R}^d)$ is radial, then u^* is weakly differentiable and*

$$\|\nabla u^*\|_{L^1(\mathbb{R}^d)} \lesssim_d \|\nabla u_0\|_{L^1(\mathbb{R}^d)}.$$

To our knowledge, this is the first instance of an affirmative result for centered maximal operators, in what concerns the boundedness of the variation, in the higher dimensional setting. The intuitive idea behind the proof of this result is as follows. First we reduce matters to the study of nonnegative functions u_0 with some degree of smoothness, say Lipschitz. We are then able to invoke one of the main results of [CFS18, CS13], that in the *detachment set* $\{u^* > |u_0|\}$ the function u^* is *subharmonic*. The proof of this fact relies on some of the qualitative properties of the underlying partial differential equations (e.g. maximum principles and semigroup property). As observed in [CFS18, Theorem 1 (iv)], this subharmonicity implies a control on the L^2-norm of ∇u^* by the L^2-norm of ∇u_0. To arrive at the L^1-control we use the fact that u^* is pointwise smaller than $\widetilde{M}u_0$. Hence, in the case of radial functions, we have a relatively well-behaved (i.e. subharmonic in the detachment set) function, namely u^*, that is trapped between u_0 and $\widetilde{M}u_0$, and the latter comes with an L^1-control of the gradient by the result of Luiro [Lui07]. As we shall see, these pieces together will ultimately imply the control of the L^1-norm of ∇u^* as well.

1.2 Proof of Theorem 1.1.1

In this section we prove Theorem 1.1.1. Without loss of generality we may assume that u_0 is *real-valued and nonnegative* (or $+\infty$). Assume also that $d \geq 2$, since the result is already known for dimension $d = 1$ from [CFS18, Theorem 1]. Throughout the proof below, with a slight abuse of notation, we identify radial functions of the variable $x \in \mathbb{R}^d$ with their one-dimensional versions of the variable $r \in (0, \infty)$, with the understanding that $r = |x|$. Naturally, if u_0 is radial, the maximal function u^* is also radial. In what follows, variables r, s, t, τ, a, b will be one-dimensional, whereas the variable x is always reserved for $\mathbb{R}^d$. We recall the fact [Ste70, Chapter III, Theorem 2] that

$$u^*(x) \leq Mu_0(x) \leq \widetilde{M}u_0(x) \tag{1.1}$$

for every $x \in \mathbb{R}^d$.

1.2.1 Lipschitz case

Let us first assume that our initial datum u_0 is a Lipschitz function. In this case u^* is also Lipschitz. Reducing matters to radial variables, we claim the following:

$$\int_0^\infty \left|(u^*)'(r)\right| r^{d-1}\, \mathrm{d}r \leq \int_0^\infty \left|u_0'(r)\right| r^{d-1}\, \mathrm{d}r + \int_0^\infty \left|\big(\widetilde{M}u_0\big)'(r)\right| r^{d-1}\, \mathrm{d}r. \tag{1.2}$$

Once we have established (1.2), the theorem follows easily by Luiro's result [Lui18], that bounds the third integral in terms of the second.

Step 1: Partial control by the uncentered maximal function

Let us define the radial detachment set (excluding the origin)

$$A_d = \big\{x \in \mathbb{R}^d \setminus \{0\} \ : \ u^*(x) > u_0(x)\big\}. \tag{1.3}$$

The one-dimensional radial version of this set will be denoted by

$$A_1 = \{|x| \ : \ x \in A_d\}.$$

These are open sets and from [CFS18, Lemma 7] we know that u^* is subharmonic on A_d. Let us write

$$A_1 = \bigcup_{i=1}^\infty (a_i, b_i) \tag{1.4}$$

as a countable union of disjoint open intervals. Let (a, b) denote a generic interval (a_i, b_i) of this union. If u^* had a strict local maximum in (a, b) (that is, a point $t_0 \in (a, b)$ for which there exist c and d with $a < c < t_0 < d < b$ such that $u^*(r) \leq u^*(t_0)$ for $r \in (c, d)$ and $u^*(c), u^*(d) < u^*(t_0)$), we could then take the average of u^* over the ball in $\mathbb{R}^d$ centered at x_0, with $|x_0| = t_0$, and radius $\min\{|t_0 - c|, |t_0 - d|\}$ to reach a contradiction to the subharmonicity of u^* in A_d. Therefore u^* has no strict local maximum in (a, b) and there exists τ with $a \leq \tau \leq b$ such that u^* is non-increasing in $[a, \tau]$ and non-decreasing in $[\tau, b]$. We then have $(u^*)'(t) \leq 0$ a.e. in $a < t < \tau$, and $(u^*)'(t) \geq 0$ a.e. in $\tau < t < b$.

Let us first consider the case $0 < a < b < \infty$. Using (1.1) and integration by parts we get

$$\begin{aligned}
\int_a^b \left|(u^*)'(r)\right| r^{d-1} \,\mathrm{d}r &= -\int_a^\tau (u^*)'(r)\, r^{d-1} \,\mathrm{d}r + \int_\tau^b (u^*)'(r)\, r^{d-1} \,\mathrm{d}r \\
&= u^*(a)\, a^{d-1} + u^*(b)\, b^{d-1} - 2\, u^*(\tau)\, \tau^{d-1} \\
&\qquad + (d-1) \int_a^\tau u^*(r)\, r^{d-2} \,\mathrm{d}r - (d-1) \int_\tau^b u^*(r)\, r^{d-2} \,\mathrm{d}r \\
&\leq u_0(a)\, a^{d-1} + u_0(b)\, b^{d-1} - 2\, u_0(\tau)\, \tau^{d-1} \\
&\qquad + (d-1) \int_a^\tau \widetilde{M} u_0(r)\, r^{d-2} \,\mathrm{d}r - (d-1) \int_\tau^b u_0(r)\, r^{d-2} \,\mathrm{d}r \\
&= u_0(a)\, a^{d-1} - u_0(\tau)\, \tau^{d-1} \\
&\qquad + (d-1) \int_a^\tau \widetilde{M} u_0(r)\, r^{d-2} \,\mathrm{d}r + \int_\tau^b u_0'(r)\, r^{d-1} \,\mathrm{d}r \\
&\leq \int_a^b \left|u_0'(r)\right| r^{d-1} \,\mathrm{d}r + (d-1) \int_a^\tau \widetilde{M} u_0(r)\, r^{d-2} \,\mathrm{d}r.
\end{aligned} \tag{1.5}$$

The last inequality holds since

$$u_0(a)\, a^{d-1} - u_0(\tau)\, \tau^{d-1} \leq -\int_a^\tau u_0'(r)\, r^{d-1} \,\mathrm{d}r \leq \int_a^\tau \left|u_0'(r)\right| r^{d-1} \,\mathrm{d}r.$$

If $b = \infty$, since $u^* \in L^{1,\infty}(\mathbb{R}^d)$ we must have $\tau = \infty$ as well (i.e. u^* non-increasing in the interval (a, ∞)) and a simple limiting argument leads to inequality (1.5) again. Note that $\lim_{r \to \infty} u_0(r)\, r^{d-1} = 0$ since $r \mapsto u_0(r)\, r^{d-1}$ is locally Lipschitz with integrable derivative in $(0, \infty)$.

Finally, if $a = 0$, the proof of (1.5) follows as above noting that $\lim_{r \to 0} u^*(r)\, r^{d-1} = 0$ (for $d \geq 2$).

If we add up (1.5) over all the intervals (a_i, b_i) of the disjoint union (1.4) we find

$$\int_{A_1} \left|(u^*)'(r)\right| r^{d-1} \,\mathrm{d}r \leq \int_{A_1} \left|u_0'(r)\right| r^{d-1} \,\mathrm{d}r + (d-1) \int_0^\infty \widetilde{M} u_0(r)\, r^{d-2} \,\mathrm{d}r\,,$$

which then leads to (note that in A_1^c we have $u^* = u_0$, and hence $(u^*)' = u_0'$ a.e. in A_1^c).

$$\int_0^\infty \left|(u^*)'(r)\right| r^{d-1} \,\mathrm{d}r \leq \int_0^\infty \left|u_0'(r)\right| r^{d-1} \,\mathrm{d}r + (d-1) \int_0^\infty \widetilde{M} u_0(r)\, r^{d-2} \,\mathrm{d}r. \tag{1.6}$$

Step 2: Control of weighted norms

As $r \mapsto \widetilde{M}u_0(r)$ is Lipschitz and its derivative is integrable (in fact $\big(\widetilde{M}u_0\big)'(r)\, r^{d-1} \in L^1(0,\infty)$ from Luiro's work [Lui18]) we have that $\lim_{r\to\infty} \widetilde{M}u_0(r)$ exists and it is equal to 0 since $\widetilde{M}u_0 \in L^{1,\infty}(\mathbb{R}^d)$. Then

$$\widetilde{M}u_0(r) = -\int_r^\infty \big(\widetilde{M}u_0\big)'(t)\ \mathrm{d}t$$

and

$$\begin{aligned}(d-1)\int_0^\infty \widetilde{M}u_0(r)\, r^{d-2}\ \mathrm{d}r &= (d-1)\int_0^\infty \left(\int_r^\infty -\big(\widetilde{M}u_0\big)'(t)\ \mathrm{d}t\right) r^{d-2}\ \mathrm{d}r \\ &\leq (d-1)\int_0^\infty \left(\int_r^\infty \big|\big(\widetilde{M}u_0\big)'(t)\big|\ \mathrm{d}t\right) r^{d-2}\ \mathrm{d}r \\ &= (d-1)\int_0^\infty \int_0^t r^{d-2}\, \big|\big(\widetilde{M}u_0\big)'(t)\big|\ \mathrm{d}r\ \mathrm{d}t \\ &= \int_0^\infty \big|\big(\widetilde{M}u_0\big)'(t)\big|\, t^{d-1}\ \mathrm{d}t.\end{aligned} \tag{1.7}$$

Finally, we combine (1.6) and (1.7) to arrive at (1.2), concluding the proof in this case.

1.2.2 General case

Let us first record a basic lemma about radial functions and weak derivatives. In what follows, when we say that a function f is weakly differentiable in a certain domain $\Omega \subset \mathbb{R}^d$, it is naturally understood that f and its weak derivatives are locally integrable in such a domain.

Lemma 1.2.1. .

(i) A radial function $f(x)$ is weakly differentiable in $\mathbb{R}^d \setminus \{0\}$ if and only if its radial restriction $f(r)$ is weakly differentiable in $(0,\infty)$.

(ii) In the situation above, if $f(x)$ and $\nabla f(x)$ are locally integrable in a neighborhood of the origin, then f is weakly differentiable in $\mathbb{R}^d$.

Proof This result is most certainly standard but we could not find an exact explicit reference. We then provide a brief proof for completeness.

Part (i). Assume that $f(x)$ is weakly differentiable in $\mathbb{R}^d \setminus \{0\}$ and let ∇f be its weak gradient. Let $\varphi \in C_c^\infty(\mathbb{R}^d \setminus \{0\})$ be a radial test function. Letting $r = |x|$ we have, by definition,

$$\begin{aligned}\int_{\mathbb{R}^d\setminus\{0\}} f(x)\left(\frac{(d-1)}{|x|}\varphi(x)+\frac{\partial\varphi}{\partial r}(x)\right)\,\mathrm{d}x &= \int_{\mathbb{R}^d\setminus\{0\}} f(x)\left(\sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left(\frac{x_i}{|x|}\varphi(x)\right)\right)\,\mathrm{d}x\\ &= -\int_{\mathbb{R}^d\setminus\{0\}}\left(\sum_{i=1}^{d}\frac{\partial f}{\partial x_i}\frac{x_i}{|x|}\varphi(x)\right)\,\mathrm{d}x = -\int_{\mathbb{R}^d\setminus\{0\}}(\nabla f(x))\cdot\frac{x}{|x|}\;\varphi(x)\,\mathrm{d}x.\end{aligned}\tag{1.8}$$

Write $x = r\omega$, with $\omega \in \mathbb{S}^{d-1}$. Letting $\Phi(r) = \varphi(r)\,r^{d-1}$, rewrite (1.8) in polar coordinates to get

$$\sigma_{d-1}\big(\mathbb{S}^{d-1}\big)\int_0^\infty f(r)\,\Phi'(r)\,\mathrm{d}r = -\int_0^\infty\left(\int_{\mathbb{S}^{d-1}}(\nabla f(r\omega))\cdot\omega\;\mathrm{d}\sigma_{d-1}(\omega)\right)\Phi(r)\,\mathrm{d}r.$$

This is the required integration by parts in $(0,\infty)$ for the generic test function Φ.

Assume now that $f(r)$ is weakly differentiable in $(0,\infty)$. If g is its weak derivative, then $f(r) = \int_1^r g(t)\,\mathrm{d}t$ almost everywhere, and hence we can modify f on a set of measure zero so that f is continuous in $(0,\infty)$; in fact absolutely continuous in each interval $[a,b]\subset(0,\infty)$. In particular, f is differentiable a.e. and $g = f'$. The radial extension $f(x)$ is then continuous in $\mathbb{R}^d\setminus\{0\}$ and differentiable almost everywhere. Let us show that integration by parts holds, say, with respect to the first coordinate x_1. Write $x = (x_1, x_1, \ldots, x_d) = r\omega = (r\cos\theta, r(\sin\theta)\xi)$, with $r\in(0,\infty)$, $\omega\in\mathbb{S}^{d-1}\subset\mathbb{R}^d$, $0\le\theta\le\pi$ and $\xi\in\mathbb{S}^{d-2}\subset\mathbb{R}^{d-1}$. Let $\psi\in C_c^\infty(\mathbb{R}^d\setminus\{0\})$ be a generic test function and consider

$$\Psi(r) = \left(\int_{\mathbb{S}^{d-1}}\psi\;\frac{x_1}{|x|}\;\mathrm{d}\sigma_{d-1}(\omega)\right)r^{d-1} = \left(\int_0^\pi\left(\int_{\mathbb{S}^{d-2}}\psi\;\mathrm{d}\sigma_{d-2}(\xi)\right)\cos\theta\,(\sin\theta)^{d-2}\;\mathrm{d}\theta\right)r^{d-1}.$$

Then

$$\Psi'(r) = \left(\int_0^\pi\int_{\mathbb{S}^{d-2}}\left(\frac{\partial\psi}{\partial r}\cos\theta - \frac{\partial\psi}{\partial\theta}\frac{\sin\theta}{r}\right)(\sin\theta)^{d-2}\;\mathrm{d}\sigma_{d-2}(\xi)\;\mathrm{d}\theta\right)\,r^{d-1}$$

where an integration by parts in the variable θ was used. Using polar coordinates one now sees that

$$\begin{aligned}\int_{\mathbb{R}^d\setminus\{0\}} f(x)\,\frac{\partial\psi}{\partial x_1}\;\mathrm{d}x = \int_0^\infty f(r)\,\Psi'(r)\;\mathrm{d}r &= -\int_0^\infty f'(r)\,\Psi(r)\;\mathrm{d}r\\ &= -\int_{\mathbb{R}^d\setminus\{0\}}\left(f'(|x|)\,\frac{x_1}{|x|}\right)\,\psi(x)\;\mathrm{d}x.\end{aligned}$$

Part (ii). Let $\psi:\mathbb{R}^d\to\mathbb{R}$ be a smooth radial non-increasing function with $\psi\equiv 1$ on $\{|x|\le 1\}$ and $\psi\equiv 0$ on $\{|x|\ge 2\}$. Let $\Psi_\alpha(x) = 1-\psi(x/\alpha)$. Let $\phi\in C_c^\infty(\mathbb{R}^d)$ be any test function. Since we know that f is weakly differentiable in $\mathbb{R}^d\setminus\{0\}$ we have, for any direction $i = 1,2,\ldots,d$ (here we denote $\partial f/\partial x_i$ simply by f_{x_i}),

$$\begin{aligned}-\int_{\mathbb{R}^d} f_{x_i}(x)\,(\phi\Psi_\alpha)(x)\;\mathrm{d}x &= \int_{\mathbb{R}^d} f(x)\,(\phi\Psi_\alpha)_{x_i}(x)\;\mathrm{d}x\\ &= \int_{\mathbb{R}^d} f(x)\,\phi_{x_i}(x)\,\Psi_\alpha(x)\;\mathrm{d}x + \int_{\mathbb{R}^d} f(x)\,\phi(x)\,(\Psi_\alpha)_{x_i}(x)\;\mathrm{d}x.\end{aligned}\tag{1.9}$$

Note that the last integral takes place inside the ball of radius 2α. In this ball we have $\phi(x) = \phi(0) + R(x)$ with $|R(x)| \le C\alpha$. Since $f(x)$ is even in the variable x_i and $(\Psi_\alpha)_{x_i}(x)$ is odd in the variable x_i we get

$$\int_{\mathbb{R}^d} f(x)(\Psi_\alpha)_{x_i}(x)\,\mathrm{d}x = 0\,, \tag{1.10}$$

and since $(\Psi_\alpha)_{x_i}(x) = -\frac{1}{\alpha}\psi_{x_i}(x/\alpha)$ we find

$$\int_{\mathbb{R}^d} f(x)R(x)(\Psi_\alpha)_{x_i}(x)\,\mathrm{d}x \to 0 \tag{1.11}$$

as $\alpha \to 0$, since f is locally integrable. Using (1.10) and (1.11) and the fact that ∇f is also locally integrable we may pass the limit as $\alpha \to 0$ in (1.9) to find

$$-\int_{\mathbb{R}^d} f_{x_i}(x)\,\phi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\phi_{x_i}(x)\,\mathrm{d}x\,,$$

as desired. ■

We now consider the case of general $u_0 \in W^{1,1}(\mathbb{R}^d)$ radial. We have seen in Lemma 1.2.1 that its radial version $u_0(r)$ is weakly differentiable in $(0,\infty)$ and

$$\int_0^\infty |u_0'(r)|\,r^{d-1}\,\mathrm{d}r < \infty.$$

In particular, after a possible redefinition on a set of measure zero, one can take $u_0(r)$ continuous in $(0,\infty)$ (in fact, absolutely continuous in each interval $[a,\infty)$ for $a > 0$). This is equivalent to assuming that $u_0(x)$ is continuous in $\mathbb{R}^d \setminus \{0\}$.

Step 3: u^* is continuous in $\mathbb{R}^d \setminus \{0\}$

With $u_0(x)$ continuous in $\mathbb{R}^d \setminus \{0\}$, the detachment set A_d defined in (1.3) is open. Throughout the rest of this section let us write

$$u_\varepsilon(x) := u(x,\varepsilon) = \big(u_0 * \phi(\cdot,\varepsilon)\big)(x), \quad x \in \mathbb{R}^d,\ \varepsilon > 0.$$

We claim that u^* is locally Lipschitz in A_d. In fact, if $x_0 \in A_d$, there exists $t_0 > 0$ such that

$$u^*(x_0) = u(x_0,t_0) > u(x_0).$$

From the continuity of $u(x,t)$, there exist a neighborhood V of x_0 and an $\varepsilon_0 > 0$ such that

$$u^*(x) = \sup_{t>0} u(x,t) = \sup_{t>\varepsilon_0} u(x,t) = \sup_{t>0} \big(u_{\varepsilon_0} * \phi(\cdot,t)\big)(x) =: u^*_{\varepsilon_0}(x) \tag{1.12}$$

for all $x \in V$. Note that in the third equality above we used the semigroup property of the family $\phi(\cdot,t)$ (i.e. the fact that $\phi(\cdot,t_1) * \phi(\cdot,t_2) = \phi(\cdot,t_1+t_2)$). Since u_{ε_0} is Lipschitz, we have that $u^* = u^*_{\varepsilon_0}$ is Lipschitz on V, which proves our claim.

Writing $\mathbb{R}^d \setminus \{0\} = A_d \cup A_d^c$, we now need so show that u^* is continuous at the points of A_d^c. Let $x_0 \in A_d^c$. If $x_0 \in \text{int}(A_d^c)$ we are done since $u^* = u_0$ is continuous in a neighborhood of x_0. Assume now that $x_0 \in A_d^c \setminus \text{int}(A_d^c)$ and that there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset A_d$ such that $x_n \to x_0$ but $u^*(x_n) \nrightarrow u^*(x_0) = u_0(x_0)$. Then there exist $t_n > 0$ and $\delta > 0$ such that $u(x_n, t_n) \geq u_0(x_0) + \delta$ for all n. From the integrability of u_0, the t_n are bounded, and passing to a subsequence we may assume that $t_n \to t \geq 0$. Then $u(x_n, t_n) \to u(x_0, t) \geq u_0(x_0) + \delta$, and we get that $t > 0$ and $x_0 \in A_d$, a contradiction. This establishes that u^* is continuous in $\mathbb{R}^d \setminus \{0\}$.

Step 4: Weak differentiability and conclusion

In the previous step we showed that $u^*(r)$ is continuous on $(0, \infty)$ and locally Lipschitz in A_1. For almost every $r \in A_1$, from (1.12) we have

$$(u^*)'(r) = \lim_{\varepsilon\to 0} (u_\varepsilon^*)'(r).$$

From Minkowski's inequality we recall that

$$\|\nabla u_\varepsilon\|_{L^1(\mathbb{R}^d)} \leq \|\nabla u_0\|_{L^1(\mathbb{R}^d)} \tag{1.13}$$

for any $\varepsilon > 0$. Using Fatou's lemma, the bound in Theorem 1.1.1 already proved for Lipschitz functions, and (1.13), we arrive at

$$\begin{aligned}\int_{A_1} \left|(u^*)'(r)\right| r^{d-1}\, \mathrm{d}r &\leq \liminf_{\varepsilon\to 0} \int_{A_1} \left|(u_\varepsilon^*)'(r)\right| r^{d-1}\, \mathrm{d}r \\ &\lesssim_d \liminf_{\varepsilon\to 0} \|\nabla u_\varepsilon\|_{L^1(\mathbb{R}^d)} \\ &\leq \|\nabla u_0\|_{L^1(\mathbb{R}^d)}.\end{aligned} \tag{1.14}$$

With this in hand, an adaptation of the argument in [CS13, §5.4] shows that $u^*(r)$ is weakly differentiable in $(0, \infty)$ with weak derivative given by $\chi_{A_1^c} u_0'(r) + \chi_{A_1}(u^*)'(r)$. This in turn implies that $u^*(x)$ is weakly differentiable in $\mathbb{R}^d \setminus \{0\}$ by Lemma 1.2.1. From (1.14), its weak gradient ∇u^* on $\mathbb{R}^d \setminus \{0\}$ verifies

$$\begin{aligned}\|\nabla u^*\|_{L^1(\mathbb{R}^d)} &= \omega_{d-1} \int_0^\infty \left|(u^*)'(r)\right| r^{d-1}\, \mathrm{d}r \\ &= \omega_{d-1} \left(\int_{A_1} \left|(u^*)'(r)\right| r^{d-1}\, \mathrm{d}r + \int_{A_1^c} \left|u_0'(r)\right| r^{d-1}\, \mathrm{d}r \right) \\ &\lesssim_d \|\nabla u_0\|_{L^1(\mathbb{R}^d)},\end{aligned} \tag{1.15}$$

with ω_{d-1} being the total surface measure of $\mathbb{S}^{d-1}$. This is our desired bound. As a final remark note that, from the Sobolev embedding, $u_0 \in L^{d/(d-1)}(\mathbb{R}^d)$ and hence so does u^*. In particular, u^* is locally integrable in $\mathbb{R}^d$. Since we already know from (1.15) that $\nabla u^* \in L^1(\mathbb{R}^d)$, an application of Lemma 1.2.1 (ii) gives us that u^* is in fact weakly differentiable in $\mathbb{R}^d$. This completes the proof of Theorem 1.1.1.

Remark: A crucial insight in the proof above was to relate the variation of u^* with the variation of the uncentered Hardy-Littlewood maximal operator $\widetilde{M}u_0$, expressed in inequality (1.2). Since $\widetilde{M}u_0(x) \lesssim_d u^*(x)$, uniformly for all $x \in \mathbb{R}^d$, we could just run the exact same proof to obtain the gradient bound for $\widetilde{M}u_0$ starting from the gradient bound for u^*, showing that these two bounds are actually equivalent to each other.

Chapter 2

Sobolev regularity for polar maximal functions

2.1 Introduction

In this chapter we consider maximal operators acting on functions defined on the sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$, in order to develop an analogous theory. First, let us establish the basic notation to be used in this context. We let $d(\zeta, \eta)$ denote the geodesic distance between two points $\zeta, \eta \in \mathbb{S}^d$. Let $\mathcal{B}_r(\zeta) \subset \mathbb{S}^d$ be the open geodesic ball of center $\zeta \in \mathbb{S}^d$ and radius $r > 0$, that is

$$\mathcal{B}_r(\zeta) = \{\eta \in \mathbb{S}^d \ : \ d(\zeta, \eta) < r\},$$

and let $\overline{\mathcal{B}_r(\zeta)}$ be the corresponding closed ball. Let $\widetilde{\mathcal{M}}$ denote the uncentered Hardy-Littlewood maximal operator on the sphere $\mathbb{S}^d$, that is, for $f \in L^1_{\text{loc}}(\mathbb{S}^d)$,

$$\widetilde{\mathcal{M}}f(\xi) = \sup_{\{\overline{\mathcal{B}_r(\zeta)} \ : \ \xi \in \overline{\mathcal{B}_r(\zeta)}\}} \frac{1}{\sigma(\mathcal{B}_r(\zeta))} \int_{\mathcal{B}_r(\zeta)} |f(\eta)| \, \mathrm{d}\sigma(\eta) = \sup_{\{\overline{\mathcal{B}_r(\zeta)} \ : \ \xi \in \overline{\mathcal{B}_r(\zeta)}\}} ⨍_{\mathcal{B}_r(\zeta)} |f(\eta)| \, \mathrm{d}\sigma(\eta),$$

where $\sigma = \sigma_d$ denotes the usual surface measure on the sphere $\mathbb{S}^d$. The centered version $\mathcal{M}$ would be defined with centered geodesic balls. Fix $\mathbf{e} = (1, 0, 0, \ldots, 0) \in \mathbb{R}^{d+1}$ to be our north pole. We say that a function $f : \mathbb{S}^d \to \mathbb{C}$ is *polar* if for every $\xi, \eta \in \mathbb{S}^d$ with $\xi \cdot \mathbf{e} = \eta \cdot \mathbf{e}$ we have $f(\xi) = f(\eta)$. This will be the analogue, in the spherical setting, of a radial function in the Euclidean setting.

When working on the circle $\mathbb{S}^1$, an adaptation of the proof of Aldaz and Pérez Lázaro [APL07] yields $\mathrm{Var}(\widetilde{\mathcal{M}}f) \leq \mathrm{Var}(f)$, where $\mathrm{Var}(f)$ denotes the total variation of the function f. This follows from the fact that $\widetilde{\mathcal{M}}f$ has no local maxima in the detachment set $\{\widetilde{\mathcal{M}}f > |f|\}$ (say, for f Lipschitz). Our first result in this chapter is the extension of this statement to the multidimensional setting, in the case of polar functions. For the basic theory of Sobolev spaces on the sphere $\mathbb{S}^d$ we refer the reader to [DX13].

Theorem 2.1.1. *If $f \in W^{1,1}(\mathbb{S}^d)$ is a polar function, then $\widetilde{\mathcal{M}}f$ is weakly differentiable and*

$$\|\nabla \widetilde{\mathcal{M}}f\|_{L^1(\mathbb{S}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbb{S}^d)}.$$

This is the analogue on the sphere $\mathbb{S}^d$ of Luiro's result [Lui18] for radial functions in the Euclidean space. The proof we present below follows broadly the strategy outlined by Luiro [Lui18]. However, due to the different geometry, several nontrivial technical points arise along the proof and must be considered carefully. A good example that such difficulties cannot be underestimated is Lemma 2.4.4 below, one of the core results used in our proof of Theorem 2.1.1. As in the case of $\mathbb{R}^d$, the analogue of Theorem 2.1.1 for the centered Hardy-Littlewood maximal operator $\mathcal{M}$ on $\mathbb{S}^d$ is an open problem.

2.1.1 Maximal operators of convolution type on $\mathbb{S}^d$

We now treat two important cases of maximal operators of convolution type on the sphere: the Poisson maximal operator and the heat flow maximal operator. We briefly recall the basic definitions and refer the reader to [CFS18, §1.4] for additional details.

Poisson maximal function on $\mathbb{S}^d$

Let $0 \leq \rho < 1$ and let $\xi, \eta \in \mathbb{S}^d$. We define the Poisson kernel $\mathcal{P}$ on the sphere by

$$\mathcal{P}(\xi, \eta, \rho) = \frac{1-\rho^2}{\omega_d\, |\rho\xi - \eta|^d} = \frac{1-\rho^2}{\omega_d\, (\rho^2 - 2\rho\, \xi \cdot \eta + 1)^{d/2}}\,,$$

with $\omega_d = \sigma(\mathbb{S}^d)$ being the total surface area of $\mathbb{S}^d$. If $u_0 \in L^1(\mathbb{S}^d)$ we let $u(\xi, \rho) = u(\rho\xi)$ be the function defined on the unit $(d+1)$-dimensional open ball $B_1 \subset \mathbb{R}^{d+1}$ by

$$u(\xi, \rho) = \int_{\mathbb{S}^d} \mathcal{P}(\xi, \eta, \rho)\, |u_0(\eta)|\, \mathrm{d}\sigma(\eta)\,,$$

and consider the associated maximal function

$$u^*(\xi) = \sup_{0 \leq \rho < 1} u(\xi, \rho). \tag{2.1}$$

Observe that $u \in C^\infty(B_1)$ and solves the Dirichlet problem

$$\begin{cases} \Delta u = 0 & \text{in } B_1\,; \\ \lim\limits_{\rho \to 1^-} u(\xi, \rho) = |u_0(\xi)| & \text{for a.e. } \xi \in \mathbb{S}^d. \end{cases}$$

Heat flow maximal function on $\mathbb{S}^d$

Let $\{Y_n^\ell\}$, $\ell = 1, 2, \ldots, \dim H_n^{d+1}$, be an orthonormal basis of the space H_n^{d+1} of spherical harmonics of degree n in the sphere $\mathbb{S}^d$. For $t \in (0, \infty)$ and $\xi, \eta \in \mathbb{S}^d$ we define the heat kernel $\mathcal{K}$ on the sphere (see [DX13, Lemma 1.2.3, Theorem 1.2.6 and Eq. 7.5.5]) by

$$\mathcal{K}(\xi, \eta, t) = \sum_{n=0}^{\infty} e^{-tn(n+d-1)} \sum_{\ell=1}^{\dim H_n^{d+1}} Y_n^\ell(\xi) Y_n^\ell(\eta) = \sum_{n=0}^{\infty} e^{-tn(n+d-1)} \frac{(n+\lambda)}{\lambda}\, C_n^\lambda(\xi \cdot \eta),$$

where $\lambda = \frac{d-1}{2}$ and $t \mapsto C_n^\beta(t)$, for $\beta > 0$, are the *Gegenbauer polynomials* defined in terms of the generating function

$$(1 - 2rt + r^2)^{-\beta} = \sum_{n=0}^{\infty} C_n^\beta(t)\, r^n.$$

If $u_0 \in L^1(\mathbb{S}^d)$ we consider

$$u(\xi, t) = \int_{\mathbb{S}^d} \mathcal{K}(\xi, \eta, t)\, |u_0(\eta)|\, \mathrm{d}\sigma(\eta)\,,$$

and consider the associated maximal function

$$u^*(\xi) = \sup_{t>0} u(\xi, t). \tag{2.2}$$

Note that u is a smooth function on $\mathbb{S}^d \times (0, \infty)$ and solves the heat equation

$$\begin{cases} \partial_t u - \Delta u = 0 & \text{in } \mathbb{S}^d \times (0,\infty)\,; \\ \lim\limits_{t\to 0^+} u(\xi, t) = |u_0(\xi)| & \text{for a.e. } \xi \in \mathbb{S}^d. \end{cases}$$

Gradient bounds

We note that the smooth kernels $\mathcal{P}$ and $\mathcal{K}$ depend only on $d(\xi, \eta)$ and are decreasing with respect to this distance. If we fix one of these two parameters, they have integral 1 on $\mathbb{S}^d$ and are approximate identities as $\rho \to 1^-$ and $t \to 0^+$, respectively. The discussion on the heat kernel can be found in [SY94, Chapter III, §2]. Also, from [DX13, Chapter 2, Theorem 2.3.6], note that the associated maximal functions u^* are dominated by the Hardy-Littlewood maximal function, that is

$$u^*(\xi) \le \mathcal{M}u_0(\xi) \le \widetilde{\mathcal{M}}u_0(\xi). \tag{2.3}$$

Our second result establishes the following.

Theorem 2.1.2. *Let u^* be the Poisson maximal function given by (2.1) or the heat flow maximal function given by (2.2). If $u_0 \in W^{1,1}(\mathbb{S}^d)$ is a polar function, then u^* is weakly differentiable and*

$$\|\nabla u^*\|_{L^1(\mathbb{S}^d)} \lesssim_d \|\nabla u_0\|_{L^1(\mathbb{S}^d)}.$$

2.1.2 The Hardy-Littlewood fractional maximal operator

In this chapter we also discuss the analogue of [LM19, Theorem 1.1] in the sphere setting. We define the uncentered fractional Hardy-Littlewood operator for $f \in L^1(\mathbb{S}^d)$:

$$\begin{aligned} \widetilde{\mathcal{M}}_\beta f(\xi) &= \sup_{\{\overline{\mathcal{B}_r(\zeta)}\ :\ \xi \in \overline{\mathcal{B}_r(\zeta)}, r \le \pi\}} \frac{r^\beta}{\sigma(\mathcal{B}_r(\zeta))} \int_{\mathcal{B}_r(\zeta)} |f(\eta)|\, \mathrm{d}\sigma(\eta) \\ &= \sup_{\{\overline{\mathcal{B}_r(\zeta)}\ :\ \xi \in \overline{\mathcal{B}_r(\zeta)}, r \le \pi\}} r^\beta ⨍_{\mathcal{B}_r(\zeta)} |f(\eta)|\, \mathrm{d}\sigma(\eta). \end{aligned}$$

We propose here the following question.

Question 2.1.1. *Let $f \in W^{1,1}(\mathbb{S}^d)$, $0 < \beta < d$. Does it hold that $\widetilde{\mathcal{M}}_\beta f$ is weakly differentiable and $\|\nabla \widetilde{\mathcal{M}}_\beta f\|_q \lesssim_{d,\beta} \|\nabla f\|_1$?*

As far as we are concerned there is no previous result in the direction of this problem. Let us notice that, for the case $\beta \geq 1$ of this question, it is not enough the argument in [CM17], in fact, by imitating their arguments we get, for all nonnegative $f \in W^{1,1}(\mathbb{S}^d)$ and almost every $\xi \in \mathbb{S}^d$, the inequality

$$|\nabla \widetilde{\mathcal{M}}_\beta f|(\xi) \lesssim_{d,\beta} \widetilde{\mathcal{M}}_{\beta-1} f(\xi). \tag{2.4}$$

Therefore, by the Sobolev embedding we get

$$\|\nabla \widetilde{\mathcal{M}}_\beta f\|_q \lesssim_{d,\beta} \|\widetilde{\mathcal{M}}_{\beta-1} f\|_q \lesssim_{d,\beta} \|f\|_{d/(d-1)} \lesssim_{d,\beta} \|f\|_{W^{1,1}(\mathbb{S}^d)}.$$

But since, differing from the Euclidean case, we cannot avoid $\|f\|_1$ in this last expression (consider, for instance, f being a positive constant), Question 2.1.1 cannot be answered directly in this case, and remains an open problem.

Concerning to the polar case, the difficulties that Carneiro and the author faced in the second section of this chapter also appear (in different ways) when dealing with this question. We go further in the methods already developed for the classical uncentered case in order to adapt the proof of [LM19]. We get the following:

Theorem 2.1.3. *Let $f \in W^{1,1}_{pol}(\mathbb{S}^d)$, $0 < \beta < d$, and $q = \frac{d}{d-\beta}$. We have*

$$\|\nabla \widetilde{\mathcal{M}}_\beta f\|_q \lesssim_{d,\beta} \|\nabla f\|_1.$$

2.2 Proof of Theorem 2.1.1

Recall that σ denotes the usual surface measure on the sphere $\mathbb{S}^d$. We denote by $\omega_d = \sigma(\mathbb{S}^d) = 2\pi^{(d+1)/2}/\Gamma((d+1)/2)$ the total surface area of $\mathbb{S}^d$. With a slight abuse of notation, we shall also write

$$\sigma(r) := \sigma\big(\mathcal{B}_r(\zeta)\big) = \omega_{d-1} \int_0^r (\sin t)^{d-1}\, \mathrm{d}t. \tag{2.5}$$

Throughout this section we assume, without loss of generality, that f *is real-valued and nonnegative* (or $+\infty$).

2.2.1 Preliminaries

If $f \in L^1(\mathbb{S}^d)$, by Lebesgue differentiation we may modify it in a set of measure zero so that

$$f(\xi) = \limsup_{\{r \to 0^+ \,:\, \xi \in \overline{\mathcal{B}_r(\zeta)}\}} \fint_{\mathcal{B}_r(\zeta)} f(\eta)\, \mathrm{d}\sigma(\eta) \tag{2.6}$$

holds everywhere. Let us assume that is the case. For $f \in L^1(\mathbb{S}^d)$ and $\xi \in \mathbb{S}^d$ let us define the set $\mathbf{B}_\xi$ as the set of closed balls that realize the supremum in the definition of the maximal function, that is

$$\mathbf{B}_\xi = \left\{ \overline{\mathcal{B}_r(\zeta)};\ \zeta \in \mathbb{S}^d\,,\ r \geq 0\,,\ \xi \in \overline{\mathcal{B}_r(\zeta)}\ :\ \widetilde{\mathcal{M}}f(\xi) = ⨍_{\overline{\mathcal{B}_r(\zeta)}} f(\eta)\ \mathrm{d}\sigma(\eta) \right\}. \tag{2.7}$$

Here we consider the slight abuse of notation

$$\overline{\mathcal{B}_0(\xi)} := \{\xi\} \quad \text{and} \quad ⨍_{\{\xi\}} f(\eta)\ \mathrm{d}\sigma(\eta) := f(\xi), \tag{2.8}$$

in order to include the closed ball of radius zero as a potential candidate in the definition of $\mathbf{B}_\xi$. In light of (2.6) we always have that $\mathbf{B}_\xi$ is non-empty. Our first lemma holds for general Sobolev functions in $W^{1,1}(\mathbb{S}^d)$ (not necessarily polar functions).

Lemma 2.2.1. *Let $f \in W^{1,1}(\mathbb{S}^d)$ be a nonnegative function that verifies (2.6) and let $\xi \in \mathbb{S}^d$ be a point such that $\widetilde{\mathcal{M}}f(\xi) > f(\xi)$. Assume that $\widetilde{\mathcal{M}}f$ is differentiable at ξ and that $\overline{\mathcal{B}} \in \mathbf{B}_\xi$. Then*

$$\nabla\widetilde{\mathcal{M}}f(\xi)v = ⨍_{\mathcal{B}} \nabla f(\eta)\big(-(\eta\cdot v)\xi + (\eta\cdot\xi)v\big)\ \mathrm{d}\sigma(\eta)$$

for every $v \in \mathbb{R}^{d+1}$ with $v \perp \xi$. In particular,

$$\big|\nabla\widetilde{\mathcal{M}}f(\xi)\big| \leq ⨍_{\mathcal{B}} |\nabla f(\eta)|\ \mathrm{d}\sigma(\eta).$$

Proof Observe first that the condition $\widetilde{\mathcal{M}}f(\xi) > f(\xi)$ implies that the ball $\mathcal{B}$ has positive radius. Without loss of generality let us assume that $|v| = 1$. Let $R_t = R_{t,\xi,v}$ be the rotation of angle t over the plane spanned by ξ and v that leaves the orthogonal complement invariant, i.e.

$$R_t(\eta) = \big((\cos t)(\eta\cdot\xi) - (\sin t)(\eta\cdot v)\big)\xi + \big((\sin t)(\eta\cdot\xi) + (\cos t)(\eta\cdot v)\big)v + z(\eta)\,,$$

where $z(\eta)$ is the component of the vector η that is orthogonal to the plane generated by ξ and v. Then

$$\begin{aligned}
\nabla\widetilde{\mathcal{M}}f(\xi)v &= \lim_{t\to0+} \frac{\widetilde{\mathcal{M}}f(R_t\xi) - \widetilde{\mathcal{M}}f(\xi)}{t} \\
&\geq \lim_{t\to0+} \frac{1}{t}\left(⨍_{R_t(\mathcal{B})} f - ⨍_{\mathcal{B}} f\right) \\
&= \lim_{t\to0+} ⨍_{\mathcal{B}} \frac{f(R_t\eta) - f(\eta)}{t}\ \mathrm{d}\sigma(\eta) \\
&= ⨍_{\mathcal{B}} \nabla f(\eta)\big(-(\eta\cdot v)\xi + (\eta\cdot\xi)v\big)\ \mathrm{d}\sigma(\eta).
\end{aligned} \tag{2.9}$$

The reverse inequality is obtained similarly by considering the limit as $t \to 0^-$. ■

Remark: The passage to the limit in (2.9) uses the fact that the difference quotients are bounded in L^1 by a multiple of L^1-norm of the gradient of f, uniformly in t. With such a uniform bound one can establish the required limit by approximating f by smooth g.

2.2.2 Lipschitz case

Throughout this subsection we assume that our polar $f \in W^{1,1}(\mathbb{S}^d)$ is a Lipschitz function. Recalling that $\mathbf{e} = (1, 0, 0, \ldots, 0) \in \mathbb{R}^{d+1}$, for $\xi \in \mathbb{S}^d$ we write

$$\cos\theta = \xi \cdot \mathbf{e}$$

with $\theta \in [0, \pi]$. Note that $\theta = \theta(\xi) = d(\mathbf{e}, \xi)$ is the polar angle. We generally write $f(\xi)$ for the function on $\mathbb{S}^d$, and $f(\theta)$ for its polar version on $(0, \pi)$. We then have

$$|\nabla f(\xi)| = |f'(\theta)|$$

for a.e. $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$, and

$$\|\nabla f\|_{L^1(\mathbb{S}^d)} = \omega_{d-1} \int_0^\pi |f'(\theta)|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

Estimates for small radii

For $\zeta \in \mathbb{S}^d$ let us define

$$w(\zeta) = \min\big\{\theta(\zeta)\,,\, \pi - \theta(\zeta)\big\} = \min\{d(\mathbf{e}, \zeta), d(-\mathbf{e}, \zeta)\}.$$

Let us define the auxiliary maximal operator $\widetilde{\mathcal{M}}^I$ by (recall convention (2.8))

$$\widetilde{\mathcal{M}}^I f(\xi) = \sup_{\{\xi \in \overline{\mathcal{B}_r(\zeta)}\ :\ 0 \le r \le w(\zeta)/4\}} ⨍_{\overline{\mathcal{B}_r(\zeta)}} f(\eta)\,\mathrm{d}\sigma(\eta). \tag{2.10}$$

For each $\xi \in \mathbb{S}^d$ we define the set of good balls

$$\mathbf{B}^I_\xi = \left\{\overline{\mathcal{B}_r(\zeta)};\ \zeta \in \mathbb{S}^d\,,\, 0 \le r \le \frac{w(\zeta)}{4}\,;\, \xi \in \overline{\mathcal{B}_r(\zeta)}\ :\ \widetilde{\mathcal{M}}^I f(\xi) = ⨍_{\overline{\mathcal{B}_r(\zeta)}} f(\eta)\,\mathrm{d}\sigma(\eta)\right\}.$$

Notice that $\widetilde{\mathcal{M}}^I f$ is also a polar function. We consider the detachment set

$$\mathcal{E}_d := \big\{\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}\ :\ \widetilde{\mathcal{M}}^I f(\xi) > f(\xi)\big\},$$

and its polar version, denoted by

$$\mathcal{E}_1 = \{\theta(\xi) = d(\mathbf{e}, \xi)\ :\ \xi \in \mathcal{E}_d\}.$$

One can check that $\widetilde{\mathcal{M}}^I f$ is a continuous function in $\mathbb{S}^d$. Further qualitative properties of $\widetilde{\mathcal{M}}^I f$ are described in the next two results.

Lemma 2.2.2. *$\widetilde{\mathcal{M}}^I f$ does not have a strict local maximum in $\mathcal{E}_1$.*

Proof The proof is identical to [Lui18, Lemma 3.10]. ∎

Lemma 2.2.3. *$\widetilde{\mathcal{M}}^I f$ is locally Lipschitz in $\mathcal{E}_d$.*

Proof Let $\xi \in \mathcal{E}_d$. Let $\overline{\mathcal{B}_r(\zeta)} \in \mathbf{B}^I_\xi$ with r minimal. Then $r > 0$ and it is possible to find a neighborhood V of ξ of the form $V = \{\eta \in \mathbb{S}^d \ : \ \theta(\xi) - \varepsilon < \theta(\eta) < \theta(\xi) + \varepsilon\}$ such that: (i) $\varepsilon < r/100$ and (ii) if $\eta \in V$ and $\overline{\mathcal{B}_s(\omega)} \in \mathbf{B}^I_\eta$ then $s > 99r/100$.

Let $\eta_1, \omega_2 \in V$. Let S be the half great circle connecting $\mathbf{e}, \eta_1, -\mathbf{e}$. If $\eta_2 \in S$ is such that $d(\mathbf{e}, \eta_2) = d(\mathbf{e}, \omega_2)$ then we have $d(\eta_1, \eta_2) \leq d(\eta_1, \omega_2)$. Since $\widetilde{\mathcal{M}}^I f(\eta_2) = \widetilde{\mathcal{M}}^I f(\omega_2)$, for the purposes of proving Lipschitz continuity it suffices to work with $\eta_1, \eta_2 \in S$. Assume without loss of generality that $\widetilde{\mathcal{M}}^I f(\eta_1) > \widetilde{\mathcal{M}}^I f(\eta_2)$. Let $\overline{\mathcal{B}_{r_1}(\zeta_1)} \in \mathbf{B}^I_{\eta_1}$ with $\zeta_1 \in S$. Then $\eta_2 \notin \overline{\mathcal{B}_{r_1}(\zeta_1)}$, and hence η_2 is not between ζ_1 and η_1. It is also easy to see that we cannot have ζ_1 between η_1 and η_2 due to conditions (i) and (ii) above. Hence we must have η_1 between ζ_1 and η_2. We now choose a ball $\mathcal{B}_{r_2}(\zeta_2)$, with $\zeta_2 \in S$ lying between ζ_1 and η_2, such that $\eta_2 \in \partial\overline{\mathcal{B}_{r_2}(\zeta_2)}$ and

$$r_2 = d(\zeta_2, \eta_2) = \min\left\{r_1, \frac{w(\zeta_2)}{4}\right\} \tag{2.11}$$

(one may think of moving the center ζ_1 along S in the direction of η_2 until finding the unique choice of ζ_2). Note that ζ_2 is in fact between ζ_1 and η_1 and hence

$$r_2 = d(\zeta_2, \eta_2) = d(\zeta_1, \eta_1) - d(\zeta_1, \zeta_2) + d(\eta_1, \eta_2) \leq r_1 - d(\zeta_1, \zeta_2) + d(\eta_1, \eta_2). \tag{2.12}$$

If $r_2 = r_1$ in (2.11) then we have $d(\zeta_1, \zeta_2) \leq d(\eta_1, \eta_2)$. In the other case we have

$$r_2 = \frac{w(\zeta_2)}{4} \geq \frac{w(\zeta_1)}{4} - \frac{d(\zeta_1, \zeta_2)}{4} \geq r_1 - \frac{d(\zeta_1, \zeta_2)}{4},$$

and combining with (2.12) we obtain $d(\zeta_1, \zeta_2) \leq \frac{4}{3} d(\eta_1, \eta_2)$, which yields $r_1 - r_2 \leq \frac{1}{3} d(\eta_1, \eta_2)$. We conclude by observing that

$$\begin{aligned} \widetilde{\mathcal{M}}^I f(\eta_1) - \widetilde{\mathcal{M}}^I f(\eta_2) &\leq \fint_{\mathcal{B}_{r_1}(\zeta_1)} f - \fint_{\mathcal{B}_{r_2}(\zeta_2)} f \\ &\leq \left(\fint_{\mathcal{B}_{r_1}(\zeta_1)} f - \fint_{\mathcal{B}_{r_2}(\zeta_1)} f\right) + \left(\fint_{\mathcal{B}_{r_2}(\zeta_1)} f - \fint_{\mathcal{B}_{r_2}(\zeta_2)} f\right) \\ &\lesssim_{d,r,f} \ d(\eta_1, \eta_2). \end{aligned}$$

∎

An adaptation of the argument in [CS13, §5.4] then shows that $\widetilde{\mathcal{M}}^I f(\theta)$ is weakly differentiable in $(0, \pi)$, with weak derivative given by $\chi_{\mathcal{E}_1^c} f'(\theta) + \chi_{\mathcal{E}_1} \big(\widetilde{\mathcal{M}}^I f\big)'(\theta)$. In fact, if $\theta \in \mathcal{E}_1^c$

is a point of differentiability of f (which are almost all points of $\mathcal{E}_1^c$) one can plainly see that $f'(\theta) = 0$, otherwise one could do better than $f(\theta)$ in the maximal function (2.10) and θ would belong to $\mathcal{E}_1$ instead. The weak derivative of $\widetilde{\mathcal{M}}^I f(\theta)$ is then simply $\chi_{\mathcal{E}_1}\big(\widetilde{\mathcal{M}}^I f\big)'(\theta)$. From Lemma 2.2.8 below we have that $\widetilde{\mathcal{M}}^I f(\xi)$ is weakly differentiable in $\mathbb{S}^d$. The next proposition establishes the desired control of the variation.

Proposition 2.2.1. *The following inequality holds*

$$\big\|\nabla \widetilde{\mathcal{M}}^I f\big\|_{L^1(\mathbb{S}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbb{S}^d)}.$$

Proof The proof follows the outline of [Lui18, Lemma 3.5] with minor changes. We need to prove that

$$\int_{\mathcal{E}_1} \big|(\widetilde{\mathcal{M}}^I f)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim_d \int_0^\pi \big|f'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

We shall prove that

$$\int_{\mathcal{E}_1\cap[0,\pi/2]} \big|(\widetilde{\mathcal{M}}^I f)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim_d \int_0^\pi \big|f'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \tag{2.13}$$

and the proposition follows by symmetry. For $k \geq 1$, we define $\mathcal{E}_1^k = \mathcal{E}_1 \cap \big[\frac{\pi}{2^{k+1}}, \frac{\pi}{2^k}\big]$, and since $\mathcal{E}_1$ is open we may write $\operatorname{int}\big(\mathcal{E}_1^k\big) = \bigcup_{i=1}^\infty (a_i^k, b_i^k)$. We observe that $\frac{(\sin 2\theta)^{d-1}}{(\sin\theta)^{d-1}} \simeq_d 1$ for $\theta \leq \frac{\pi}{4}$. When $a_i^k = \frac{\pi}{2^{k+1}}$ or $b_i^k = \frac{\pi}{2^k}$ we observe, from the definition of the auxiliary operator in (2.10), that

$$\widetilde{\mathcal{M}}^I f(\pi/2^{k+1})\ ,\ \widetilde{\mathcal{M}}^I f(\pi/2^k) \leq \sup_{\theta(\xi)\,\in\,[\pi/2^{k+2},\pi/2^{k-1}]} f(\xi)$$

for $k \geq 2$. These are the ingredients needed to run the argument in [Lui18, Lemma 3.5] in order to get

$$\int_{\mathcal{E}_1^k} \big|(\widetilde{\mathcal{M}}^I f)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim_d \int_{\pi/2^{k+2}}^{\pi/2^{k-1}} \big|f'(\theta)\big|(\sin\theta)^{d-1}\,\mathrm{d}\theta \tag{2.14}$$

for $k \geq 2$. In the case $k = 1$ we must be a bit more careful when $b_i^1 = \pi/2$ by using the bound

$$\widetilde{\mathcal{M}}^I f(\pi/2) \leq \sup_{\theta(\xi)\,\in\,[\pi/4,3\pi/4]} f(\xi)\,,$$

which then yields

$$\begin{aligned}\int_{\mathcal{E}_1^1} \big|(\widetilde{\mathcal{M}}^I f)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta &\leq \int_{\mathcal{E}_1^1} \big|(\widetilde{\mathcal{M}}^I f)'(\theta)\big|\,\mathrm{d}\theta \lesssim \int_{\pi/8}^{3\pi/4} \big|f'(\theta)\big|\,\mathrm{d}\theta \\ &\lesssim_d \int_{\pi/8}^{3\pi/4} \big|f'(\theta)\big|(\sin\theta)^{d-1}\,\mathrm{d}\theta.\end{aligned} \tag{2.15}$$

Finally, we add up (2.14) and (2.15) to get (2.13). ■

Estimates for large radii - preliminary lemmas

The other crucial ingredient in the proof of Luiro [Lui18, Lemma 2.2 (v)] is the bound

$$\left|\nabla \widetilde{M} f(x)\right| \leq \frac{1}{|x|} \fint_B |\nabla f(y)|\, |y| \; \mathrm{d}y,$$

where $\overline{B} \ni x$ is a ball in which the maximal function is realized. The main difficulty in the case of $\mathbb{S}^d$ is in establishing a bound that will serve a similar purpose. This is accomplished in Lemma 2.2.7 below but before we actually get there we need a few preliminary lemmas. Recall the definition of $\sigma(r)$ in (2.5), and observe that $\sigma'(r) = \omega_{d-1}(\sin r)^{d-1}$ is equal to the $(d-1)$-dimensional area of $\partial\mathcal{B}_r(\zeta)$.

Lemma 2.2.4. *Let $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ and let $\overline{\mathcal{B}_r(\zeta)} \in \mathbf{B}_\xi$, with ζ in the half great circle determined by $\mathbf{e}$, ξ and $-\mathbf{e}$. Assume that $0 \leq \theta(\zeta) < \theta(\xi)$, that $\xi \in \partial\mathcal{B}_r(\zeta)$, that $\widetilde{\mathcal{M}} f(\xi) > f(\xi)$ and that $\widetilde{\mathcal{M}} f$ is differentiable at ξ. Then*

$$\nabla \widetilde{\mathcal{M}} f(\xi)(v(\xi, \mathbf{e})) = \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}_r(\zeta)} \nabla f(\eta)(v(\eta, \zeta))\, \frac{\sigma(d(\zeta, \eta))}{\sigma'(d(\zeta, \eta))} \; \mathrm{d}\sigma(\eta),$$

where

$$v(\eta, \zeta) = \frac{\zeta - (\eta \cdot \zeta)\eta}{|\zeta - (\eta \cdot \zeta)\eta|}$$

is the unit vector, tangent to η, in the direction of the geodesic that goes from η to ζ.

Proof Since $\widetilde{\mathcal{M}} f(\xi) > f(\xi)$ we have $r > 0$. Let S be the great circle determined by $\mathbf{e}$ and ξ. For small $h \in \mathbb{R}$ we consider a rotation R_h of angle h in this circle (in the direction from ξ to $\mathbf{e}$) leaving the orthogonal complement in $\mathbb{R}^{d+1}$ invariant, and write $\zeta - h := R_h(\zeta)$. The idea is to look at the following quantity

$$\lim_{h\to 0} \frac{\fint_{\mathcal{B}_{r+h}(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta)} f}{h} = \lim_{h\to 0} \frac{\fint_{\mathcal{B}_{r+h}(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta-h)} f + \fint_{\mathcal{B}_r(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta)} f}{h}. \tag{2.16}$$

In principle we do not know that the limit above exists. We shall prove that it in fact exists using the right-hand side of (2.16). Once this is established, the left-hand side of (2.16) tells us that this limit must be zero, since the numerator is always nonpositive regardless of the sign of h.

From Lemma 2.2.1 (in particular, see computation (2.9)) we note that

$$\lim_{h\to 0} \frac{\fint_{\mathcal{B}_r(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta)} f}{h} = \nabla \widetilde{\mathcal{M}} f(\xi)(v(\xi, \mathbf{e})). \tag{2.17}$$

Note also that

$$\begin{aligned} \frac{\fint_{\mathcal{B}_{r+h}(\zeta-h)} f - \fint_{\mathcal{B}_r(\zeta-h)} f}{h} &= \frac{\frac{1}{\sigma(r+h)} - \frac{1}{\sigma(r)}}{h} \int_{\mathcal{B}_{r+h}(\zeta-h)} f \; + \; \frac{1}{\sigma(r)} \frac{\int_{\mathcal{B}_{r+h}(\zeta-h)} f - \int_{\mathcal{B}_r(\zeta-h)} f}{h} \\ &\to -\frac{\sigma'(r)}{\sigma(r)^2} \int_{\mathcal{B}_r(\zeta)} f \; + \; \frac{1}{\sigma(r)} \int_{\partial\mathcal{B}_r(\zeta)} f \end{aligned} \tag{2.18}$$

as $h \to 0$. Hence the limit in (2.16) exists and is zero. Now we consider momentarily ζ as the north pole in the computation below and proceed with the standard polar coordinates on the sphere. Writing $\eta = (\cos\theta, \omega\sin\theta)$, with $\omega \in \mathbb{S}^{d-1}$ we use integration by parts to get

$$\begin{aligned}
\int_{\mathcal{B}_r(\zeta)} \nabla f(\eta)\,(-v(\eta,\zeta))\,\frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\,\mathrm{d}\sigma(\eta) &= \int_{\mathbb{S}^{d-1}}\int_0^r \frac{\partial f}{\partial\theta}(\theta,\omega)\left(\int_0^\theta (\sin t)^{d-1}\,\mathrm{d}t\right)\mathrm{d}\theta\;\mathrm{d}\sigma_{d-1}(\omega)\\
&= \int_{\mathbb{S}^{d-1}} f(r,\omega)\left(\int_0^r (\sin t)^{d-1}\,\mathrm{d}t\right)\mathrm{d}\sigma_{d-1}(\omega) - \int_{\mathbb{S}^{d-1}}\int_0^r f(\theta,\omega)(\sin\theta)^{d-1}\,\mathrm{d}\theta\;\mathrm{d}\sigma_{d-1}(\omega)\\
&= \frac{\sigma(r)}{\sigma'(r)}\int_{\partial\mathcal{B}_r(\zeta)} f - \int_{\mathcal{B}_r(\zeta)} f.
\end{aligned} \tag{2.19}$$

The lemma then plainly follows from (2.16), (2.17), (2.18) and (2.19). ■

We now state a basic geometric lemma.

Lemma 2.2.5. *Denote by $\triangle ABC$ a geodesic triangle with vertices A, B, C, opposite geodesic side lengths a, b, c, and (geodesic) angles $\hat{A}, \hat{B}, \hat{C}$.*

(i) There exist universal constants $\gamma > 1$ and $\rho > 0$ such that for every $\triangle ABC \subset \overline{\mathcal{B}_\rho(\mathbf{e})}$ we have

$$a\sin\hat{B} \le \gamma\, b.$$

(ii) Under the same hypotheses, if $\hat{B} \le \frac{\pi}{2}$ we have

$$\left|c - a\,\cos\hat{B}\right| \le b.$$

Proof Part (i). By the triangle inequality we have $a \le 2\rho$. Then, for any $\gamma > 1$ we can choose ρ small so that $\sin\theta \le \theta \le \gamma\sin\theta$ for $0 \le \theta \le 2\rho$. Using the spherical law of sines we have

$$a\sin\hat{B} \le \gamma\sin a\sin\hat{B} = \gamma\sin b\sin\hat{A} \le \gamma\sin b \le \gamma b.$$

Part (ii). Assume that ρ is small. We shall prove that $\cos(c - a\,\cos\hat{B}) \ge \cos b$, which shall imply that $|c - a\,\cos\hat{B}| \le b$. By the spherical law of cosines we have

$$\cos b = \cos c\cos a + \sin c\sin a\cos\hat{B}.$$

Note that

$$\cos(c - a\,\cos\hat{B}) = \cos c\,\cos(a\,\cos\hat{B}) + \sin c\,\sin(a\,\cos\hat{B}).$$

Since $0 \le a\,\cos\hat{B} \le a$ we have that $\cos(a\,\cos\hat{B}) \ge \cos a$. Also, by elementary calculus we have $\sin(a\,\cos\hat{B}) \ge \sin a\cos\hat{B}$, and the result plainly follows from these estimates. ■

We conclude this part with another elementary fact.

Lemma 2.2.6. *We have*

$$u(t) := \frac{\int_0^t (\sin s)^{d-1}\, \mathrm{d}s}{t\,(\sin t)^{d-1}} = \frac{\sigma(t)}{t\,\sigma'(t)} \simeq_d 1$$

for $0 \le t \le 1/4$. *Moreover, u is a C^∞-function in this range.*

Proof Note that

$$\frac{\int_0^t (\sin s)^{d-1}\, \mathrm{d}s}{t(\sin t)^{d-1}} = \frac{1}{t}\int_0^t \left(\frac{\sin s}{\sin t}\right)^{d-1} \mathrm{d}s = \frac{\sin t}{t}\int_0^1 a^{d-1}\frac{1}{(1-a^2(\sin t)^2)^{1/2}}\, \mathrm{d}a\,,$$

and both $t \mapsto \frac{\sin t}{t}$ and $t \mapsto \int_0^1 a^{d-1}\frac{1}{(1-a^2(\sin t)^2)^{1/2}}\, \mathrm{d}a$ are smooth functions bounded above and below in the proposed range. ■

Estimates for large radii - main lemma

We are now in position to prove the key result of this subsection.

Lemma 2.2.7. *Let* $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ *and let* $\overline{\mathcal{B}_r(\zeta)} \in \mathbf{B}_\xi$, *with* ζ *in the half great circle determined by* $\mathbf{e}$, ξ *and* $-\mathbf{e}$. *Assume that* $0 \le \theta(\zeta) < \theta(\xi)$, *that* $\xi \in \partial\mathcal{B}_r(\zeta)$, *that* $\widetilde{\mathcal{M}}f(\xi) > f(\xi)$ *and that* $\widetilde{\mathcal{M}}f$ *is differentiable at* ξ. *There is a universal constant* $\rho > 0$ *such that if* $\mathcal{B} = \mathcal{B}_r(\zeta) \subset \overline{\mathcal{B}_\rho(\mathbf{e})}$ *then*

$$\left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| \lesssim_d \frac{1}{\theta(\xi)}⨍_{\mathcal{B}} |\nabla f(\eta)|\,\theta(\eta)\, \mathrm{d}\sigma(\eta) + \frac{r\,\theta(\zeta)}{\theta(\xi)}⨍_{\mathcal{B}} |\nabla f(\eta)|\, \mathrm{d}\sigma(\eta). \tag{2.20}$$

Proof From Lemma 2.4.3 we have

$$\nabla\widetilde{\mathcal{M}}f(\xi)(-v(\xi,\mathbf{e})) = \frac{\sigma'(r)}{\sigma(r)}⨍_{\mathcal{B}} \nabla f(\eta)(-v(\eta,\zeta))\,\frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\, \mathrm{d}\sigma(\eta). \tag{2.21}$$

In the case $\zeta = \mathbf{e}$, estimate (2.20) follows directly from (2.21) and Lemma 2.2.6. From now on we assume that $\zeta \ne \mathbf{e}$. From Lemma 2.2.1 we also know that

$$\nabla\widetilde{\mathcal{M}}f(\xi)(-v(\xi,\mathbf{e})) = ⨍_{\mathcal{B}} \nabla f(\eta)\, S(\eta)\, \mathrm{d}\sigma(\eta), \tag{2.22}$$

with $S(\eta) = (\eta\cdot v(\xi,\mathbf{e}))\xi - (\eta\cdot\xi)v(\xi,\mathbf{e})$. The idea is to compare the identities (2.21) and (2.22) in order to bound $\left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| = \left|\nabla\widetilde{\mathcal{M}}f(\xi)(-v(\xi,\mathbf{e}))\right|$. To do so, we write the right-hand side of (2.22) as a sum of three terms, one being comparable to $\left|\nabla\widetilde{\mathcal{M}}f(\xi)\right|$, the second one being small, and the third one being close to the right-hand side of (2.21) in a suitable sense. We start by writing

$$1 = \frac{\theta(\xi) - \theta(\zeta)}{r} = \frac{d(\mathbf{e},\xi) - d(\mathbf{e},\zeta)}{r}.$$

Let us define $v_1(\eta) = S(\eta)/|S(\eta)|$. We then have

$$
\begin{aligned}
\fint_{\mathcal{B}} \nabla f(\eta)\, S(\eta)\, \mathrm{d}\sigma(\eta) = & \fint_{\mathcal{B}} \nabla f(\eta)\, |S(\eta)| \left(\frac{\theta(\xi)-\theta(\zeta)}{r}\right) v_1(\eta)\, \mathrm{d}\sigma(\eta) \\
= & \fint_{\mathcal{B}} \nabla f(\eta)\, |S(\eta)|\, \frac{\theta(\xi)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta) \\
& - \fint_{\mathcal{B}} \nabla f(\eta) \big(|S(\eta)|-1\big) \frac{\theta(\zeta)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta) - \fint_{\mathcal{B}} \nabla f(\eta) \frac{\theta(\zeta)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta).
\end{aligned} \tag{2.23}
$$

Step 1. Let us start by bounding the quantity

$$
\frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(-v(\eta,\zeta))\, \frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\, \mathrm{d}\sigma(\eta) + \fint_{\mathcal{B}} \nabla f(\eta) \frac{\theta(\zeta)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta).
$$

This last expression is equal to (recall the definition of u in Lemma 2.2.6)

$$
\frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta) \big[d(\zeta,\eta)\, u(d(\zeta,\eta))\, (-v(\eta,\zeta)) + d(\mathbf{e},\zeta)\, u(r)\, v_1(\eta)\big]\, \mathrm{d}\sigma(\eta). \tag{2.24}
$$

Note now that

$$
\begin{aligned}
d(\zeta,\eta)\, u(d(\zeta,\eta))\, (-v(\eta,\zeta)) + d(\mathbf{e},\zeta)\, u(r)\, v_1(\eta) = & \, u(d(\zeta,\eta)) \big[d(\zeta,\eta)(-v(\eta,\zeta)) + d(\mathbf{e},\zeta) v_1(\eta)\big] \\
& - d(\mathbf{e},\zeta) \big[u(d(\zeta,\eta)) - u(r)\big] v_1(\eta).
\end{aligned}
$$

From Lemma 2.2.6 we know that $u(t)$ is Lipschitz for $0 \leq t \leq 1/4$. We then have $|u(d(\zeta,\eta)) - u(r)| \lesssim_d r$ and another application of Lemma 2.2.6 yields

$$
\frac{\sigma'(r)}{\sigma(r)} \left| \fint_{\mathcal{B}} \nabla f(\eta)\, d(\mathbf{e},\zeta) \big[u(d(\zeta,\eta)) - u(r)\big] v_1(\eta)\, \mathrm{d}\sigma(\eta) \right| \lesssim_d \fint_{\mathcal{B}} |\nabla f(\eta)|\, d(\mathbf{e},\zeta)\, \mathrm{d}\sigma(\eta). \tag{2.25}
$$

Let us now deal with the remaining piece. Observe that

$$
\begin{aligned}
& d(\zeta,\eta)\, (-v(\eta,\zeta)) + d(\mathbf{e},\zeta)\, v_1(\eta) = d(\zeta,\eta) \big(v_1(\eta) \cos\alpha + v_1(\eta)^* \sin\alpha\big) + d(\mathbf{e},\zeta)\, v_1(\eta) \\
& \quad = \big[d(\zeta,\eta) v_1(\eta) \cos\beta + d(\mathbf{e},\zeta)\, v_1(\eta)\big] + \big[d(\zeta,\eta) v_1(\eta)^* \sin\alpha\big] + \big[d(\zeta,\eta) v_1(\eta) (\cos\alpha - \cos\beta)\big] \\
& \quad = [I] + [II] + [III],
\end{aligned} \tag{2.26}
$$

where $\cos\alpha = -v(\eta,\zeta) \cdot v_1(\eta)$ $(0 \leq \alpha \leq \pi)$, $v_1(\eta)^*$ is unitary and orthogonal to $v_1(\eta)$ (in the plane determined by $v_1(\eta)$ and $v(\eta,\zeta)$), and $\cos\beta = v(\zeta,\eta) \cdot (-v(\zeta,\mathbf{e}))$ $(0 \leq \beta \leq \pi)$. Naturally, we may assume without loss of generality that $\eta \neq \zeta$. We now proceed with the analysis of the three terms in (2.26).

Analysis of $[I]$. Observe that

$$
|d(\zeta,\eta) v_1(\eta) \cos\beta + d(\mathbf{e},\zeta)\, v_1(\eta)| = |d(\zeta,\eta) \cos\beta + d(\mathbf{e},\zeta)|.
$$

Consider the geodesic triangle with vertices $\mathbf{e}, \zeta, \eta$ (that has angle $\angle \mathbf{e}\zeta\eta = \pi - \beta$). Assuming ρ small, if $\beta > \pi/2$ we may use Lemma 2.2.5 (ii) to find

$$|d(\zeta, \eta) \cos \beta + d(\mathbf{e}, \zeta)| \le d(\mathbf{e}, \eta).$$

In case $0 \le \beta \le \pi/2$ we have

$$0 \le \operatorname{sgn}(\cos \beta) = \operatorname{sgn}\big[(\eta - (\zeta \cdot \eta)\zeta) \cdot (-\mathbf{e} + (\zeta \cdot \mathbf{e})\zeta)\big] = \operatorname{sgn}\big[-(\eta \cdot \mathbf{e}) + (\zeta \cdot \mathbf{e})(\zeta \cdot \eta)\big],$$

which implies that

$$\cos(\theta(\zeta)) = (\zeta \cdot \mathbf{e}) \ge (\zeta \cdot \mathbf{e})(\zeta \cdot \eta) \ge (\eta \cdot \mathbf{e}) = \cos(\theta(\eta)).$$

From this we conclude that $d(\mathbf{e}, \zeta) = \theta(\zeta) \le \theta(\eta) = d(\mathbf{e}, \eta)$ and hence

$$|d(\zeta, \eta) \cos \beta + d(\mathbf{e}, \zeta)| \le d(\zeta, \eta) + d(\mathbf{e}, \zeta) \le (d(\mathbf{e}, \zeta) + d(\mathbf{e}, \eta)) + d(\mathbf{e}, \zeta) \le 3d(\mathbf{e}, \eta).$$

Analysis of $[II]$ *and* $[III]$. We note that the angles α and β are close, and it is important for our purposes to actually quantify this discrepancy. In order to do this, let us parametrize the points as follows. We write $\zeta = (\cos \theta, \sin \theta, \mathbf{0})$, with $\mathbf{0} \in \mathbb{R}^{d-1}$, and $\eta = (\cos \theta_1, \sin \theta_1 \cos \varphi, \sin \theta_1 \sin \varphi \; \omega)$ with $\omega \in \mathbb{S}^{d-2} \subset \mathbb{R}^{d-1}$. Here we set $0 \le \theta, \theta_1, \varphi \le \pi$. Recall that in this notation we have $\mathbf{e} = (1, 0, \mathbf{0})$. We then have $-v(\zeta, \mathbf{e}) = (-\sin \theta, \cos \theta, \mathbf{0})$. Recall also that the vector $v_1(\eta)$ is the unitary vector tangent to η in the direction of the derivative of the curve that takes the point η along the rotation in the first two coordinates (in the direction from $\mathbf{e}$ to ζ). A direct computation yields

$$S(\eta) = (-\sin \theta_1 \cos \varphi, \cos \theta_1, \mathbf{0}) \tag{2.27}$$

and

$$v_1(\eta) = \frac{1}{\sqrt{1 - \sin^2 \theta_1 \sin^2 \varphi}} (-\sin \theta_1 \cos \varphi, \cos \theta_1, \mathbf{0}).$$

Using that $v(\zeta, \mathbf{e}) \perp \zeta$ and $v_1(\eta) \perp \eta$ we then find

$$\cos \beta = v(\zeta, \eta) \cdot (-v(\zeta, \mathbf{e})) = \frac{\eta - (\eta \cdot \zeta)\zeta}{|\eta - (\eta \cdot \zeta)\zeta|} \cdot (-v(\zeta, \mathbf{e})) = \frac{-\sin \theta \cos \theta_1 + \cos \theta \sin \theta_1 \cos \varphi}{|\eta - (\eta \cdot \zeta)\zeta|}$$

and

$$\cos \alpha = -v(\eta, \zeta) \cdot v_1(\eta) = \frac{-\zeta + (\eta \cdot \zeta)\eta}{|-\zeta + (\eta \cdot \zeta)\eta|} \cdot v_1(\eta) = \frac{-\sin \theta \cos \theta_1 + \cos \theta \sin \theta_1 \cos \varphi}{\sqrt{1 - \sin^2 \theta_1 \sin^2 \varphi}\; |-\zeta + (\eta \cdot \zeta)\eta|}.$$

Since $|\eta - (\eta \cdot \zeta)\zeta| = |-\zeta + (\eta \cdot \zeta)\eta| = \sqrt{1 - (\eta \cdot \zeta)^2}$, we plainly obtain that $|\cos \beta| \le |\cos \alpha|$ and hence $\sin \alpha \le \sin \beta$. Using Lemma 2.2.5 (i) we then find

$$|d(\zeta, \eta) v_1(\eta)^* \sin \alpha| \le d(\zeta, \eta) \sin \beta \lesssim d(\mathbf{e}, \eta).$$

This takes care of the term $[II]$ in (2.26). Finally, we recall that all the action takes place inside a small ball $\mathcal{B}_\rho(\mathbf{e})$, which means that the angles θ and θ_1 are small. This yields an estimate for the term $[III]$ of the form

$$\begin{aligned}
|d(\zeta,\eta)v_1(\eta)(\cos\alpha-\cos\beta)| &\lesssim |\zeta-\eta||\cos\alpha-\cos\beta| \\
&= \frac{\sqrt{2(1-(\eta\cdot\zeta))}}{\sqrt{1-(\eta\cdot\zeta)^2}}\,|-\sin\theta\cos\theta_1+\cos\theta\sin\theta_1\cos\varphi|\left(\frac{1}{\sqrt{1-\sin^2\theta_1\sin^2\varphi}}-1\right) \\
&\lesssim \sin^2\theta_1 \\
&\lesssim \theta_1 = d(\mathbf{e},\eta).
\end{aligned}$$

Combining (2.24), (2.25) and the bounds for the terms $[I],[II],[III]$ in (2.26), and using Lemma 2.2.6, we arrive at

$$\begin{aligned}
&\left|\frac{\sigma'(r)}{\sigma(r)}\fint_{\mathcal{B}}\nabla f(\eta)(-v(\eta,\zeta))\,\frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\,\mathrm{d}\sigma(\eta)+\fint_{\mathcal{B}}\nabla f(\eta)\frac{\theta(\zeta)}{r}\,v_1(\eta)\,\mathrm{d}\sigma(\eta)\right| \\
&\qquad\lesssim_d \fint_{\mathcal{B}}|\nabla f(\eta)|\,\theta(\zeta)\,\mathrm{d}\sigma(\eta)+\frac{1}{r}\fint_{\mathcal{B}}|\nabla f(\eta)|\,\theta(\eta)\,\mathrm{d}\sigma(\eta).
\end{aligned} \tag{2.28}$$

Step 2. We continue our analysis with the term

$$\fint_{\mathcal{B}}\nabla f(\eta)\,\big(|S(\eta)|-1\big)\frac{\theta(\zeta)}{r}\,v_1(\eta)\,\mathrm{d}\sigma(\eta).$$

From (2.27) we know that $|S(\eta)|^2=\eta\cdot p(\eta)$, where $p(\eta)$ is the projection of η over the plane generated by ζ and $\mathbf{e}$. Therefore

$$\begin{aligned}
\left|\fint_{\mathcal{B}}\nabla f(\eta)\,\big(|S(\eta)|-1\big)\frac{\theta(\zeta)}{r}\,v_1(\eta)\,\mathrm{d}\sigma(\eta)\right| &\leq \fint_{\mathcal{B}}|\nabla f(\eta)|\,\big(1-|S(\eta)|^2\big)\frac{\theta(\zeta)}{r}\,\mathrm{d}\sigma(\eta) \\
&\leq \fint_{\mathcal{B}}\big|\nabla f(\eta)\big|\,\big|\eta\cdot(\eta-p(\eta))\big|\frac{\theta(\zeta)}{r}\,\mathrm{d}\sigma(\eta) \\
&\leq \fint_{\mathcal{B}}\big|\nabla f(\eta)\big|\,|\eta-p(\eta)|\frac{\theta(\zeta)}{r}\,\mathrm{d}\sigma(\eta) \\
&\leq \fint_{\mathcal{B}}\big|\nabla f(\eta)\big|\,\theta(\zeta)\,\mathrm{d}\sigma(\eta).
\end{aligned} \tag{2.29}$$

Step 3. Combining (2.21), (2.22), (2.23), (2.28) and (2.29) we find that

$$\left|\fint_{\mathcal{B}}\nabla f(\eta)\,|S(\eta)|\frac{\theta(\xi)}{r}\,v_1(\eta)\,\mathrm{d}\sigma(\eta)\right| \lesssim_d \fint_{\mathcal{B}}|\nabla f(\eta)|\,\theta(\zeta)\,\mathrm{d}\sigma(\eta)+\frac{1}{r}\fint_{\mathcal{B}}|\nabla f(\eta)|\,\theta(\eta)\,\mathrm{d}\sigma(\eta)\,,$$

and therefore

$$\begin{aligned}
\left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| &= \left|\fint_{\mathcal{B}}\nabla f(\eta)\,S(\eta)\,\mathrm{d}\sigma(\eta)\right| \\
&\lesssim_d \frac{1}{\theta(\xi)}\fint_{\mathcal{B}}|\nabla f(\eta)|\,\theta(\eta)\,\mathrm{d}\sigma(\eta)+\frac{r\,\theta(\zeta)}{\theta(\xi)}\fint_{\mathcal{B}}|\nabla f(\eta)|\,\mathrm{d}\sigma(\eta).
\end{aligned}$$

This concludes the proof of the lemma. ■

Proof of Theorem 2.1.1 - Lipschitz case

We are now in position to move on to the proof of Theorem 2.1.1 when our initial datum f is a Lipschitz function. In this case we also have $\widetilde{\mathcal{M}}f$ Lipschitz. Consider the set $\mathcal{H}_d = \{\xi \in \mathbb{S}^d \ : \ \widetilde{\mathcal{M}}f(\xi) > \widetilde{\mathcal{M}}^I f(\xi)\}$. In light of Proposition 2.2.1 it suffices to show that

$$\int_{\mathcal{H}_d} \left|\nabla \widetilde{\mathcal{M}}f(\xi)\right| \, \mathrm{d}\sigma(\xi) \lesssim_d \int_{\mathbb{S}^d} |\nabla f(\xi)| \, \mathrm{d}\sigma(\xi).$$

For each $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ let us choose a ball $\overline{\mathcal{B}_{r_\xi}(\zeta_\xi)} \in \mathbf{B}_\xi$ with r_ξ minimal and, subject to this condition, with ζ_ξ in the half great circle connecting $\mathbf{e}, \xi, -\mathbf{e}$ in a way that $w(\zeta_\xi) = \min\{d(\mathbf{e}, \zeta_\xi), d(-\mathbf{e}, \zeta_\xi)\}$ is minimal. If there are two potential choices for ζ_ξ we choose the one with $0 \leq \theta(\zeta_\xi) \leq \theta(\xi)$.

First let us observe that we can restrict our attention to small balls. For $c > 0$, define the set $\mathcal{R}_c = \{\xi \in \mathbb{S}^d \ : \xi \in \mathcal{H}_d \text{ and } r_\xi \geq c\}$. By Lemma 2.2.1 we find

$$\int_{\mathcal{R}_c} \left|\nabla \widetilde{\mathcal{M}}f(\xi)\right| \, \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{R}_c} \frac{1}{\sigma(\mathcal{B}_{r_\xi}(\zeta_\xi))} \int_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta) \, \mathrm{d}\sigma(\xi) \lesssim_{c,d} \int_{\mathbb{S}^d} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta).$$

If $\xi \in \mathcal{H}_d$ and r_ξ is small we must have $w(\zeta_\xi) < 4r_\xi$ (otherwise we would fall in the regime of the operator $\widetilde{\mathcal{M}}^I$). Assuming that $\xi \in \mathcal{H}_d$, that $\widetilde{\mathcal{M}}f$ is differentiable at ξ, and that $\nabla \widetilde{\mathcal{M}}f(\xi) \neq 0$ (which implies that $\xi \in \partial \mathcal{B}_{r_\xi}(\zeta_\xi)$), we may restrict ourselves to the situation where $d(\mathbf{e}, \xi) \leq \rho$ or $d(-\mathbf{e}, \xi) \leq \rho$ (where ρ is given by Lemma 2.2.7). By symmetry let us assume that $\theta(\xi) = d(\mathbf{e}, \xi) \leq \rho$. We call such set $\mathcal{G}_d$ and further decompose it in $\mathcal{G}_d^- = \{\xi \in \mathcal{G}_d \ : \ 0 \leq \theta(\zeta_\xi) < \theta(\xi)\}$ and $\mathcal{G}_d^+ = \{\xi \in \mathcal{G}_d \ : \ 0 < \theta(\xi) < \theta(\zeta_\xi)\}$. We bound the integrals over these two sets separately.

Step 1. For $\mathcal{G}_d^+$ we use Lemma 2.2.1 and proceed as follows:

$$\begin{aligned} \int_{\mathcal{G}_d^+} \left|\nabla \widetilde{\mathcal{M}}f(\xi)\right| \, \mathrm{d}\sigma(\xi) &\leq \int_{\mathcal{G}_d^+} ⨍_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)| \, \mathrm{d}\sigma(\eta) \, \mathrm{d}\sigma(\xi) \\ &= \int_{\mathbb{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^+} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)}{\sigma(\mathcal{B}_{r_\xi}(\zeta_\xi))} \, \mathrm{d}\sigma(\xi) \, \mathrm{d}\sigma(\eta). \end{aligned} \tag{2.30}$$

Note that $\theta(\eta) \geq \theta(\xi)$ in this case. Observe that

$$r_\xi > \frac{w(\zeta_\xi)}{4} = \frac{\theta(\zeta_\xi)}{4} \geq \frac{\theta(\xi)}{4}, \tag{2.31}$$

and also, by triangle inequality,

$$r_\xi \geq \frac{d(\eta, \xi)}{2} \geq \frac{\theta(\eta)}{2} - \frac{\theta(\xi)}{2}. \tag{2.32}$$

Dividing (2.32) by 2 and adding up to (2.31) we get

$$r_\xi \geq \frac{\theta(\eta)}{6}.$$

Returning to the computation (2.30) we have, for a fixed η,

$$\int_{\mathcal{G}_d^+} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)}{\sigma(\mathcal{B}_{r_\xi}(\zeta_\xi))}\, \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{B}_{\theta(\eta)}(\mathbf{e})} \frac{1}{\sigma\left(\frac{\theta(\eta)}{6}\right)}\, \mathrm{d}\sigma(\xi) \simeq_d 1,$$

from which the required bound follows.

Step 2. We now bound the integral over $\mathcal{G}_d^-$ using Lemma 2.2.7. If $\xi \in \mathcal{G}_d^-$ then

$$r_\xi \leq \theta(\xi) < 5r_\xi. \tag{2.33}$$

We then have

$$\begin{aligned}
&\int_{\mathcal{G}_d^-} \left|\nabla \widetilde{\mathcal{M}} f(\xi)\right| \mathrm{d}\sigma(\xi) \\
&\lesssim_d \int_{\mathcal{G}_d^-} \left(\frac{1}{\theta(\xi)} \fint_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)|\, \theta(\eta)\, \mathrm{d}\sigma(\eta) + \frac{r_\xi\, \theta(\zeta_\xi)}{\theta(\xi)} \fint_{\mathcal{B}_{r_\xi}(\zeta_\xi)} |\nabla f(\eta)|\, \mathrm{d}\sigma(\eta) \right) \mathrm{d}\sigma(\xi) \\
&\lesssim_d \int_{\mathbb{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)\, \theta(\eta)}{r_\xi\, \sigma(r_\xi)}\, \mathrm{d}\sigma(\xi)\, \mathrm{d}\sigma(\eta) \\
&\quad + \int_{\mathbb{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)\, \theta(\zeta_\xi)}{\sigma(r_\xi)}\, \mathrm{d}\sigma(\xi)\, \mathrm{d}\sigma(\eta).
\end{aligned} \tag{2.34}$$

Using (2.33) and the fact that $\theta(\zeta_\xi) \leq \theta(\xi)$ in this case, we have, for a fixed η,

$$\int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)\, \theta(\zeta_\xi)}{\sigma(r_\xi)}\, \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)\, \theta(\xi)}{\sigma(r_\xi)}\, \mathrm{d}\sigma(\xi) \lesssim_d \int_0^\rho \frac{\theta\, (\sin\theta)^{d-1}}{\sigma(\theta)}\, \mathrm{d}\theta \lesssim_d 1\,, \tag{2.35}$$

where we used Lemma 2.2.6 in the last inequality. For the other integral, we use (2.33), the fact that $\theta(\eta) \leq \theta(\xi)$ in this case, and Lemma 2.2.6 again to get

$$\int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_{r_\xi}(\zeta_\xi)}(\eta)\, \theta(\eta)}{r_\xi\, \sigma(r_\xi)}\, \mathrm{d}\sigma(\xi) \leq \theta(\eta) \int_{\theta(\eta)}^\rho \frac{(\sin\theta)^{d-1}}{r_\xi\, \sigma(r_\xi)}\, \mathrm{d}\theta \lesssim_d \theta(\eta) \int_{\theta(\eta)}^\rho \frac{1}{\theta^2}\, \mathrm{d}\theta \lesssim 1. \tag{2.36}$$

Our desired inequality plainly follows from inserting the bounds given by (2.35) and (2.36) into (2.34). This completes the proof of Theorem 2.1.1 in the Lipschitz case.

2.2.3 Passage to the general case

We will be brief here since the outline is the same as in §1.2.2. The following lemma is the analogue of Lemma 1.2.1 in the case of the sphere and we omit its proof.

Lemma 2.2.8. .

(i) A polar function $f(\xi)$ is weakly differentiable in $\mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ if and only if its polar restriction $f(\theta)$ is weakly differentiable in $(0, \pi)$.

(ii) In the situation above, if $f(\xi)$ and $\nabla f(\xi)$ are locally integrable in neighborhoods of $\mathbf{e}$ *and* $-\mathbf{e}$*, then* f *is weakly differentiable in* $\mathbb{S}^d$*.*

Consider now a (nonnegative) polar function $f(\xi)$ in $W^{1,1}(\mathbb{S}^d)$. Then, by Lemma 2.2.8, its polar version $f(\theta)$ is weakly differentiable in $(0,\pi)$ and verifies

$$\int_0^\pi |f'(\theta)| (\sin\theta)^{d-1} \, \mathrm{d}\theta < \infty.$$

In particular, after a possible redefinition on a set of measure zero, one can take $f(\theta)$ continuous in $(0,\pi)$ (in fact, absolutely continuous in each compact interval of $(0,\pi)$). This is equivalent to assuming that $f(\xi)$ is continuous in $\mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$.

In this case the detachment set

$$\mathcal{D}_d := \{\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\} \ : \ \widetilde{\mathcal{M}}f(\xi) > f(\xi)\}$$

is an open set. One can also show that $\widetilde{\mathcal{M}}f$ is continuous in $\mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$, being indeed locally Lipschitz in $\mathcal{D}_d$ and the remark thereafter). In particular, $\widetilde{\mathcal{M}}f$ is differentiable almost everywhere in $\mathcal{D}_d$.

Let $\{f_n\} \subset C^\infty(\mathbb{S}^d)$ be a sequence of nonnegative smooth functions such that $f_n \to f$ in $W^{1,1}(\mathbb{S}^d)$. We may simply assume that f_n is given by the spherical convolution of f with a smooth polar kernel φ_n (say, non-increasing in the polar angle) of integral 1 supported in the geodesic ball of radius $1/n$ centered at the north pole; see [DX13, Chapter 2, §2.1 and §2.3, and Proposition 2.6.4] for details on the spherical convolution. We may also assume that $f_n \to f$ and $\nabla f_n \to \nabla f$ pointwise almost everywhere in $\mathbb{S}^d$ (say, outside a set $\mathcal{X} \subset \mathbb{S}^d$ of measure zero). Let $\xi \in \mathcal{D}_d \setminus \mathcal{X}$ be a point at which $\widetilde{\mathcal{M}}f$ is differentiable and all $\widetilde{\mathcal{M}}f_n$ are differentiable (this is still almost everywhere in $\mathcal{D}_d$). Note that for n large we shall have $\xi \in \{\widetilde{\mathcal{M}}f_n(\xi) > f_n(\xi)\}$. We now observe that if $\mathcal{B}_n = \mathcal{B}_{r_n}(\zeta_n)$ is a ball that realizes the maximal function $\widetilde{\mathcal{M}}f_n(\xi)$ with $r_n \to r$ and $\zeta_n \to \zeta$, then we must have $r > 0$ and the limiting ball $\mathcal{B}_r(\zeta)$ realizing the maximal function $\widetilde{\mathcal{M}}f(\xi)$. This plainly implies that

$$\widetilde{\mathcal{M}}f_n(\xi) \to \widetilde{\mathcal{M}}f(\xi)$$

as $n \to \infty$, and also, by Lemma 2.4.2,

$$\nabla\widetilde{\mathcal{M}}f_n(\xi) \to \nabla\widetilde{\mathcal{M}}f(\xi)$$

as $n \to \infty$.

Since we have proved Theorem 2.1.1 for Lipschitz functions, using Fatou's lemma we have

$$\int_{\mathcal{D}_d} \left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| \, \mathrm{d}\sigma(\xi) \leq \liminf_{n\to\infty} \int_{\mathcal{D}_d} \left|\nabla\widetilde{\mathcal{M}}f_n(\xi)\right| \, \mathrm{d}\sigma(\xi) \lesssim_d \liminf_{n\to\infty} \|\nabla f_n\|_{L^1(\mathbb{S}^d)} = \|\nabla f\|_{L^1(\mathbb{S}^d)}. \tag{2.37}$$

This places us in position to adapt the one-dimensional argument of [CS13, §5.4] to show that $\widetilde{\mathcal{M}}f(\theta)$ is weakly differentiable in $(0,\pi)$, with weak derivative given by

$$\chi_{\mathcal{D}_1^c} f'(\theta) + \chi_{\mathcal{D}_1} (\widetilde{\mathcal{M}}f)'(\theta)\,, \tag{2.38}$$

where $\mathcal{D}_1 = \{\theta(\xi) \ : \ \xi \in \mathcal{D}_d\}$ is the polar version of $\mathcal{D}_d$. In fact, if $\theta \in \mathcal{D}_1^c$ is a point of differentiability of f (which are almost all points of $\mathcal{D}_1^c$) one can verify that $f'(\theta) = 0$, otherwise θ would belong to $\mathcal{D}_1$ instead. The weak derivative of $\widetilde{\mathcal{M}}f(\theta)$ is then simply $\chi_{\mathcal{D}_1}\big(\widetilde{\mathcal{M}}f\big)'(\theta)$. This in turn implies that $\widetilde{\mathcal{M}}f$ is weakly differentiable in $\mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ by Lemma 2.2.8. From (2.37) and (2.38) we have

$$\big\|\nabla \widetilde{\mathcal{M}}f\big\|_{L^1(\mathbb{S}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbb{S}^d)}, \tag{2.39}$$

which is our desired bound. From the Sobolev embedding we know that $f \in L^{d/(d-1)}(\mathbb{S}^d)$, and hence so does $\widetilde{\mathcal{M}}f$. In particular, $\widetilde{\mathcal{M}}f$ is locally integrable in $\mathbb{S}^d$. From (2.39) we already know that $\nabla\widetilde{\mathcal{M}}f$ is locally integrable in $\mathbb{S}^d$, and a further application of Lemma 2.2.8 shows that $\widetilde{\mathcal{M}}f$ is in fact weakly differentiable in $\mathbb{S}^d$, which completes our proof.

2.3 Proof of Theorem 2.1.2

We now turn our attention to the proof of Theorem 2.1.2. As presented in the introduction, the notation here is slightly different, as we denote our initial datum by u_0 and our maximal function by u^*. As usual, throughout this section, we assume that u_0 is real-valued and nonnegative (or $+\infty$).

2.3.1 Lipschitz case

Now, we address first the case when our polar $u_0 \in W^{1,1}(\mathbb{S}^d)$ is a Lipschitz function. In this case we have that u^* is a polar function that is also Lipschitz (see [CFS18, Lemma 16 (ii)]).

A preliminary lemma

The following result will be important for our purposes.

Lemma 2.3.1. *Let $u_0 : \mathbb{S}^d \to \mathbb{R}^+$ be a polar and Lipschitz function. Then, in polar coordinates,*

$$\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) - u_0\big(\tfrac{\pi}{2}\big) \lesssim_d \|\nabla u_0\|_{L^1(\mathbb{S}^d)}.$$

Proof Let us assume that $\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) > u_0\big(\tfrac{\pi}{2}\big)$. First observe that

$$\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) - u_0\big(\tfrac{\pi}{2}\big) = \left(\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) - \sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta)\right) + \left(\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta) - u_0\big(\tfrac{\pi}{2}\big)\right),$$

and

$$\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta)-u_0\big(\tfrac{\pi}{2}\big)\le\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}}|u_0'(\theta)|\ \mathrm{d}\theta\lesssim_d\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}}|u_0'(\theta)|\,(\sin\theta)^{d-1}\ \mathrm{d}\theta\lesssim_d\|\nabla u_0\|_{L^1(\mathbb{S}^d)}.$$

Therefore it suffices to bound $\widetilde{\mathcal{M}}u_0\big(\frac{\pi}{2}\big)-\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]}u_0(\theta)$. Bringing things back to the notation of §2.2.1, let $\xi\in\mathbb{S}^d$ be such that $\theta(\xi)=\frac{\pi}{2}$ and let $\overline{\mathcal{B}}=\overline{\mathcal{B}_r(\zeta)}\in\mathbf{B}_\xi$. Let $\mathcal{Z}=\{\eta\in\mathbb{S}^d\ :\ \frac{\pi}{4}\le\theta(\eta)\le\frac{3\pi}{4}\}$. If $\mathcal{B}\subset\mathcal{Z}$, then $\widetilde{\mathcal{M}}u_0\big(\frac{\pi}{2}\big)-\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]}u_0(\theta)\le 0$ and we are done. Assume henceforth that $\mathcal{B}\not\subset\mathcal{Z}$ and that $\widetilde{\mathcal{M}}u_0\big(\frac{\pi}{2}\big)-\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]}u_0(\theta)\ge 0$. Writing $\eta=(\cos\theta,(\sin\theta)\,\omega)$, with $\omega\in\mathbb{S}^{d-1}$, we define

$$\ell(\theta)=\int_{\mathbb{S}^{d-1}}\chi_{\mathcal{B}}(\eta)\,(\sin\theta)^{d-1}\ \mathrm{d}\sigma_{d-1}(\omega)$$

(that is, the $(d-1)$-dimensional measure of the intersection of $\mathcal{B}$ with the level set $d(\mathbf{e},\eta)=\theta$). We then have

$$\begin{aligned}\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big)&=\fint_{\mathcal{B}}u_0(\eta)\ \mathrm{d}\sigma(\eta)=\frac{1}{\sigma(\mathcal{B})}\int_0^\pi u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta\\&=\frac{1}{\sigma(\mathcal{B})}\left(\int_0^{\frac{\pi}{4}}u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta+\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}}u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta+\int_{\frac{3\pi}{4}}^{\pi}u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta\right)\\&\le\left(\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]}u_0(\theta)\right)\frac{1}{\sigma(\mathcal{B})}\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}}\ell(\theta)\ \mathrm{d}\theta+\frac{1}{\sigma(\mathcal{B})}\left(\int_0^{\frac{\pi}{4}}u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta+\int_{\frac{3\pi}{4}}^{\pi}u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta\right).\end{aligned}\tag{2.40}$$

Now observe that

$$\begin{aligned}\int_{\frac{3\pi}{4}}^{\pi}u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta&=\int_{\frac{3\pi}{4}}^{\pi}\left(\int_{\frac{3\pi}{4}}^{\theta}u_0'(\tau)\ \mathrm{d}\tau+u_0(\tfrac{3\pi}{4})\right)\ell(\theta)\ \mathrm{d}\theta\\&=u_0(\tfrac{3\pi}{4})\int_{\frac{3\pi}{4}}^{\pi}\ell(\theta)\ \mathrm{d}\theta+\int_{\frac{3\pi}{4}}^{\pi}u_0'(\tau)\left(\int_\tau^\pi\ell(\theta)\ \mathrm{d}\theta\right)\mathrm{d}\tau.\end{aligned}\tag{2.41}$$

Plugging the bound

$$\int_\tau^\pi\ell(\theta)\ \mathrm{d}\theta\lesssim_d\int_\tau^\pi(\sin\theta)^{d-1}\ \mathrm{d}\theta\ \lesssim\ (\sin\tau)^{d-1}$$

into (2.41) we get

$$\int_{\frac{3\pi}{4}}^{\pi}u_0(\theta)\,\ell(\theta)\ \mathrm{d}\theta\le\left(\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]}u_0(\theta)\right)\int_{\frac{3\pi}{4}}^{\pi}\ell(\theta)\ \mathrm{d}\theta+C_d\,\|\nabla u_0\|_{L^1(\mathbb{S}^d)},\tag{2.42}$$

where C_d is a universal constant. In an analogous way we obtain

$$\int_0^{\frac{\pi}{4}} u_0(\theta)\,\ell(\theta)\,\mathrm{d}\theta \leq \left(\sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta)\right)\int_0^{\frac{\pi}{4}} \ell(\theta)\,\mathrm{d}\theta + C_d\,\|\nabla u_0\|_{L^1(\mathbb{S}^d)}. \tag{2.43}$$

Combining (2.40), (2.42) and (2.43) we get

$$\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) \leq \sup_{\theta\in[\frac{\pi}{4},\frac{3\pi}{4}]} u_0(\theta) + C_d\,\|\nabla u_0\|_{L^1(\mathbb{S}^d)},$$

from where our result follows. ■

Proof of Theorem 2.1.2 - Lipschitz case

Assume $d \geq 2$ since the case $d = 1$ has already been treated in [CFS18, Theorem 3]. Define the detachment set (excluding the poles)

$$\mathcal{A}_d = \{\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}\ :\ u^*(\xi) > u_0(\xi)\}$$

and its one-dimensional polar version

$$\mathcal{A}_1 = \{\theta(\xi)\ :\ \xi \in \mathcal{A}_d\} \subset (0,\pi).$$

These sets are open and from [CFS18, Lemma 17] we know that u^* is subharmonic on $\mathcal{A}_d$. We write

$$\mathcal{A}_1 = \bigcup_{i=0}^{\infty}(a_i,b_i)$$

as a countable union of disjoint open intervals. If $\frac{\pi}{2} \in \mathcal{A}_1$ we let $\frac{\pi}{2} \in (a_0,b_0)$ and let

$$\mathcal{A}_1^- = \bigcup_{(a_i,b_i)\subset\left(0,\frac{\pi}{2}\right)} (a_i,b_i) \quad \text{and} \quad \mathcal{A}_1^+ = \bigcup_{(a_i,b_i)\subset\left(\frac{\pi}{2},\pi\right)} (a_i,b_i).$$

If $\frac{\pi}{2} \notin \mathcal{A}_1$ we just regard (a_0,b_0) as empty, and keep $\mathcal{A}_1^{\pm}$ as above.

Let (a,b) denote a generic interval (a_i,b_i) of this union. As in the proof of Theorem 4.1.1, the subharmonicity implies that u^* has no strict local maximum in (a,b) and then there exists τ with $a \leq \tau \leq b$ such that u^* is non-increasing in $[a,\tau]$ and non-decreasing in $[\tau,b]$. We then have $(u^*)'(\theta) \leq 0$ a.e. in $a < \theta < \tau$, and $(u^*)'(\theta) \geq 0$ a.e. in $\tau < \theta < b$.

An important idea of this proof is to proceed via the comparison (2.3) to the uncentered Hardy-Littlewood maximal function when appropriate, and make use of the gradient bound established in Theorem 2.1.1. We consider first the case when $(a,b) \subset \mathcal{A}_1^-$. Using integration by parts we get

$$\begin{aligned}
\int_a^b \left|(u^*)'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta &= -\int_a^\tau (u^*)'(\theta)\,(\sin\theta)^{d-1}\,\mathrm{d}\theta + \int_\tau^b (u^*)'(\theta)\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \\
&= u^*(a)\,(\sin a)^{d-1} + u^*(b)\,(\sin b)^{d-1} - 2\,u^*(\tau)\,(\sin\tau)^{d-1} \\
&\quad + \int_a^\tau u^*(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta - \int_\tau^b u^*(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\
&\le u_0(a)\,(\sin a)^{d-1} + u_0(b)\,(\sin b)^{d-1} - 2\,u_0(\tau)\,(\sin\tau)^{d-1} \qquad (2.44) \\
&\quad + \int_a^\tau \widetilde{\mathcal{M}}u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta - \int_\tau^b u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\
&\le \int_a^b \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta + \int_a^\tau \left(\widetilde{\mathcal{M}}u_0(\theta) - u_0(\theta)\right)\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta.
\end{aligned}$$

In the computation above we have taken advantage of the fact that $\frac{\partial}{\partial\theta}(\sin\theta)^{d-1} \ge 0$. Note also that we have no problem if $a = 0$ since $\lim_{a\to 0} u^*(a)\,(\sin a)^{d-1} = 0$ as $d \ge 2$. If we sum (2.44) over all the intervals $(a,b) \subset \mathcal{A}_1^-$ we find

$$\begin{aligned}
\int_{\mathcal{A}_1^-} \left|(u^*)'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta &\le \int_0^{\frac{\pi}{2}} \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta + \int_0^{\frac{\pi}{2}} \left(\widetilde{\mathcal{M}}u_0(\theta) - u_0(\theta)\right)\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\
= \int_0^{\frac{\pi}{2}} \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta &- \int_0^{\frac{\pi}{2}} \left(\left(\widetilde{\mathcal{M}}u_0\right)'(\theta) - u_0'(\theta)\right)(\sin\theta)^{d-1}\,\mathrm{d}\theta + \left(\widetilde{\mathcal{M}}u_0\left(\tfrac{\pi}{2}\right) - u_0\left(\tfrac{\pi}{2}\right)\right) \qquad (2.45) \\
\lesssim_d \int_0^\pi \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta,
\end{aligned}$$

where we have used Theorem 2.1.1 and Lemma 2.3.1.

Finally we have to consider the case when $\frac{\pi}{2} \in \mathcal{A}_1$ and bound the integral

$$\int_{a_0}^{\pi/2} \left|(u^*)'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

Let τ_0 be the corresponding local minimum over the interval (a_0, b_0). Let $c_0 = \min\{\tau_0, \frac{\pi}{2}\}$. Proceeding as in (2.44) and (2.45) we obtain

$$\begin{aligned}
-\int_{a_0}^{c_0} (u^*)'(\theta)(\sin\theta)^{d-1}\,\mathrm{d}\theta &= u^*(a_0)(\sin a_0)^{d-1} - u^*(c_0)(\sin c_0)^{d-1} + \int_{a_0}^{c_0} u^*(\theta)\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\
&\le u_0(a_0)\,(\sin a_0)^{d-1} - u_0(c_0)\,(\sin c_0)^{d-1} + \int_{a_0}^{c_0} \widetilde{\mathcal{M}}u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\
&= -\int_{a_0}^{c_0} u_0'(\theta)\,(\sin\theta)^{d-1}\,\mathrm{d}\theta + \int_{a_0}^{c_0} \left(\widetilde{\mathcal{M}}u_0(\theta) - u_0(\theta)\right)\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\
&\lesssim_d \int_0^\pi \left|u_0'(\theta)\right| (\sin\theta)^{d-1}\,\mathrm{d}\theta. \qquad (2.46)
\end{aligned}$$

The last estimate we need is the following

$$\begin{aligned}\int_{c_0}^{\frac{\pi}{2}} (u^*)'(\theta)\,(\sin\theta)^{d-1}\,\mathrm{d}\theta &= u^*\big(\tfrac{\pi}{2}\big) - u^*(c_0)(\sin c_0)^{d-1} - \int_{c_0}^{\frac{\pi}{2}} u^*(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\ &\leq \widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) - u_0(c_0)(\sin c_0)^{d-1} - \int_{c_0}^{\frac{\pi}{2}} u_0(\theta)\,\frac{\partial}{\partial\theta}(\sin\theta)^{d-1}\,\mathrm{d}\theta \\ &= \Big(\widetilde{\mathcal{M}}u_0\big(\tfrac{\pi}{2}\big) - u_0\big(\tfrac{\pi}{2}\big)\Big) + \int_{c_0}^{\frac{\pi}{2}} u_0'(\theta)\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \qquad (2.47) \\ &\lesssim_d \int_0^{\pi} \big|u_0'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.\end{aligned}$$

By combining (2.45), (2.46) and (2.47), and adding the integral over the set $\{u^* = u_0\}$ we find

$$\int_0^{\frac{\pi}{2}} \big|(u^*)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim_d \int_0^{\pi} \big|u_0'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

By symmetry we then have

$$\int_{\frac{\pi}{2}}^{\pi} \big|(u^*)'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta \lesssim_d \int_0^{\pi} \big|u_0'(\theta)\big|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta,$$

and the proof is complete by adding these two estimates.

2.3.2 Passage to the general case

The passage to the general case of a polar $f \in W^{1,1}(\mathbb{S}^d)$ follows closely the outline of §1.2.2, with Lemma 1.2.1 replaced by Lemma 2.2.8 when appropriate. We omit the details.

2.4 Proof of Theorem 2.1.3

Now we move into the proof of our result for the fractional Hardy-Littlewood maximal operator.

2.4.1 Preliminaries

For the sake of simplicity we henceforth assume that f belongs to the set of interest for our main theorems, that is $f \in W^{1,1}_{\text{pol}}(\mathbb{S}^d)$. We define the one dimensional version of f (that we also call f), for $r \in [0,\pi]$, as $f(r) = f(\xi)$, where $\theta(\xi) = r$. By Lemma 2.2.6 we know that after modifying f in a set of measure zero we can assume f (the one dimensional version) absolutely continuous in compacts not containing 0 or π. In the following we continue with this assumption.

For $f \in W^{1,1}_{\text{pol}}(\mathbb{S}^d)$ and $\xi \in \mathbb{S}^d$ let us define the set $\mathbf{B}^{\beta}_{\xi}$ as the set of closed balls that realize the supremum in the definition of the maximal function (since we assume f continuous outside $\mathbf{e}$ and $-\mathbf{e}$ these balls have positive radius outside these points), that is

$$\mathbf{B}^{\beta}_{\xi} = \left\{ \overline{\mathcal{B}_r(\zeta)};\ \zeta \in \mathbb{S}^d\,,\, \pi \geq r \geq 0\,,\, \xi \in \overline{\mathcal{B}_r(\zeta)}\ :\ \widetilde{\mathcal{M}}_{\beta} f(\xi) = r^{\beta} \fint_{\mathcal{B}_r(\zeta)} f(\eta)\ \mathrm{d}\sigma(\eta) \right\}.$$

Observe that $\mathbf{B}^{\beta}_{\xi}$ is non-empty for $\xi \notin \{\mathbf{e}, -\mathbf{e}\}$.

We are mostly interested in the case where $|\nabla \widetilde{\mathcal{M}}_{\beta} f(\xi)| \neq 0$ and that can only happen in the case where $\xi \in \partial \mathcal{B}_r(\zeta)$ for every $\mathcal{B}_r(\zeta) \in \mathbf{B}^{\beta}_{\xi}$ (otherwise we would have that ξ is a local minimum of $\widetilde{\mathcal{M}_{\beta}} f$). Moreover, since f is polar, we can conclude that ξ, ζ and $\mathbf{e}$ belong to the same great circle of $\mathbb{S}^d$, and that $\mathbf{e}$ is not between ξ and ζ. Otherwise we may rotate the ball $\mathcal{B}_r(\zeta)$ with respect to the north pole $\mathbf{e}$ in order to get $\mathbf{e}$, the new center and ξ in the same great circle. The crucial observation is that in this context we would have $\xi \in \text{int}(\mathcal{B}_r(\zeta))$, reaching a contradiction. We first state an adaptation to the sphere setup of [BM19, Lemma 2.1]. The proof is a straightforward adaptation, we omit it.

Lemma 2.4.1. *Let $f \in W^{1,1}_{pol}(\mathbb{S}^d)$ and $\{f_j\}_{j\in\mathbb{N}} \subset W^{1,1}_{pol}(\mathbb{S}^d)$ such that $\|f - f_j\|_{W^{1,1}(\mathbb{S}^d)} \to 0$ as $j \to \infty$. For every $\xi \in \mathbb{S}^d$, choose $\mathcal{B}_{r_j}(\zeta_j) \in \mathbf{B}^{\beta}_{\xi,j}$ (where $\mathbf{B}^{\beta}_{\xi,j}$ is defined analogously to $\mathbf{B}^{\beta}_{\xi}$, for each $j \in \mathbb{N}$). Then, for a.e. ξ, if (ζ, r) is an accumulation point of $\{(\zeta_j, r_j)\}_{j\in\mathbb{N}}$, we have $\mathcal{B}_r(\zeta) \in \mathbf{B}^{\beta}_{\xi}$.*

Here we state the fractional version of Lemma 2.2.1, the proof is similar, we omit it.

Lemma 2.4.2. *Let $f \in W^{1,1}_{pol}(\mathbb{S}^d)$ be a nonnegative function. Assume that $\widetilde{\mathcal{M}}_{\beta} f$ is differentiable at ξ and that $\mathcal{B}_r(\zeta) = \mathcal{B} \in \mathbf{B}^{\beta}_{\xi}$. Then*

$$\nabla \widetilde{\mathcal{M}}_{\beta} f(\xi) v = r^{\beta} \fint_{\mathcal{B}} \nabla f(\eta) \big(-(\eta \cdot v)\xi + (\eta \cdot \xi) v \big)\ \mathrm{d}\sigma(\eta)$$

for every $v \in \mathbb{R}^{d+1}$ with $v \perp \xi$. In particular,

$$\big|\nabla \widetilde{\mathcal{M}}_{\beta} f(\xi)\big| \leq r^{\beta} \fint_{\mathcal{B}} |\nabla f(\eta)|\ \mathrm{d}\sigma(\eta).$$

Since f is polar, we can prove that $\widetilde{\mathcal{M}}_{\beta} f$ is polar and locally Lipschitz outside the poles, so Lemma 2.4.2 holds almost everywhere. The proof of this fact relies on the continuity of f outside the poles, that implies that near every point the radius of the maximal function is bounded by below.

Now a comment about the weak differentiability. In Lemma 2.2.8, Carneiro and the author stated the equivalence between g polar being weakly differentiable in $\mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ and g being weakly differentiable in $(0, \pi)$. Moreover, we stated that if that is the case and g and ∇g are locally integrable in the poles then g is weakly differentiable in the sphere. This result, Sobolev embedding and the previous remark joint with Theorem 2.1.3 will imply that $\widetilde{\mathcal{M}}_{\beta} f$ is weakly differentiable in $\mathbb{S}^d$ when $f \in W^{1,1}_{\text{pol}}(\mathbb{S}^d)$.

2.4.2 Lipschitz case

We assume now that our $f \in W^{1,1}_{\text{pol}}(\mathbb{S}^d)$ is a Lipschitz function. We then have

$$|\nabla f(\xi)| = |f'(\theta)|$$

for a.e. $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$, and

$$\|\nabla f\|_{L^1(\mathbb{S}^d)} = \omega_{d-1} \int_0^\pi |f'(\theta)|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

Estimates for large radii - preliminary lemmas

We start with the following result.

Lemma 2.4.3. *Let $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ and let $\mathcal{B}(\zeta, r) \in \mathbf{B}^\beta_\xi$, with ζ in the half great circle determined by $\mathbf{e}$, ξ and $-\mathbf{e}$. Assume that $0 \le \theta(\zeta) < \theta(\xi)$, that $\xi \in \partial\mathcal{B}_r(\zeta)$ and that $\widetilde{\mathcal{M}}_\beta f$ is differentiable at ξ. Then*

$$\nabla\widetilde{\mathcal{M}}_\beta f(\xi)(v(\xi,\mathbf{e})) = r^\beta \frac{\sigma'(r)}{\sigma(r)} ⨍_{\mathcal{B}_r(\zeta)} \nabla f(\eta)(v(\eta,\zeta))\,\frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\,\mathrm{d}\sigma(\eta) - \beta r^{\beta-1} ⨍_{\mathcal{B}_r(\zeta)} f,$$

where

$$v(\eta,\zeta) = \frac{\zeta - (\eta\cdot\zeta)\eta}{|\zeta - (\eta\cdot\zeta)\eta|}$$

is the unit vector, tangent to η, in the direction of the geodesic that goes from η to ζ. In particular, since $\nabla\widetilde{\mathcal{M}}_\beta f(\xi)(v(\xi,\mathbf{e})) \ge 0$, we have

$$\left|\nabla\widetilde{\mathcal{M}}_\beta f(\xi)\right| \le r^\beta \frac{\sigma'(r)}{\sigma(r)} ⨍_{\mathcal{B}_r(\zeta)} \nabla f(\eta)(v(\eta,\zeta))\,\frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\,\mathrm{d}\sigma(\eta) = r^\beta\frac{\sigma'(r)}{\sigma(r)}\left(⨍_{\mathcal{B}_r(\zeta)} f - ⨍_{\partial\mathcal{B}_r(\zeta)} f\right).$$

Proof Follows as a variation of the proof of Lemma 2.2.4, taking into consideration that

$$\lim_{h\to 0} \frac{(r+h)^\beta ⨍_{\mathcal{B}_{r+h}(\zeta-h)} f - r^\beta ⨍_{\mathcal{B}_{r+h}(\zeta-h)} f}{h} = \beta r^{\beta-1} ⨍_{\mathcal{B}_r(\zeta)} f.$$

■

Estimates for large radii - main lemma

Now we prove an important estimate.

Lemma 2.4.4. *Let $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ and let $\mathcal{B}_r(\zeta) \in \mathbf{B}^\beta_\xi$, with ζ in the half great circle determined by $\mathbf{e}$, ξ and $-\mathbf{e}$. Assume that $0 \le \theta(\zeta) < \theta(\xi)$, that $\xi \in \partial\mathcal{B}_r(\zeta)$ and that $\widetilde{\mathcal{M}}_\beta f$ is differentiable at ξ. There is a universal constant $\rho > 0$ such that if $\mathcal{B} = \mathcal{B}_r(\zeta) \subset \overline{\mathcal{B}_\rho(\mathbf{e})}$ then*

$$\left|\nabla\widetilde{\mathcal{M}}_\beta f(\xi)\right| \lesssim_d \frac{r^\beta}{\theta(\xi)} ⨍_{\mathcal{B}} |\nabla f(\eta)|\,\theta(\eta)\,\mathrm{d}\sigma(\eta) + \frac{r^{\beta+1}\,\theta(\zeta)}{\theta(\xi)} ⨍_{\mathcal{B}} |\nabla f(\eta)|\,\mathrm{d}\sigma(\eta). \tag{2.48}$$

Proof In the following we choose ρ such that both estimates in Lemma 2.2.5 hold. From Lemma 2.4.3 (and considering that $\left|\nabla\widetilde{\mathcal{M}}f(\xi)\right| = \nabla\widetilde{\mathcal{M}}f(\xi)(v(\xi,\mathbf{e}))$) we have

$$\nabla\widetilde{\mathcal{M}}_\beta f(\xi)(-v(\xi,\mathbf{e})) = r^\beta \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(-v(\eta,\zeta))\, \frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\, \mathrm{d}\sigma(\eta) + \beta r^{\beta-1} \fint_{\mathcal{B}} f. \tag{2.49}$$

In the case $\zeta = \mathbf{e}$, estimate (2.48) follows directly from (2.49) and Lemma 2.2.6 (this is just the smoothness of the function $\frac{t\sigma'(t)}{\sigma(t)}$ near 0). From now on we assume that $\zeta \neq \mathbf{e}$. From Lemma 2.4.2 we also know that

$$\nabla\widetilde{\mathcal{M}}_\beta f(\xi)(-v(\xi,\mathbf{e})) = r^\beta \fint_{\mathcal{B}} \nabla f(\eta)(S(\eta))\, \mathrm{d}\sigma(\eta), \tag{2.50}$$

with S defined as in the previous section. The idea is to compare the identities (2.49) and (2.50) in order to bound $\left|\nabla\widetilde{\mathcal{M}}_\beta f(\xi)\right| = \left|\nabla\widetilde{\mathcal{M}}_\beta f(\xi)(-v(\xi,\mathbf{e}))\right|$. To do so, we write the right-hand side of (2.50) as a sum of three terms, one being comparable to $\left|\nabla\widetilde{\mathcal{M}}_\beta f(\xi)\right|$, the second one being small, and the third one being close to the right-hand side of (2.49) in a suitable sense. We start by writing

$$1 = \frac{\theta(\xi) - \theta(\zeta)}{r} = \frac{d(\mathbf{e},\xi) - d(\mathbf{e},\zeta)}{r}.$$

We then have

$$\begin{aligned} r^\beta \fint_{\mathcal{B}} \nabla f(\eta)\, S(\eta)\, \mathrm{d}\sigma(\eta) &= r^\beta \fint_{\mathcal{B}} \nabla f(\eta)\, |S(\eta)| \left(\frac{\theta(\xi)-\theta(\zeta)}{r}\right) v_1(\eta)\, \mathrm{d}\sigma(\eta) \\ &= r^\beta \fint_{\mathcal{B}} \nabla f(\eta)\, |S(\eta)|\, \frac{\theta(\xi)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta) \\ &\quad - r^\beta \fint_{\mathcal{B}} \nabla f(\eta)\, \big(|S(\eta)| - 1\big) \frac{\theta(\zeta)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta) \\ &\quad - r^\beta \fint_{\mathcal{B}} \nabla f(\eta) \frac{\theta(\zeta)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta). \end{aligned} \tag{2.51}$$

By Lemma 2.2.7 we have

$$\begin{aligned} &\left|\frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(-v(\eta,\zeta))\, \frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\, \mathrm{d}\sigma(\eta) + \fint_{\mathcal{B}} \nabla f(\eta) \frac{\theta(\zeta)}{r}\, v_1(\eta)\, \mathrm{d}\sigma(\eta)\right| \\ &\qquad \lesssim_d \fint_{\mathcal{B}} |\nabla f(\eta)|\, \theta(\zeta)\, \mathrm{d}\sigma(\eta) + \frac{1}{r} \fint_{\mathcal{B}} |\nabla f(\eta)|\, \theta(\eta)\, \mathrm{d}\sigma(\eta). \end{aligned}$$

So, we have

$$\begin{aligned} &\left| r^\beta \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(-v(\eta,\zeta))\, \frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\, \mathrm{d}\sigma(\eta) + r^{\beta-1} \fint_{\mathcal{B}} \nabla f(\eta) \theta(\zeta)\, v_1(\eta)\, \mathrm{d}\sigma(\eta)\right| \\ &\qquad \lesssim_d r^\beta \fint_{\mathcal{B}} |\nabla f(\eta)|\, \theta(\zeta)\, \mathrm{d}\sigma(\eta) + r^{\beta-1} \fint_{\mathcal{B}} |\nabla f(\eta)|\, \theta(\eta)\, \mathrm{d}\sigma(\eta). \end{aligned} \tag{2.52}$$

Also, by the estimate (2.29) we have:

$$\left| r^{\beta} \fint_{\mathcal{B}} \nabla f(\eta) \left(|S(\eta)|-1\right) \frac{\theta(\zeta)}{r}\, v_1(\eta)\ \mathrm{d}\sigma(\eta) \right| \lesssim_d r^{\beta} \fint_{\mathcal{B}} \left|\nabla f(\eta)\right| \theta(\zeta)\ \mathrm{d}\sigma(\eta). \tag{2.53}$$

We notice that

$$\begin{aligned}
-\frac{\theta(\xi)}{r}\left|\nabla \widetilde{\mathcal{M}}_{\beta} f(\xi)\right| &= r^{\beta} \fint_{\mathcal{B}} \nabla f(\eta)\, |S(\eta)|\, \frac{\theta(\xi)}{r}\, v_1(\eta)\ \mathrm{d}\sigma(\eta) \\
&= r^{\beta} \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(-v(\eta,\zeta))\, \frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\ \mathrm{d}\sigma(\eta) + \beta r^{\beta-1} \fint_{\mathcal{B}} f \\
&\quad + r^{\beta} \fint_{\mathcal{B}} \nabla f(\eta) \left(|S(\eta)|-1\right) \frac{\theta(\zeta)}{r}\, v_1(\eta)\ \mathrm{d}\sigma(\eta) \\
&\quad + r^{\beta} \fint_{\mathcal{B}} \nabla f(\eta) \frac{\theta(\zeta)}{r}\, v_1(\eta)\ \mathrm{d}\sigma(\eta),
\end{aligned} \tag{2.54}$$

where the last equality is obtained by comparing identities (2.49), (2.50) and (2.51). So, combining (2.52), (2.53) and (2.54), we get

$$\begin{aligned}
\frac{\theta(\xi)}{r}\left|\nabla \widetilde{\mathcal{M}}_{\beta} f(\xi)\right| &\leq r^{\beta} \frac{\sigma'(r)}{\sigma(r)} \fint_{\mathcal{B}} \nabla f(\eta)(v(\eta,\zeta))\, \frac{\sigma(d(\zeta,\eta))}{\sigma'(d(\zeta,\eta))}\ \mathrm{d}\sigma(\eta) \\
&\quad + \left| r^{\beta} \fint_{\mathcal{B}} \nabla f(\eta) \left(|S(\eta)|-1\right) \frac{\theta(\zeta)}{r}\, v_1(\eta)\ \mathrm{d}\sigma(\eta) \right| - r^{\beta} \fint_{\mathcal{B}} \nabla f(\eta) \frac{\theta(\zeta)}{r}\, v_1(\eta)\ \mathrm{d}\sigma(\eta) \\
&\lesssim_d r^{\beta} \fint_{\mathcal{B}} |\nabla f(\eta)|\, \theta(\zeta)\ \mathrm{d}\sigma(\eta) + r^{\beta-1} \fint_{\mathcal{B}} |\nabla f(\eta)|\, \theta(\eta)\ \mathrm{d}\sigma(\eta).
\end{aligned}$$

And finally

$$\left|\nabla \widetilde{\mathcal{M}}_{\beta} f(\xi)\right| \lesssim_d \frac{r^{\beta}}{\theta(\xi)} \fint_{\mathcal{B}} |\nabla f(\eta)|\, \theta(\eta)\ \mathrm{d}\sigma(\eta) + \frac{r^{\beta+1}\,\theta(\zeta)}{\theta(\xi)} \fint_{\mathcal{B}} |\nabla f(\eta)|\ \mathrm{d}\sigma(\eta).$$

This conclude the proof of the lemma. ■

Estimates for small radii

We also need another estimate, similar to the one obtained in [LM19, Lemma 2.10]. Given a ball $\mathcal{B} = \mathcal{B}_r(\zeta)$ we define $2\mathcal{B} = \mathcal{B}_{2r}(\zeta)$. We use the following estimate (the analogous in the polar case to [LM19, Proposition 2.8]), its verification is left to the interested reader:

Proposition 2.4.1. *Suppose that $g \in L^1(\mathbb{S}^d)$ is polar, $\mathcal{B} := \mathcal{B}_r(\zeta) \subset \mathbb{S}^d \setminus \mathcal{B}_{2r}(\mathbf{e}) \cup \mathcal{B}_{2r}(-\mathbf{e})$, then we have that*

$$\fint_{[\theta(\zeta)-r,\theta(\zeta)+r]} |g| \lesssim_d \fint_{2\mathcal{B}_r(\zeta)} |g|,$$

where in the first integral we consider the one dimensional function corresponding to g.

We also need the following proposition. We say that $\mathcal{B}_r(\zeta) \subset \mathbb{S}^d$, with $r \leq \pi$, is a best ball for $\widetilde{\mathcal{M}}_\beta f$, if there exists $\xi \in \mathcal{B}_r(\zeta)$ with $\widetilde{\mathcal{M}}_\beta f(\xi) = r^\beta \fint_{\mathcal{B}_r(\zeta)} f$.

Proposition 2.4.2. *Suppose that $0 < \beta < d$, $f \in L^1(\mathbb{S}^d)$, $\mathcal{B}_1 := \mathcal{B}_{r_1}(\zeta_1)$ and $\mathcal{B}_2 = \mathcal{B}_{r_2}(\zeta_2)$ are best balls for $\widetilde{\mathcal{M}}_\beta f$ such that $\mathcal{B}_2 \subset \mathcal{B}_{cr_1}(\zeta_1)$ with $c > 1$, then we have that:*

$$\left(\frac{r_1}{r_2}\right)^\beta \fint_{\mathcal{B}_1} f \lesssim_{c,d} \fint_{\mathcal{B}_2} f.$$

Proof *The proof is analogous to the proof of [LM19, Proposition 2.11], by using the fact that $1 \lesssim_{c,d} \frac{\sigma(r_1)}{\sigma(cr_1)}$.* ■

Now, we define $w(\xi) := \min\{\theta(\xi), \pi - \theta(\xi)\}$. Then, we have the following result.

Lemma 2.4.5. *Suppose that $f \in W^{1,1}(\mathbb{S}^d)$ is polar, $0 < \beta < d$, $\mathcal{B} \in \mathbf{B}_\xi^\beta$ for some $\xi \in \mathbb{S}^d$, $\mathcal{B} = \mathcal{B}_r(\zeta)$, $r \leq \frac{w(\zeta)}{4}$ and*

$$\mathcal{E} := \left\{\eta \in 2\mathcal{B} : \frac{1}{2}\fint_{\mathcal{B}} f \leq f(\eta) \leq 2\fint_{\mathcal{B}} f\right\}.$$

Then

$$\left|\fint_{\mathcal{B}} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta)\right| \lesssim_{d,\beta} \fint_{2\mathcal{B}} |\nabla f(\eta)| \chi_{\mathcal{E}}(\eta)\, \mathrm{d}\sigma(\eta).$$

Proof We know by Lemma 2.4.3 and 2.4.2 that:

$$\left|\fint_{\mathcal{B}} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta)\right| \leq \frac{\sigma'(r)}{\sigma(r)} \left(\fint_{\mathcal{B}} f - \fint_{\partial\mathcal{B}} f\right).$$

Let us define $a := \theta(\zeta) - r$, $b := \theta(\zeta) + r$ and

$$A := \left\{t \in 2[a,b] : \frac{1}{2}\fint_{\mathcal{B}} f \leq f(t) \leq 2\fint_{\mathcal{B}} f\right\}.$$

Now we show that

$$\fint_{\mathcal{B}} f - \fint_{\partial\mathcal{B}} f \leq 2 \int_{[a,b]} |f'(t)| \chi_A(t)\, \mathrm{d}t$$

in an analogous way to [LM19, Lemma 2.10]. We conclude using that $|\nabla f(\eta)|\, \chi_{\mathcal{E}}(\eta) = |f'(\theta(\eta))|\, \chi_A(\theta(\eta))$ for $\eta \in \mathcal{E}$, Proposition 2.4.1 and the fact that $\frac{r\sigma'(r)}{\sigma(r)}$ is bounded. ■

Proof of Theorem 2.1.3-Lipschitz case

We are now in position to move on to the proof of Theorem 2.1.3 when our initial datum f is a Lipschitz function. In this case we also have $\widetilde{\mathcal{M}}_\beta f$ Lipschitz.

For each $\xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\}$ let us choose a ball $\mathcal{B}_\xi := \mathcal{B}_{r_\xi}(\zeta_\xi) \in \mathbf{B}_\xi^\beta$ with r_ξ minimal and, subject to this condition, with ζ_ξ in the half great circle connecting $\mathbf{e}, \xi, -\mathbf{e}$ in a way that $w(\zeta_\xi) = \min\{d(\mathbf{e}, \zeta_\xi), d(-\mathbf{e}, \zeta_\xi)\}$ is minimal. If there are two potential choices for ζ_ξ we choose the one with $0 \leq \theta(\zeta_\xi) \leq \theta(\xi)$.

Proof [Proof of Theorem 2.1.3, Lipschitz case] First let us observe that by Lemma 2.4.2 we have:

$$\begin{aligned}
\int_{\mathbb{S}^d} |\nabla \widetilde{\mathcal{M}}_\beta f|^q &= \int_{\mathbb{S}^d} \left| r_\xi^\beta \fint_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) \right|^q \mathrm{d}\sigma(\xi) \\
&= \int_{\mathbb{S}^d} \frac{r_\xi^{q\beta}}{\sigma(r_\xi)^{q-1}} \left| \int_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) \right|^{q-1} \left| \fint_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) \right| \mathrm{d}\sigma(\xi) \qquad (2.55) \\
&\lesssim_{d,\beta} \|\nabla f\|_1{}^{q-1} \int_{\mathbb{S}^d} \left| \fint_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) \right| \mathrm{d}\sigma(\xi),
\end{aligned}$$

where we use the fact that $q\beta = d(q-1)$ and that $\frac{r^d}{\sigma(r)}$ is bounded. So we need to bound the integral term. This is done in four steps.

Step 1: Let us observe that we can restrict our attention to small balls. Define the set $\mathcal{R}_c = \left\{\xi \in \mathbb{S}^d \ : \xi \in \mathbb{S}^d \setminus \{\mathbf{e}, -\mathbf{e}\} \text{ and } r_\xi \geq c\right\}$.We find that

$$\int_{\mathcal{R}_c} \left| \fint_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta)\, \mathrm{d}\eta \right| \mathrm{d}\sigma(\xi) \leq \int_{\mathcal{R}_c} \frac{1}{\sigma(\mathcal{B}_\xi)} \int_{\mathcal{B}_\xi} |\nabla f(\eta)|\, \mathrm{d}\sigma(\eta)\, \mathrm{d}\sigma(\xi) \lesssim_{c,d} \int_{\mathbb{S}^d} |\nabla f(\eta)|\, \mathrm{d}\sigma(\eta). \tag{2.56}$$

Step 2: Let us define $\mathcal{W}_d = \{\xi \in \mathbb{S}^d ; r_\xi \leq \frac{w(\xi)}{4}\}$. We show that we can restrict our attention to $\xi \in \mathbb{S}^d \setminus \mathcal{W}_d$. For this, we use Lemma 2.4.5. For every $\xi \in \mathcal{W}_d$ we define:

$$\mathcal{A}_\xi := \left\{ \eta \in 2\mathcal{B}_\xi : \frac{1}{2} \fint_{\mathcal{B}_\xi} f \leq f(\eta) \leq 2 \fint_{\mathcal{B}_\xi} f \right\}.$$

So, by Lemma 2.4.5 we have:

$$\left| \fint_{B_\xi} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) \right| \lesssim_{d,\beta} \fint_{2B_\xi} |\nabla f(\eta)| \chi_{\mathcal{A}_\xi}(\eta)\, \mathrm{d}\sigma(\eta),$$

therefore:

$$\begin{aligned}
\int_{\mathcal{W}_d} \left| \fint_{B_\xi} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) \right| \mathrm{d}\sigma(\xi) &\lesssim_{d,\beta} \int_{\mathcal{W}_d} \fint_{2B_\xi} |\nabla f(\eta)| \chi_{\mathcal{A}_\xi}(\eta)\, \mathrm{d}\sigma(\eta)\, \mathrm{d}\sigma(\xi) \\
&\lesssim_{d,\beta} \int_{\mathbb{S}^d} |\nabla f(\eta)| \left(\int_{\mathbb{S}^d} \frac{\chi_{2\mathcal{B}_\xi}(\eta) \chi_{\mathcal{A}_\xi}(\eta) \chi_{\mathcal{W}_d}(\xi)}{\sigma(2\mathcal{B}_\xi)}\, \mathrm{d}\sigma(\xi) \right) \mathrm{d}\sigma(\eta).
\end{aligned}$$

We want to bound the inner integral for fixed $\eta \in \mathbb{S}^d$. Now suppose that $\chi_{\mathcal{A}_{\xi_1}}(\eta) \neq 0$ and $\chi_{\mathcal{A}_{\xi_2}}(\eta) \neq 0$ for some $\xi_1, \xi_2 \in \mathbb{S}^d$. If these points do not exist, the estimates are obvious. By definition, the above means that $\frac{1}{2}\fint_{B_{\xi_1}} f \leq f(\eta) \leq 2\fint_{B_{\xi_1}} f$ and $\frac{1}{2}\fint_{B_{\xi_2}} f \leq f(\eta) \leq 2\fint_{B_{\xi_2}} f$. In particular, we have

$$\frac{1}{4}\fint_{B_{\xi_1}} f \leq \fint_{B_{\xi_2}} f \leq 4\fint_{B_{\xi_1}} f.$$

Let $r_1 := rad(\mathcal{B}_{\xi_1})$ and $r_2 := rad(\mathcal{B}_{\xi_2})$. First, assume $r_2 \leq r_1$. Since $\eta \in 2\mathcal{B}_{\xi_2} \cap 2\mathcal{B}_{\xi_1}$ it follows that $\mathcal{B}_{\xi_2} \subset 8\mathcal{B}_{\xi_1}$. And then, by Proposition 2.4.2:

$$\left(\frac{r_1}{r_2}\right)^{\beta} \fint_{\mathcal{B}_{\xi_1}} f \lesssim_d \fint_{\mathcal{B}_{\xi_2}} f \lesssim_d \fint_{\mathcal{B}_{\xi_1}} f,$$

then it follows that $r_1 \lesssim_{\beta,d} r_2$. And then, by symmetry, we have

$$\frac{r_1}{r_2} \simeq_{d,\beta} 1$$

and that implies that if $\eta \in \mathcal{A}_\xi$ then $d(\xi,\eta) \lesssim_{d,\beta} rad(\mathcal{B}_{\xi_1})$ and $\sigma(\mathcal{B}_{\xi_1}) \lesssim_{d,\beta} \sigma(\mathcal{B}_\xi)$. Combining these estimates we have the following

$$\int_{\mathbb{S}^d} \frac{\chi_{2\mathcal{B}_\xi}(\eta)\chi_{\mathcal{A}_\xi}(\eta)\chi_{\mathcal{W}_d}(\eta)}{\sigma(2\mathcal{B}_\xi)} \,\mathrm{d}\sigma(\xi) \lesssim_{d,\beta} \int_{B_{C(d,\beta)rad(\mathcal{B}_{\xi_1})}(\eta)} \frac{\mathrm{d}\xi}{\sigma(\mathcal{B}_{\xi_1})} \lesssim_{d,\beta} 1.$$

From where we have that

$$\int_{\mathcal{W}_d} \left| \fint_{B_\xi} \nabla f(\eta) S(\eta) \,\mathrm{d}\sigma(\eta) \right| \mathrm{d}\sigma(\xi) \lesssim_{d,\beta} \|\nabla f\|_1. \tag{2.57}$$

So, we need to prove a similar estimate for the remaining points. Using (2.56) we can see that we may restrict ourselves to the situation where $d(\mathbf{e},\xi) \leq \rho$ or $d(-\mathbf{e},\xi) \leq \rho$ (where ρ is given by Lemma 2.4.4), we can do that because there exist r_ρ such that if $r \leq r_\rho$ and $\mathcal{B}_r(\zeta) \in \mathbf{B}_\xi^\beta$ then $w(\xi) \leq \rho$ or $\xi \in \mathcal{W}_d$. By symmetry let us assume that $\theta(\xi) = d(\mathbf{e},\xi) \leq \rho$. Then we define the set

$$\mathcal{G}_d = \left\{\xi \notin \mathcal{W}_d \cup \mathcal{R}_{r_\rho} : \xi \in \mathcal{B}_\rho(\mathbf{e})\right\},$$

and further decompose it in $\mathcal{G}_d^- = \{\xi \in \mathcal{G}_d \ : \ 0 \leq \theta(\zeta_\xi) < \theta(\xi)\}$ and $\mathcal{G}_d^+ = \{\xi \in \mathcal{G}_d \ : \ 0 < \theta(\xi) < \theta(\zeta_\xi)\}$. We bound the integrals over these two sets separately.

Step 3 (Bounding the integral on $\mathcal{G}_d^+$). For $\mathcal{G}_d^+$ we proceed as follows.

$$\begin{aligned}\int_{\mathcal{G}_d^+} \left| \fint_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta) \,\mathrm{d}\sigma(\eta) \right| \mathrm{d}\sigma(\xi) &\leq \int_{\mathcal{G}_d^+} \fint_{\mathcal{B}_\xi} |\nabla f(\eta)| \,\mathrm{d}\sigma(\eta) \,\mathrm{d}\sigma(\xi) \\ &= \int_{\mathbb{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^+} \frac{\chi_{\mathcal{B}_\xi}(\eta)}{\sigma(\mathcal{B}_\xi)} \,\mathrm{d}\sigma(\xi) \,\mathrm{d}\sigma(\eta).\end{aligned}$$

Note that $\theta(\eta) \geq \theta(\xi)$ in this case. So, we get

$$\int_{\mathcal{G}_d^+} \frac{\chi_{\mathcal{B}_\xi}(\eta)}{\sigma(\mathcal{B}_\xi)} \,\mathrm{d}\sigma(\xi) \lesssim_d 1,$$

and conclude that

$$\int_{\mathcal{G}_d^+} \left| \fint_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta) \,\mathrm{d}\sigma(\eta) \right| \mathrm{d}\sigma(\xi) \lesssim_d \|\nabla f\|_1. \tag{2.58}$$

Step 4 (Bounding the integral on $\mathcal{G}_d^-$). We now bound the integral over $\mathcal{G}_d^-$ using Lemma 2.4.4. If $\xi \in \mathcal{G}_d^-$ we then have

$$\begin{aligned}
&\int_{\mathcal{G}_d^-} \left| \fint_{B_\xi} \nabla f(\eta) S(\eta) \,\mathrm{d}\sigma(\eta) \right| \,\mathrm{d}\sigma(\xi) \\
&\lesssim_d \int_{\mathcal{G}_d^-} \left(\frac{1}{\theta(\xi)} \fint_{\mathcal{B}_\xi} |\nabla f(\eta)|\, \theta(\eta) \,\mathrm{d}\sigma(\eta) + \frac{r_\xi\, \theta(\zeta_\xi)}{\theta(\xi)} \fint_{\mathcal{B}_\xi} |\nabla f(\eta)| \,\mathrm{d}\sigma(\eta) \right) \mathrm{d}\sigma(\xi) \\
&\lesssim_d \int_{\mathbb{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_\xi}(\eta)\, \theta(\eta)}{r_\xi\, \sigma(r_\xi)} \,\mathrm{d}\sigma(\xi) \,\mathrm{d}\sigma(\eta) + \int_{\mathbb{S}^d} |\nabla f(\eta)| \int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_\xi}(\eta)\, \theta(\zeta_\xi)}{\sigma(r_\xi)} \,\mathrm{d}\sigma(\xi) \,\mathrm{d}\sigma(\eta).
\end{aligned} \tag{2.59}$$

Now, we notice that

$$\int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_\xi}(\eta)\theta(\zeta)}{\sigma(\mathcal{B}_\xi)} \,\mathrm{d}\sigma(\xi) \lesssim_d 1 \tag{2.60}$$

and

$$\int_{\mathcal{G}_d^-} \frac{\chi_{\mathcal{B}_\xi}(\eta)\theta(\eta)}{r_\xi \sigma(\mathcal{B}_\xi)} \,\mathrm{d}\sigma(\xi) \lesssim_d 1. \tag{2.61}$$

Our desired inequality

$$\int_{\mathcal{G}_d^-} \left| \fint_{\mathcal{B}_\xi} \nabla f(\eta) S(\eta) \,\mathrm{d}\sigma(\eta) \right| \mathrm{d}\sigma(\xi) \lesssim_d \|\nabla f\|_1 \tag{2.62}$$

follows combining (2.59), (2.60) and (2.61). Then, by combining (2.55), (2.56), (2.57), (2.58) and (2.62) we conclude Theorem 2.1.3 in this case. ∎

2.4.3 Passage to the general case

Preliminaries of the reduction

Lemma 2.4.6. *Let $f \in W^{1,1}(\mathbb{S}^d)$ be such that $\|f - f_j\|_{W^{1,1}} \to 0$ as $j \to \infty$. Then $\||f_j| - |f|\|_{W^{1,1}} \to 0$ as $j \to \infty$.*

Proof The proof is exactly as [BM19, Lemma 2.3]. ■

Lemma 2.4.7. *Let $f \in W^{1,1}_{pol}(\mathbb{S}^d)$ and $\{f_j\}_{j\in\mathbb{N}} \subset W^{1,1}_{pol}(\mathbb{S}^d)$ be such that $\|f_j - f\|_{W^{1,1}} \to 0$ as $j \to \infty$. Then*

$$\nabla\widetilde{\mathcal{M}_\beta} f_j(\xi) \to \nabla\widetilde{\mathcal{M}_\beta} f(\xi)$$

a.e. as $j \to \infty$.

Proof By Lemma 2.4.6 we can assume that the functions f and f_j are nonnegative. We consider the set $E \subset [0,\pi]$ that consists of the points $\theta(\xi)$ where $\widetilde{\mathcal{M}}_\beta f, \widetilde{\mathcal{M}}_\beta f_j$ are all differentiable at ξ. By Lemma 2.4.2 and the almost everywhere differentiability of $\widetilde{\mathcal{M}}_\beta f$ and $\widetilde{\mathcal{M}}_\beta f_j$, we have that $m(E^c) = 0$ and that

$$\nabla\widetilde{\mathcal{M}_\beta} f_j(\xi)(-v(\xi,\mathbf{e})) = r_{\xi,j}^{\beta} ⨍_{\mathcal{B}_{r_{\xi,j}}(\zeta_{\xi,j})} \nabla f_j(\eta) S(\eta)\, \mathrm{d}\sigma(\eta),$$

for every $\xi \in \mathbb{S}^d$ with $\theta(\xi) \in E$. So, we just need to prove that

$$\lim_{j\to\infty} r_{\xi,j}^{\beta} ⨍_{\mathcal{B}_{r_{\xi,j}}(\zeta_{\xi,j})} \nabla f_j(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) = r_\xi^\beta ⨍_{\mathcal{B}_{r_\xi}(\zeta_\xi)} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta).$$

Let us assume that there exists $\varepsilon > 0$ and $(j_k)_{k\in\mathbb{N}}$ such that

$$\left| r_{\xi,j_k}^{\beta} ⨍_{\mathcal{B}_{r_{\xi,j_k}}(\zeta_{\xi,j_k})} \nabla f_{j_k}(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) - r_\xi^\beta ⨍_{\mathcal{B}_{r_\xi}(\zeta_\xi)} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) \right| > \varepsilon. \tag{2.63}$$

Then, by compactness, there exists a subsequence of $(j_k)_{k\in\mathbb{N}}$ (we write this subsequence also by $(j_k)_{k\in\mathbb{N}}$) such that $\lim\limits_{k\to\infty} r_{\xi,j_k} = r_0$ and $\lim\limits_{k\to\infty} \zeta_{\xi,j_k} = \zeta_0$. By Lemma 2.4.1 we conclude that $r_0 > 0$ and that $\mathcal{B}_{r_0}(\zeta_0) \in \mathbf{B}_\xi^\beta$ for almost every ξ with $\theta(\xi) \in E$, so we have that

$$\lim_{k\to\infty} r_{\xi,j_k}^{\beta} ⨍_{\mathcal{B}_{\xi,j_k}} \nabla f_{j_k}(\eta) S(\eta)\, \mathrm{d}\sigma(\eta) = r_0^\beta ⨍_{\mathcal{B}_{r_0}(\zeta_0)} \nabla f(\eta) S(\eta)\, \mathrm{d}\sigma(\eta),$$

reaching a contradiction with (2.63). This concludes the proof of the lemma. ■
We can conclude, in a similar way, the following proposition.

Proposition 2.4.3. *If $f_j \to f$ in $W^{1,1}_{pol}(\mathbb{S}^d)$ and $0 < \beta < d$, we have that*

$$\lim_{j\to\infty} \widetilde{\mathcal{M}}_\beta f_j(\xi) = \widetilde{\mathcal{M}}_\beta f(\xi)$$

for a.e. $\xi \in \mathbb{S}^d$.

Now we conclude the passage to the general case:

Proof of general case of Theorem 2.1.3

Proof [Proof of Theorem 2.1.3]

Consider a sequence $f_n \in W^{1,1}_{\text{pol}}(\mathbb{S}^d)$ with $f_n \geq 0$ Lipschitz and $\|f_n - f\|_{W^{1,1}(\mathbb{S}^d)} \to 0$. By Fatou's lemma, Lemma 2.4.7 and Theorem 2.1.3 in the Lipschitz case we conclude:

$$\|\nabla \widetilde{\mathcal{M}}_\beta f\|_q \leq \liminf_{n\to\infty} \|\nabla \widetilde{\mathcal{M}}_\beta f_n\|_q \lesssim_{d,\beta} \lim_{n\to\infty} \|\nabla f_n\|_1 = \|\nabla f\|_1.$$

This concludes the proof of the theorem in the general case. ∎

Chapter 3

Continuity of the gradient of the fractional maximal operator on $W^{1,1}(\mathbb{R}^d)$

The study of regularity for $\mathfrak{M}_\beta$ (where $\mathfrak{M}$ denotes either the centered or uncentered Hardy-Littlewood maximal operator (8)) started with the influential work of Kinnunen and Saksman [KS03], where it was established that if $1 \leq \beta < d$ and $f \in L^p(\mathbb{R}^d)$ with $1 \leq p \leq d/\beta$, then

$$|\nabla \mathfrak{M}_\beta f(x)| \leq (d-\beta)\mathfrak{M}_{\beta-1} f(x) \tag{3.1}$$

a.e. in $\mathbb{R}^d$.

In this chapter, we explore the continuity of the map $f \mapsto |\nabla \mathfrak{M}_\beta f|$ for $\beta > 0$. Here, we establish the following complete result for $\beta > 0$, which in particular yields the continuity in the remaining open cases, that is, for $d > 1$, $0 < \beta < 1$ and general functions $f \in W^{1,1}(\mathbb{R}^d)$.

Theorem 3.0.1. *Let $\mathfrak{M}_\beta \in \{\widetilde{M}_\beta, M_\beta\}$. If $0 < \beta < d$, the operator $f \mapsto |\nabla \mathfrak{M}_\beta f|$ maps continuously $W^{1,1}(\mathbb{R}^d)$ into $L^{d/(d-\beta)}(\mathbb{R}^d)$.*

As observed by Beltran and Madrid in [BM19], it suffices to establish the continuity for any compact set $K \subset \mathbb{R}^d$. For any given $\delta > 0$, we consider two types of points in K, depending on whether the ball with maximal average has *large* radius (larger than δ) or *small* radius (smaller than δ). The techniques from [BM19, BM20] immediately apply to prove the continuity for the points whose maximal ball has *large* radius: the radiality assumption was not used in that situation. Thus, in order to establish continuity in Theorem 3.0.1, it suffices to bound contributions coming from points whose maximal ball has *small* radius, i.e. radius smaller than δ, and show that they go to zero for $\delta \to 0$. This is the main novelty of this chapter. To obtain this bound for points with *small* radius, we first note that on any compact set K, $\mathfrak{M}_\beta f$ is bounded away from 0. Then we use the Poincaré–Sobolev inequality, which becomes stronger the smaller the radius is and the larger the average of the function is. Then we apply a refined version of (3.1) which allows us to invoke a local version of the main theorem in [Wei21a] on the subset of points with *small* radius. This yields the desired result.

The proof of Theorem 3.0.1 is presented in §3.3. Auxiliary results which feature prominently in the proof are presented in §3.1 and §3.2.

3.1 Families of good balls

In this section we develop some estimates and identities regarding the weak derivative of the maximal functions of interest. We shall only be concerned with $0 < \beta < d$, although many of the arguments can also be extended to $\beta = 0$.

3.1.1 The truncated fractional maximal function

An important object for our purposes are the truncated fractional maximal operators which, for a given $\delta > 0$, are defined as

$$M_\beta^\delta f(x) := \sup_{r>\delta} r^\beta \fint_{B(x,r)} |f(y)|\,\mathrm{d}y \qquad \text{and} \qquad \widetilde{M}_\beta^\delta f(x) := \sup_{\substack{\bar{B}(z,r)\ni x \\ r>\delta}} r^\beta \fint_{B(z,r)} |f(y)|\,\mathrm{d}y.$$

We use $\mathfrak{M}_\beta^\delta$ to denote either M_β^δ or $\widetilde{M}_\beta^\delta$. Note that if $\delta = 0$, we recover the original operators $\mathfrak{M}_\beta = \mathfrak{M}_\beta^0$. The following is a well-known and elementary result; see for instance [BM20, Lemma 2.4] and [HM10, Lemma 8].

Proposition 3.1.1. *Let $0 < \beta < d$ and $\delta > 0$. If $f \in L^1(\mathbb{R}^d)$, then $\mathfrak{M}_\beta^\delta f$ is Lipschitz continuous (in particular, a.e. differentiable).*

3.1.2 Weak derivative and approximate derivative

As mentioned in the introduction, Weigt proved in [Wei21a], after partial contributions by many, the following result.

Theorem 3.1.1 ([Wei21a, Theorem 1.1 and Remark 1.3])**.** *Let $0 < \beta < d$ and $f \in W^{1,1}(\mathbb{R}^d)$. Then $\mathfrak{M}_\beta f$ is weakly differentiable and there exists a constant $C_{d,\beta} > 0$ such that*

$$\|\nabla \mathfrak{M}_\beta f\|_{L^{d/(d-\beta)}(\mathbb{R}^d)} \leq C_{d,\beta} \|\nabla f\|_{L^1(\mathbb{R}^d)}.$$

It will be convenient in our arguments to also recall the concept of approximate derivative. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *approximately differentiable* at a point $x_0 \in \mathbb{R}$ if there exists a vector $Df(x_0) \in \mathbb{R}^d$ such that, for any $\varepsilon > 0$, the set

$$A_\varepsilon := \left\{ x \in \mathbb{R} : \frac{|f(x) - f(x_0) - \langle Df(x_0), x - x_0\rangle|}{|x - x_0|} < \varepsilon \right\} \tag{3.2}$$

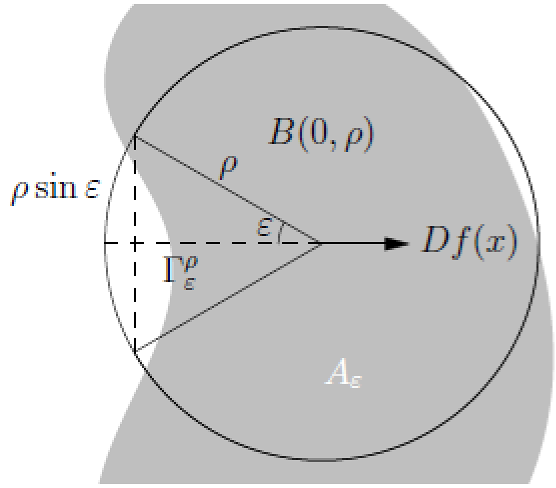

Figure 3.1: The sets $\Gamma_{\varepsilon,\rho}$ and A_ε intersect.

has x_0 as a density point. In this case, $Df(x_0)$ is called the *approximate derivative* of f at x_0 and it is uniquely determined. It is well-known that if f is weakly differentiable, then f is approximate differentiable a.e. and the weak and approximate derivatives coincide [EG92, Theorem 6.4].

The approximate derivative satisfies the following property, which will play a rôle in Propositions 3.1.2 and 3.1.3 below.

Lemma 3.1.1. *Let f be approximately differentiable at a point $x \in \mathbb{R}^d$. Then there exists a sequence $\{h_n\}_{n\in\mathbb{N}}$ with $|h_n| \to 0$ such that*

$$|Df(x)| = -\lim_{n\to\infty} \frac{f(x+h_n) - f(x)}{|h_n|},$$

where $Df(x)$ denotes the approximate derivative of f at x.

Proof Let $0 < \varepsilon < \pi/2$. By the definition of the approximate derivative, there exists $0 < \rho < \varepsilon$ such that

$$|A_\varepsilon \cap B(0,\rho)| \geq \Big(1 - \frac{\omega_{d-1}}{d\,\omega_d}(\sin\varepsilon)^{d-1}(\cos\varepsilon)^d\Big)|B(0,\rho)| \tag{3.3}$$

where A_ε is as in (3.2).

If $Df(x) = 0$, the result simply follows by the definition of A_ε and taking $\varepsilon = 1/n$.

Assume next $Df(x) \neq 0$. For each $h \in \mathbb{R}^d$, let α_h denote the angle formed by h and $-Df(x)$, so that

$$-\langle Df(x), h\rangle = |Df(x)||h| \cos\alpha_h.$$

The set

$$\Gamma_{\varepsilon,\rho} := \{h \in B(0,\rho) : \alpha_h \leq \varepsilon\}$$

has measure

$$|\Gamma_{\varepsilon,\rho}| > \int_0^{\rho\cos\varepsilon} \omega_{d-1}(r\sin\varepsilon)^{d-1}\,\mathrm{d}r = \frac{\omega_{d-1}}{d}(\sin\varepsilon)^{d-1}(\cos\varepsilon)^d\rho^d.$$

Thus, it follows from (3.3) that $\Gamma_{\varepsilon,\rho} \cap A_\varepsilon \neq \emptyset$, so by the definition of A_ε there is an $h \in \mathbb{R}^d$ such that

$$\frac{|f(x+h) - f(x) - \langle Df(x), h\rangle|}{|h|} < \varepsilon, \qquad \alpha_h \leq \varepsilon \qquad \text{and} \qquad 0 < |h| < \rho < \varepsilon. \tag{3.4}$$

By the triangle inequality, for h satisfying (3.4),

$$\begin{aligned}\left||Df(x)| + \frac{f(x+h)-f(x)}{|h|}\right| &\leq \left||Df(x)| + \frac{\langle Df(x),h\rangle}{|h|}\right| + \left|\frac{f(x+h)-f(x)}{|h|} - \frac{\langle Df(x),h\rangle}{|h|}\right| \\ &< |Df(x)||1-\cos\alpha_h| + \varepsilon \\ &\leq |Df(x)||1-\cos\varepsilon| + \varepsilon.\end{aligned}$$

As $|Df(x)| \neq 0$, the result now follows taking $\varepsilon = \min\{1/2n, 1/\sqrt{|Df(x)n}\}$ and the corresponding $h_n = h$ from the previous display. ∎

The approximate derivative of Mf for a.e. approximately differentiable functions $f \in L^1(\mathbb{R}^d)$ was studied by Hajłasz and Maly [HM10]. In particular, their arguments show that if $f \in L^1$ is a.e. approximately differentiable, then $\mathfrak{M}_\beta f$ is a.e. approximately differentiable.

3.1.3 The families of good balls

Let $0 < \beta < d$ and $\delta \geq 0$. For the uncentered maximal operator, given a function $f \in W^{1,1}(\mathbb{R}^d)$ and a point $x \in \mathbb{R}^d$, define the family of *good balls* for f at x as

$$\widetilde{B}^\delta_{\beta,x} \equiv \widetilde{B}^\delta_{\beta,x}(f) := \left\{B(z,r) : r \geq \delta,\ x \in \overline{B(z,r)},\ M^\delta_\beta f(x) = r^\beta \fint_{B(z,r)} |f(y)|\,\mathrm{d}y\right\}.$$

For the centered maximal operator we use the same definition (using $B_{\beta,x}$ instead), except that $z = x$. Note that $\mathfrak{B}^\delta_{\beta,x} \neq \emptyset$ for all $x \in \mathbb{R}^d$ if $\delta > 0$, where $\mathfrak{B}_{\beta,x}$ denotes either $\widetilde{B}_{\beta,x}$ or $B_{\beta,x}$. Moreover, by the Lebesgue differentiation theorem $\mathfrak{B}_{\beta,x} \equiv \mathfrak{B}^0_{\beta,x} \neq \emptyset$ for a.e. $x \in \mathbb{R}^d$, and if $B(z,r) \in \mathfrak{B}^0_{\beta,x}$, then $r > 0$. This immediately implies that for a.e. x there exists $\delta_x > 0$ such that if $0 \leq \delta < \delta_x$, then

$$\mathfrak{M}^\delta_\beta f(x) = \mathfrak{M}_\beta f(x).$$

This type of observation will be used at the derivative level in the forthcoming Lemma 3.2.3.

3.1.4 Luiro's formula

An important tool for our purposes is the so called Luiro's formula, which relates the derivative of the maximal function with the derivative of the original function. This has its roots in [Lui07, Theorem 3.1].

Proposition 3.1.2. *Let $0 < \beta < d$, $\delta \geq 0$ and $f \in W^{1,1}(\mathbb{R}^d)$. Then, for a.e. $x \in \mathbb{R}^d$ and $B = B(z,r) \in \mathfrak{B}^{\delta}_{\beta,x}$, the weak derivative $\nabla \mathfrak{M}^{\delta}_{\beta} f$ satisfies*

$$\nabla \mathfrak{M}^{\delta}_{\beta} f(x) = r^{\beta} \fint_B \nabla |f|(y)\, \mathrm{d}y. \tag{3.5}$$

Proof This essentially follows from an argument of Hajłasz and Maly [HM10, Theorem 2], which we include for completeness. By §3.1.2 the weak gradient of $\mathfrak{M}^{\delta}_{\beta} f$ equals its approximate gradient almost everywhere, so it suffices to show (3.5) at a point x at which $\mathfrak{M}^{\delta}_{\beta} f$ is approximately differentiable and for which there exists $B = B(z_x, r_x) \in \mathfrak{B}^{\delta}_{\beta,x}$. Define the function $\varphi : \mathbb{R}^d \to \mathbb{R}$ by

$$\varphi(y) := \mathfrak{M}^{\delta}_{\beta} f(y) - r^{\beta} \fint_{B(z_x+y-x,r_x)} |f(t)|\, \mathrm{d}t = \mathfrak{M}^{\delta}_{\beta} f(y) - r^{\beta} \fint_{B(z_x-x,r_x)} |f(y+t)|\, \mathrm{d}t,$$

which satisfies $\varphi \geq 0$ and $\varphi(x) = 0$. Thus, φ has a minimum at x. Furthermore, φ is approximately differentiable at x (note that one can differentiate under the integral sign) and by Lemma 3.1.1 there exists a sequence $\{h_n\}_{n\in\mathbb{N}}$ with $|h_n| \to 0$ such that

$$|D\varphi(x)| = -\lim_{n\to\infty} \frac{\varphi(x+h_n) - \varphi(x)}{|h_n|}.$$

As φ has a minimum at x, the right-hand side is nonpositive and thus $D\varphi(x) = 0$, which yields the desired result. ■

Remark 3.1.1. Proposition 3.1.2 continues to hold for $\beta = 0$, replacing the weak derivative by the approximate derivative in the cases where the weak differentiability of $\mathfrak{M}$ is currently unknown.

3.1.5 Refined Kinnunen–Saksman inequality

The Kinnunen–Saksman inequality (3.1) admits a refinement in terms of the good balls. It is noted that further refinements involving boundary terms (that is, averages along spheres) have been obtained in [LM19] and [BM20] for $\widetilde{M}_{\beta}$ and M_{β} respectively, although these are not required for the purposes of this chapter.

Proposition 3.1.3. *Let $0 < \beta < d$, $\delta \geq 0$ and $f \in W^{1,1}(\mathbb{R}^d)$. Then, for a.e. $x \in \mathbb{R}^d$ and $B = B(z,r) \in \mathfrak{B}^{\delta}_{x,\beta}$, the weak derivative $\nabla \mathfrak{M}^{\delta}_{\beta} f$ satisfies*

$$|\nabla \mathfrak{M}^{\delta}_{\beta} f(x)| \leq (d-\beta) r^{\beta-1} \fint_B |f(y)|\, \mathrm{d}y. \tag{3.6}$$

Proof By §3.1.2 the weak gradient of $\mathfrak{M}^\delta_\beta f$ equals its approximate gradient almost everywhere, so it suffices to show (3.6) at a point x at which $\mathfrak{M}^\delta_\beta f$ is approximately differentiable and for which there exists $B = B(z_x, r_x) \in \mathfrak{B}^\delta_{\beta,x}$. By Lemma 3.1.1 there is a sequence $\{h_n\}_{n\in\mathbb{N}}$ with $|h_n| \to 0$ and

$$|\nabla \mathfrak{M}^\delta_\beta f(x)| = \lim_{n\to\infty} \frac{\mathfrak{M}^\delta_\beta f(x) - \mathfrak{M}^\delta_\beta f(x+h_n)}{|h_n|}.$$

Now the proof follows from the classical Kinnunen–Saksman [KS03] reasoning, which we include for completeness. Note that $x + h_n \in \overline{B(z+h_n, r+|h_n|)}$, and that for the centered maximal operator we have $z = x$. This implies

$$\mathfrak{M}^\delta_\beta f(x+h_n) \geq (r+|h_n|)^\beta \fint_{B(z+h_n, r+|h_n|)} |f(y)| \,\mathrm{d}y.$$

Therefore

$$\begin{aligned}
&\frac{\mathfrak{M}^\delta_\beta f(x) - \mathfrak{M}^\delta_\beta f(x+h_n)}{|h_n|} \\
&\quad \leq \frac{1}{\omega_d |h_n|}\Big(r^{\beta-d}\int_{B(z,r)} |f(y)|\,\mathrm{d}y - (r+h_n)^{\beta-d}\int_{B(z+h_n,r+|h_n|)} |f(y)|\,\mathrm{d}y\Big) \\
&\quad \leq \frac{1}{\omega_d |h_n|}\Big(r^{\beta-d}\int_{B(z+h_n,r+|h_n|)} |f(y)|\,\mathrm{d}y - (r+|h_n|)^{\beta-d}\int_{B(z+h_n,r+|h_n|)} |f(y)|\,\mathrm{d}y\Big) \\
&\quad = \frac{r^{\beta-d} - (r+|h_n|)^{\beta-d}}{\omega_d |h_n|}\int_{B(z+h_n,r+|h_n|)} |f(y)|\,\mathrm{d}y \\
&\quad \to \frac{(d-\beta) r^{\beta-d-1}}{\omega_d}\int_{B(z,r)} |f(y)|\,\mathrm{d}y
\end{aligned}$$

for $n \to \infty$, which concludes the proof. ■

Remark 3.1.2. Proposition 3.1.3 continues to hold for $\beta = 0$, replacing the weak derivative by the approximate derivative in the cases where the weak differentiability of $\mathfrak{M}$ is currently unknown.

3.1.6 A refined fractional maximal function

In view of the Kinnunen–Saksman type inequality (3.6), it is instructive to define the operator

$$\mathfrak{M}_{\beta,-1} f(x) = \sup_{B\in\mathfrak{B}_{\beta,x}(f)} r(B)^{\beta-1} \fint_B |f(y)|\,\mathrm{d}y,$$

so that for any $0 < \beta < d$,

$$|\nabla \mathfrak{M}_\beta f(x)| \leq (d-\beta)\mathfrak{M}_{\beta,-1} f(x) \qquad \text{for a.e. } x \in \mathbb{R}^d. \tag{3.7}$$

Furthermore, this extends to the case $\delta > 0$, that is,

$$|\nabla \mathfrak{M}_\beta^\delta f(x)| \leq (d-\beta)\mathfrak{M}_{\beta,-1} f(x) \qquad \text{for a.e. } x \in \mathbb{R}^d. \tag{3.8}$$

Indeed, let $\delta > 0$ and $B \in \mathfrak{B}^\delta_{\beta,x}$. Then, there exists $C \in \mathfrak{B}_{\beta,x}$ such that $r(C) \leq r(B)$. This immediately yields

$$r(B)^{\beta-1} \fint_B |f| \leq r(C)^{\beta-1} \fint_C |f| \leq \mathfrak{M}_{\beta,-1} f(x),$$

which implies (3.8) via Proposition 3.1.3.

The proof of Theorem 3.1.1 in [Wei21a] is obtained through the analogous bound on $\mathfrak{M}_{\beta,-1}$. Indeed, such a bound is of local nature. The following is a local version of [Wei21a, Theorem 1.2]; see [Wei21a, Remark 1.9].

Theorem 3.1.2. *Let $0 < \beta < d$ and $E \subseteq \mathbb{R}^d$. There exist constants $c > 1$ and $C_{d,\beta} > 0$ such that the inequality*

$$\|\mathfrak{M}_{\beta,-1} f\|_{L^{d/(d-\beta)}(E)} \leq C_{d,\beta} \|\nabla f\|_{L^1(D)}$$

holds for all $f \in W^{1,1}(\mathbb{R}^d)$, where

$$D = \bigcup_{B \in \mathfrak{I}_E} cB \qquad \textit{and} \qquad \mathfrak{I}_E := \{B \in \mathfrak{B}_{\beta,x}\,;\, \textit{for some } x \in E\}.$$

Remark 3.1.3. For $0 < \beta < d$ one has, combining (3.7) and 3.1.2, that

$$\|\nabla \mathfrak{M}_\beta f\|_{L^{d/(d-\beta)}(E)} \leq (d-\beta) C_{d,\beta} \|\nabla f\|_{L^1(D)},$$

where $C_{d,\beta}$ is the constant in Theorem 3.1.2.

3.1.7 Poincaré–Sobolev inequality

Another important tool for our purposes is the following.

Lemma 3.1.2. *Let $0 < \beta < d$, $f \in W^{1,1}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, $B = B(z,r) \in \mathfrak{B}_{\beta,x}(f)$ and $c > 1$. Then there is a constant $C_{d,\beta,c}$ such that*

$$\fint_{cB} |f(y)| \,\mathrm{d}y \leq C_{d,\beta,c}\, r \fint_{cB} |\nabla f(y)| \,\mathrm{d}y.$$

Proof By the triangle inequality and the Poincaré-Sobolev inequality there is a C_d such that

$$\fint_{cB} \big||f(y)| - |f|_{cB}\big| \,\mathrm{d}y \leq \fint_{cB} |f(y) - f_{cB}| \,\mathrm{d}y \leq C_d\, r \fint_{cB} |\nabla f(y)| \,\mathrm{d}y.$$

Since $B \in \mathfrak{B}_{\beta,x}$ we have $c^\beta |f|_{cB} < |f|_B$. This and the triangle inequality yield

$$c^d \fint_{cB} \big||f(y)| - |f|_{cB}\big| \,\mathrm{d}y \geq \fint_B \big||f(y)| - |f|_{cB}\big| \,\mathrm{d}y \geq |f|_B - |f|_{cB} \geq (c^\beta - 1) \fint_{cB} |f(y)| \,\mathrm{d}y.$$

Then, combining the above, we obtain

$$\fint_{cB} |f(y)| \,\mathrm{d}y \leq \frac{c^d C_d}{c^\beta - 1} r \fint_{cB} |\nabla f(y)| \,\mathrm{d}y,$$

as desired. ■

3.2 Convergences

In this section we review some auxiliary convergence results established in the series of papers [CMP17, BM19] which reduce the proof of Theorem 3.0.1 to the convergence of the difference $\mathfrak{M}_\beta f_j - \mathfrak{M}_\beta^\delta f_j$ on a compact set.

3.2.1 A Sobolev space lemma

We start recalling an auxiliary result concerning the convergence of the modulus of a sequence in $W^{1,1}(\mathbb{R}^d)$. This is useful in view of the identity (3.5).

Lemma 3.2.1 ([BM19, Lemma 2.3])**.** *Let $f \in W^{1,1}(\mathbb{R}^d)$ and $\{f_j\}_{j\in\mathbb{N}} \subset W^{1,1}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. Then $\big\||f_j| - |f|\big\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$.*

3.2.2 Convergence outside a compact set

By Theorem 3.1.1 and the work of the first and third author in [BM19] we have that it suffices to study the convergence in a compact a set.

Proposition 3.2.1 ([BM19, Proposition 4.10])**.** *Let $0 < \beta < d$, $f \in W^{1,1}(\mathbb{R}^d)$ and $\{f_j\}_{j\in\mathbb{N}} \subset W^{1,1}(\mathbb{R}^d)$ such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$. Then, for any $\varepsilon > 0$ there exists a compact set K and $j_\varepsilon > 0$ such that*

$$\|\nabla \mathfrak{M}_\beta f_j - \nabla \mathfrak{M}_\beta f\|_{L^{d/(d-\beta)}((3K)^c)} < \varepsilon$$

for all $j \geq j_\varepsilon$.

3.2.3 Continuity of $\mathfrak{M}_\beta^\delta$ in $W^{1,1}(\mathbb{R}^d)$, $\delta > 0$

A key observation is the a.e. convergence of the maximal function $\mathfrak{M}_\beta^\delta f_j$ at the derivative level.

Lemma 3.2.2. *Let $0 < \beta < d$, $\delta \geq 0$, $f \in W^{1,1}(\mathbb{R}^d)$ and $\{f_j\}_{j\in\mathbb{N}} \subset W^{1,1}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. Then*

$$\nabla \mathfrak{M}^\delta_\beta f_j(x) \to \nabla \mathfrak{M}^\delta_\beta f(x) \quad a.e. \quad as\ j \to \infty.$$

A version of this result for the full $\mathfrak{M}_\beta$ is given in [BM19, Lemma 2.4]. The proof for $\mathfrak{M}^\delta_\beta$ is identical (in fact, it slightly simplifies), and relies on Luiro's formula for $\mathfrak{M}^\delta_\beta$, that is, 3.1.2. We omit further details. For $\delta > 0$, we have the following norm convergence.

Proposition 3.2.2. *Let $0 < \beta < d$, $\delta > 0$, $f \in W^{1,1}(\mathbb{R}^d)$ and $\{f_j\}_{j\in\mathbb{N}} \subset W^{1,1}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. Let $K \subset \mathbb{R}^d$ be a compact set.*

$$\|\nabla \mathfrak{M}^\delta_\beta f - \nabla \mathfrak{M}^\delta_\beta f_j\|_{L^{d/(d-\beta)}(K)} \to 0 \ as\ j \to \infty.$$

Proof By Proposition 3.1.2 and Lemma 3.2.1 there exists $j_0 \in \mathbb{N}$ such that

$$|\nabla \mathfrak{M}^\delta_\beta f_j(x)| \leq \frac{1}{\omega_d\, \delta^{d-\beta}} \|\nabla |f_j|\|_1 \leq \frac{1}{\omega_d\, \delta^{d-\beta}} \|\nabla |f|\|_1 + 1 \qquad \text{for all } j \geq j_0, \quad \text{a.e. } x \in K.$$

Furthermore, by Lemma 3.2.2

$$\nabla \mathfrak{M}^\delta_\beta f_j(x) \to \nabla \mathfrak{M}^\delta_\beta f(x) \quad \text{a.e.} \quad as\ j \to \infty.$$

The convergence on $L^{d/(d-\beta)}(K)$ then follows from the dominated convergence theorem. ∎

3.2.4 δ-convergence of $\nabla \mathfrak{M}^\delta_\beta f$

Here we establish that $\nabla \mathfrak{M}^\delta_\beta f$ provides a good approximation for $\nabla \mathfrak{M}_\beta f$ in $L^{d/(d-\beta)}(K)$ when $\delta \to 0$. This relies on the Theorem 3.1.1.

Lemma 3.2.3. *Let $0 < \beta < d$ and $f \in W^{1,1}(\mathbb{R}^d)$. Then*

$$\|\nabla \mathfrak{M}_\beta f - \nabla \mathfrak{M}^\delta_\beta f\|_{L^{d/(d-\beta)}(K)} \to 0 \qquad as \quad \delta \to 0.$$

Proof *Recall from §3.1.3 that for a.e. $x \in \mathbb{R}^d$ one has that if $B(z, r) \in \mathfrak{B}^\delta_{\beta,x}$, then $r > 0$. This and Luiro's formula (3.5) imply that for a.e. $x \in \mathbb{R}^d$ there exists $\delta_x > 0$ such that*

$$\nabla \mathfrak{M}^\delta_\beta f(x) = \nabla \mathfrak{M}_\beta f(x) \qquad for\ all\ 0 \leq \delta < \delta_x,$$

and thus $\nabla \mathfrak{M}^\delta_\beta f(x) \to \nabla \mathfrak{M}_\beta f(x)$ for a.e. $x \in \mathbb{R}^d$ as $\delta \to 0$. Furthermore, as proven in (3.8), for a.e. $x \in \mathbb{R}^d$ we have that

$$|\nabla \mathfrak{M}^\delta_\beta f(x)| \leq \mathfrak{M}_{\beta,-1} f(x) \qquad for\ all\ \delta \geq 0.$$

Since $f \in W^{1,1}(\mathbb{R}^d)$, Theorem 3.1.2 ensures that $\mathfrak{M}_{\beta,-1} f \in L^{d/(d-\beta)}(\mathbb{R}^d)$ and we can then conclude the result by the dominated convergence theorem. ∎

3.3 Proof of Theorem 3.0.1

Let $f \in W^{1,1}(\mathbb{R}^d)$ and $\{f_j\}_{j\in\mathbb{N}} \subset W^{1,1}(\mathbb{R}^d)$ be a sequence of functions such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. If $f = 0$ then the result follows directly from the boundedness, that is 3.1.1. From now on we assume that $f \neq 0$. Let $\varepsilon > 0$. Then by Proposition 3.2.1 it is sufficient to prove that there exists $j^* \in \mathbb{N}$ such that

$$\|\nabla \mathfrak{M}_\beta f - \nabla \mathfrak{M}_\beta f_j\|_{L^{d/(d-\beta)}(K)} < 3\varepsilon \tag{3.9}$$

for all $j \geq j^*$. To this end, for any $\delta > 0$, use the triangle inequality to bound

$$\begin{aligned} \|\nabla \mathfrak{M}_\beta f - \nabla \mathfrak{M}_\beta f_j\|_{L^{d/(d-\beta)}(K)} \leq & \|\nabla \mathfrak{M}_\beta f - \nabla \mathfrak{M}_\beta^\delta f\|_{L^{d/(d-\beta)}(K)} \\ & + \|\nabla \mathfrak{M}_\beta^\delta f - \nabla \mathfrak{M}_\beta^\delta f_j\|_{L^{d/(d-\beta)}(K)} \\ & + \|\nabla \mathfrak{M}_\beta^\delta f_j - \nabla \mathfrak{M}_\beta f_j\|_{L^{d/(d-\beta)}(K)}. \end{aligned} \tag{3.10}$$

To finish the proof, it suffices to show that for $\varepsilon > 0$ fixed, there exist a δ^* and a j^* such that for $\delta = \delta^*$ and all $j \geq j^*$, each of the summands on the right hand side of (3.10) is bounded by ε. We choose δ^* depending on ε, K and f, and j^* depending on δ^*, ε, K, f and the sequence $\{f_j\}_{j\in\mathbb{N}}$.

For the first term, we know by Lemma 3.2.3 that there exists a $\delta' > 0$ such that

$$\|\nabla \mathfrak{M}_\beta f - \nabla \mathfrak{M}_\beta^\delta f\|_{L^{d/(d-\beta)}(K)} < \varepsilon$$

for all $0 \leq \delta \leq \delta'$. For the second term, we have by Proposition 3.2.2 that for every $\delta > 0$ there exists a $j(\delta) \in \mathbb{N}$ such that

$$\|\nabla \mathfrak{M}_\beta^\delta f - \nabla \mathfrak{M}_\beta^\delta f_j\|_{L^{d/(d-\beta)}(K)} < \varepsilon$$

for all $j \geq j(\delta)$. The rest of the section is devoted to proving a favourable bound for the third term. More precisely, we will show that there are $\tilde{\delta} > 0$ and $\tilde{j} \in \mathbb{N}$ such that for all $0 \leq \delta \leq \tilde{\delta}$ and $j \geq \tilde{j}$,

$$\|\nabla \mathfrak{M}_\beta^\delta f_j - \nabla \mathfrak{M}_\beta f_j\|_{L^{d/(d-\beta)}(K)} < \varepsilon. \tag{3.11}$$

Temporarily assuming this, we can then conclude that for $\delta = \delta^* := \min\{\delta', \tilde{\delta}\}$ and $j \geq j^* := \max\{j(\delta^*), \tilde{j}\}$, the right-hand side of (3.10) is bounded by at most 3ε, as desired for (3.9).

We now turn to the proof of (3.11). We start by noting that there exists a $\lambda_0 > 0$ and a $j_0 \in \mathbb{N}$ such that for all $j \geq j_0$ and $x \in K$ we have $\mathfrak{M}_\beta f_j(x) > \lambda_0$. Indeed, as $f \in L^1(\mathbb{R}^d)$, there exists a ball B_0 that contains K with $\int_{B_0} |f| > \frac{1}{2} \int_{\mathbb{R}^d} |f|$. As $\|f_j - f\|_1 \to 0$ as $j \to 0$, by the triangle inequality, there exists $j_0 > 0$ such that for all $j \geq j_0$ we have $\int_{B_0} |f_j| > \frac{1}{2} \int_{B_0} |f| > \frac{1}{4} \int_{\mathbb{R}^d} |f|$. Then, for every $j \geq j_0$ and $x \in K$ we have

$$\mathfrak{M}_\beta f_j(x) \geq 2^\beta r(B_0)^\beta \fint_{B(x,2r(B_0))} |f_j| > \frac{(2r(B_0))^{\beta-d}}{4\,\omega_d} \int_{\mathbb{R}^d} |f|,$$

where in the last inequality we have used that $B(x, 2r(B_0)) \supset B_0$ for all $x \in K$. Thus, we can take λ_0 to be the right-hand side of the inequality above. Furthermore by (3.1.2), if there exists a $B \in \mathfrak{B}_{\beta,x}(f_j)$ such that $r(B) \geq \delta$ then $\nabla \mathfrak{M}_\beta f_j(x) = \nabla \mathfrak{M}_\beta^\delta f_j(x)$. Define

$$E_{\lambda_0,\delta,j} := \Big\{ x \in K : \text{ if } B \in \mathfrak{B}_{\beta,x}(f_j), \text{ then } \ r(B) < \delta \ \text{ and } \ r(B)^\beta \fint_B |f_j| > \lambda_0 \Big\}.$$

By the previous two observations, Proposition 3.1.3 and a crude application of the triangle inequality, one has

$$\begin{aligned} \|\nabla \mathfrak{M}_\beta^\delta f_j - \nabla \mathfrak{M}_\beta f_j\|_{L^{d/(d-\beta)}(K)} &= \|\nabla \mathfrak{M}_\beta^\delta f_j - \nabla \mathfrak{M}_\beta f_j\|_{L^{d/(d-\beta)}(E_{\lambda_0,\delta,j})} \\ &\leq 2(d-\beta)\, \|\mathfrak{M}_{\beta,-1} f_j\|_{L^{d/(d-\beta)}(E_{\lambda_0,\delta,j})}. \end{aligned}$$

for all $j \geq j_0$. Define the indexing set

$$\mathfrak{I}_{\lambda_0,\delta,j} := \Big\{ B \in \mathfrak{B}_{\beta,x}(f_j) \ : x \in K, \ r(B) < \delta \ \text{ and } \ r(B)^\beta \fint_B |f_j| > \lambda_0 \Big\}$$

and consider the set

$$D_{\lambda_0,\delta,j} := \bigcup_{B \in \mathfrak{I}_{\lambda_0,\delta,j}} cB,$$

where c is the constant from (3.1.2). Then, by Theorem 3.1.2, we have

$$\|\mathfrak{M}_{\beta,-1} f_j\|_{L^{d/(d-\beta)}(E_{\lambda_0,\delta,j})} \leq C_{d,\beta} \|\nabla f_j\|_{L^1(D_{\lambda_0,\delta,j})}$$

for any $\delta > 0$. Thus, the proof of (3.11) is reduced to showing that there exist a $\tilde{\delta} > 0$ and a $j_1 \in \mathbb{N}$ such that for all $j \geq j_1$ and $0 \leq \delta \leq \tilde{\delta}$ we have

$$\|\nabla f_j\|_{L^1(D_{\lambda_0,\delta,j})} < \frac{\varepsilon}{2(d-\beta)C_{d,\beta}}, \tag{3.12}$$

as one can then take $\tilde{j} := \max\{j_0, j_1\}$.

In order to prove (3.12), we first use the triangle inequality and that $\|\nabla f_j - \nabla f\|_{L^1(\mathbb{R}^d)} \to 0$ as $j \to \infty$ to find a $j_2 \in \mathbb{N}$ such that

$$\|\nabla f_j\|_{L^1(D_{\lambda_0,\delta,j})} \leq \|\nabla f\|_{L^1(D_{\lambda_0,\delta,j})} + \frac{\varepsilon}{4(d-\beta)C_{d,\beta}}. \tag{3.13}$$

for any $\delta > 0$ and $j \geq j_2$.

Next, let $x \in D_{\lambda_0,\delta,j}$. Then there is a $B \in \mathfrak{I}_{\lambda_0,\delta,j}$ with $x \in cB$. So, by 3.1.2, we have

$$\begin{aligned} \lambda_0 \leq c^d r(B)^\beta \fint_{cB} |f_j| &\leq C_{d,\beta,c}\, c^{d+1} r(B)^{\beta+1} \fint_{cB} |\nabla f_j| \\ &\leq C_{d,\beta,c}\, c^{d-\beta+1} \delta\, \widetilde{M}_\beta |\nabla f_j|(x), \end{aligned}$$

where $\widetilde{M}_\beta$ in the above inequality denotes the uncentered fractional maximal operator. Hence, by the weak $(1, d/(d-\beta))$ inequality for $\widetilde{M}_\beta$,

$$\begin{aligned} |D_{\lambda_0,\delta,j}| &\leq \left| \left\{ x : \widetilde{M}_\beta |\nabla f_j|(x) \geq \frac{\lambda_0}{C_{d,\beta,c} c^{d-\beta+1} \delta} \right\} \right| \\ &\leq C_{d,\beta,c,\lambda_0} \delta^{d/(d-\beta)} \|\nabla f_j\|_1^{d/(d-\beta)} \\ &\leq C_{d,\beta,c,\lambda_0} \delta^{d/(d-\beta)} \left(1 + \|\nabla f\|_1^{d/(d-\beta)} \right) \end{aligned} \tag{3.14}$$

if $j \geq j_3$ for some $j_3 \in \mathbb{N}$, using that $\|\nabla f_j - \nabla f\|_{L^1(\mathbb{R}^d)} \to 0$ as $j \to \infty$.

Finally, note that as $\nabla f \in L^1(\mathbb{R}^d)$, there exists $\rho > 0$ such that for all $A \subseteq \mathbb{R}^d$ satisfying $|A| < \rho$, one has

$$\|\nabla f\|_{L^1(A)} < \frac{\varepsilon}{4(d-\beta) C_{d,\beta}}. \tag{3.15}$$

As the right-hand side of (3.14) goes to zero for $\delta \to 0$ uniformly in j, there exists $\tilde{\delta} > 0$ such that $|D_{\lambda_0,\delta,j}| < \rho$ for all $j \geq j_3$ and $\delta < \tilde{\delta}$. Thus, taking $j_1 := \max\{j_2, j_3\}$, (3.12) follows from combining (3.13) and (3.15) with $A = D_{\lambda_0,\delta,j}$. This implies the claimed inequality (3.11) and therefore finishes the proof of Theorem 3.0.1. ■

Remark 3.3.1. Note that in the above proof, instead of using 3.2.3 to bound the first term in (3.10), we could have also bounded it running the same scheme as for the third term.

Chapter 4

Sunrise strategy for the continuity of maximal operators

4.1 Introduction

In this chapter we aim to continue developing the continuity theory for maximal operators at the derivative level. Our purpose here is to develop a strategy to approach the $W^{1,1}$–continuity problem for a certain class of maximal operators of general interest. Our first result, a model case for our global strategy, complements the recent boundedness result of Luiro [Lui18].

Theorem 4.1.1. *The map $f \mapsto \nabla \widetilde{M} f$ is continuous from $W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$ for $d \geq 2$.*

Recall that we use $\widetilde{M} f$ as the uncentered version of (1). Despite the innocence of the statement in Theorem 4.1.1, one should not underestimate the subtlety of the problem, as it will become evident as the proof unfolds and we find ourselves in a beautiful maze of possibilities. It is worth mentioning a few words on the difficulties that one faces when trying to prove this theorem, in direct comparison to the core papers in the literature that deal with similar continuity issues. First, the original proof of Luiro [Lui07] to show the continuity of M (or $\widetilde{M}$) in $W^{1,p}(\mathbb{R}^d)$ $(1 < p \leq \infty)$ relies decisively on the boundedness of M in $L^p(\mathbb{R}^d)$, which is not available in our situation. This was already an issue in the work of Carneiro, Madrid and Pierce [CMP17, Theorem 1] to prove the continuity of $f \mapsto \big(\widetilde{M} f\big)'$ from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$, and a new path was developed. A crucial element in the proof of [CMP17, Theorem 1] was the ability to decompose $\widetilde{M}$ as a maximum of two operators, namely,

$$\widetilde{M} f(x) = \max\big\{M_R f(x), M_L f(x)\big\} \quad \text{for all} \quad x \in \mathbb{R}, \tag{4.1}$$

where M_R and M_L are the one-sided maximal operators, to the right and left, respectively. The monotonicity properties of these one-sided operators in the connecting and disconnecting sets played a very important role [CMP17, §5.4.1]. In our situation of Theorem 4.1.1, when dealing with radial functions on $\mathbb{R}^d$, there is no obvious way to decompose $\widetilde{M}$ into two "lateral" operators with similar monotonicity properties, and this is a major obstacle.

The proof of Theorem 4.1.1 is carefully developed in §4.2 to §4.5, where each section addresses an independent aspect of the overall strategy. In §4.2 we provide the preliminaries about maximal operators and radial Sobolev functions, and treat some basic regularity and convergence issues in this setup. In §4.3 we establish a control of the convergence in a neighborhood of the origin, where potential singularities may appear, thus making it possible to concentrate our efforts in the complement of such neighborhood. §4.4 develops what is really the main insight of our study: a suitable decomposition in replacement of (4.1), inspired in the classical sunrise lemma in harmonic analysis. Finally, §4.5 brings the proof itself, in which we put together all the pieces in our board, and conclude by carefully analyzing a dichotomy that naturally arises.

Once the work in §4.2 to §4.5 is complete, and we are able to fully see the strategy working in the model case of Theorem 4.1.1, we take a moment in §4.6 to reflect on what really are the abstract core elements that make the method work. In fact, the reach of our sunrise strategy goes way beyond the situation of Theorem 4.1.1, and these abstract guidelines pave the way for further applications that we now describe.

4.1.1 Further applications

In the sphere set up, we establish the following.

Theorem 4.1.2. *The map $f \mapsto \nabla\widetilde{\mathcal{M}}f$ is continuous from $W^{1,1}(\mathbb{S}^1)$ to $L^1(\mathbb{S}^1)$ and from $W^{1,1}_{\text{pol}}(\mathbb{S}^d)$ to $L^1(\mathbb{S}^d)$ for $d \geq 2$.*

The proof of this result is given in §4.7.1.

Non-tangential Hardy-Littlewood maximal operator

For $\alpha \geq 0$ and $f \in L^1_{\text{loc}}(\mathbb{R})$ we define the non-tangential Hardy-Littlewood maximal operator M^α by

$$M^\alpha f(x) = \sup_{|x-y|\leq \alpha r} \frac{1}{2r}\int_{y-r}^{y+r} |f(t)| \, \mathrm{d}t. \tag{4.2}$$

With our previous notation, note that when $\alpha = 0$ we have $M^0 = M$ (the centered Hardy-Littlewood maximal operator) and when $\alpha = 1$ we have $M^1 = \widetilde{M}$ (the uncentered one). In [Ram19], J. P. Ramos established a beautiful regularity result for such operators: for $\alpha \geq \frac{1}{3}$ and $f : \mathbb{R} \to \mathbb{R}$ of bounded variation, one has

$$\mathrm{Var}(M^\alpha f) \leq \mathrm{Var}(f). \tag{4.3}$$

The interesting feature of (4.3) is the variation contractivity property (i.e. the constant $C = 1$ on the right-hand side of the inequality). The mechanism that implies the contractivity in (4.3) is the fact that $M^\alpha f$ has no local maxima in the disconnecting set (say, with f slightly smoother, and then one approximates). The threshold $\alpha = \frac{1}{3}$ is geometrically relevant for this absence of local maxima, and we will review later how it comes into play. From (4.3) one can show that when $\alpha \geq \frac{1}{3}$ and $f \in W^{1,1}(\mathbb{R})$ then $M^\alpha f$ is weakly differentiable and

$$\|(M^\alpha f)'\|_{L^1(\mathbb{R})} \leq \|f'\|_{L^1(\mathbb{R})}.$$

We now consider an extension of this operator to several variables. Let $\mathcal{Q}$ be the family of all closed cubes in $\mathbb{R}^d$ (with any possible center and any possible orientation, not necessarily with sides parallel to the original axes). If $Q \in \mathcal{Q}$ we let αQ be the cube that is the dilation of Q by a factor α with the same center. For $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ we now define

$$M^\alpha f(x) = \sup_{x \in \alpha Q} \fint_Q |f(y)| \, \mathrm{d}y. \tag{4.4}$$

Note that in dimension $d = 1$ definitions (4.2) and (4.4) agree. We establish here the following result.

Theorem 4.1.3. *Let $\alpha \geq \frac{1}{3}$ and M^α be defined by (4.4).*

(i) If $d = 1$ the map $f \mapsto (M^\alpha f)'$ is continuous from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$.

(ii) If $d \geq 2$ and $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ then $M^\alpha f$ is weakly differentiable. Moreover, the map $f \mapsto \nabla M^\alpha f$ is bounded and continuous from $W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$.

The proof of this result is given in §4.7.2. The boundedness in Theorem 4.1.3 (ii) is also a novelty in the theory. We give a self-contained argument that, *en passant*, provides an alternative approach to [Ram19] in order to prove (4.3); see Proposition 4.7.1 for details.

Non-tangential heat flow maximal operator

For $t > 0$ and $x \in \mathbb{R}^d$ let

$$\varphi_t(x) = \frac{1}{(4\pi t)^{d/2}} \, e^{-|x|^2/4t}$$

be the heat kernel. For $\alpha \geq 0$, consider the following maximal operator

$$M^\alpha_\varphi f(x) = \sup_{t>0\,;\,|y-x| \leq \alpha\sqrt{t}} (|f| * \varphi_t)(y). \tag{4.5}$$

If we write

$$u(x,t) := (|f| * \varphi_t)(x)$$

then we know that u verifies the heat equation $u_t - \Delta u = 0$ in $\mathbb{R}^d \times (0,\infty)$ with $\lim_{t\to 0^+} u(x,t) = |f(x)|$ for a.e. $x \in \mathbb{R}^d$ (provided f has some minimal regularity, say $f \in L^p(\mathbb{R}^d)$ for any $1 \leq p \leq \infty$). In this sense, when $\alpha = 0$, $M^0_\varphi f(x)$ is just the sup of $u(x,t)$ over the vertical fiber over x (the heat flow maximal operator) and, when $\alpha > 0$, $M^\alpha_\varphi f(x)$ is a sup of $u(y,t)$ within a parabolic region with lower vertex in x (the non-tangential heat flow maximal operator). Here we consider the non-tangential case and prove the following.

Theorem 4.1.4. *Let $\alpha > 0$ and M_φ^α defined by (4.5). The map $f \mapsto \nabla M_\varphi^\alpha f$ is bounded and continuous from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$ and from $W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$ for $d \geq 2$.*

The proof of this result is given in §4.7.3. The boundedness part will follow from the circle of ideas in Chapter 1, and the main novelty here is the continuity part that will follow from our sunrise strategy. The continuity in the centered case $\alpha = 0$ is not exactly currently accessible with our methods, and we comment a bit on the difficulties for this and other operators of convolution type (e.g. with the Poisson kernel) in §4.7.4. However, the one-dimensional case of this problem will be revisited in Chapter 6.

4.2 Preliminaries: regularity and convergence

4.2.1 Basic regularity

Let us first make some generic considerations about radial functions in $\mathbb{R}^d$ and weak derivatives. Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{\pm\infty\}$ be a radial function. With a (hopefully) harmless abuse of notation, throughout the text we write $f(x)$ when we referring to this function in $\mathbb{R}^d$, and $f(r)$ when referring to its radial restriction in $(0, \infty)$, where $r = |x|$.

A radial function $f(x)$ is weakly differentiable in $\mathbb{R}^d \setminus \{0\}$ if and only if its radial restriction $f(r)$ is weakly differentiable in $(0, \infty)$. In this case, the weak gradient ∇f of $f(x)$ and the weak derivative f' of $f(r)$ are related by $\nabla f(x) = f'(|x|)\frac{x}{|x|}$. Hence $f(x) \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ if and only if $f(r) \in W^{1,1}((0, \infty), r^{d-1}\,\mathrm{d}r)$, and

$$\int_{\mathbb{R}^d} |\nabla f(x)|\ \mathrm{d}x = \omega_{d-1} \int_0^\infty |f'(r)|\, r^{d-1}\ \mathrm{d}r < \infty. \tag{4.6}$$

In particular, after a possible redefinition on a set of measure zero, one can take $f(r)$ continuous in $(0, \infty)$; in fact, absolutely continuous in each interval $[\delta, \infty) \subset (0, \infty)$, and hence differentiable a.e. in $(0, \infty)$. This is equivalent to saying that $f(x)$ is continuous in $\mathbb{R}^d \setminus \{0\}$ and differentiable a.e. in $\mathbb{R}^d \setminus \{0\}$. *It is henceforth agreed that we will always work under such regularity assumptions.* Note that this is essentially the best regularity one can expect, since at the origin our function $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ may have a singularity like $|x|^\alpha$ with $-d+1 < \alpha < 0$.

If $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ is continuous in $\mathbb{R}^d \setminus \{0\}$, it is not hard to show that $\widetilde{M}f$ is also continuous in $\mathbb{R}^d \setminus \{0\}$ (and, of course, radial). From [Lui18] we know that $\widetilde{M}f$ is weakly differentiable in $\mathbb{R}^d$ and

$$\|\nabla \widetilde{M}f\|_{L^1(\mathbb{R}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbb{R}^d)}. \tag{4.7}$$

As in (4.6), it follows that $\widetilde{M}f(r)$ is absolutely continuous in each interval $[\delta, \infty) \subset (0, \infty)$, and hence differentiable a.e. in $(0, \infty)$. Observe that both f and $\widetilde{M}f$ vanish at infinity (recall that $\widetilde{M}f \in L^{1,\infty}(\mathbb{R}^d)$). In fact, a bit more can be said. Since

$$\widetilde{M}f(r) = -\int_r^\infty \big(\widetilde{M}f\big)'(t)\ \mathrm{d}t,$$

we have

$$\begin{aligned}(d-1)\int_0^\infty \widetilde{M}f(r)\, r^{d-2}\,\mathrm{d}r &= (d-1)\int_0^\infty \left(\int_r^\infty -\big(\widetilde{M}f\big)'(t)\,\mathrm{d}t\right) r^{d-2}\,\mathrm{d}r\\ &\le (d-1)\int_0^\infty \left(\int_r^\infty \big|\big(\widetilde{M}f\big)'(t)\big|\,\mathrm{d}t\right) r^{d-2}\,\mathrm{d}r\\ &= (d-1)\int_0^\infty \int_0^t r^{d-2}\big|\big(\widetilde{M}f\big)'(t)\big|\,\mathrm{d}r\,\mathrm{d}t\\ &= \int_0^\infty \big|\big(\widetilde{M}f\big)'(t)\big|\, t^{d-1}\,\mathrm{d}t<\infty.\end{aligned} \tag{4.8}$$

The latter is finite from (4.7). An analogous computation holds with $|f(r)|$ replacing $\widetilde{M}f(r)$. Hence $r\mapsto |f(r)|\,r^{d-1}$ and $r\mapsto \widetilde{M}f(r)\,r^{d-1}$ have integrable derivatives in $(0,\infty)$, and by the fundamental theorem of calculus the limits $\lim_{r\to\infty}|f(r)|\,r^{d-1}$ and $\lim_{r\to\infty}\widetilde{M}f(r)\,r^{d-1}$ must exist. If these limits were not zero, one would plainly contradict (4.8) and therefore

$$\lim_{r\to\infty}|f(r)|\,r^{d-1}=\lim_{r\to\infty}\widetilde{M}f(r)\,r^{d-1}=0. \tag{4.9}$$

Another application of the fundamental theorem of calculus shows that the limits

$$\lim_{r\to 0^+}|f(r)|\,r^{d-1}$$

and

$$\lim_{r\to 0^+}\widetilde{M}f(r)\,r^{d-1}$$

must also exist. If these were not zero, one would contradict the fact that f and $\widetilde{M}f$ belong to $L^{d/(d-1)}(\mathbb{R}^d)$ (the former by Sobolev embedding, and the latter by the boundedness of $\widetilde{M}$ in $L^{d/(d-1)}(\mathbb{R}^d)$). Hence

$$\lim_{r\to 0^+}|f(r)|\,r^{d-1}=\lim_{r\to 0^+}\widetilde{M}f(r)\,r^{d-1}=0. \tag{4.10}$$

4.2.2 Splitting and nonnegative functions

The following result will be very useful in our strategy. We state it here in a more general version, having in mind the additional applications given in the forthcoming §4.7.

Lemma 4.2.1 (Divide and conquer)**.** *Let $I\subset\mathbb{R}$ be an open interval and let μ be a nonnegative measure on I such that μ and the Lebesgue measure are mutually absolutely continuous. Let $\mathcal{X}$ be the space of functions $\psi: I\to\mathbb{R}$ satisfying the following conditions:*

(i) ψ is absolutely continuous in each compact interval of I;

(ii) $\psi'\in L^1(I,\,\mathrm{d}\mu)$.

Let h and g be two functions in $\mathcal{X}$ and let $\{h_j\}_{j\geq 1}$ and $\{g_j\}_{j\geq 1}$ be two sequences in $\mathcal{X}$ such that

(a) $h_j(x) \to h(x)$ and $g_j(x) \to g(x)$ as $j \to \infty$, for all $x \in I$;

(b) $\|h_j' - h'\|_{L^1(I,\,\mathrm{d}\mu)} \to 0$ and $\|g_j' - g'\|_{L^1(I,\,\mathrm{d}\mu)} \to 0$ as $j \to \infty$.

Define $f_j := \max\{g_j, h_j\}$ for each $j \geq 1$ and $f := \max\{g, h\}$. Then $f \in \mathcal{X}$, $\{f_j\}_{j\geq 1} \subset \mathcal{X}$, and

$$\|f_j' - f'\|_{L^1(I,\,\mathrm{d}\mu)} \to 0 \quad \text{as} \quad j \to \infty.$$

Proof This is essentially [CMP17, Lemma 11], with minor modifications in the proof. ■
Remark: For the proof of Theorem 4.1.1 we shall use Lemma 4.2.1 with $I = (0, \infty)$ and $\mathrm{d}\mu(r) = r^{d-1}\,\mathrm{d}r$. A basic modification of Lemma 4.2.1 allows us to also consider the situation where $I = \mathbb{S}^1$ and μ is the arclength measure. This shall be used in §4.7.

We now perform a basic reduction. The next result holds for general sequences of functions in $W^{1,1}(\mathbb{R}^d)$, see [MM79]. We state it and give a brief proof in the case of radial functions, that will be sufficient for our purposes here.

Proposition 4.2.1 (Reduction to nonnegative functions)**.** *Let $f \in W^{1,1}_{\mathrm{rad}}(\mathbb{R}^d)$ and $\{f_j\}_{j\geq 1} \subset W^{1,1}_{\mathrm{rad}}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. Then $\||f_j| - |f|\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$.*

Proof Since $\big||f_j| - |f|\big| \leq |f_j - f|$ pointwise, it follows that $\||f_j| - |f|\|_{L^1(\mathbb{R}^d)} \to 0$ as $j \to \infty$. By the fundamental theorem of calculus, for each $r \geq \delta$,

$$|f(r) - f_j(r)| = \left|\int_r^\infty (f' - f_j')(t)\ \mathrm{d}t\right| \lesssim_\delta \int_0^\infty \left|(f' - f_j')(t)\right| t^{d-1}\ \mathrm{d}t \ \to\ 0, \tag{4.11}$$

as $j \to \infty$. Noting that $|f| = \max\{f, -f\}$, the fact that $\|\nabla|f_j| - \nabla|f|\|_{L^1(\mathbb{R}^d)} = \omega_{d-1}\||f_j|' - |f|'\|_{L^1((0,\infty),\, r^{d-1}\,\mathrm{d}r)} \to 0$ follows directly from Lemma 4.2.1. ■

Since the maximal operator only sees the absolute value of a function, in light on Proposition 4.2.1 *we can assume for the rest of the proof of Theorem 4.1.1 that all the functions considered are nonnegative.*

4.2.3 Connecting and disconnecting sets, and local maxima

Let $f \in W^{1,1}_{\mathrm{rad}}(\mathbb{R}^d)$ be continuous in $\mathbb{R}^d \setminus \{0\}$ and nonnegative. Define the d-dimensional disconnecting set by

$$\mathcal{D}(f) = \left\{x \in \mathbb{R}^d \setminus \{0\}\ :\ \widetilde{M}f(x) > f(x)\right\},$$

and its corresponding one-dimensional radial version

$$D(f) = \{|x|\ :\ x \in \mathcal{D}(f)\}.$$

Analogously, we define the connecting set

$$\mathcal{C}(f) = \big\{x \in \mathbb{R}^d \setminus \{0\} \ : \ \widetilde{M}f(x) = f(x)\big\},$$

and its one-dimensional radial version

$$C(f) = \{|x| \ : \ x \in \mathcal{C}(f)\}.$$

Note that the sets $\mathcal{D}(f) \subset \mathbb{R}^d \setminus \{0\}$ and $D(f) \subset (0, \infty)$ are open. Note also that if $r \in C(f)$ is a point of differentiability of f, then we must have $f'(r) = 0$; otherwise one could find a small ball over which the average beats $f(r)$, and r would belong to $D(f)$ instead. We now recall a basic result of the theory, that will be crucial for our sunrise construction later in §4.4.

Proposition 4.2.2 (Absence of local maxima)**.** *Let $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$. The function $\widetilde{M}f(r)$ does not have a strict local maximum in $D(f)$.*

Proof By a strict local maximum we mean a point $r_0 \in D(f)$ for which there exist s_0 and t_0 with $s_0 < r_0 < t_0$, $[s_0, t_0] \subset D(f)$, such that $\widetilde{M}f(r) \leq \widetilde{M}f(r_0)$ for all $r \in [s_0, t_0]$ and $\widetilde{M}f(s_0)$, $\widetilde{M}f(t_0) < \widetilde{M}f(r_0)$. Let $x_0 \in \mathbb{R}^d$ be such that $|x_0| = r_0$, and consider a closed ball $\overline{B}$ such that $x_0 \in \overline{B}$ and $\widetilde{M}f(x_0) = ⨍_{\overline{B}} f$ (observe that such a ball exists and has a strictly positive radius since $x_0 \in \mathcal{D}(f)$). From the above we see that $\{|y| \ : \ y \in \overline{B}\} \subset (s_0, t_0)$. Since $[s_0, t_0] \subset D(f)$ we obtain

$$\widetilde{M}f(x_0) = ⨍_{\overline{B}} f \ < ⨍_{\overline{B}} \widetilde{M}f \leq \widetilde{M}f(x_0)\,,$$

a contradiction. ■

4.2.4 Pointwise convergence

For $x \in \mathbb{R}^d \setminus \{0\}$, let us define $\mathcal{B}(f;x)$ as the set of closed balls $\overline{B}$ that realize the supremum in the definition of the maximal function at the point x, that is

$$\mathcal{B}(f;x) = \left\{\overline{B}\,;\ x \in \overline{B} \ : \ \widetilde{M}f(x) = ⨍_{\overline{B}} f(y)\ \mathrm{d}y\right\}. \tag{4.12}$$

Note that we include possibility that $\overline{B} = \{x\}$ (we may think of radius zero here), with the understanding that $⨍_{\{x\}} f(y)\ \mathrm{d}y := f(x)$. Therefore, we note that $\mathcal{B}(f;x)$ is always non-empty. The next proposition qualitatively describes the derivative of the maximal function.

Proposition 4.2.3 (The derivative of the maximal function)**.** *Let $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ be a non-negative function and let $x \in \mathbb{R}^d \setminus \{0\}$ be a point of differentiability of $\widetilde{M}f$. Then, for any ball $\overline{B} \in \mathcal{B}(f;x)$ of strictly positive radius, we have*

$$\nabla\widetilde{M}f(x) = ⨍_{\overline{B}} \nabla f(y)\ \mathrm{d}y.$$

Proof This is contained in [Lui18, Lemma 2.2]. ■

This leads us to our considerations on pointwise convergence issues.

Proposition 4.2.4 (Pointwise convergence for $\widetilde{M}$)**.** *Let $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ and $\{f_j\}_{j\geq 1} \subset W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. The following statements hold.*

(i) For each $\delta > 0$, we have $f_j(r) \to f(r)$ and $\widetilde{M}f_j(r) \to \widetilde{M}f(r)$ uniformly in the set $\{r \geq \delta\}$ as $j \to \infty$.

(ii) If $x \in \mathbb{R}^d \setminus \{0\}$, $\overline{B_{s_j}(z_j)} \in \mathcal{B}(f_j; x)$[1] and $(s, z) \in [0,\infty) \times \mathbb{R}^d$ is an accumulation point of the sequence $\{(s_j, z_j)\}_{j\geq 1}$, then $\overline{B_s(z)} \in \mathcal{B}(f; x)$.

(iii) For almost all $r \in D(f)$ we have $\big(\widetilde{M}f_j\big)'(r) \to \big(\widetilde{M}f\big)'(r)$ as $j \to \infty$.

Proof Part (i). The uniform convergence $f_j(r) \to f(r)$ as $j \to \infty$ follows from (4.11). Using the sublinearity of $\widetilde{M}$ we also have

$$
\begin{aligned}
\big|\widetilde{M}f(r) - \widetilde{M}f_j(r)\big| \leq \widetilde{M}(f - f_j)(r) &= -\int_r^\infty \big(\widetilde{M}(f-f_j)\big)'(t)\;\mathrm{d}t \\
&\lesssim_\delta \int_0^\infty \big|\big(\widetilde{M}(f-f_j)\big)'(t)\big|\, t^{d-1}\;\mathrm{d}t \qquad (4.13)\\
&\lesssim_{\delta,d} \int_0^\infty \big|(f' - f_j')(t)\big|\, t^{d-1}\;\mathrm{d}t \;\to\; 0
\end{aligned}
$$

as $j \to \infty$. Note the use of (4.7) in the last passage above.

Part (ii). This follows by using part (i). One may divide in the cases $s > 0$ and $s = 0$.

Part (iii). Assume that $D(f) \subset (0,\infty)$ has positive measure, otherwise we are done (in particular we may assume that $f \not\equiv 0$). Let $E(f) \subset (0,\infty)$ (resp. $E(f_j) \subset (0,\infty)$) be the set of measure zero where $\widetilde{M}f(r)$ (resp. $\widetilde{M}f_j(r)$) is not differentiable. Let us prove the statement for any $r \in D(f) \setminus \big(E(f) \cup \big(\cup_{j=1}^\infty E(f_j)\big)\big)$.

Let $x \in \mathbb{R}^d \setminus \{0\}$ be such that $|x| = r$. Then $\widetilde{M}f$ and all $\big\{\widetilde{M}f_j\big\}_{j\geq 1}$ are differentiable at x. From part (i) we find that $x \in \mathcal{D}(f_j)$ for $j \geq j_0$. Using parts (i) and (ii), and the fact that $\{\|f_j\|_{L^1(\mathbb{R}^d)}\}_{j\geq 1}$ is bounded we find that there exist $\varepsilon > 0$, $N > 0$ and $j_1 \geq j_0$ such that if $\overline{B_{s_j}(z_j)} \in \mathcal{B}(f_j; x)$ for $j \geq j_1$ then $\varepsilon \leq s_j \leq N$. The result now follows from part (ii) and Proposition 4.2.3. ■

[1] Recall that we allow for the possibility $\overline{B_0(x)} = \{x\}$.

4.3 Control near the origin

In this section we develop the first part of our overall strategy of the proof of Theorem 4.1.1, by establishing a control of the convergence near the origin. This is inspired in an argument of [A2].

Proposition 4.3.1 (Control near the origin). *Let $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ and $\{f_j\}_{j\geq 1} \subset W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. Then for every $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon) > 0$ such that*

$$\int_{B_\eta} \big|\nabla \widetilde{M} f\big| < \varepsilon \quad \text{and} \quad \int_{B_\eta} \big|\nabla \widetilde{M} f_j\big| < \varepsilon$$

for all $j \geq j_1(\varepsilon, \eta)$.

Proof If $f = 0$ the result follows directly from (4.7). So let us assume that $f \not\equiv 0$. Recall that we may assume that all our functions are nonnegative. For a generic $g \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ nonnegative we claim that for any $\eta > 0$ and $\ell > 2$ we have

$$\int_{B_\eta} \big|\nabla \widetilde{M} g\big| \lesssim_d \int_{B_{\ell\eta}} |\nabla g| + \frac{1}{\ell^d} \int_{\mathbb{R}^d} |\nabla g| + g(\ell\eta)(\ell\eta)^{d-1}. \tag{4.14}$$

The conclusion of Proposition 4.3.1 plainly follows from this claim by taking ℓ large, η small (with the product $\ell\eta$ still small), and using (4.10) and the fact that $f_j(\ell\eta)$ converges pointwise to $f(\ell\eta)$ given by Proposition 4.2.4 (i).

Let us then prove the claim (4.14). For each $x \in \mathbb{R}^d \setminus \{0\}$ let r_x be the maximal radius of a closed ball in $\mathcal{B}(g;x)$. Define the set

$$\mathcal{A} := \Big\{x \in B_\eta \setminus \{0\} \ : \ r_x \geq \frac{\ell\eta}{4}\Big\}.$$

Using Proposition 4.2.3 we find that

$$\int_{\mathcal{A}} \big|\nabla \widetilde{M} g\big| \lesssim_d \int_{B_\eta} \frac{\big\|\nabla g\big\|_{L^1(\mathbb{R}^d)}}{(\ell\eta)^d} \lesssim_d \frac{\big\|\nabla g\big\|_{L^1(\mathbb{R}^d)}}{\ell^d}. \tag{4.15}$$

We now take care of the integral over $B_\eta \backslash \mathcal{A}$. For every $\beta > 0$ define a function $g_\beta \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ by

$$g_\beta(r) = \begin{cases} g(r) & \text{for} \quad 0 < r < \ell\eta; \\ \frac{-g(\ell\eta)}{\beta} r + \frac{(\ell\eta+\beta)g(\ell\eta)}{\beta} & \text{for} \quad \ell\eta \leq r \leq \ell\eta + \beta; \\ 0 & \text{for} \quad \ell\eta + \beta < r. \end{cases}$$

Assume for a moment that $\ell\eta$ is a point of differentiability of $g(r)$. Then, for β small enough, we have that $g_\beta \leq g$, and hence $\widetilde{M} g_\beta \leq \widetilde{M} g$. If $x \in B_\eta \setminus \mathcal{A}$, then $r_x < \ell\eta/4$ and any ball $\overline{B} \in \mathcal{B}(g;x)$ will be entirely contained in $\overline{B_{\eta+\frac{\ell\eta}{2}}} \subset \overline{B_{\ell\eta}}$. This implies that $\widetilde{M} g(x) \leq \widetilde{M} g_\beta(x)$

for such x, and hence $\widetilde{M}g_\beta = \widetilde{M}g$ in the set $B_\eta \setminus \mathcal{A}$ (note also that this set is open by Proposition 4.2.4 (ii)). Using (4.7) we then find

$$\int_{B_\eta \setminus \mathcal{A}} \big|\nabla \widetilde{M}g\big| = \int_{B_\eta \setminus \mathcal{A}} \big|\nabla \widetilde{M}g_\beta\big| \leq \int_{\mathbb{R}^d} \big|\nabla \widetilde{M}g_\beta\big| \lesssim_d \int_{\mathbb{R}^d} |\nabla g_\beta|$$
$$= \int_{B_{\ell\eta}} \big|\nabla g\big| + \omega_{d-1}\frac{g(\ell\eta)}{\beta}\int_{\ell\eta}^{\ell\eta+\beta} t^{d-1}\,\mathrm{d}t.$$

Sending $\beta \to 0$ we obtain

$$\int_{B_\eta \setminus \mathcal{A}} \big|\nabla \widetilde{M}g\big| \lesssim_d \int_{B_{\ell\eta}} \big|\nabla g\big| + \omega_{d-1}\, g(\ell\eta)\,(\ell\eta)^{d-1}. \tag{4.16}$$

By adding (4.15) and (4.16) we arrive at (4.14). For any fixed $\eta > 0$, the right-hand side of (4.14) is continuous in ℓ, and hence the inequality holds also if $\ell\eta$ is not a point of differentiability of $g(r)$. ■

4.4 The sunrise construction

The purpose of this section is to present a decomposition that will play the role of (4.1) in our multidimensional radial case, and understand its basic properties.

4.4.1 Definition

Let $f \in W^{1,1}_{\mathrm{rad}}(\mathbb{R}^d)$ be continuous in $\mathbb{R}^d \setminus \{0\}$ and nonnegative. From (4.9) we henceforth denote $f(+\infty) = \widetilde{M}f(+\infty) := 0$. For technical reasons that will become clearer later (e.g. see Proposition 4.4.2 below), it will be convenient to avoid a neighborhood of the origin in our discussion, and we let $\rho > 0$ be a fixed parameter throughout this section. *It should be clear from the start that all the new constructions in this section depend on such parameter $\rho > 0$,* and we shall excuse ourselves from an explicit mention to it in some of the passages and definitions below in order to simplify the notation.

We start by decomposing the open set $D(f) \cap (\rho, \infty)$ into a countable union of open intervals

$$D(f) \cap (\rho,\infty) = \bigcup_{i=1}^{\infty} \big(a_i(f;\rho), b_i(f;\rho)\big). \tag{4.17}$$

When the dependence on f and ρ is clear, we simply write (a_i, b_i) instead of $\big(a_i(f;\rho), b_i(f;\rho)\big)$. Let (a_i, b_i) be a generic interval of this decomposition. Proposition 4.2.2 guarantees the existence of $\tau_i^- = \tau_i^-(f;\rho)$ and $\tau_i^+ = \tau_i^+(f;\rho)$ such that $a_i \leq \tau_i^- \leq \tau_i^+ \leq b_i$ and

$$[\tau_i^-, \tau_i^+] = \big\{ r \in [a_i, b_i] \ : \ \widetilde{M}f(r) = \min\big\{ \widetilde{M}f(s)\ ;\ s \in [a_i, b_i]\big\}\big\}.$$

That is, $[\tau_i^-, \tau_i^+]$ is the interval of points of minima of $\widetilde{M}f$ in $[a_i, b_i]$. Note that possibilities like $\tau_i^- = \tau_i^+$, $\tau_i^- = a_i = \rho$ or $\tau_i^+ = b_i = +\infty$ are all duly accounted for. From Proposition 4.2.2 we know that $\widetilde{M}f(r)$ is non-increasing in $[a_i, \tau_i^-]$ and non-decreasing in $[\tau_i^+, b_i]$.

Inspired by the classical construction of the sunrise lemma in harmonic analysis we now consider the following functions. For $r \in (a_i, \tau_i^-)$ (this interval may be empty) define

$$W_R^i f(r) = \max\left\{ \max_{r \leq t \leq \tau_i^-} f(t)\ ,\ \widetilde{M}f(\tau_i^-) \right\}, \tag{4.18}$$

and for $r \in (\tau_i^+, b_i)$ (this interval may be empty) define

$$W_L^i f(r) = \max\left\{ \max_{\tau_i^+ \leq t \leq r} f(t)\ ,\ \widetilde{M}f(\tau_i^+) \right\}.$$

We are now in position to define our analogues of the lateral maximal functions in (4.1). For each $r \in (\rho, \infty)$ we define the functions $\widetilde{M}_R f = \widetilde{M}_R(f;\rho)$ and $\widetilde{M}_L f = \widetilde{M}_L(f;\rho)$ at the point r by

$$\widetilde{M}_R f(r) = \begin{cases} \widetilde{M}f(r) & \text{if} \quad r \in C(f) \text{ or } r \in [\tau_i^-, b_i) \text{ for some } i \geq 1. \\ W_R^i f(r) & \text{if} \quad r \in (a_i, \tau_i^-) \text{ for some } i \geq 1; \end{cases}$$

and

$$\widetilde{M}_L f(r) = \begin{cases} \widetilde{M}f(r) & \text{if} \quad r \in C(f) \text{ or } r \in (a_i, \tau_i^+] \text{ for some } i \geq 1; \\ W_L^i f(r) & \text{if} \quad r \in (\tau_i^+, b_i) \text{ for some } i \geq 1. \end{cases}$$

Remark: Note that we are not defining these functions in the interval $(0, \rho]$.

Before moving on to discuss the basic properties of these new functions, let us point out two important facts. First, in dimension $d = 1$ it is not necessarily true that $\widetilde{M}_R f = M_R f$ and $\widetilde{M}_L f = M_L f$ in the interval (ρ, ∞), where M_R and M_L are the classical one-sided maximal operators, to the right and left, respectively (consider, for instance, f being two sharp bumps to the right of ρ). Second, note that $\widetilde{M}_R f$ and $\widetilde{M}_L f$ are generated from f indirectly, i.e. passing through $\widetilde{M}f$, and it is not in principle true that the operators $f \mapsto \widetilde{M}_R f$ and $f \mapsto \widetilde{M}_L f$ are sublinear. This is a source of technical difficulty in the proof, especially in the upcoming Proposition 4.4.2, that will be carefully handled.

4.4.2 Basic properties

From the definition, for all $r \in (\rho, \infty)$ one plainly sees that

$$f(r) \leq \widetilde{M}_R f(r) \leq \widetilde{M}f(r) \ \text{ and } \ f(r) \leq \widetilde{M}_L f(r) \leq \widetilde{M}f(r), \tag{4.19}$$

and

$$\widetilde{M}f(r) = \max\left\{ \widetilde{M}_R f(r)\, ,\ \widetilde{M}_L f(r) \right\}. \tag{4.20}$$

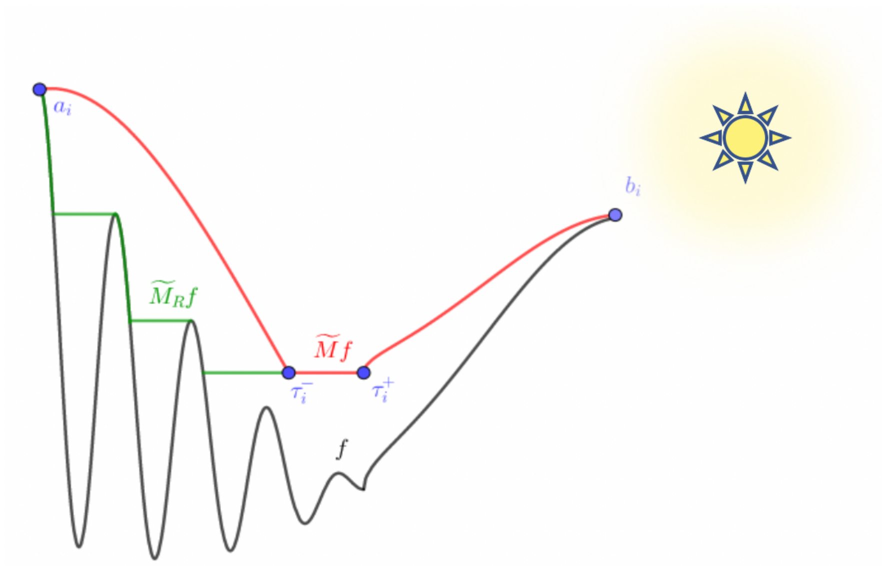

Figure 4.1: The sunrise lateral maximal function $\widetilde{M}_R f$ in a disconnecting interval (a_i, b_i).

Also, for any $\rho < r < s < \infty$, one can show that

$$\left| \widetilde{M}_R f(r) - \widetilde{M}_R f(s) \right| \leq \int_r^s |f'(t)| \, \mathrm{d}t + \int_r^s \left| (\widetilde{M} f)'(t) \right| \, \mathrm{d}t \, ,$$

and the same holds for $\widetilde{M}_L f$. For this one may consider the different cases when r and s belong to $C(f)$ or $D(f)$. This plainly implies that $\widetilde{M}_R f$ and $\widetilde{M}_L f$ are absolutely continuous in (ρ, ∞). In particular, $\widetilde{M}_R f$ and $\widetilde{M}_L f$ are differentiable a.e. in (ρ, ∞).

As before, let us define the disconnecting set $D_R(f) = D_R(f;\rho)$ and the connecting set $C_R(f) = C_R(f;\rho)$ by

$$D_R(f) = \left\{ r \in (\rho, \infty) \, : \, \widetilde{M}_R f(r) > f(r) \right\} \text{ and } C_R(f) = \left\{ r \in (\rho, \infty) \, : \, \widetilde{M}_R f(r) = f(r) \right\}, \tag{4.21}$$

and, analogously, we define $D_L(f) = D_L(f;\rho)$ and $C_L(f) = C_L(f;\rho)$ by

$$D_L(f) = \left\{ r \in (\rho, \infty) \, : \, \widetilde{M}_L f(r) > f(r) \right\} \text{ and } C_L(f) = \left\{ r \in (\rho, \infty) \, : \, \widetilde{M}_L(f)(r) = f(r) \right\}.$$

We now prove a fundamental property of our construction.

Proposition 4.4.1 (Monotonicity)**.** *The following monotonicity properties hold:*

$$\big(\widetilde{M}_R f\big)'(r) \geq 0 \;\text{ a.e. in } D_R(f) \quad \text{and} \quad \big(\widetilde{M}_R f\big)'(r) \leq 0 \;\text{ a.e. in } C_R(f),$$

and

$$\big(\widetilde{M}_L f\big)'(r) \leq 0 \;\text{ a.e. in } D_L(f) \quad \text{and} \quad \big(\widetilde{M}_L f\big)'(r) \geq 0 \;\text{ a.e. in } C_L(f).$$

Proof We consider $\widetilde{M}_R f$. The proof for $\widetilde{M}_L f$ is essentially analogous. Let us consider the disjoint decomposition

$$D(f) \cap (\rho, \infty) = D^-(f) \cup D^0(f) \cup D^+(f)\,, \tag{4.22}$$

where $D^-(f) = D^-(f;\rho)$, $D^0(f) = D^0(f;\rho)$ and $D^+(f) = D^+(f;\rho)$ are defined by

$$D^-(f) = \bigcup_{i=1}^{\infty}(a_i, \tau_i^-)\ ;\ D^0(f) = \bigcup_{i=1}^{\infty}\Big([\tau_i^-, \tau_i^+] \cap D(f) \cap (\rho,\infty)\Big) \ \text{ and } \ D^+(f) = \bigcup_{i=1}^{\infty}(\tau_i^+, b_i). \tag{4.23}$$

Note that $\big(D^0(f) \cup D^+(f)\big) \subset D_R(f) \subset D(f)$ and hence

$$D_R(f) = D^0(f) \cup D^+(f) \cup \big(D_R(f) \cap D^-(f)\big). \tag{4.24}$$

Also,

$$C_R(f) = \big(C(f) \cap (\rho,\infty)\big) \cup \big(C_R(f) \cap D^-(f)\big).$$

We claim that the derivative of $\widetilde{M}_R f$ in (ρ,∞) is given by

$$\big(\widetilde{M}_R f\big)'(r) = \begin{cases} \big(\widetilde{M} f\big)'(r) \geq 0 & \text{for a.e. } r \in D^+(f); \\ \big(\widetilde{M} f\big)'(r) = 0 & \text{for a.e. } r \in D^0(f); \\ 0 & \text{for all } r \in D_R(f) \cap D^-(f); \\ f'(r) = 0 & \text{for a.e. } r \in C(f) \cap (\rho,\infty); \\ f'(r) \leq 0 & \text{for a.e. } r \in C_R(f) \cap D^-(f). \end{cases} \tag{4.25}$$

Let us look at the disconnecting set first. Since $\widetilde{M}_R f(r) = \widetilde{M} f(r)$ is non-decreasing in each (τ_i^+, b_i), we find that $\big(\widetilde{M}_R f\big)'(r) = \big(\widetilde{M} f\big)'(r) \geq 0$ a.e. in $D^+(f)$. In each point $r \in (\tau_i^-, \tau_i^+)$ (if this set is non-empty) we have $\widetilde{M}_R f = \widetilde{M} f$ being constant in a neighborhood of r, and hence $\big(\widetilde{M} f\big)'(r) = 0$. If $r \in D_R(f) \cap D^-(f)$, then $\widetilde{M}_R f(r)$ is also constant in a neighborhood of r, and we have $\big(\widetilde{M}_R f\big)'(r) = 0$.

As for the connecting set, if $r \in C(f) \cap (\rho,\infty)$ is a point of differentiability of $\widetilde{M}_R f$, $\widetilde{M} f$ and f, and is not an isolated point of $C(f) \cap (\rho,\infty)$ (note that this is still a.e. in $C(f) \cap (\rho,\infty)$), we observe that $\big(\widetilde{M}_R f\big)'(r) = \big(\widetilde{M} f\big)'(r) = f'(r) = 0$; see the discussion in §4.2.3. We are left with analyzing $C_R(f) \cap D^-(f)$. Note that $W_R^i f$ is non-increasing in $\big(a_i, \tau_i^-\big)$, which means that $\big(\widetilde{M}_R f\big)'(r) = \big(W_R^i f\big)'(r) \leq 0$ a.e. in (a_i, τ_i^-) for each $i \geq 1$, and hence for a.e. $r \in C_R(f) \cap D^-(f)$. Then, if $r \in C_R(f) \cap D^-(f)$ is a point of differentiability of $\widetilde{M}_R f$ and f, and is not an isolated point of $C_R(f) \cap D^-(f)$ (which is still a.e. in $C_R(f) \cap D^-(f)$) we have $\big(\widetilde{M}_R f\big)'(r) = f'(r) \leq 0$. ■

Remark: From the description (4.25) note that $\big(\widetilde{M}_R f\big)' \in L^1((\rho,\infty), r^{d-1}\,\mathrm{d}r)$, and so does $\big(\widetilde{M}_L f\big)'$.

4.4.3 Pointwise convergence

We now move to a crucial and delicate result in our strategy, the analogue of Proposition 4.2.4 for the lateral operators $\widetilde{M}_R$ and $\widetilde{M}_L$. Note how the use of the sublinearity of $\widetilde{M}$ allows for a relatively simple proof of Proposition 4.2.4 (i). Unfortunately, sublinearity is a tool we do not possess here, and we must handle the situation differently. Our approach will be more of a *tour-de-force* one, in which we carefully study the many different building blocks and possibilities of the sunrise construction. We will split the content now into two propositions, as the proofs will be more elaborate. Recall that we assume that all functions considered here are nonnegative.

Proposition 4.4.2 (Pointwise convergence for $\widetilde{M}_R$ and $\widetilde{M}_L$)**.** *Let $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ and $\{f_j\}_{j\geq 1} \subset W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. Then, for each $r \in (\rho, \infty)$, we have $\widetilde{M}_R f_j(r) \to \widetilde{M}_R f(r)$ and $\widetilde{M}_L f_j(r) \to \widetilde{M}_L f(r)$ as $j \to \infty$.*

Proof Let us prove the statement for $\widetilde{M}_R$. The proof for $\widetilde{M}_L$ is essentially analogous. Recall the decomposition given by (4.22) - (4.23). Given $\varepsilon > 0$, from Proposition 4.2.4 (i) there exists $j_0 = j_0(\varepsilon)$ such that

$$|f_j(t) - f(t)| \leq \varepsilon \quad \text{and} \quad \big|\widetilde{M} f_j(t) - \widetilde{M} f(t)\big| \leq \varepsilon \tag{4.26}$$

for all $j \geq j_0$ and all $t \in (\rho, \infty)$. Any mention of $j_0(\varepsilon)$ below refers to this uniform convergence. We divide our analysis into the following exhaustive list of cases.

Case 1: $r \in C(f)$. In this case $\widetilde{M}_R f(r) = \widetilde{M} f(r) = f(r)$. From Proposition 4.2.4 (i) we know that $\widetilde{M} f_j(r) \to \widetilde{M} f(r)$ and that $f_j(r) \to f(r)$ as $j \to \infty$. The desired result follows from (4.19).

Case 2: $r \in D^+(f)$. In this case $r \in (\tau_i^+, b_i)$ for some $i \geq 1$, and we know that $\widetilde{M}_R f(r) = \widetilde{M} f(r) > \max\big\{f(r),\, \widetilde{M} f\big(\tau_i^+\big)\big\}$. Let s be such that $\tau_i^+ < s < r$ and $\widetilde{M} f(s) < \widetilde{M} f(r)$. Then $[s, r] \subset D^+(f)$ and by Proposition 4.2.4 (i) we have that $[s, r] \subset D(f_j)$ and $\widetilde{M} f_j(s) < \widetilde{M} f_j(r)$ for $j \geq j_1$. This plainly implies that $r \in D^+(f_j)$ and hence $\widetilde{M}_R f_j(r) = \widetilde{M} f_j(r)$ for $j \geq j_1$. The result follows from another application of Proposition 4.2.4 (i).

Case 3: $r \in D^0(f)$. In this case $r \in [\tau_i^-, \tau_i^+] \cap D(f) \cap (\rho, \infty)$ for some $i \geq 1$ and we have $\widetilde{M}_R f(r) = \widetilde{M} f(r) > f(r)$. In particular note that $f \not\equiv 0$. Hence, we cannot have $b_i = +\infty$ since this would plainly imply $\tau_i^- = \tau_i^+ = b_i = +\infty$, contradicting our situation. We then have two subcases to consider:

Subcase 3.1: $\tau_i^+ < b_i < +\infty$. Given $\varepsilon > 0$ sufficiently small, let $\tau_i^+ < u < b_i$ be such that

$$\widetilde{M} f(r) \geq \max\big\{f(t) \;:\; t \in [r, u]\big\} + 3\varepsilon.$$

Then $[r, u] \subset D(f)$ and by Proposition 4.2.4 (i) we have that $[r, u] \subset D(f_j)$ and $\widetilde{M} f_j(u) > \widetilde{M} f_j(r)$ for $j \geq j_1 \geq j_0(\varepsilon)$. We now observe two possibilities for each $j \geq j_1$:

(i) if $r \in D^0(f_j) \cup D^+(f_j)$ we have that $\widetilde{M}_R f_j(r) = \widetilde{M} f_j(r)$ and hence

$$\big| \widetilde{M}_R f_j(r) - \widetilde{M}_R f(r) \big| = \big| \widetilde{M} f_j(r) - \widetilde{M} f(r) \big| \leq \varepsilon.$$

(ii) if $r \in D^-(f_j)$ then, by the considerations above, the corresponding left minimum $\tau^-_{i_j}(f_j) = \tau^-_{i_j}(f_j;\rho)$ of $\widetilde{M} f_j$ in the disconnecting open interval of $D(f_j) \cap (\rho, \infty)$ that contains r is such that $r < \tau^-_{i_j}(f_j) < u$ and $\widetilde{M}_R f_j(r) = W_R^{i_j} f_j(r) = \widetilde{M} f_j(\tau^-_{i_j}(f_j))$. In this situation we have

$$\widetilde{M} f(r) + \varepsilon \geq \widetilde{M} f_j(r) \geq \widetilde{M} f_j(\tau^-_{i_j}(f_j)) \geq \widetilde{M} f(\tau^-_{i_j}(f_j)) - \varepsilon \geq \widetilde{M} f(r) - \varepsilon,$$

which implies that

$$\big| \widetilde{M}_R f_j(r) - \widetilde{M}_R f(r) \big| = \big| \widetilde{M} f_j(\tau^-_{i_j}(f_j)) - \widetilde{M} f(r) \big| \leq \varepsilon.$$

Subcase 3.2: $\tau_i^+ = b_i < +\infty$. Let $\varepsilon > 0$ be given. From Proposition 4.2.4 (i) we have that $r \in D(f_j)$ for $j \geq j_1 \geq j_0(\varepsilon)$. We now observe three possibilities for each $j \geq j_1$:

(i) if $r \in D^0(f_j) \cup D^+(f_j)$ we have that $\widetilde{M}_R f_j(r) = \widetilde{M} f_j(r)$ and hence

$$\big| \widetilde{M}_R f_j(r) - \widetilde{M}_R f(r) \big| = \big| \widetilde{M} f_j(r) - \widetilde{M} f(r) \big| \leq \varepsilon.$$

(ii) if $r \in D^-(f_j)$ and the corresponding left minimum $\tau^-_{i_j}(f_j) = \tau^-_{i_j}(f_j;\rho)$ is such that $r < \tau^-_{i_j}(f_j) \leq b_i$ we have

$$\widetilde{M} f(r) + \varepsilon \geq \widetilde{M} f_j(r) \geq \widetilde{M}_R f_j(r) \geq \widetilde{M} f_j(\tau^-_{i_j}(f_j)) \geq \widetilde{M} f(\tau^-_{i_j}(f_j)) - \varepsilon = \widetilde{M} f(r) - \varepsilon,$$

from which we conclude that

$$\big| \widetilde{M}_R f_j(r) - \widetilde{M}_R f(r) \big| = \big| \widetilde{M}_R f_j(r) - \widetilde{M} f(r) \big| \leq \varepsilon.$$

(iii) if $r \in D^-(f_j)$ and the corresponding left minimum $\tau^-_{i_j}(f_j) = \tau^-_{i_j}(f_j;\rho)$ is such that $b_i < \tau^-_{i_j}(f_j)$ we have (recall that $\widetilde{M} f(b_i) = f(b_i)$ in this situation)

$$\widetilde{M} f(r) + \varepsilon \geq \widetilde{M} f_j(r) \geq \widetilde{M}_R f_j(r) \geq f_j(b_i) \geq f(b_i) - \varepsilon = \widetilde{M} f(b_i) - \varepsilon = \widetilde{M} f(r) - \varepsilon,$$

and again we conclude that

$$\big| \widetilde{M}_R f_j(r) - \widetilde{M}_R f(r) \big| = \big| \widetilde{M}_R f_j(r) - \widetilde{M} f(r) \big| \leq \varepsilon.$$

Case 4: $r \in D^-(f)$. In this case $r \in (a_i, \tau_i^-)$ for some $i \geq 1$ and we have $\widetilde{M}f(r) > \widetilde{M}_R f(r) = W_R^i f(r)$ defined in (4.18). In particular $f \not\equiv 0$. We consider the following subcases:

Subcase 4.1: $W_R^i f(r) = \max_{r \leq t \leq \tau_i^-} f(t) \geq \widetilde{M}f(\tau_i^-)$. Given $\varepsilon > 0$ sufficiently small, let s be such that $r \leq s < \tau_i^-$ and $0 \leq W_R^i f(r) - f(s) \leq \varepsilon$. Let u and v be such that $r \leq s < u < v < \tau_i^-$ and

$$\min\{\widetilde{M}f(r), \widetilde{M}f(\tau_i^-) + \varepsilon\} > \widetilde{M}f(u) > \widetilde{M}f(v) > \widetilde{M}f(\tau_i^-).$$

Then $[r, v] \subset D^-(f)$ and by Proposition 4.2.4 (i) we have that $[r, v] \subset D(f_j)$ and $\widetilde{M}f_j(r) > \widetilde{M}f_j(u) > \widetilde{M}f_j(v)$ for $j \geq j_1 \geq j_0(\varepsilon)$. This implies that $r \in D^-(f_j)$ for $j \geq j_1$, and we again let $\tau_{i_j}^-(f_j) = \tau_{i_j}^-(f_j; \rho)$ be the corresponding left minimum. Observe that $u < \tau_{i_j}^-(f_j)$. From this we get

$$\widetilde{M}f_j(\tau_{i_j}^-(f_j)) < \widetilde{M}f_j(u) \leq \widetilde{M}f(u) + \varepsilon \leq \widetilde{M}f(\tau_i^-) + 2\varepsilon \leq W_R^i f(r) + 2\varepsilon, \tag{4.27}$$

and using (4.27) we also get

$$\begin{aligned} \max_{r \leq t \leq \tau_{i_j}^-(f_j)} f_j(t) &\leq \max\left\{\max_{r \leq t \leq u} f_j(t)\,,\ \max_{u \leq t \leq \tau_{i_j}^-(f_j)} \widetilde{M}f_j(t)\right\} \\ &\leq \max\left\{\max_{r \leq t \leq u} f(t) + \varepsilon\,,\ \widetilde{M}f_j(u)\right\} \leq W_R^i f(r) + 2\varepsilon. \end{aligned} \tag{4.28}$$

From (4.27) and (4.28) we have, for $j \geq j_1$,

$$W_R^i f(r) + 2\varepsilon \geq W_R^{i_j} f_j(r) \geq f_j(s) \geq f(s) - \varepsilon \geq W_R^i f(r) - 2\varepsilon\,,$$

which implies

$$\left|\widetilde{M}_R f_j(r) - \widetilde{M}_R f(r)\right| = \left|W_R^{i_j} f_j(r) - W_R^i f(r)\right| \leq 2\varepsilon.$$

Subcase 4.2: $W_R^i f(r) = \widetilde{M}f(\tau_i^-) > \max_{r \leq t \leq \tau_i^-} f(t)$. Note that $b_i < +\infty$, otherwise we would have $\tau_i^- = \tau_i^+ = b_i = +\infty$, contradicting our situation. We analyze here the two possibilities:

§4.2.1: $\tau_i^+ < b_i < +\infty$. Given $\varepsilon > 0$ sufficiently small, let u be such that $r < \tau_i^- \leq \tau_i^+ < u < b_i$ and

$$\widetilde{M}f(\tau_i^-) \geq \max\left\{f(t)\ :\ t \in [r, u]\right\} + 3\varepsilon.$$

Then $[r, u] \subset D(f)$ and by Proposition 4.2.4 (i) we have that $[r, u] \subset D(f_j)$, $\widetilde{M}f_j(r) > \widetilde{M}f_j(\tau_i^-)$ and $\widetilde{M}f_j(\tau_i^-) < \widetilde{M}f_j(u)$ for $j \geq j_1 \geq j_0(\varepsilon)$. This implies that $r \in D^-(f_j)$ for $j \geq j_1$ and the corresponding left minimum $\tau_{i_j}^-(f_j) = \tau_{i_j}^-(f_j; \rho)$ is such that $r < \tau_{i_j}^-(f_j) < u$. In this scenario, note that $\widetilde{M}_R f_j(r) = W_R^{i_j} f_j(r) = \widetilde{M}f_j(\tau_{i_j}^-(f_j))$ and, for $j \geq j_1$,

$$\widetilde{M}f(\tau_i^-) + \varepsilon \geq \widetilde{M}f_j(\tau_i^-) \geq \widetilde{M}f_j(\tau_{i_j}^-(f_j)) \geq \widetilde{M}f(\tau_{i_j}^-(f_j)) - \varepsilon \geq \widetilde{M}f(\tau_i^-) - \varepsilon,$$

which implies

$$\left|\widetilde{M}_R f_j(r) - \widetilde{M}_R f(r)\right| = \left|\widetilde{M} f_j(\tau_{i_j}^-(f_j)) - \widetilde{M} f(\tau_i^-)\right| \leq \varepsilon.$$

§4.2.2: $\tau_i^+ = b_i < +\infty$. Given $\varepsilon > 0$ sufficiently small, let u and v be such that $r < u < v < \tau_i^-$ and

$$\min\{\widetilde{M} f(r), \widetilde{M} f(\tau_i^-) + \varepsilon\} > \widetilde{M} f(u) > \widetilde{M} f(v) > \widetilde{M} f(\tau_i^-).$$

Then $[r, v] \subset D^-(f)$ and by Proposition 4.2.4 (i) we have that $[r, v] \subset D(f_j)$ and $\widetilde{M} f_j(r) > \widetilde{M} f_j(u) > \widetilde{M} f_j(v)$ for $j \geq j_1 \geq j_0(\varepsilon)$. This implies that $r \in D^-(f_j)$ for $j \geq j_1$, and we again let $\tau_{i_j}^-(f_j) = \tau_{i_j}^-(f_j; \rho)$ be the corresponding left minimum. Observe that $u < \tau_{i_j}^-(f_j)$ and hence

$$\widetilde{M} f_j(\tau_{i_j}^-(f_j)) < \widetilde{M} f_j(u) \leq \widetilde{M} f(u) + \varepsilon \leq \widetilde{M} f(\tau_i^-) + 2\varepsilon. \tag{4.29}$$

Using (4.29) we get

$$\begin{aligned} \max_{r \leq t \leq \tau_{i_j}^-(f_j)} f_j(t) &\leq \max\left\{\max_{r \leq t \leq u} f_j(t)\,,\, \max_{u \leq t \leq \tau_{i_j}^-(f_j)} \widetilde{M} f_j(t)\right\} \\ &\leq \max\left\{\widetilde{M} f(\tau_i^-) + \varepsilon\,,\, \widetilde{M} f_j(u)\right\} \leq \widetilde{M} f(\tau_i^-) + 2\varepsilon. \end{aligned} \tag{4.30}$$

From (4.29) and (4.30) we conclude that

$$W_R^{i_j} f_j(r) \leq \widetilde{M} f(\tau_i^-) + 2\varepsilon. \tag{4.31}$$

For the other inequality we proceed as follows. If $\tau_{i_j}^-(f_j) \leq b_i$ we have

$$W_R^{i_j} f_j(r) \geq \widetilde{M} f_j(\tau_{i_j}^-(f_j)) \geq \widetilde{M} f(\tau_{i_j}^-(f_j)) - \varepsilon \geq \widetilde{M} f(\tau_i^-) - \varepsilon.$$

If $\tau_{i_j}^-(f_j) > b_i$ we have (recall that $\widetilde{M} f(b_i) = f(b_i)$ in this situation)

$$W_R^{i_j} f_j(r) \geq f_j(b_i) \geq f(b_i) - \varepsilon = \widetilde{M} f(b_i) - \varepsilon \geq \widetilde{M} f(\tau_i^-) - \varepsilon.$$

In either case we conclude that

$$W_R^{i_j} f_j(r) \geq \widetilde{M} f(\tau_i^-) - \varepsilon. \tag{4.32}$$

Finally, from (4.31) and (4.32) we reach the desired conclusion

$$\left|\widetilde{M}_R f_j(r) - \widetilde{M}_R f(r)\right| = \left|W_R^{i_j} f_j(r) - \widetilde{M} f(\tau_i^-)\right| \leq 2\varepsilon.$$

This completes the proof. ■

Proposition 4.4.3 (Pointwise convergence for the derivatives of $\widetilde{M}_R$ and $\widetilde{M}_L$)**.** *Let $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ and $\{f_j\}_{j\geq 1} \subset W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ be such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$. Then, for almost all $r \in D_R(f)$, we have $\big(\widetilde{M}_R f_j\big)'(r) \to \big(\widetilde{M}_R f\big)'(r)$ as $j \to \infty$, and for almost all $r \in D_L(f)$ we have $\big(\widetilde{M}_L f_j\big)'(r) \to \big(\widetilde{M}_L f\big)'(r)$ as $j \to \infty$.*

Proof We prove the statement for $\widetilde{M}_R$ as the proof for $\widetilde{M}_L$ is essentially analogous. For each $\varepsilon > 0$ we keep defining $j_0(\varepsilon)$ by (4.26). Recalling decomposition (4.24) we divide again our analysis into cases.

Case 1: $D^+(f)$. Let us consider an interval (τ_i^+, b_i) for some $i \geq 1$. For each $\tau_i^+ < u < v < b_i$ we may choose s with $\tau_i^+ < s < u < v < b_i$ such that $\widetilde{M}f(s) < \widetilde{M}f_j(u)$. Then $[s,v] \subset D^+(f)$ and by Proposition 4.2.4 (i) we have that $[s,v] \subset D(f_j)$ and $\widetilde{M}f_j(s) < \widetilde{M}f_j(u)$ for $j \geq j_1$. This plainly implies that $[u,v] \subset D^+(f_j)$ and hence $\widetilde{M}_R f_j(r) = \widetilde{M}f_j(r)$ for all $r \in [u,v]$ and $j \geq j_1$. The result follows from Proposition 4.2.4 (iii).

Case 2: $D^0(f)$. Since we want to prove the result almost everywhere, it is sufficient to consider only the intervals $[\tau_i^-, \tau_i^+] \cap D(f) \cap (\rho, \infty)$ where $\tau_i^- < \tau_i^+$ (in particular, this implies that $b_i < +\infty$). Let u and v be such that $\tau_i^- < u < v < \tau_i^+$. We consider two subcases:

Subcase 2.1: $\tau_i^+ < b_i < +\infty$. Given $\varepsilon > 0$ sufficiently small, let $\tau_i^+ < s < b_i$ be such that

$$\widetilde{M}f(u) = \widetilde{M}f(v) \geq \max\big\{f(t) \ : \ t \in [u,s]\big\} + 3\varepsilon.$$

From Proposition 4.2.4 (i) we know that $[u,s] \subset D(f_j)$ and $\widetilde{M}f_j(s) > \widetilde{M}f_j(\tau_i^+)$ for $j \geq j_1 \geq j_0(\varepsilon)$. Let $\tau_{i_j}^-(f_j) = \tau_{i_j}^-(f_j;\rho)$ be the corresponding left minimum of $\widetilde{M}f_j$ in the disconnecting open interval of $D(f_j) \cap (\rho,\infty)$ that contains $[u,s]$. Note that $\tau_{i_j}^-(f_j) < s$ and for $r \in [u,v]$ we have

$$\widetilde{M}_R f_j(r) = \begin{cases} \widetilde{M}f_j(r) & \text{if} \quad \tau_{i_j}^-(f_j) \leq r \leq v; \\ W_R^{i_j} f_j(r) = \widetilde{M}f_j(\tau_{i_j}^-(f_j)) & \text{if} \quad u \leq r < \tau_{i_j}^-(f_j). \end{cases}$$

Then, for a.e. $r \in [u,v]$ we have

$$\big(\widetilde{M}_R f_j\big)'(r) = \begin{cases} \big(\widetilde{M}f_j\big)'(r) & \text{if} \quad \tau_{i_j}^-(f_j) \leq r \leq v; \\ 0 & \text{if} \quad u \leq r < \tau_{i_j}^-(f_j). \end{cases} \tag{4.33}$$

We conclude from (4.25) and Proposition 4.2.4 (iii).

Subcase 2.2: $\tau_i^+ = b_i < +\infty$. Let $\varepsilon > 0$ be sufficiently small so that

$$\widetilde{M}f(u) = \widetilde{M}f(v) \geq \max\big\{f(t) \ : \ t \in [u,v]\big\} + 3\varepsilon.$$

From Proposition 4.2.4 (i) we know that $[u,v] \subset D(f_j)$ for $j \geq j_1 \geq j_0(\varepsilon)$, and we again let $\tau_{i_j}^-(f_j) = \tau_{i_j}^-(f_j;\rho)$ be the corresponding left minimum. As before we have, for $r \in [u,v]$,

$$\widetilde{M}_R f_j(r) = \begin{cases} \widetilde{M}f_j(r) & \text{if} \quad \tau_{i_j}^-(f_j) \leq r \leq v; \\ W_R^{i_j} f_j(r) & \text{if} \quad u \leq r < \tau_{i_j}^-(f_j). \end{cases} \tag{4.34}$$

Let us take a closer look at the second possibility in (4.34). Observe that if $u \le r < \tau_{i_j}^-(f_j) \le b_i$ we have

$$W_R^{i_j} f_j(r) \ge \widetilde{M} f_j(\tau_{i_j}^-(f_j)),$$

and if $u \le r < b_i < \tau_{i_j}^-(f_j)$ we have

$$W_R^{i_j} f_j(r) \ge f_j(b_i) \ge f(b_i) - \varepsilon = \widetilde{M} f(u) - \varepsilon.$$

In either case what matters is that

$$W_R^{i_j} f_j(r) = \max\left\{ \max_{r \le t \le \tau_{i_j}^-(f_j)} f_j(t)\ ,\ \ \widetilde{M} f_j(\tau_{i_j}^-(f_j)) \right\} = \max\left\{ \max_{v < t \le \tau_{i_j}^-(f_j)} f_j(t)\ ,\ \ \widetilde{M} f_j(\tau_{i_j}^-(f_j)) \right\},$$

and the expression on the right-hand side is independent of r. Then (4.34) implies (4.33) and we conclude from (4.25) and Proposition 4.2.4 (iii) as before.

Case 3: $D_R(f) \cap D^-(f)$. Let $r \in D_R(f) \cap D^-(f)$. Then $r \in (a_i, \tau_i^-)$ for some $i \ge 1$. Let s be such that $r < s < \tau_i^-$ and $\widetilde{M} f(r) > \widetilde{M} f(s)$. Then $[r, s] \subset D^-(f)$ and by Proposition 4.2.4 (i) we have that $[r, s] \subset D(f_j)$ and $\widetilde{M} f_j(r) > \widetilde{M} f_j(s)$ for $j \ge j_1$. In particular, this implies that $r \in D^-(f_j)$ for $j \ge j_1$. We have already seen in Proposition 4.4.2 that if $r \in D_R(f)$ then $r \in D_R(f_j)$ for $j \ge j_2 \ge j_1$. Hence $r \in D_R(f_j) \cap D^-(f_j)$ for $j \ge j_2$ and we conclude by using (4.25). ■

4.5 The proof

We are now in position to move on to the proof of Theorem 4.1.1.

4.5.1 Setup

Given $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ and $\{f_j\}_{j \ge 1} \subset W^{1,1}_{\text{rad}}(\mathbb{R}^d)$, all nonnegative, and such that $\|f_j - f\|_{W^{1,1}(\mathbb{R}^d)} \to 0$ as $j \to \infty$, we want to show that

$$\left\|\nabla \widetilde{M} f_j - \nabla \widetilde{M} f\right\|_{L^1(\mathbb{R}^d)} \to 0 \quad \text{as} \quad j \to \infty.$$

Given $\varepsilon > 0$, let $\eta > 0$ be given by Proposition 4.3.1. Then

$$\left\|\nabla \widetilde{M} f_j - \nabla \widetilde{M} f\right\|_{L^1(B_\eta)} < 2\varepsilon$$

for $j \ge j_1(\varepsilon, \eta)$. It is then enough to prove that

$$\left\|\nabla \widetilde{M} f_j - \nabla \widetilde{M} f\right\|_{L^1(\mathbb{R}^d \setminus B_\eta)} \to 0 \quad \text{as} \quad j \to \infty,$$

which is equivalent to

$$\big\|\big(\widetilde{M}f_j\big)' - \big(\widetilde{M}f\big)'\big\|_{L^1((\eta,\infty),\, r^{d-1}\,\mathrm{d}r)} \to 0 \quad \text{as} \quad j \to \infty. \tag{4.35}$$

From now on we fix $\rho = \eta/2$ and consider the sunrise construction of the lateral operators $\widetilde{M}_R$ and $\widetilde{M}_L$ in §4.4 with respect to this parameter ρ. We have seen in §4.4.2 that the functions $\widetilde{M}_R f$, $\widetilde{M}_L f$, $\big\{\widetilde{M}_R f_j\big\}_{j\geq 1}, \big\{\widetilde{M}_L f_j\big\}_{j\geq 1}$ are all contained in the space $\mathcal{X}$ of Lemma 4.2.1 and hence, by the same lemma and identity (4.20), in order to prove (4.35) is is sufficient to show that

$$\big\|\big(\widetilde{M}_R f_j\big)' - \big(\widetilde{M}_R f\big)'\big\|_{L^1((\eta,\infty),\, r^{d-1}\,\mathrm{d}r)} \to 0 \quad \text{and} \quad \big\|\big(\widetilde{M}_L f_j\big)' - \big(\widetilde{M}_L f\big)'\big\|_{L^1((\eta,\infty),\, r^{d-1}\,\mathrm{d}r)} \to 0$$

as $j \to \infty$. This is what we are going to do in the remaining of this section. We shall prove it for $\widetilde{M}_R$ and the proof for $\widetilde{M}_L$ is essentially analogous.

4.5.2 Splitting into the connecting and disconnecting sets

Recall definition (4.21). For the rest of the section let us adopt a simple notation by writing

$$D = D_R(f) \cap (\eta,\infty)\ ;\ D_j = D_R(f_j) \cap (\eta,\infty)\ ;\ C = C_R(f) \cap (\eta,\infty)\ ;\ C_j = C_R(f_j) \cap (\eta,\infty).$$

Also in the spirit of easing the notation, we sometimes omit the argument of the functions in the integrals below when the context is clear (e.g. writing f' for $f'(r)$) and sometimes use the "little o" notation for limits (i.e. writing $\lambda_j = o(1)$ when $\lim_{j\to\infty} \lambda_j = 0$). We split our original integral into the following four pieces:

$$\begin{aligned}\int_\eta^\infty \big|\big(\widetilde{M}_R f_j\big)' - \big(\widetilde{M}_R f\big)'\big|\, r^{d-1}\,\mathrm{d}r &= \int_{C\cap C_j} + \int_{D\cap C_j} + \int_{C\cap D_j} + \int_{D\cap D_j} \\ &=: (I)_j + (II)_j + (III)_j + (IV)_j.\end{aligned}$$

Our objective is to show that each of these pieces is $o(1)$ as $j \to \infty$ (note that each of these pieces is nonnegative). In what follows the reader should have in mind all times the description (4.25) for the derivative of $\widetilde{M}_R f$. Two of the integral pieces above are particularly simple to analyze, and we clear them out first.

The term $(I)_j$

By our hypotheses we have

$$(I)_j = \int_{C\cap C_j} \big|\big(\widetilde{M}_R f_j\big)' - \big(\widetilde{M}_R f\big)'\big|\, r^{d-1}\,\mathrm{d}r = \int_{C\cap C_j} |f_j' - f'|\, r^{d-1}\,\mathrm{d}r = o(1).$$

The term $(II)_j$

From Proposition 4.4.2, if $r \in D$ then $r \in D_j$ for j large, and hence $\chi_{D\cap C_j}(r) \to 0$ as $j \to \infty$. Therefore, by our hypotheses and dominated convergence we have

$$\begin{aligned}(II)_j &= \int_{D\cap C_j} \big|\big(\widetilde{M}_R f_j\big)' - \big(\widetilde{M}_R f\big)'\big|\, r^{d-1}\, \mathrm{d}r = \int_\eta^\infty \big|f_j' - \big(\widetilde{M}_R f\big)'\big|\, \chi_{D\cap C_j}(r)\, r^{d-1}\, \mathrm{d}r \\ &\leq \int_\eta^\infty |f_j' - f'|\, \chi_{D\cap C_j}(r)\, r^{d-1}\, \mathrm{d}r + \int_\eta^\infty \Big(|f'| + \big|\big(\widetilde{M}_R f\big)'\big|\Big)\, \chi_{D\cap C_j}(r)\, r^{d-1}\, \mathrm{d}r = o(1).\end{aligned} \tag{4.36}$$

4.5.3 Brezis-Lieb reduction and some useful identities

Using the convergence of the derivatives

The *raison d'être* of Proposition 4.4.3 is to allow for an application of the classical Brezis-Lieb lemma [BL83] to conclude that

$$(II)_j + (IV)_j = \int_D \big|\big(\widetilde{M}_R f_j\big)' - \big(\widetilde{M}_R f\big)'\big|\, r^{d-1}\, \mathrm{d}r \to 0 \tag{4.37}$$

as $j \to \infty$ if and only if

$$\int_D \big|\big(\widetilde{M}_R f_j\big)'\big|\, r^{d-1}\, \mathrm{d}r \to \int_D \big|\big(\widetilde{M}_R f\big)'\big|\, r^{d-1}\, \mathrm{d}r = \int_D \big(\widetilde{M}_R f\big)'\, r^{d-1}\, \mathrm{d}r \tag{4.38}$$

as $j \to \infty$. The equality on the right-hand side of (4.38) is due to Proposition 4.4.1. From Proposition 4.4.3 and Fatou's lemma we already have

$$\int_D \big(\widetilde{M}_R f\big)'\, r^{d-1}\, \mathrm{d}r = \int_D \big|\big(\widetilde{M}_R f\big)'\big|\, r^{d-1}\, \mathrm{d}r \leq \liminf_{j\to\infty} \int_D \big|\big(\widetilde{M}_R f_j\big)'\big|\, r^{d-1}\, \mathrm{d}r. \tag{4.39}$$

Let us decompose the open set $D \subset (\eta, \infty)$ into a disjoint union of open intervals:

$$D = \bigcup_{i=1}^{\infty} (\alpha_i, \beta_i). \tag{4.40}$$

We may have one of the left endpoints in (4.40) being η and, if that is the case, let us agree that $\eta = \alpha_1$. Note that, as in (4.8), we have

$$(d-1)\int_\eta^\infty \widetilde{M}_R f(r)\, r^{d-2}\, \mathrm{d}r \leq \int_\eta^\infty \big|\big(\widetilde{M}_R f\big)'(t)\big|\, t^{d-1}\, \mathrm{d}t < \infty. \tag{4.41}$$

Recall also (4.9). Using integration by parts (and dominated convergence with (4.41) to properly justify the limiting process in the potentially infinite sum) we have

$$\begin{aligned}\int_D \big(\widetilde{M}_R f\big)' r^{d-1}\, \mathrm{d}r &= \sum_{i=1}^{\infty} \int_{\alpha_i}^{\beta_i} \big(\widetilde{M}_R f\big)' r^{d-1}\, \mathrm{d}r \\ &= \sum_{i=1}^{\infty} \left(\Big(\widetilde{M}_R f(\beta_i)\, \beta_i^{d-1} - \widetilde{M}_R f(\alpha_i)\, \alpha_i^{d-1} \Big) - (d-1) \int_{\alpha_i}^{\beta_i} \widetilde{M}_R f \; r^{d-2}\, \mathrm{d}r \right) \\ &= \gamma(f) + \sum_{i=1}^{\infty} \left(\big(f(\beta_i)\, \beta_i^{d-1} - f(\alpha_i)\, \alpha_i^{d-1} \big) - (d-1) \int_{\alpha_i}^{\beta_i} f \; r^{d-2}\, \mathrm{d}r \right) \\ &\qquad + (d-1) \int_D \big(f - \widetilde{M}_R f\big)\, r^{d-2}\, \mathrm{d}r \\ &= \gamma(f) + \int_D f' \; r^{d-1}\, \mathrm{d}r + (d-1) \int_D \big(f - \widetilde{M}_R f\big)\, r^{d-2}\, \mathrm{d}r,\end{aligned} \tag{4.42}$$

where we introduced the term

$$\gamma(f) := \begin{cases} f(\eta)\, \eta^{d-1} - \widetilde{M}_R f(\eta)\, \eta^{d-1} & \text{if } \eta = \alpha_1; \\ 0 & \text{otherwise.} \end{cases} \tag{4.43}$$

Similarly, we may decompose $D_j = \bigcup_{i=1}^{\infty} (\alpha_i^j, \beta_i^j)$, with the agreement that if η is a left endpoint in this decomposition then $\eta = \alpha_1^j$. We define $\gamma(f_j)$ as in (4.43) and proceed as in (4.42) to find

$$\int_{D_j} \big(\widetilde{M}_R f_j\big)' r^{d-1}\, \mathrm{d}r = \gamma(f_j) + \int_{D_j} f_j' \; r^{d-1}\, \mathrm{d}r + (d-1) \int_{D_j} \big(f_j - \widetilde{M}_R f_j\big)\, r^{d-2}\, \mathrm{d}r. \tag{4.44}$$

Combining (4.42) and (4.44) we arrive at the following identity

$$\int_{D_j} \big(\widetilde{M}_R f_j\big)' r^{d-1}\, \mathrm{d}r - \int_{D_j} f_j' \; r^{d-1}\, \mathrm{d}r = \int_D \big(\widetilde{M}_R f\big)' r^{d-1}\, \mathrm{d}r - \int_D f' \; r^{d-1}\, \mathrm{d}r + \lambda_j, \tag{4.45}$$

where

$$\lambda_j := \big(\gamma(f_j) - \gamma(f)\big) + \left(\int_{D_j} \big(f_j - \widetilde{M}_R f_j\big)\, r^{d-2}\, \mathrm{d}r - \int_D \big(f - \widetilde{M}_R f\big)\, r^{d-2}\, \mathrm{d}r \right). \tag{4.46}$$

Smallness of the remainder: analysis of λ_j

We now claim that λ_j defined in (4.46) verifies

$$\lambda_j = o(1). \tag{4.47}$$

Note first that $\gamma(f_j) \to \gamma(f)$ as $j \to \infty$. This is an immediate consequence of the pointwise convergences $f_j(\eta) \to f(\eta)$ and $\widetilde{M}_R f_j(\eta) \to \widetilde{M}_R f(\eta)$ as $j \to \infty$. The second observation is that

$$\int_{D_j} \left(f_j - \widetilde{M}_R f_j\right) r^{d-2} \, \mathrm{d}r \to \int_D \left(f - \widetilde{M}_R f\right) r^{d-2} \, \mathrm{d}r. \tag{4.48}$$

as $j \to \infty$. This requires some work to verify. Start by writing the difference in the following form

$$\begin{aligned} &\int_{D_j} \left(f_j - \widetilde{M}_R f_j\right) r^{d-2} \, \mathrm{d}r - \int_D \left(f - \widetilde{M}_R f\right) r^{d-2} \, \mathrm{d}r \\ &= \int_\eta^\infty \Big(\left(f_j - \widetilde{M}_R f_j\right) - \left(f - \widetilde{M}_R f\right)\Big) \chi_D(r)\, r^{d-2} \, \mathrm{d}r + \int_\eta^\infty \left(f_j - \widetilde{M}_R f_j\right) \chi_{C \cap D_j}(r)\, r^{d-2} \, \mathrm{d}r. \end{aligned} \tag{4.49}$$

Let $N > 0$ be large. Using (4.19) and the sublinearity of $\widetilde{M}$, the portion of each of the two integrals on the right-hand side of (4.49) evaluated from N to ∞ is bounded in absolute value by

$$4 \int_N^\infty \left(\widetilde{M} f + \widetilde{M}(f - f_j)\right) r^{d-2} \, \mathrm{d}r. \tag{4.50}$$

A computation as in (4.8), together with (4.7), shows that (4.50) is bounded by

$$\begin{aligned} &\lesssim_d \int_N^\infty \Big(\big|\big(\widetilde{M} f\big)'(t)\big| + \big|\big(\widetilde{M}(f - f_j)\big)'(t)\big|\Big)\, t^{d-1} \, \mathrm{d}t \\ &\lesssim_d \int_N^\infty \big|\big(\widetilde{M} f\big)'(t)\big|\, t^{d-1} \, \mathrm{d}t + \int_0^\infty |(f - f_j)'(t)|\, t^{d-1} \, \mathrm{d}t, \end{aligned}$$

and by our hypotheses this is small if N is large and j is large. In the interval $[\eta, N]$ all the functions $\widetilde{M} f_j$ are uniformly bounded (by Proposition 4.2.4 (i)). By applying Proposition 4.2.4 (i), Proposition 4.4.2 and dominated convergence, we find that the portion of each of the two integrals on the right-hand side of (4.49) evaluated from η to N converges to zero. This establishes (4.48) and hence (4.47).

Final preparation

We need yet another useful identity to run our upcoming dichotomy scheme. We use Proposition 4.4.1 (multiple times) to remove the absolute values when the quantities inside have a well-defined sign, and identity (4.45) - (4.47) (in the third line below), to get

$$
\begin{aligned}
&\int_D \left|\left(\widetilde{M}_R f_j\right)'\right| r^{d-1}\, \mathrm{d}r = \int_{D\cap D_j} \left|\left(\widetilde{M}_R f_j\right)'\right| r^{d-1}\, \mathrm{d}r + \int_{D\cap C_j} \left|\left(\widetilde{M}_R f_j\right)'\right| r^{d-1}\, \mathrm{d}r + (III)_j - (III)_j \\
&= \int_{D\cap D_j} \left(\widetilde{M}_R f_j\right)' r^{d-1}\, \mathrm{d}r - \int_{D\cap C_j} f_j'\, r^{d-1}\, \mathrm{d}r + \int_{C\cap D_j} \left(\left(\widetilde{M}_R f_j\right)' - f'\right) r^{d-1}\, \mathrm{d}r - (III)_j \\
&= \int_{D_j} \left(\widetilde{M}_R f_j\right)' r^{d-1}\, \mathrm{d}r - \int_{D\cap C_j} f_j'\, r^{d-1}\, \mathrm{d}r - \int_{C\cap D_j} f'\, r^{d-1}\, \mathrm{d}r - (III)_j \\
&= \int_D \left(\widetilde{M}_R f\right)' r^{d-1}\, \mathrm{d}r - \int_D f'\, r^{d-1}\, \mathrm{d}r + \int_{D_j} f_j'\, r^{d-1}\, \mathrm{d}r + o(1) \qquad (4.51) \\
&\qquad\qquad - \int_{D\cap C_j} f_j'\, r^{d-1}\, \mathrm{d}r - \int_{C\cap D_j} f'\, r^{d-1}\, \mathrm{d}r - (III)_j \\
&= \int_D \left(\widetilde{M}_R f\right)' r^{d-1}\, \mathrm{d}r + \int_{D_j} (f_j' - f')\, r^{d-1}\, \mathrm{d}r - \int_{D\cap C_j} (f_j' + f')\, r^{d-1}\, \mathrm{d}r - (III)_j + o(1) \\
&= \int_D \left(\widetilde{M}_R f\right)' r^{d-1}\, \mathrm{d}r - (III)_j + o(1).
\end{aligned}
$$

Note that in the last passage above we used the fact that $\chi_{D\cap C_j}(r) \to 0$ and dominated convergence as in (4.36).

4.5.4 Finale: the dichotomy

Let us take a closer look at identity (4.45). For each $j \geq 1$ we have the following dichotomy: either

$$\int_{C\cap D_j} \left(\widetilde{M}_R f_j\right)' r^{d-1}\, \mathrm{d}r \leq \int_{C\cap D_j} f_j'\, r^{d-1}\, \mathrm{d}r \tag{4.52}$$

or

$$\int_{D\cap D_j} \left(\widetilde{M}_R f_j\right)' r^{d-1}\, \mathrm{d}r \leq \int_{D\cap D_j} f_j'\, r^{d-1}\, \mathrm{d}r + \int_D \left(\widetilde{M}_R f\right)' r^{d-1}\, \mathrm{d}r - \int_D f'\, r^{d-1}\, \mathrm{d}r + \lambda_j. \tag{4.53}$$

Case 1

Assume that we go over the subsequence of j's such that (4.52) holds. Using Proposition 4.4.1 to remove the absolute value in the first equality below, and (4.52), we get

$$(III)_j = \int_{C\cap D_j} \left(\left(\widetilde{M}_R f_j\right)' - f'\right) r^{d-1}\, \mathrm{d}r \leq \int_{C\cap D_j} \left(f_j' - f'\right) r^{d-1}\, \mathrm{d}r = o(1).$$

Then, from (4.37), (4.38) and (4.51) we find that

$$(II)_j + (IV)_j = o(1).$$

Then $(IV)_j = o(1)$ and the proof is complete in this case.

Case 2

Assume now that we go over the subsequence of j's such that (4.53) holds. Using Proposition 4.4.1, (4.47) and (4.53), we get

$$\begin{aligned}\int_D \big|\big(\widetilde{M}_R f_j\big)'\big|\, r^{d-1}\,\mathrm{d}r &= \int_{D\cap D_j} \big(\widetilde{M}_R f_j\big)'\, r^{d-1}\,\mathrm{d}r - \int_{D\cap C_j} f_j'\, r^{d-1}\,\mathrm{d}r \\ &\le \int_{D\cap D_j} f_j'\, r^{d-1}\,\mathrm{d}r + \int_D \big(\widetilde{M}_R f\big)'\, r^{d-1}\,\mathrm{d}r - \int_D f'\, r^{d-1}\,\mathrm{d}r - \int_{D\cap C_j} f_j'\, r^{d-1}\,\mathrm{d}r + o(1) \\ &= \int_D \big(\widetilde{M}_R f\big)'\, r^{d-1}\,\mathrm{d}r + \int_{D\cap D_j} \big(f_j' - f'\big)\, r^{d-1}\,\mathrm{d}r - \int_{D\cap C_j} \big(f_j' + f'\big)\, r^{d-1}\,\mathrm{d}r + o(1) \quad (4.54) \\ &= \int_D \big(\widetilde{M}_R f\big)'\, r^{d-1}\,\mathrm{d}r + o(1).\end{aligned}$$

Note in the last passage the use of $\chi_{D\cap C_j}(r) \to 0$ and dominated convergence as in (4.36). It follows from (4.54) that, along our subsequence of j's,

$$\limsup_{j\to\infty} \int_D \big|\big(\widetilde{M}_R f_j\big)'\big|\, r^{d-1}\,\mathrm{d}r \le \int_D \big|\big(\widetilde{M}_R f\big)'\big|\, r^{d-1}\,\mathrm{d}r. \qquad (4.55)$$

From (4.39) and (4.55) we arrive at (4.38), and hence at (4.37). That is,

$$(II)_j + (IV)_j = o(1).$$

Then $(IV)_j = o(1)$, and from (4.38) and (4.51) we find that $(III)_j = o(1)$ along this subsequence. This completes the proof.

4.6 Sunrise strategy reviewed: the core abstract elements

A posteriori, let us take a moment to reflect on some of the main ingredients of our sunrise strategy in general terms. It should be clear by now that it is a one-dimensional mechanism, but part of its power relies on the fact that it can be applied to multidimensional maximal operators, when these act of subspaces of $W^{1,1}$ that can be identified with one-dimensional spaces.

Assume that we are working on a space $W^{1,1}(I,\,\mathrm{d}\mu)$, where $I \subset \mathbb{R}$ is an open interval or $I = \mathbb{S}^1$, and μ is a nonnegative measure on I such that μ and the Lebesgue measure (or arclength measure in the case of $\mathbb{S}^1$) are mutually absolutely continuous. It will be also convenient to assume that the Radon-Nikodym derivative $\frac{\mathrm{d}\mu}{\mathrm{d}x}$ is an absolutely continuous function on I. The cases we have in mind are: $(I,\,\mathrm{d}\mu) = (\mathbb{R},\,\mathrm{d}x)$; $\big((0,\infty), r^{d-1}\,\mathrm{d}r\big)$ for $d \ge 2$; $\big(\mathbb{S}^1,\,\mathrm{d}\theta\big)$; and $\big((0,\pi), (\sin\theta)^{d-1}\,\mathrm{d}\theta\big)$ for $d \ge 2$. The second option, as we have seen, appears associated to the subspace $W^{1,1}_{\mathrm{rad}}(\mathbb{R}^d)$ while the fourth option is associated to the subspace $W^{1,1}_{\mathrm{pol}}(\mathbb{S}^d)$.

For $f \in W^{1,1}(I,\, \mathrm{d}\mu)$, that we assume nonnegative and absolutely continuous in compact subsets of I, we let $\mathfrak{M}$ be a maximal operator acting on f such that $\mathfrak{M}f$ is a continuous function defined on I. We make the additional assumption that $\mathfrak{M}f$ is weakly differentiable and verifies the a priori bound

$$\left\|(\mathfrak{M}f)'\right\|_{L^1(I,\,\mathrm{d}\mu)} \lesssim_{I,\mu} \|f\|_{W^{1,1}(I,\,\mathrm{d}\mu)}. \tag{4.56}$$

In particular, by (4.56), $\mathfrak{M}f$ is also absolutely continuous in compact subsets of I, and hence differentiable a.e. in I.

The sunrise strategy aims to establish the continuity of the map $f \mapsto (\mathfrak{M}f)'$, from $W^{1,1}(I,\, \mathrm{d}\mu)$ to $L^1(I,\, \mathrm{d}\mu)$. Assume that $f_j \to f$ in $W^{1,1}(I,\, \mathrm{d}\mu)$ as $j \to \infty$ (all f_j's nonnegative and absolutely continuous in compact subsets of I). As we have seen in the proof of Theorem 4.1.1, the following five properties are the core elements that make the method work:

(P1) *Absence of local maxima in the disconnecting set*: $\mathfrak{M}f$ does not have strict local maxima in the set $\{\mathfrak{M}f > f\}$ (analogue of Proposition 4.2.2).

(P2) *Convergence properties*: we have $f_j \to f$ and $\mathfrak{M}f_j \to \mathfrak{M}f$ pointwise in I (uniformly, away from the potential singularities) and $(\mathfrak{M}f_j)' \to (\mathfrak{M}f)'$ pointwise a.e. in $\{\mathfrak{M}f > f\}$ (analogue of Proposition 4.2.4 (i) and (iii)).

(P3) *Flatness in the connecting set*: we have $f' = 0$ for a.e. point in the set $\{\mathfrak{M}f = f\}$. This is necessary for the lateral sunrise operators to have the desired monotonicity properties of Proposition 4.4.1.

(P4) *Singularity control*: uniform control of $(\mathfrak{M}f_j)'$ near the potential singularities (analogue of Proposition 4.3.1).

(P5) *Smallness of the remainder*: control of the remainder terms coming from the integration by parts in the final part of the proof (analogue of (4.46) - (4.47)).

If these five core abstract elements are in place, the proof of Theorem 4.1.1 can be adapted to this situation. Note that Lemma 4.2.1 is already in place to absorb the general setup, and our sunrise construction of the lateral operators in §4.4 can be performed with respect to any open interval (ρ_1, ρ_2) whose closure is contained in $I \subset \mathbb{R}$ (this includes the whole $\mathbb{R}$ itself if $I = \mathbb{R}$), and with respect to the whole I in the case $I = \mathbb{S}^1$.

4.7 Further applications

In this section we briefly discuss how our sunrise strategy can be applied to establish the endpoint Sobolev continuity of the other maximal operators discussed in §4.1.1. For simplicity, the presentation here will be kept on a broad level, and we shall only indicate the major steps or changes required for each adaptation in order to verify properties (P1) - (P5) above. We omit some of the routine details.

4.7.1 Proof of Theorem 4.1.2

We start by recalling that the space $W^{1,1}_{\text{pol}}(\mathbb{S}^d)$ can be naturally associated to

$$W^{1,1}\big((0,\pi),(\sin\theta)^{d-1}\,\mathrm{d}\theta\big),$$

where $\theta = \theta(\xi) = d(\mathbf{e},\xi)$ is the polar angle. For $d \geq 2$, we shall refer to $f(\xi)$ when viewing $f \in W^{1,1}_{\text{pol}}(\mathbb{S}^d)$ on $\mathbb{S}^d$ and to $f(\theta)$ when viewing it on $(0,\pi)$. In this sense we may write

$$\|\nabla f\|_{L^1(\mathbb{S}^d)} = \omega_{d-1}\int_0^\pi |f'(\theta)|\,(\sin\theta)^{d-1}\,\mathrm{d}\theta.$$

Properties (P1) and (P3) can be proved exactly as in §4.2.3.

In order to verify the remaining properties, let us first consider the case $d \geq 2$. Let $g \in W^{1,1}_{\text{pol}}(\mathbb{S}^d) \simeq W^{1,1}\big((0,\pi),(\sin\theta)^{d-1}\,\mathrm{d}\theta\big)$ be a given nonnegative function, absolutely continuous in compact subsets of $(0,\pi)$. We start with a suitable replacement for (4.8) since we do not have the "vanishing at infinity" situation anymore. For $0 < \theta \leq \pi/4$ we have

$$\begin{aligned}
\int_0^{\frac{\pi}{4}} \widetilde{\mathcal{M}}g(\theta)\,(\sin\theta)^{d-2}\,\cos\theta\,\mathrm{d}\theta &= \int_0^{\frac{\pi}{4}}\left(\int_\theta^{\frac{\pi}{4}} -\big(\widetilde{\mathcal{M}}g\big)'(t)\,\mathrm{d}t + \widetilde{\mathcal{M}}g(\tfrac{\pi}{4})\right)(\sin\theta)^{d-2}\,\cos\theta\,\mathrm{d}\theta\\
&\lesssim_d \int_0^{\frac{\pi}{4}}\left(\int_\theta^{\frac{\pi}{4}} \big|\big(\widetilde{\mathcal{M}}g\big)'(t)\big|\,\mathrm{d}t\right)(\sin\theta)^{d-2}\,\cos\theta\,\mathrm{d}\theta + \widetilde{\mathcal{M}}g(\tfrac{\pi}{4})\\
&= \int_0^{\frac{\pi}{4}}\int_0^t (\sin\theta)^{d-2}\,\cos\theta\,\big|\big(\widetilde{\mathcal{M}}g\big)'(t)\big|\,\mathrm{d}\theta\,\mathrm{d}t + \widetilde{\mathcal{M}}g(\tfrac{\pi}{4})\\
&\simeq_d \int_0^{\frac{\pi}{4}} \big|\big(\widetilde{\mathcal{M}}g\big)'(t)\big|\,(\sin t)^{d-1}\,\mathrm{d}t + \widetilde{\mathcal{M}}g(\tfrac{\pi}{4})\\
&< \infty.
\end{aligned}$$

An analogous computation holds in the interval $(\frac{3\pi}{4},\pi)$, and also if $\widetilde{\mathcal{M}}g(\theta)$ is replaced by $g(\theta)$. If follows that the functions $\theta \mapsto g(\theta)(\sin\theta)^{d-1}$ and $\theta \mapsto \widetilde{\mathcal{M}}g(\theta)(\sin\theta)^{d-1}$ have integrable derivatives in $(0,\pi)$ and hence, by the fundamental theorem of calculus, the limits of these functions as $\theta \to 0^+$ or $\theta \to \pi^-$ must exist. If any of these limits were not zero, we would have a contradiction to the fact that g and $\widetilde{\mathcal{M}}g$ belong to $L^{d/(d-1)}(\mathbb{S}^d)$ (the former by Sobolev embedding, and the latter by the boundedness of $\widetilde{\mathcal{M}}$ in $L^{d/(d-1)}(\mathbb{S}^d)$). Therefore

$$\lim_{\theta\to0^+} g(\theta)(\sin\theta)^{d-1} = \lim_{\theta\to\pi^-} g(\theta)(\sin\theta)^{d-1} = \lim_{\theta\to0^+}\widetilde{\mathcal{M}}g(\theta)(\sin\theta)^{d-1} = \lim_{\theta\to\pi^-}\widetilde{\mathcal{M}}g(\theta)(\sin\theta)^{d-1} = 0. \tag{4.57}$$

Given $\lambda > 0$ recall now the weak-type estimate

$$\sigma\{\xi \in \mathbb{S}^d \;:\; \widetilde{\mathcal{M}}g(\xi) \geq \lambda\} \lesssim_d \frac{\|g\|_{L^1(\mathbb{S}^d)}}{\lambda}. \tag{4.58}$$

Fix an interval $J_\eta := [\eta, \pi - \eta] \subset (0, \pi)$, say with $\eta < \frac{\pi}{4}$. Let $\theta_\eta \in J_\eta$ be such that $\widetilde{\mathcal{M}}g(\theta_\eta) = \min_{\theta \in J_\eta} \widetilde{\mathcal{M}}g(\theta)$. Then, taking $\lambda = \widetilde{\mathcal{M}}g(\theta_\eta)$ in (4.58), we find

$$\widetilde{\mathcal{M}}g(\theta_\eta) \lesssim_d \|g\|_{L^1(\mathbb{S}^d)}.$$

Hence, for any $\theta \in J_\eta$, we have

$$\begin{aligned}\widetilde{\mathcal{M}}g(\theta) &= \int_{\theta_\eta}^{\theta} \big(\widetilde{\mathcal{M}}g\big)'(t)\, \mathrm{d}t + \widetilde{\mathcal{M}}g(\theta_\eta)\\ &\lesssim_{\eta,d} \int_0^{\pi} \big|\big(\widetilde{\mathcal{M}}g\big)'(t)\big|\, (\sin t)^{d-1}\, \mathrm{d}t + \widetilde{\mathcal{M}}g(\theta_\eta)\\ &\lesssim_{\eta,d} \|\nabla g\|_{L^1(\mathbb{S}^d)} + \|g\|_{L^1(\mathbb{S}^d)}.\end{aligned} \tag{4.59}$$

Of course, estimates (4.58) and (4.59) also hold with g replacing $\widetilde{\mathcal{M}}g$. Then, if $f_j \to f$ in $W^{1,1}(\mathbb{S}^d)$, an application of (4.59) with $g = f_j - f$ yields (note the sublinearity of $\widetilde{\mathcal{M}}$) that $f_j \to f$ and $\widetilde{\mathcal{M}}f_j \to \widetilde{\mathcal{M}}f$ uniformly in the interval $J_\eta := [\eta, \pi - \eta]$. This is the analogue of Proposition 4.2.4 (i). Parts (ii) and (iii) of Proposition 4.2.4 can be proved in the same way as we did in §4.2.4. This builds up to property (P2).

The analogue of Proposition 4.3.1, the uniform control of $\nabla\widetilde{\mathcal{M}}f_j$ near the potential singularities (in this case, the poles $\mathbf{e}$ and $-\mathbf{e}$), can be proved in the exact same way using (4.57) and the pointwise convergence. This is property (P4). Then we proceed with the sunrise construction with respect to an open interval $(\rho, \pi - \rho)$, with ρ small, and adapt the scheme of proof in §4.5. Note the presence of potentially two remainder terms in (4.43) coming from the integration by parts, and the proof of (4.47) will follow from directly from dominated convergence and the fact that all quantities involved are uniformly bounded in the considered interval by another application of (4.59). This is property (P5), which completes the skeleton of the proof. We omit the remaining details of the adaptation.

The case $d = 1$ is in fact simpler. Here our functions f_j and f will be absolutely continuous in the whole $\mathbb{S}^1$, and so will $\widetilde{\mathcal{M}}f_j$ and $\widetilde{\mathcal{M}}f$. Proceeding as in (4.58) and (4.59) we deduce the pointwise convergence, which is now uniform in $\mathbb{S}^1$. The analogues of Proposition 4.2.4 (ii) and (iii) also hold. There is no need for Proposition 4.3.1 (property (P4)) since we do not have any singularities. We can carry out the sunrise construction with respect to the whole space $\mathbb{S}^1$ (here we must choose an orientation a priori, say clockwise, to read the decomposition (4.17); note that the set $\widetilde{\mathcal{M}}f = f$ is always non-empty) and proceed smoothly as in §4.5.

4.7.2 Proof of Theorem 4.1.3

The $\alpha = \frac{1}{3}$ threshold: a geometric argument

If $d \geq 2$ and $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ we have seen in §4.2.1 and §4.2.2 that we may assume f is continuous in $\mathbb{R}^d \setminus \{0\}$ (and nonnegative for our purposes). In this case, one can verify that $M^\alpha f$ is also continuous in $\mathbb{R}^d \setminus \{0\}$, and we may also consider a degenerate cube of side zero, that is, just the point x itself, in our definition of M^α. As in §4.2.3 we may define the d-dimensional disconnecting set

$$\mathcal{D}^\alpha(f) = \{x \in \mathbb{R}^d \setminus \{0\} \ : \ M^\alpha f(x) > f(x)\},$$

and its corresponding one-dimensional radial version

$$D^\alpha(f) = \{|x| \ : \ x \in \mathcal{D}^\alpha(f)\}.$$

These are open sets in $\mathbb{R}^d \setminus \{0\}$ and $(0,\infty)$, respectively. We define the connecting sets $\mathcal{C}^\alpha(f) := \big(\mathbb{R}^d \setminus \{0\}\big) \setminus \mathcal{D}^\alpha(f)$ and $C^\alpha(f) := (0,\infty) \setminus D^\alpha(f)$. In dimension $d = 1$ we define the sets $D^\alpha(f)$ and its complement $C^\alpha(f)$ over the whole $\mathbb{R}$, for $f \in W^{1,1}(\mathbb{R})$. With start by proving the analogue of Proposition 4.2.2 in this case, a result that involves some insightful geometric considerations coming from the fact that $\alpha \geq \frac{1}{3}$.

Proposition 4.7.1. *Let $d \geq 2$ and $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$. The function $M^\alpha f(r)$ does not have a strict local maximum in $D^\alpha(f)$.*

Proof Assume there is a point $r_0 \in D^\alpha(f)$ for which there exist s_0 and t_0 with $s_0 < r_0 < t_0$, $[s_0, t_0] \subset D^\alpha(f)$, such that $\widetilde{M}f(r) \leq \widetilde{M}f(r_0)$ for all $r \in [s_0, t_0]$ and $\widetilde{M}f(s_0), \widetilde{M}f(t_0) < \widetilde{M}f(r_0)$. Let $x_0 \in \mathbb{R}^d$ be such that $|x_0| = r_0$. Let Q_0 be a cube such that $x_0 \in \alpha Q_0$ and

$$M^\alpha f(x_0) = \fint_{Q_0} f(y) \, \mathrm{d}y.$$

Observe that Q_0 has a positive side since $x_0 \in D^\alpha(f)$. Note that for any $x \in \alpha Q_0$ we have $M^\alpha f(x) \geq M^\alpha f(x_0)$, and hence $|x| \in [s_0, t_0]$ and $M^\alpha f(x) = M^\alpha f(x_0)$. This is due to the fact that the set $\{|x| \ : \ x \in \alpha Q_0\}$ contains $|x_0| = r_0$ and is connected. In particular, this implies that f is not constant in Q_0, since this would contradict the fact that $M^\alpha f(x) > f(x)$ when x is the center of Q_0.

Throughout the rest of the proof we only consider cubes with sides parallel to those of Q_0 (in fact, only dyadic cubes starting from Q_0). Let $\mathcal{A}_0 = \{Q_0\}$ and proceed inductively by defining $\mathcal{A}_k$ as the family obtained by partitioning each cube in $\mathcal{A}_{k-1}$ into 2^d dyadic cubes. Then $\mathcal{A}_k$ has 2^{dk} cubes of side 2^{-k} times the original side of Q_0. Since f is continuous in $\mathbb{R}^d \setminus \{0\}$ and not constant in Q_0, there exists $k \geq 1$ such that the family $\mathcal{A}_k$ has a cube Q_k over which we have

$$\fint_{Q_k} f(y) \, \mathrm{d}y > \fint_{Q_0} f(y) \, \mathrm{d}y = M^\alpha f(x_0). \tag{4.60}$$

Choose such k minimal. We consider the genealogical sequence

$$Q_k \subset Q_{k-1} \subset \ldots Q_1 \subset Q_0,$$

where $Q_i \in \mathcal{A}_i$, and Q_i is the parent of Q_{i+1} for $i = 0, 1, \ldots, k-1$. From the minimality of k, note that for $i = 0, 1, \ldots k-1$ we have

$$\fint_{Q_i} f(y) \, \mathrm{d}y = \fint_{Q_0} f(y) \, \mathrm{d}y = M^\alpha f(x_0). \tag{4.61}$$

Observe that we could not have a strictly smaller average in (4.61), otherwise another average in the same family would be strictly larger, contradicting the minimality of k.

If $\alpha \geq \frac{1}{3}$ we have the following relevant geometric property (recall our cubes are closed):

$$\alpha Q_i \cap \alpha Q_{i+1} \neq \emptyset$$

for any $i = 0, 1, \ldots, k-1$. This means that the set $\mathcal{Y} = \cup_{i=0}^{k} \alpha Q_i$ is connected in $\mathbb{R}^d$ and hence its one-dimensional version, excluding the origin, $Y = \{|x| : x \in \mathcal{Y} \setminus \{0\}\}$ is also connected in $(0, \infty)$. If $x \in \mathcal{Y} \setminus \{0\}$ is such that $x \in \alpha Q_i$ for some $i = 0, 1, \ldots k-1$, by (4.61) we have

$$M^\alpha f(x) \geq \fint_{Q_i} f(y) \, \mathrm{d}y = \fint_{Q_0} f(y) \, \mathrm{d}y = M^\alpha f(x_0).$$

If $x \in \alpha Q_k$, by (4.60) we have

$$M^\alpha f(x) \geq \fint_{Q_k} f(y) \, \mathrm{d}y > \fint_{Q_0} f(y) \, \mathrm{d}y = M^\alpha f(x_0).$$

Hence Y is a connected set in $(0, \infty)$ (i.e. an interval) such that: (i) it contains $r_0 = |x_0|$; (ii) $M^\alpha f(r) \geq M^\alpha f(r_0)$ for every $r \in Y$; (iii) there is a point $r_k = |x|$ (with $x \in \alpha Q_k$) in Y such that $M^\alpha f(r_k) > M^\alpha f(r_0)$. This contradicts the fact that r_0 was a strict local maximum. ■

Remark: The proof of Proposition 4.7.1 can be modified to the case of dimension $d = 1$ and a function $f : \mathbb{R} \to \mathbb{R}$ that is continuous and of bounded variation. In this case we also have $M^\alpha f$ continuous and a strict local maximum in the disconnecting set would have $M^\alpha f$ realized in a bounded and non-denegerate interval. This provides an alternative approach to [Ram19] in order to prove (4.3).

We now proceed to the proof of Theorem 4.1.3.

Proof of Theorem 4.1.3: boundedness

We first briefly consider the boundedness claim in part (ii). Here $d \geq 2$. Observe first that

$$M^\alpha f(x) \lesssim_{d,\alpha} \widetilde{M} f(x). \tag{4.62}$$

One now proceeds via the following steps:

Step 1. Show that $M^\alpha f$ is locally Lipschitz in the disconnecting set $\mathcal{D}^\alpha(f)$. For this, note that every $x \in \mathcal{D}^\alpha(f)$ has a neighborhood $x \in U_x \subset \mathcal{D}^\alpha(f)$ in which the cubes that realize the maximal function for any $y \in U_x$ are of size bounded by below. Take two points $y, z \in U_x$ and compare their maximal functions by using translated cubes and the fact that the difference quotients are uniformly bounded in L^1 by a multiple of the L^1-norm of the gradient of f. Hence $M^\alpha f$ is differentiable a.e. in $\mathcal{D}^\alpha(f)$.

Step 2. Follow line-by-line the mechanism of the main theorem in Chapter 1, to prove that

$$\int_{D^\alpha(f)} \left| \left(M^\alpha f\right)'(r) \right| r^{d-1} \, \mathrm{d}r \lesssim_{d,\alpha} \int_0^\infty |f'(r)| \, r^{d-1} \, \mathrm{d}r. \tag{4.63}$$

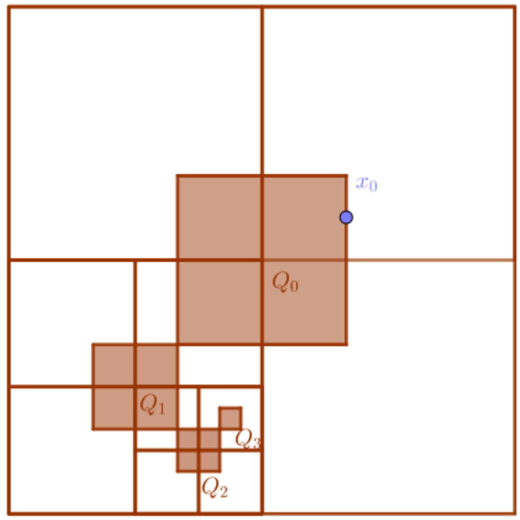

Figure 4.2: Illustration of the construction in the case $\alpha = \frac{1}{3}$. The dyadic cubes Q_0, Q_1, Q_2, Q_3 are in white, and the colored cubes represent αQ_i $(i = 0, 1, 2, 3)$.

This scheme, which in Chapter 1 is used for maximal functions of convolution type, only requires the control (4.62), the bound (4.7), and the absence of local maxima in the disconnecting set given by Proposition 4.7.1.

Step 3. Follow line-by-line the argument in [CS13, §5.4] to show that $M^\alpha f(r)$ is weakly differentiable in $(0, \infty)$ with weak derivative given by $\chi_{C^\alpha(f)} f' + \chi_{D^\alpha(f)} (M^\alpha f)'$. Conclude that $M^\alpha f(x)$ is weakly differentiable in $\mathbb{R}^d$ by the discussion in §4.2.1 and that the desired bound

$$\|\nabla M^\alpha f\|_{L^1(\mathbb{R}^d)} \lesssim_{d,\alpha} \|\nabla f\|_{L^1(\mathbb{R}^d)}$$

follows from (4.63).

Proof of Theorem 4.1.3: continuity

Let us look at properties (P1) - (P5) described in §4.6. We have already established (P1). Let us move to property (P2). The uniform pointwise convergence $M^\alpha f_j(r) \to M^\alpha f(r)$ follows from the sublinearity of M^α, together with (4.62) and (4.13). For the convergence of the derivatives a.e. in the disconnecting set $D^\alpha(f)$ one may start establishing an analogue of Proposition 4.2.3 to move the derivative inside an average over a "good" cube; this follows with the same proof, that only uses translations in $\mathbb{R}^d$. One also needs the analogue of Proposition 4.2.4 (ii) on accumulating sequences of "good cubes". Here the proof is also the same, and one may think of parametrizing the cubes by its center, its side and its orientation (say, with a set of d orthogonal vectors in $\mathbb{S}^{d-1}$). This leads to the desired analogue of Proposition 4.2.4 (iii).

Establishing (P3) requires a brief computation and we do it for $d \geq 2$ in the next proposition (the case $d \geq 1$ and $I = \mathbb{R}$ being easier and following via the same reasoning).

Proposition 4.7.2. *Let $\alpha > 0$, $d \geq 2$ and $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$. Let $r_0 > 0$ be a point of differentiability of $f(r)$ such that $f'(r_0) \neq 0$. Then $r_0 \in D^\alpha(f)$.*

Proof Assume first that $f'(r_0) = c > 0$. Take a point $x_0 = (r_0, 0, \ldots, 0) \in \mathbb{R}^d$. For $h > 0$, we consider a cube Q_h with sides parallel to the usual axes, with side length $2h$, and center $z_0 = (r_0 + \alpha h, 0, \ldots, 0)$. Note that x_0 belongs to the boundary of αQ_h. The idea is to have Q_h "to the right of x_0" as much as possible. If $\alpha \geq 1$, we see that this cube is completely to the right of x_0 and for h small we can easily infer that $⨍_{Q_h} f > f(x_0)$. If $\alpha < 1$, part of this cube will be "to the left of x_0" and we must be a bit more careful. Fix $\varepsilon > 0$ small (say, with $\varepsilon < \min\{c, 1\}$ to begin with). Then

$$f(r_0 + s) \geq f(r_0) + (c - \varepsilon)s \quad \text{and} \quad f(r_0 - s) \geq f(r_0) - (c + \varepsilon)s \tag{4.64}$$

for $|s| \leq s_0(\varepsilon)$. Assume h is sufficiently small so that $\big||y| - r_0\big| \leq s_0(\varepsilon)$ for all $y \in Q_h$. Then, letting $y = (y_1, y_2, \ldots, y_d)$ be our variable in $\mathbb{R}^d$, using (4.64) and the basic fact that $|y| \geq y_1$ we get

$$\begin{aligned}
⨍_{Q_h} f(y)\, \mathrm{d}y - f(x_0) &= \frac{1}{|Q_h|}\left(\int_{y\in Q_h\,:\,|y|\geq r_0} f(y)\, \mathrm{d}y + \int_{y\in Q_h\,:\,|y|<r_0} f(y)\, \mathrm{d}y\right) - f(x_0)\\
&\geq \frac{1}{|Q_h|}\left(\int_{y\in Q_h\,:\,|y|\geq r_0} (c-\varepsilon)(|y| - r_0)\, \mathrm{d}y + \int_{y\in Q_h\,:\,|y|<r_0} (c+\varepsilon)(|y| - r_0)\, \mathrm{d}y\right)\\
&\geq \frac{1}{|Q_h|}\left(\int_{y\in Q_h\,:\,y_1\geq r_0} (c-\varepsilon)(y_1 - r_0)\, \mathrm{d}y + \int_{y\in Q_h\,:\,y_1<r_0} (c+\varepsilon)(y_1 - r_0)\, \mathrm{d}y\right)\\
&= \frac{1}{2h}\left(\int_{r_0}^{r_0+\alpha h+h} (c-\varepsilon)(y_1 - r_0)\, \mathrm{d}y_1 + \int_{r_0+\alpha h-h}^{r_0} (c+\varepsilon)(y_1 - r_0)\, \mathrm{d}y_1\right)\\
&= \frac{1}{2h}\left((c-\varepsilon)\frac{(\alpha h + h)^2}{2} - (c+\varepsilon)\frac{(\alpha h - h)^2}{2}\right)\\
&= \frac{h(2\alpha c - \varepsilon\alpha^2 - \varepsilon)}{2}.
\end{aligned} \tag{4.65}$$

The latter is strictly positive as long as we choose $\varepsilon < 2\alpha c/(\alpha^2 + 1)$, which is clearly possible if $\alpha > 0$. A similar argument shows that if $f'(r_0) = c < 0$ then $r_0 \in D^\alpha(f)$. Here we choose $z_0 = (r_0 - \alpha h, 0, \ldots, 0)$, and choose h small so that if $y \in Q_h$ then $y_1 \geq r_0 - \alpha h - h \geq \frac{r_0}{2}$ and

$$|y| \leq \sqrt{y_1^2 + (d-1)h^2} \leq \left(y_1 + \frac{(d-1)h^2}{2y_1}\right) \leq y_1 + \varepsilon h. \tag{4.66}$$

We use (4.66) in the analogue of passage (4.65). ■

Property (P4) is not needed in the case $d = 1$, whereas in the case $d \geq 2$ we can prove it following the same outline of Proposition 4.3.1, with minor adjustments to allow for a dependence on α. The we perform the surnrise construction, in the case $d = 1$ with respect to the whole $\mathbb{R}$, and in the case $d \geq 2$ as we already did, in an interval (ρ, ∞). The proof in §4.5 goes through identically, as (4.62) can be used to prove the analogue of (4.47) (property (P5)).

4.7.3 Proof of Theorem 4.1.4

We start by observing that, for any $\alpha \geq 0$, we have the pointwise bound (see [SW72, Chapter II, Eq. (3.18)])

$$M_\varphi^\alpha f(x) \lesssim_{d,\alpha} \widetilde{M} f(x). \tag{4.67}$$

In the rest of the proof we focus in the case $d \geq 2$. The case $d = 1$ is simpler and requires only minor modifications. We start with the usual setup, in which our $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$ is nonnegative and continuous in $\mathbb{R}^d \setminus \{0\}$, and one can verify that $M_\varphi^\alpha f$ is also radial and continuous in $\mathbb{R}^d \setminus \{0\}$.

Absence of local maxima

Define the disconnecting sets $\mathcal{D}_\varphi^\alpha(f)$ (in $\mathbb{R}^d \setminus \{0\}$) and $D_\varphi^\alpha(f)$ (in $(0, \infty)$), and the connecting sets $\mathcal{C}_\varphi^\alpha(f)$ and $C_\varphi^\alpha(f)$ as we already did in §4.2.3 or §4.7.2. We first establish property (P1), the analogue of Proposition 4.2.2.

Proposition 4.7.3. *Let $\alpha > 0$, $d \geq 2$ and $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$. The function $M_\varphi^\alpha f(r)$ does not have a strict local maximum in $D_\varphi^\alpha(f)$.*

Proof Assume there is a point $r_0 \in D_\varphi^\alpha(f)$ for which there exist s_0 and t_0 with $s_0 < r_0 < t_0$, $[s_0, t_0] \subset D_\varphi^\alpha(f)$, such that $M_\varphi^\alpha f(r) \leq M_\varphi^\alpha f(r_0)$ for all $r \in [s_0, t_0]$ and $M_\varphi^\alpha f(s_0), M_\varphi^\alpha f(t_0) < M_\varphi^\alpha f(r_0)$. Let $x_0 \in \mathbb{R}^d$ be such that $|x_0| = r_0$. Assume that $M_\varphi^\alpha f(x_0) = f * \varphi_t(z_0)$ with $|z_0 - x_0| \leq \alpha\sqrt{t}$. For any $y \in \overline{B_{\alpha\sqrt{t}}(z_0)}$ note that the pair (z_0, t) is an admissible choice for the maximal function $M_\varphi^\alpha f$ at y, hence $M_\varphi^\alpha f(y) \geq M_\varphi^\alpha f(x_0)$. Since x_0 is a strict local maximum, in our setup we must then have $\left\{|y| \,:\, y \in \overline{B_{\alpha\sqrt{t}}(z_0)}\right\} \subset [s_0, t_0]$ and $M_\varphi^\alpha f(y) = M_\varphi^\alpha f(x_0) = f * \varphi_t(z_0)$ for such y. In particular this implies that $z_0 \neq 0$ and that $M_\varphi^\alpha f(z_0) = M_\varphi^0 f(z_0) = f * \varphi_t(z_0) > f(z_0)$. Hence $|z_0|$ is a strict local maximum of $M_\varphi^0 f(r)$ in the disconnecting set $D_\varphi^0(f)$. This contradicts [CS13, Lemma 8], i.e. the fact that $M_\varphi^0 f$ is subharmonic in the disconnecting set (which is the case $\alpha = 0$ of this proposition). Note that [CS13, Lemma 8] is originally stated for continuous functions f but its proof only uses such continuity in a neighborhood of z_0 whose closure is contained in the disconnecting set $\mathcal{D}_\varphi^0(f)$ (which serves our purposes here). ■

Proof of Theorem 4.1.4: boundedness

Once we have (4.67) and Proposition 4.7.3 in our hands, the proof of the boundedness follows the exact same outline with three steps of §4.7.2 (in Step 1, one would think of the time t being bounded by below).

Having gone through the three steps above and established the gradient bound, it will be useful to take a closer look at the second step, for it provides, as a corollary, a local estimate that will imply our desired property (P4). Let $\rho > 0$ and write

$$D_{\varphi}^{\alpha}(f) \cap (0,\rho) = \bigcup_{i=1}^{\infty}(a_i, b_i).$$

For each i, let $\tau_i \in [a,b]$ be a point of minimum of $M_{\varphi}^{\alpha} f$ is such interval (then $M_{\varphi}^{\alpha} f$ is non-increasing in $[a_i,\tau_i]$ and non-decreasing in $[\tau_i, b_i]$). Assuming for a moment that $b_i \neq \rho$, using integration by parts we get

$$\begin{aligned}
\int_{a_i}^{b_i} \left|\left(M_{\varphi}^{\alpha} f\right)'(r)\right| r^{d-1}\,\mathrm{d}r &= -\int_{a_i}^{\tau_i} \left(M_{\varphi}^{\alpha} f\right)'(r)\, r^{d-1}\,\mathrm{d}r + \int_{\tau_i}^{b_i} \left(M_{\varphi}^{\alpha} f\right)'(r)\, r^{d-1}\,\mathrm{d}r \\
&= M_{\varphi}^{\alpha} f(a_i)\, {a_i}^{d-1} + M_{\varphi}^{\alpha} f(b_i)\, {b_i}^{d-1} - 2M_{\varphi}^{\alpha} f(\tau_i)\, \tau_i^{d-1} \\
&\qquad + (d-1)\int_{a_i}^{\tau_i} M_{\varphi}^{\alpha} f(r)\, r^{d-2}\,\mathrm{d}r - (d-1)\int_{\tau_1}^{b_i} M_{\varphi}^{\alpha} f(r)\, r^{d-2}\,\mathrm{d}r \\
&\lesssim_{d,\alpha} f(a_i)\, {a_i}^{d-1} + f(b_i)\, {b_i}^{d-1} - 2f(\tau_i)\, \tau_i^{d-1} \\
&\qquad + (d-1)\int_{a_i}^{\tau_i} \widetilde{M} f(r)\, r^{d-2}\,\mathrm{d}r - (d-1)\int_{\tau_i}^{b_i} f(r)\, r^{d-2}\,\mathrm{d}r \\
&= f(a_i)\, {a_i}^{d-1} - f(\tau_i)\, \tau_i^{d-1} + (d-1)\int_{a_i}^{\tau_i} \widetilde{M} f(r)\, r^{d-2}\,\mathrm{d}r + \int_{\tau_i}^{b_i} f'(r)\, r^{d-1}\,\mathrm{d}r \\
&\leq (d-1)\int_{a_i}^{\tau_i} \widetilde{M} f(r)\, r^{d-2}\,\mathrm{d}r + \int_{a_i}^{b_i} |f'(r)|\, r^{d-1}\,\mathrm{d}r.
\end{aligned} \tag{4.68}$$

The last inequality holds since

$$f(a_i)\, {a_i}^{d-1} - f(\tau_i)\, \tau_i^{d-1} \leq -\int_{a_i}^{\tau_i} f'(r)\, r^{d-1}\,\mathrm{d}r \leq \int_{a_i}^{\tau_i} |f'(r)|\, r^{d-1}\,\mathrm{d}r.$$

From (4.67) and (4.9) note that there is no issue in (4.68) if $a_i = 0$. If $b_i = \rho$, the inequality (4.68) continues to hold if we add a term $\widetilde{M}(\rho)\,\rho^{d-1} - f(\rho)\,\rho^{d-1}$ on the right hand-side. If we sum over all intervals (and take also the connecting set into consideration) we arrive at

$$\int_0^{\rho} \left|\left(M_{\varphi}^{\alpha} f\right)'(r)\right| r^{d-1}\,\mathrm{d}r \lesssim_{d,\alpha} \int_0^{\rho} |f'(r)|\, r^{d-1}\,\mathrm{d}r + \int_0^{\rho} \widetilde{M} f(r)\, r^{d-2}\,\mathrm{d}r + \widetilde{M}(\rho)\,\rho^{d-1}. \tag{4.69}$$

On the other hand, a similar computation to (4.8) yields

$$\begin{aligned}
\int_0^{\rho} \widetilde{M} f(r)\, r^{d-2}\,\mathrm{d}r &= \int_0^{\rho} \left(\widetilde{M} f(\rho) - \int_r^{\rho} \left(\widetilde{M} f\right)'(t)\,\mathrm{d}t\right) r^{d-2}\,\mathrm{d}r \\
&\lesssim_d \widetilde{M} f(\rho)\,\rho^{d-1} + \int_0^{\rho} \left|\left(\widetilde{M} f\right)'(t)\right| t^{d-1}\,\mathrm{d}t.
\end{aligned} \tag{4.70}$$

Combining (4.69) and (4.70) we arrive at

$$\int_0^\rho \left|\left(M_\varphi^\alpha f\right)'(r)\right| r^{d-1}\, \mathrm{d}r \lesssim_{d,\alpha} \int_0^\rho |f'(r)|\, r^{d-1}\, \mathrm{d}r + \int_0^\rho \left|\left(\widetilde{M} f\right)'(r)\right| r^{d-1}\, \mathrm{d}r + \widetilde{M}(\rho)\, \rho^{d-1}.$$

Observe that this estimate, combined with Proposition 4.3.1, plainly yields the analogue of Proposition 4.3.1 for the non-tangential operators M_φ^α. This is property (P4) in our to-do list (which is not needed for the case $d = 1$).

Proof of Theorem 4.1.4: continuity

We have already established properties (P1) and (P4) of our sunrise strategy outlined in §4.6. Property (P2) follows pretty much as in Proposition 4.2.4, using (4.67) and the sublinearity of $M_\varphi^\alpha f$ for the convergences at the function level, and verifying that one can move the gradient inside the integral as in Proposition 4.2.3 in the disconnecting set. The sunrise construction will be identical to §4.4 when $d \geq 2$ (and over $I = \mathbb{R}$ when $d = 1$) and one shall use (4.67) to prove the analogue of (4.47) (property (P5)). The proof will be complete once we establish property (P3). This is the content of our final proposition (which also holds for $d = 1$ and $I = \mathbb{R}$ with the same reasoning).

Proposition 4.7.4. *Let $\alpha > 0$, $d \geq 2$ and $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$. Let $r_0 > 0$ be a point of differentiability of $f(r)$ such that $f'(r_0) \neq 0$. Then $r_0 \in D_\varphi^\alpha(f)$.*

Proof The proof here is similar in spirit to the proof of Proposition 4.7.2, but technically slightly more involved. We first consider the case $f'(r_0) = c > 0$ and let $x_0 := (r_0, 0, \ldots, 0) \in \mathbb{R}^d$. Fix $\varepsilon > 0$ small (say, with $\varepsilon < \min\{c, 1\}$ to begin with). Then we have

$$f(r_0 + s) \geq f(r_0) + (c - \varepsilon)s \quad \text{and} \quad f(r_0 - s) \geq f(r_0) - (c + \varepsilon)s \tag{4.71}$$

for $|s| \leq s_0(\varepsilon)$.

For $t < 1$ small we set $N := (t\varepsilon)^{-1/8}$ and consider the cube Q_t of center at the origin and side $2N\sqrt{t}$ (with sides parallel to the usual axes). We let $z_0 = (r_0 + \alpha\sqrt{t}, 0, 0, \ldots, 0)$ and we want to show that $f * \varphi_t(z_0) > f(x_0)$ when t and ε are small enough (note that we are trying to place the mass of the heat kernel "to the right" of r_0). Since the heat kernel is radial we may write

$$f * \varphi_t(z_0) - f(x_0) = \int_{\mathbb{R}^d} \big(f(z_0 + y) - f(x_0)\big)\, \varphi_t(y)\, \mathrm{d}y = \int_{Q_t} + \int_{Q_t^c} =: (I) + (II).$$

We first verify that the integral (II) is small. By the Sobolev embedding, recall that $f \in L^{d/(d-1)}(\mathbb{R}^d)$. Observe also that

$$\begin{aligned}\int_{Q_t^c} \varphi_t(y)^d\, \mathrm{d}y &\leq \int_{|y| \geq N\sqrt{t}} \varphi_t(y)^d\, \mathrm{d}y = \omega_{d-1} \int_{N\sqrt{t}}^\infty \frac{s^{d-1}}{(4\pi t)^{d^2/2}}\, e^{\frac{-d\, s^2}{4t}}\, \mathrm{d}s \\ &= \frac{\omega_{d-1}}{t^{d(d-1)/2}} \int_N^\infty \frac{u^{d-1}}{(4\pi)^{d^2/2}}\, e^{\frac{-d\, u^2}{4}}\, \mathrm{d}u \lesssim_d \frac{e^{\frac{-dN^2}{8}}}{t^{d(d-1)/2}} \lesssim_d \frac{N^{4d(-d(d-1)-1)}}{t^{d(d-1)/2}} \leq \left(\varepsilon\sqrt{t}\right)^d.\end{aligned} \tag{4.72}$$

Hence, using Hölder's inequality we get

$$\int_{Q_t^c} f(z_0+y)\,\varphi_t(y)\;\mathrm{d}y \le \|f\|_{L^{d/(d-1)}(\mathbb{R}^d)} \left(\int_{Q_t^c} \varphi_t(y)^d\;\mathrm{d}y\right)^{1/d} \lesssim_d \|f\|_{L^{d/(d-1)}(\mathbb{R}^d)}\,\varepsilon\sqrt{t}. \tag{4.73}$$

Similarly, one can show that

$$\int_{Q_t^c} f(x_0)\,\varphi_t(y)\;\mathrm{d}y \lesssim_d f(x_0)\,\varepsilon\sqrt{t}. \tag{4.74}$$

Combining (4.73) and (4.74) we arrive at

$$(II) = \varepsilon\sqrt{t}\,O(1), \tag{4.75}$$

where the implicit constant depends only on $d, \|f\|_{L^{d/(d-1)}(\mathbb{R}^d)}$ and $f(x_0)$.

We then move to the analysis of the term (I). Let $Q_t = Q_t^+ \cup Q_t^-$, where $Q_t^+ = \{y \in Q_t : |z_0+y| \ge r_0\}$ and $Q_t^- = \{y \in Q_t : |z_0+y| < r_0\}$. Assume t is sufficiently small so that $\big||z_0+y| - r_0\big| \le s_0(\varepsilon)$ for all $y \in Q_t$. Then, letting $y = (y_1, y_2, \ldots, y_d)$, using (4.71) and the fact that $|z_0+y| \ge r_0 + \alpha\sqrt{t} + y_1$, we get

$$\begin{aligned}
&\int_{Q_t} \big(f(z_0+y) - f(x_0)\big)\,\varphi_t(y)\;\mathrm{d}y = \int_{Q_t^+} + \int_{Q_t^-}\\
&\ge \int_{Q_t^+} (c-\varepsilon)\big(|z_0+y| - r_0\big)\,\varphi_t(y)\;\mathrm{d}y - \int_{Q_t^-} (c+\varepsilon)\big(-|z_0+y| + r_0\big)\,\varphi_t(y)\;\mathrm{d}y\\
&\ge \int_{Q_t^+} (c-\varepsilon)\big((r_0+\alpha\sqrt{t}+y_1) - r_0\big)\varphi_t(y)\;\mathrm{d}y - \int_{Q_t^-} (c+\varepsilon)\big(-(r_0+\alpha\sqrt{t}+y_1) + r_0\big)\varphi_t(y)\;\mathrm{d}y
\end{aligned} \tag{4.76}$$

$$= c\,\alpha\sqrt{t}\int_{Q_t} \varphi_t(y)\;\mathrm{d}y + \varepsilon\left(-\int_{Q_t^+} (\alpha\sqrt{t}+y_1)\,\varphi_t(y)\;\mathrm{d}y + \int_{Q_t^-} (\alpha\sqrt{t}+y_1)\,\varphi_t(y)\;\mathrm{d}y\right).$$

Note that we used above the fact that $\int_{Q_t} y_1\varphi_t(y)\;\mathrm{d}y = 0$, since φ_t is even. Proceeding as in (4.72) and (4.74) we find that

$$1 - O_d(\varepsilon) \le \int_{Q_t} \varphi_t(y)\;\mathrm{d}y \le 1. \tag{4.77}$$

and

$$\int_{Q_t} |y_1|\,\varphi_t(y)\;\mathrm{d}y \le \int_{-N\sqrt{t}}^{N\sqrt{t}} \frac{|y_1|}{(4\pi t)^{1/2}}\,e^{\frac{-y_1^2}{4t}}\;\mathrm{d}y_1 = \sqrt{t}\int_{-N}^{N} \frac{|u|}{(4\pi)^{1/2}}\,e^{\frac{-u^2}{4}}\;\mathrm{d}u \lesssim \sqrt{t}. \tag{4.78}$$

Using (4.77) and (4.78) in (4.76) we arrive at

$$(I) = \int_{Q_t} \big(f(z_0+y) - f(x_0)\big)\,\varphi_t(y)\;\mathrm{d}y \ge \sqrt{t}\,\Big(c\,\alpha\big(1 - O_d(\varepsilon)\big) - \varepsilon\big(\alpha + O(1)\big)\Big). \tag{4.79}$$

Note that the work in (4.75) and (4.79) had the intention of leaving things in the same scale $\sqrt{t}$. Combining (4.75) and (4.79) we arrive at

$$(I) + (II) \geq \sqrt{t}\Big(c\,\alpha\big(1 - O_d(\varepsilon)\big) - \varepsilon\big(\alpha + O(1)\big)\Big), \tag{4.80}$$

where the implicit constant in the $O(1)$ depends only on d, $\|f\|_{L^{d/(d-1)}(\mathbb{R}^d)}$ and $f(x_0)$. Since $c > 0$ and $\alpha > 0$, the conclusion is that for our initial choice of ε sufficiently small we will have (4.80) strictly positive, as we wanted.

The case $f'(r_0) = c < 0$ follows along the same lines. Given our initial $\varepsilon > 0$, we will now choose $z_0 = (r_0 - \alpha\sqrt{t}, 0, 0, \ldots, 0)$. We start with t small so that $r_0 - \alpha\sqrt{t} - N\sqrt{t} \geq \frac{r_0}{2}$. Then we can go to t even smaller such that for every $y \in Q_t$ we have

$$\begin{aligned}|z_0 + y| &\leq \Big((r_0 - \alpha\sqrt{t} + y_1)^2 + (d-1)N^2 t\Big)^{1/2} \leq (r_0 - \alpha\sqrt{t} + y_1) + \frac{(d-1)N^2 t}{2(r_0 - \alpha\sqrt{t} + y_1)} \\ &\leq (r_0 - \alpha\sqrt{t} + y_1) + \varepsilon\sqrt{t}.\end{aligned}$$

We use this inequality in the analogue of (4.76). ■

4.7.4 Concluding remarks

We briefly comment on the obstructions towards the endpoint $W^{1,1}$–continuity via the sunrise strategy for some maximal operators mentioned, or at least hinted at, in our text (and for which the corresponding boundedness result is already established). The non-tangential Hardy-Littlewood maximal operator M^α, in the case of dimension $d = 1$ and $0 < \alpha < \frac{1}{3}$, does not necessarily verify property (P1) as exemplified in [Ram19, Theorem 2] (think of f being two high bumps far apart). Still in dimension $d = 1$, for the centered Hardy-Littlewood maximal operator, on top of obstruction (P1), property (P3) may also not be verified. The centered heat flow maximal function M^0_φ (in dimension $d = 1$ for general $f \in W^{1,1}(\mathbb{R})$ and if $d \geq 2$ for $f \in W^{1,1}_{\text{rad}}(\mathbb{R}^d)$) verifies (P1) but does not necessarily verify the flatness property (P3) (just think of f being the Gaussian φ_1).

Another standard maximal function of convolution type is the one associated to the Poisson kernel

$$\Psi_t(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}}\,\frac{t}{(|x|^2 + t^2)^{(d+1)/2}}.$$

Similarly to (4.81), for $\alpha \geq 0$ we may consider

$$M^\alpha_\Psi f(x) = \sup_{t>0\,;\,|y-x|\leq\alpha t} (|f| * \Psi_t)(y). \tag{4.81}$$

The boundedness of the map $f \to \big(M^\alpha_\Psi f\big)'$ from $W^{1,1}(\mathbb{R}) \to L^1(\mathbb{R})$ was established for $\alpha = 0$ in [CS13, Theorem 2] and for $\alpha > 0$ in [CFS18, Theorem 4]. When $d \geq 2$ and $\alpha = 0$ the boundedness of the map $f \to \nabla M^0_\Psi f$ from $W^{1,1}_{\text{rad}}(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ was established in Chapter

1. Following the exact same argument of our Theorem 4.1.3 we can extend this boundedness result in dimension $d \geq 2$ for $\alpha > 0$ as well (this has not been recorded in the literature before). In all of the cases above, property (P1) holds; and this is actually an important ingredient in such boundedness proofs. One may be naturally led to think that the analogue of Proposition 4.7.4, i.e. property (P3), would be somewhat reasonable for such an operator, at least in the non-tangential case $\alpha > 0$. This turns out to be false. The flatness property (P3) is not necessarily verified for any $\alpha \geq 0$.

In dimension $d \geq 2$, it is shown in [CFS18, §5.3] that the function

$$f(x) = (1 + |x|^2)^{\frac{-d+1}{2}}$$

is such that $M_\Psi^\alpha f(x) = f(x)$ for $|x| \leq \frac{1}{\alpha}$. Such f is not in $W^{1,1}(\mathbb{R}^d)$, but we could simply multiply f by a smooth and radially non-increasing function ϕ with $\phi(x) = 1$ if $|x| \leq 1$, and $\phi(x) = 0$ if $|x| \geq 2$, that the property $M_\Psi^\alpha f(x) = f(x)$ would continue to hold in a neighborhood of the origin. In dimension $d = 1$ we may consider the function

$$f(x) = \log\left(\frac{4 + x^2}{1 + x^2}\right) = 2\int_1^2 \frac{s}{(s^2 + x^2)}\,\mathrm{d}s = 2\pi\int_1^2 \Psi_s(x)\,\mathrm{d}s$$

This function belongs to $W^{1,1}(\mathbb{R})$. Using the semigroup property of the Poisson kernel we get

$$\begin{aligned} v(y,t) := (f * \Psi_t)(y) &= 2\pi\int_{-\infty}^{\infty}\int_1^2 \Psi_s(y-x)\,\Psi_t(x)\,\mathrm{d}s\,\mathrm{d}x = 2\pi\int_1^2\int_{-\infty}^{\infty} \Psi_s(y-x)\,\Psi_t(x)\,\mathrm{d}x\,\mathrm{d}s \\ &= 2\pi\int_1^2 \Psi_{t+s}(y)\,\mathrm{d}s = \log\left(\frac{(t+2)^2 + y^2}{(t+1)^2 + y^2}\right). \end{aligned}$$

For a fixed $x \in \mathbb{R}$, by the maximum principle (recall that v verifies $\Delta v = 0$ in $\mathbb{R} \times (0,\infty)$), the supremum of $v(y,t)$ in the cone $|y - x| \leq \alpha t$ is attained at a point $y = x \pm \alpha t$. We want to show that, for x in a neighborhood of the origin we have

$$\log\left(\frac{4 + x^2}{1 + x^2}\right) \geq \log\left(\frac{(t+2)^2 + (x \pm \alpha t)^2}{(t+1)^2 + (x \pm \alpha t)^2}\right)$$

for all $t \geq 0$. After removing the log and multiplying out, this is equivalent to

$$t\,(-2x^2 \pm 6x\alpha + 3\alpha^2 t + 3t + 4) \geq 0\,,$$

which is clearly true if $|x|$ is small.

Chapter 5

BV continuity for the uncentered Hardy-Littlewood maximal operator

5.1 Introduction

In this chapter we are interested in another extension for the continuity of the map

$$f \mapsto \nabla \widetilde{M} f$$

at the endpoint $p = 1$, from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$. Recall that we write $\widetilde{M} f$ as the uncentered version of (1). In [CMP17] the authors also consider the space $BV(\mathbb{R})$ endowed with the norm $\|f\|_{BV} := |f(-\infty)| + \mathrm{Var}\,(f)$. About this, they asked the following question:

Question. *(Question B in [CMP17]) Is the map* $\widetilde{M} : BV(\mathbb{R}) \to BV(\mathbb{R})$ *continuous?*

This question, in case of being answered affirmatively, would provide a generalization of [CMP17, Theorem 1] (since $W^{1,1}(\mathbb{R})$ embeds isometrically in $BV(\mathbb{R})$). It is important to notice that, in general, the continuity in the $BV(\mathbb{R})$ setting is more delicate than in the $W^{1,1}(\mathbb{R})$ setting. An example of this is that in the fractional setting the analogue of [CMP17, Theorem 1] holds (see [Mad19]) but the answer to the analogue of the previous question is negative (see [CMP17, Theorem 3]).

The main goal of the present chapter is to answer this question. We prove the following.

Theorem 5.1.1. *The map* $\widetilde{M} : BV(\mathbb{R}) \to BV(\mathbb{R})$ *is continuous.*

Our general strategy is similar to that proposed in [CMP17, Theorem 1]. Indeed, Lemma 5.3.5 in §5.3 and the proof of Theorem 5.1.1 in §5.4 together constitute a suitable variant of the analysis presented in [CMP17, Section 5.4], as they show that the main result follows provided that two specific properties regarding the behavior of the maximal function hold. One of the two ingredients is a weaker version of Theorem 5.1.1 (see Proposition 5.3.1) saying that the map $f \mapsto \mathrm{Var}\,(\widetilde{M} f)$ is continuous from $BV(\mathbb{R})$ to $\mathbb{R}$. The remaining one concerns good properties of the derivative of the maximal functions on the so-called disconnecting and connecting sets (see Lemmas 5.3.3 and 5.3.4).

We notice that several of the arguments in [CMP17] (also the arguments in Chapter 4) rely on the regularity of the original function, therefore they are not enough to conclude Theorem 5.1.1. In fact, the authors of [CMP17] (also the authors of [A4]) used in their work a reduction of the problem to the analogous question stated for "lateral" maximal operators, but in our case that reduction causes several problems. For instance, the one-sided maximal function of a function in $BV(\mathbb{R})$ is not necessarily continuous, which makes it much less useful in our approach. Instead, we provide detailed studies on the variation of the maximal function, where the sets of discontinuity points of initial functions receive particular attention (see §5.2). Another difficulty is due to the fact that for $f \in BV(\mathbb{R})$ its value at a given point is not determined by the values around this point, contrary to the $W^{1,1}(\mathbb{R})$ case. To gain more control on the local behavior of initial functions, several times we replace $|f|$ by its adjusted version $\overline{|f|}$ (see §5.3), which is upper semicontinuous. In particular, the disconnecting and connecting sets mentioned above are defined with the aid of $\overline{|f|}$ instead of $|f|$.

5.2 Preliminaries

In this section we develop some preliminary tools required in our work. We start by stating the following result which describes the behavior of the maximal function at infinity.

Lemma 5.2.1. *Given $f \in BV(\mathbb{R})$ let $|f|(\infty) := \lim_{x\to\infty} |f|(x)$ and $|f|(-\infty) := \lim_{x\to-\infty} |f|(x)$. Then*

$$\lim_{x\to\infty} \widetilde{M}f(x) = \lim_{x\to-\infty} \widetilde{M}f(x) = c,$$

where $c = \max\{|f|(\infty), |f|(-\infty)\}$.

Proof Without loss of generality we assume that $f \geq 0$ and $c = f(\infty)$. Observe that

$$\widetilde{M}f(x) \geq \lim_{r\to\infty} \fint_{(x-1,x+r)} f = c$$

holds for any $x \in \mathbb{R}$. Fix $\varepsilon > 0$ and let $N_0 > 0$ be such that $f(x) \leq c + \frac{\varepsilon}{2}$ for $|x| > N_0$. We choose $N_1 > N_0$ satisfying

$$\frac{2N_0\|f\|_\infty}{N_1 - N_0} \leq \frac{\varepsilon}{2}.$$

Consider x_0 satisfying $|x_0| > N_1$ and any interval $I \ni x_0$. If $|I| < N_1 - N_0$, then clearly

$$\fint_I f \leq c + \frac{\varepsilon}{2}.$$

On the other hand, if $|I| \geq N_1 - N_0$, then

$$\fint_I f \leq \frac{1}{|I|} \int_{I\cap[-N_0,N_0]^c} f(x)\, \mathrm{d}x + \frac{1}{N_1 - N_0} \int_{[-N_0,N_0]} f(x)\, \mathrm{d}x \leq c + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the claim follows. ■

The next goal is to use the $BV(\mathbb{R})$ norm to control the difference between two $BV(\mathbb{R})$ functions or between their maximal functions at a given point x. The following estimates, although very basic, will be extremely useful later on.

Lemma 5.2.2. *Let $f, g \in BV(\mathbb{R})$. Then*

$$|f(x) - g(x)| \leq 2\|f-g\|_{BV} \quad \text{and} \quad \left|\widetilde{M}f(x) - \widetilde{M}g(x)\right| \leq 2\|f-g\|_{BV}$$

hold for any $x \in \mathbb{R}$.

Proof The first inequality follows since

$$|f(x) - g(x)| \leq |(f(x) - g(x)) - (f(-\infty) - g(-\infty))| + |f(-\infty) - g(-\infty)|.$$

Now, assume $\widetilde{M}f(x) \geq \widetilde{M}g(x)$. By the first part of the lemma for any $I \ni x$ we have

$$⨍_I |g| \geq ⨍_I |f| - ⨍_I |g-f| \geq ⨍_I |f| - 2\|f-g\|_{BV}.$$

Thus, $\widetilde{M}g(x) \geq \widetilde{M}f(x) - 2\|f-g\|_{BV}$ and the second part follows as well. ■

Contrasting with the $W^{1,1}(\mathbb{R})$ setting (see [CMP17, Lemma 14]), in our context to make the reduction to the case $f \geq 0$ is much more problematic. In order to deal with this issue we require several results describing the relations between f and $|f|$.

In the following, for given $g \in BV(\mathbb{R})$ we define $\lim\limits_{y\uparrow x} g(y) =: g(x^-)$ and $\lim\limits_{y\downarrow x} g(y) =: g(x^+)$. Also, for each $-\infty \leq a < b \leq \infty$ we introduce the quantity

$$\mathrm{Var}_{(a,b)}(g) := \sup\left\{\sum_{i=1}^{K} |g(a_i) - g(a_{i-1})|\right\},$$

where the supremum is taken over all $K \in \mathbb{N}$ and all sequences $a < a_0 < \cdots < a_K < b$ (notice that if g is continuous at a and b, then the sequences satisfying $a = a_0 < \cdots < a_K = b$ can be considered instead and the supremum will not change). For a given partition $\mathcal{P} = \{a_0 < a_1 < \cdots < a_K\}$ we denote $\mathrm{Var}\,(g, \mathcal{P}) := \sum_{i=1}^{K} |g(a_i) - g(a_{i-1})|$. Finally, we write $E_l(g) := \{x \in \mathbb{R}; g(x) \neq g(x^-)\}$ and $E_r(g) := \{x \in \mathbb{R}; g(x) \neq g(x^+)\}$.

Lemma 5.2.3. *Fix $f \in BV(\mathbb{R})$. Then for any $-\infty \leq a < b \leq \infty$ we have*

$$\begin{aligned}\mathrm{Var}_{(a,b)}(f) - \mathrm{Var}_{(a,b)}(|f|) = &\sum_{x \in E_l(f)\cap(a,b)} |f(x) - f(x^-)| - \big||f|(x) - |f|(x^-)\big| \\ &+ \sum_{x \in E_r(f)\cap(a,b)} |f(x) - f(x^+)| - \big||f|(x) - |f|(x^+)\big|.\end{aligned}$$

Proof Fix $-\infty \leq a < b \leq \infty$. We write $E_l(f) \cap (a,b) =: \{x_{l,n}; n \in \mathbb{N}\}$ and $E_r(f) \cap (a,b) =: \{x_{r,n}; n \in \mathbb{N}\}$, assuming that both of these sets are infinite (the other cases can be treated very similarly). Given $\varepsilon > 0$ we choose a partition $\mathcal{P} \subset (a,b)$ such that

$$\mathrm{Var}\,(f, \mathcal{P}) > \mathrm{Var}_{\,(a,b)}(f) - \varepsilon$$

and

$$\mathrm{Var}\,(|f|, \mathcal{P}) > \mathrm{Var}_{\,(a,b)}(|f|) - \varepsilon.$$

Then for fixed $N \in \mathbb{N}$ we construct $\widetilde{\mathcal{P}} = \widetilde{\mathcal{P}}(N) \subset (a,b)$ by adding to $\mathcal{P}$ (if needed) some extra points. The procedure consists of the following three steps.

- We set $\mathcal{P}_1 = \mathcal{P} \cup \{x_{l,n}, x_{r,n}; n \leq N\}$.
- For each $n \leq N$ we choose $\widetilde{x}_{l,n} < x_{l,n}$ such that

  $$\mathcal{P}_1 \cap (\widetilde{x}_{l,n}, x_{l,n}) = \emptyset$$

  and $|f(x_{l,n}^-) - f(\widetilde{x}_{l,n})| < 2^{-n}\varepsilon$. Similarly, we choose $\widetilde{x}_{r,n} > x_{r,n}$ such that

  $$\mathcal{P}_1 \cap (x_{r,n}, \widetilde{x}_{r,n}) = \emptyset$$

  and $|f(x_{r,n}^+) - f(\widetilde{x}_{r,n})| < 2^{-n}\varepsilon$. Then we set $\mathcal{P}_2 = \mathcal{P}_1 \cup \{\widetilde{x}_{l,n}, \widetilde{x}_{r,n}; n \leq N\}$.
- For $K = K(\mathcal{P}_2)$, we let $\{\{x_k, y_k\}; k \leq K\}$ be the set of all pairs $\{x, y\} \subset \mathcal{P}_2$ satisfying $x < y$ with $(x,y) \cap \mathcal{P}_2 = \emptyset$ and $f(x)f(y) < 0$, which are not of the form $\{\widetilde{x}_{l,n}, x_{l,n}\}$ or $\{x_{r,n}, \widetilde{x}_{r,n}\}$. Let $k \leq K$. If there exists $z_k^\circ \in (x_k, y_k)$ such that $|f(z_k^\circ)| < 2^{-k}\varepsilon$, then we just add z_k° to $\mathcal{P}_2$. If not, then at least one of the sets

  $$I_{k,l} := (x_k, y_k] \cap \{z; \,\mathrm{sign}(f(z)f(y_k)) = \,\mathrm{sign}(f(z^-)f(x_k)) = 1\}$$

  and

  $$I_{k,r} := [x_k, y_k) \cap \{z; \,\mathrm{sign}(f(z^+)f(y_k)) = \,\mathrm{sign}(f(z)f(x_k)) = 1\}$$

  must be non-empty (here $\mathrm{sign}(x)$ is the usual sign function taking the value of -1, 0, or 1, if $x < 0$, $x = 0$, or $x > 0$, respectively). Assume $I_{k,l} \neq \emptyset$ (the other case is similar) and choose $z_k \in I_{k,l}$. Then $z_k = x_{l,n}$ for some $n > N$. We find $\widetilde{z}_k \in (x_k, z_k)$ such that $|f(\widetilde{z}_k) - f(z_k^-)| < 2^{-k}\varepsilon$ (in particular, we have $\mathrm{sign}(f(x_k)) = \,\mathrm{sign}(f(\widetilde{z}_k))$), and add both z_k and $\widetilde{z}_k$ to $\mathcal{P}_2$. The above process terminates after K steps and we denote the final collection of points by $\widetilde{\mathcal{P}}$.

Having constructed $\widetilde{\mathcal{P}}$ we see that

$$\begin{aligned}\text{Var}_{(a,b)}(f) - \text{Var}_{(a,b)}(|f|) &\geq \text{Var}(f, \widetilde{\mathcal{P}}) - \text{Var}(|f|, \widetilde{\mathcal{P}}) - \varepsilon \\ &\geq \sum_{n=1}^{N} |f(x_{l,n}) - f(\widetilde{x}_{l,n})| - \big||f|(x_{l,n}) - |f|(\widetilde{x}_{l,n})\big| \\ &\quad + \sum_{n=1}^{N} |f(x_{r,n}) - f(\widetilde{x}_{r,n})| - \big||f|(x_{r,n}) - |f|(\widetilde{x}_{r,n})\big| - \varepsilon \\ &\geq \sum_{n=1}^{N} |f(x_{l,n}) - f(x_{l,n}^{-})| - \big||f|(x_{l,n}) - |f|(x_{l,n}^{-})\big| \\ &\quad + \sum_{n=1}^{N} |f(x_{r,n}) - f(x_{r,n}^{+})| - \big||f|(x_{r,n}) - |f|(x_{r,n}^{+})\big| - 5\varepsilon.\end{aligned}$$

Also, we obtain

$$\begin{aligned}\text{Var}_{(a,b)}(f) - \text{Var}_{(a,b)}(|f|) &\leq \text{Var}(f, \widetilde{\mathcal{P}}) - \text{Var}(|f|, \widetilde{\mathcal{P}}) + \varepsilon \\ &\leq \sum_{n=1}^{\infty} |f(x_{l,n}) - f(x_{l,n}^{-})| - \big||f|(x_{l,n}) - |f|(x_{l,n}^{-})\big| \\ &\quad + \sum_{n=1}^{\infty} |f(x_{r,n}) - f(x_{r,n}^{+})| - \big||f|(x_{r,n}) - |f|(x_{r,n}^{+})\big| + 6\varepsilon,\end{aligned}$$

since the only terms that contributes to $\text{Var}(f, \widetilde{\mathcal{P}}) - \text{Var}(|f|, \widetilde{\mathcal{P}})$ are those corresponding to the pairs $\{\widetilde{x}_{l,n}, x_{l,n}\}$, $\{x_{r,n}, \widetilde{x}_{r,n}\}$, $\{x_k, z_k^{\circ}\}$, $\{z_k^{\circ}, y_k\}$ and $\{z_k, \widetilde{z}_k\}$. Letting $N \to \infty$ and $\varepsilon \to 0$, we obtain the claim. ■

Now, we use Lemma 5.2.3 to show that the map $f \mapsto \text{Var}_{(a,b)}(|f|)$ is continuous from $BV(\mathbb{R})$ to $[0, \infty)$.

Lemma 5.2.4. *Fix $f \in BV(\mathbb{R})$ and let $\{f_j; j \in \mathbb{N}\} \subset BV(\mathbb{R})$ be such that $\lim_{j\to\infty} \|f_j - f\|_{BV} = 0$. Then for any $-\infty \leq a < b \leq \infty$ we have*

$$\lim_{j\to\infty} \text{Var}_{(a,b)}(|f_j|) = \text{Var}_{(a,b)}(|f|).$$

Proof It is possible to verify that $f_j \to f$ implies $\text{Var}_{(a,b)}(f_j) \to \text{Var}_{(a,b)}(f)$. Thus, it remains to show

$$\lim_{j\to\infty} \text{Var}_{(a,b)}(f_j) - \text{Var}_{(a,b)}(|f_j|) = \text{Var}_{(a,b)}(f) - \text{Var}_{(a,b)}(|f|).$$

We define $\{x_{l,n}; n \in \mathbb{N}\}$ and $\{x_{r,n}; n \in \mathbb{N}\}$ as in the previous lemma. Given $\varepsilon > 0$ we choose $N \in \mathbb{N}$ such that

$$\sum_{n=N+1}^{\infty} |f(x_{l,n}) - f(x_{l,n}^{-})| + |f(x_{r,n}) - f(x_{r,n}^{+})| < \varepsilon.$$

We also denote $E_l^{j,N} := E_l(f_j) \cap \{x_{l,1}, \dots, x_{l,N}\}$ and $E_r^{j,N} := E_r(f_j) \cap \{x_{r,1}, \dots, x_{r,N}\}$. By Lemma 5.2.2 we have that for j big enough

$$\Big|\Big(\sum_{E_l^{j,N}} |f_j(x) - f_j(x^-)| - \big||f_j|(x) - |f_j|(x^-)\big| - |f(x) - f(x^-)| + \big||f|(x) - |f|(x^-)\big|\Big)\Big| < \varepsilon$$

and

$$\Big|\Big(\sum_{E_r^{j,N}} |f_j(x) - f_j(x^+)| - \big||f_j|(x) - |f_j|(x^+)\big| - |f(x) - f(x^+)| + \big||f|(x) - |f|(x^+)\big|\Big)\Big| < \varepsilon.$$

Moreover, we have

$$\begin{aligned} 0 &\le \sum_{x \in \big(E_l(f_j) \cap (a,b)\big) \setminus E_l^{j,N}} |f_j(x) - f_j(x^-)| - \big||f_j|(x) - |f_j|(x^-)\big| \\ &\le \sum_{x \in \big(E_l(f_j) \cap (a,b)\big) \setminus E_l^{j,N}} |f_j(x) - f_j(x^-)| \\ &\le \sum_{x \in \big(E_l(f_j) \cap (a,b)\big) \setminus E_l^{j,N}} |f(x) - f(x^-)| + 4\|f - f_j\|_{BV} < 2\varepsilon \end{aligned}$$

and, similarly,

$$0 \le \sum_{x \in \big(E_r(f_j) \cap (a,b)\big) \setminus E_r^{j,N}} |f_j(x) - f_j(x^+)| - \big||f_j|(x) - |f_j|(x^+)\big| < 2\varepsilon.$$

Finally, we observe that $\{x_{l,1}, \dots, x_{l,N}\} \subset E_l(f_j)$ and $\{x_{r,1}, \dots, x_{r,N}\} \subset E_r(f_j)$ for j big enough, by the uniform convergence. Letting $\varepsilon \to 0$ (and thus $N \to \infty$) and applying Lemma 5.2.3, we obtain the claim. ■

Let us now take a closer look at the properties of the maximal operator. Recall that the total variation of $\widetilde{M}f$ can be controlled by the total variation of f. There is also a local version of this principle, where we focus on an interval (a, b). However in this case some boundary terms must be included. Thus, to avoid the possibility that f behaves badly at a or b, we use its adjusted version $\overline{|f|}$ defined by

$$\overline{|f|}(x) := \limsup_{I \ni x; |I| \to 0} \fint_I |f|.$$

It is known that $\overline{|f|}$ is upper semicontinuous and that $\overline{|f|} \le \widetilde{M}f$ (see [APL07, Lemma 3.3]).

Lemma 5.2.5. *Fix $f \in BV(\mathbb{R})$. Given $-\infty \le a < b \le \infty$, we have*

$$\mathrm{Var}_{(a,b)}\big(\widetilde{M}f\big) \le \mathrm{Var}_{(a,b)}\big(\overline{|f|}\big) + \big|\widetilde{M}f(a) - \overline{|f|}(a)\big| + \big|\widetilde{M}f(b) - \overline{|f|}(b)\big|.$$

Proof This follows by a slight modification of the proof of [APL07, Lemma 3.9]. ∎

The next result gives us the uniform control (with respect to j) on the behavior of $\big(\widetilde{M}f_j\big)'$ near infinity, provided that $\{f_j\}_{j\in\mathbb{N}}$ is a converging sequence in $BV(\mathbb{R})$. This, in turn, allows one to restrict the attention to a bounded interval, while dealing with the total variations of the maximal functions $\widetilde{M}f_j$. We point out that it is also possible to proceed without this reduction, but then for all considered functions the extended domain $[-\infty,\infty]$ should be used instead of $\mathbb{R}$.

Lemma 5.2.6. *Fix $f \in BV(\mathbb{R})$ and let $\{f_j; j\in\mathbb{N}\} \subset BV(\mathbb{R})$ be such that $\lim_{j\to\infty}\|f_j - f\|_{BV} = 0$. Then for any $\varepsilon > 0$ there exist $-\infty < a < b < \infty$ such that*

$$\int_{\mathbb{R}\setminus(a,b)} \left|\big(\widetilde{M}f\big)'\right| < \varepsilon$$

and

$$\int_{\mathbb{R}\setminus(a,b)} \left|\big(\widetilde{M}f_j\big)'\right| < \varepsilon,$$

for every j big enough.

Proof We prove that there exists $b<\infty$ such that $\int_{(b,\infty)} \big|\big(\widetilde{M}f\big)'\big| < \varepsilon$, the symmetric case is treated analogously. First we deal with the case where $\widetilde{M}f(\infty) > |f|(\infty)$. Assume that $\widetilde{M}f(\infty) - |f|(\infty) > 4\varepsilon$. Let us take b big enough such that (we use here Lemma 5.2.2) we have $\big|\widetilde{M}f(x) - \widetilde{M}f(\infty)\big| < \varepsilon$ and $\big||f|(x) - |f|(\infty)\big| < \varepsilon$ for every $x \in (b,\infty)$. Therefore, for j big enough such that $\||f_j| - |f|\|_\infty \le \frac{\varepsilon}{2}$ and $\|\widetilde{M}f_j - \widetilde{M}f\|_\infty \le \frac{\varepsilon}{2}$, for each $y\in(b,x)$ we have $\widetilde{M}f_j(y) \ge \widetilde{M}f_j(x)$. This is the case because any interval $I \ni x$ satisfying $⨍_I |f_j| > \widetilde{M}f_j(x) - \frac{\varepsilon}{2}$ contains y, since if $I\subset(y,\infty)$, in particular $I \subset (b,\infty)$, and then

$$⨍_I |f_j| \le ⨍_I |f| + \frac{\varepsilon}{2} \le |f|(\infty) + \frac{3\varepsilon}{2} \le \widetilde{M}f(\infty) - 2\varepsilon \le \widetilde{M}f(x) - \varepsilon \le \widetilde{M}f_j(x) - \frac{\varepsilon}{2}.$$

Therefore

$$\begin{aligned}\int_{(b,\infty)} \left|\big(\widetilde{M}f_j\big)'\right| &= \widetilde{M}f_j(b) - \widetilde{M}f_j(\infty)\\ &\le \big|\widetilde{M}f(b) - \widetilde{M}f_j(b)\big| + \big|\widetilde{M}f_j(\infty) - \widetilde{M}f(\infty)\big| + \big|\widetilde{M}f(b) - \widetilde{M}f(\infty)\big| \le 3\varepsilon\end{aligned}$$

for j big enough, from where we conclude this case.

Now, we deal with the case where $|f|(\infty) = \widetilde{M}f(\infty)$. By Lemma 5.2.5 and [APL07, Lemma 3.3], assuming that b is a continuity point for f_j, we obtain

$$\begin{aligned}\int_{(b,\infty)} \left|\big(\widetilde{M}f_j\big)'\right| &\le \mathrm{Var}_{(b,\infty)}\big(\overline{|f_j|}\big) + \big|\widetilde{M}f_j(b) - \overline{|f_j|}(b)\big| + \big|\widetilde{M}f_j(\infty) - \overline{|f_j|}(\infty)\big|\\ &\le \mathrm{Var}_{(b,\infty)}(|f_j|) + \big|\widetilde{M}f_j(b) - |f_j|(b)\big| + \big|\widetilde{M}f_j(\infty) - |f_j|(\infty)\big|.\end{aligned}$$

The analogous is obtained for f instead of f_j. Let us assume that b is a continuity point for f and every f_j, such that $\text{Var}_{(b,\infty)}\big(|f|\big) < \varepsilon$. By Lemma 5.2.4 we have

$$\text{Var}_{(b,\infty)}(|f_j|) < 2\varepsilon, \tag{5.1}$$

for j big enough. Also,

$$\big|\widetilde{M}f_j(b) - |f_j|(b)\big| \leq \big|\widetilde{M}f(b) - |f|(b)\big| + |\widetilde{M}f_j(b) - \widetilde{M}f(b)| + \big||f_j|(b) - |f|(b)\big|.$$

If b is big enough to have $\big|\widetilde{M}f(b) - |f|(b)\big| < \varepsilon$, then by Lemma 5.2.2 we get $\big|\widetilde{M}f_j(b) - |f_j|(b)\big| < 2\varepsilon$ and $\big|\widetilde{M}f_j(\infty) - |f_j|(\infty)\big| < \varepsilon$ for j big enough. Combining this with (5.1) concludes the proof. ■

5.3 Main tools: variation convergence and pointwise derivative analysis

This section is the core of this chapter, here we develop the main tools that lead us to our desired result. Before we prove our key result regarding the variation of the maximal functions, we need the following definition. A given partition $\mathcal{P} = \{a_0 < a_1 < \cdots < a_n\}$, with $n \geq 2$, has property (V) with respect to f if for each $i \in \{0, 1, \ldots, n-2\}$, we have $\text{sign}(f(a_{i+2}) - f(a_{i+1})) \cdot \text{sign}(f(a_{i+1}) - f(a_i)) < 0$.

Proposition 5.3.1. *Fix $f \in BV(\mathbb{R})$ and let $\{f_j; j \in \mathbb{N}\} \subset BV(\mathbb{R})$ be such that $\lim_{j\to\infty} \|f_j - f\|_{BV} = 0$. Then*

$$\text{Var}_{(-\infty,\infty)}\big(\widetilde{M}f_j\big) \to \text{Var}_{(-\infty,\infty)}\big(\widetilde{M}f\big).$$

Proof By Lemma 5.2.6 it is enough to prove that $\text{Var}_{(a,b)}\big(\widetilde{M}f_j\big) \to \text{Var}_{(a,b)}\big(\widetilde{M}f\big)$ for every interval $(a,b) \subset \mathbb{R}$ with both a and b being points of continuity for f and every f_j. In the following we fix $-\infty < a < b < \infty$ satisfying such assumption. Observe that Lemma 5.2.2 and Fatou's lemma imply

$$\liminf_{j\to\infty} \text{Var}_{(a,b)}\big(\widetilde{M}f_j\big) \geq \text{Var}_{(a,b)}\big(\widetilde{M}f\big).$$

Now, we prove the remaining inequality, that is,

$$\limsup_{j\to\infty} \text{Var}_{(a,b)}\big(\widetilde{M}f_j\big) \leq \text{Var}_{(a,b)}\big(\widetilde{M}f\big).$$

Given $\varepsilon > 0$ we show that

$$\text{Var}_{(a,b)}\big(\widetilde{M}f_j\big) < \text{Var}_{(a,b)}\big(\widetilde{M}f\big) + 4\varepsilon$$

holds if j is big enough. Let $\mathcal{P} = \{a = a_0 < a_1 < \cdots < a_K = b\} \subset \mathbb{R}$, $K \in \mathbb{N}$, be a partition satisfying

$$\mathrm{Var}\,\big(|f|, \mathcal{P}\big) > \mathrm{Var}_{\,(a,b)}\big(|f|\big) - \varepsilon$$

and

$$\mathrm{Var}\,\big(\widetilde{M}f, \mathcal{P}\big) > \mathrm{Var}_{\,(a,b)}\big(\widetilde{M}f\big) - \varepsilon.$$

Also, by the uniform convergence and Lemma 5.2.4 we conclude that

$$\mathrm{Var}\,\big(|f_j|, \mathcal{P}\big) > \mathrm{Var}_{\,(a,b)}\big(|f_j|\big) - 2\varepsilon \tag{5.2}$$

and

$$\mathrm{Var}\,\big(\widetilde{M}f_j, \mathcal{P}\big) > \mathrm{Var}_{\,(a,b)}\big(\widetilde{M}f\big) - 2\varepsilon \tag{5.3}$$

hold for j big enough. Now, we take $\widetilde{\mathcal{P}} = \widetilde{\mathcal{P}}(j)$ such that $\mathcal{P} \subset \widetilde{\mathcal{P}} \subset [a, b]$ and

$$\mathrm{Var}\,\big(\widetilde{M}f_j, \widetilde{\mathcal{P}}\big) > \mathrm{Var}_{\,(a,b)}\big(\widetilde{M}f_j\big) - \varepsilon.$$

Without loss of generality we can assume that $\widetilde{\mathcal{P}}$ is such that for each $i \in \{1, \ldots, K\}$ the set $\widetilde{\mathcal{P}} \cap [a_{i-1}, a_i] = \{a_{i-1} = a_{i,0} < \cdots < a_{i,n_i} = a_i\}$ satisfies property (V) with respect to $\widetilde{M}f_j$ unless it consists of two elements. For each such i we denote $\widetilde{\mathcal{P}}_i = \{a_{i,1}, \ldots, a_{i,n_i-1}\}$ and claim that it is possible to find a partition $\widetilde{\mathcal{P}}_i^* = \{a^*_{i,1}, \ldots, a^*_{i,n_i-1}\} \subset (a_{i-1}, a_i)$ such that

$$\mathrm{Var}\,\Big(|f_j|, \widetilde{\mathcal{P}}_i^*\Big) - \mathrm{Var}\,\Big(|f_j|, \{a^*_{i,1}, a^*_{i,n_i-1}\}\Big) > \mathrm{Var}\,\Big(\widetilde{M}f_j, \widetilde{\mathcal{P}}_i\Big) - \mathrm{Var}\,\Big(\widetilde{M}f_j, \{a_{i,1}, a_{i,n_i-1}\}\Big) - \frac{\varepsilon}{K}.$$

Indeed, for $n_i \leq 2$ we use the convention that all the variation terms above are equal to 0, so the inequality holds (we set $\widetilde{\mathcal{P}}_i^* = \emptyset$ or $\widetilde{\mathcal{P}}_i^* = \{a_{i,1}\}$ if $n = 1$ or $n = 2$, respectively). It remains to consider the case $n_i \geq 3$ in which property (V) is guaranteed. We assume that $\widetilde{M}f_j(a_{i,0}) < \widetilde{M}f_j(a_{i,1})$ (the opposite case can be treated analogously). Then $\widetilde{\mathcal{P}}_i^*$ shall be chosen in such a way that given $k \in \{1, \ldots, n_i - 1\}$ we have

$$|f_j|(a^*_{i,k}) > \max\Big\{\widetilde{M}f_j(a_{i,k-1}), \widetilde{M}f_j(a_{i,k}) - \frac{\varepsilon}{2n_iK}, \widetilde{M}f_j(a_{i,k+1})\Big\},$$

for k odd, and

$$|f_j|(a^*_{i,k}) \leq \widetilde{M}f_j(a_{i,k})$$

for k even. We describe in detail the procedure for selecting the points $a^*_{i,k}$. If k is odd, then we find an interval $I \ni a_{i,k}$ such that

$$⨍_I |f_j| > \max\Big\{\widetilde{M}f_j(a_{i,k-1}), \widetilde{M}f_j(a_{i,k}) - \frac{\varepsilon}{2n_iK}, \widetilde{M}f_j(a_{i,k+1})\Big\}.$$

Clearly, $I \subset (a_{i,k-1}, a_{i,k+1})$ and we can find $a_{i,k}^* \in I$ satisfying $|f_j|(a_{i,k}^*) \geq \fint_I |f_j|$. For k even we take $I = (a_{i,k-1}^*, a_{i,k+1}^*)$ if $k \neq n_i - 1$ or $I = (a_{i,n_i-2}^*, a_i)$ otherwise. Since $\fint_I |f_j| \leq \widetilde{M} f_j(a_{i,k})$, there exists $a_{i,k}^* \in I$ satisfying $|f_j|(a_{i,k}^*) \leq \widetilde{M} f_j(a_{i,k})$. We note that the appropriate configuration of the sets I guarantees that the inequalities $a_{i-1} < a_{i,1}^* < \cdots < a_{i,n_i-1}^* < a_i$ hold.

Observe, also, that the partition $\{a_{i,1}^* < \cdots < a_{i,n_i-1}^*\}$ either consists of 2 elements or satisfies property (V) with respect to $|f_j|$. Thus, regardless of which case occurs, we obtain

$$\mathrm{Var}\left(|f_j|, \widetilde{\mathcal{P}}_i^*\right) - \mathrm{Var}\left(|f_j|, \{a_{i,1}^*, a_{i,n_i-1}^*\}\right) = \sum_{k=1}^{n_i-1} \alpha_k |f_j|(a_{i,k}^*),$$

where $\alpha_k = 2(-1)^{k+1}$ for $k \in \{2, \ldots, n_i - 2\}$ and $\alpha_k \in \left\{0, 2(-1)^{k+1}\right\}$ for $k \in \{1, n_i - 1\}$ (the boundary values depend on $\operatorname{sign}\big(|f_j|(a_{i,1}^*) - |f_j|(a_{i,n_i}^*)\big)$ and the parity of n_i). Similarly,

$$\mathrm{Var}\left(\widetilde{M} f_j, \widetilde{\mathcal{P}}_i\right) - \mathrm{Var}\left(\widetilde{M} f_j, \{a_{i,1}, a_{i,n_i-1}\}\right) \leq \sum_{k=1}^{n_i-1} \alpha_k \widetilde{M} f_j(a_{i,k})$$

(we eventually change the sign of the second term on the left-hand side in order to get the boundary coefficients equal to α_1 and α_{n_i-1}). Consequently, the claim follows since for each k we have

$$\alpha_k \left(|f_j|(a_{i,k}^*) - \widetilde{M} f_j(a_{i,k})\right) \geq \frac{-\varepsilon}{n_i K}.$$

Now, we apply the claim in order to get the following estimate

$$\begin{aligned}
\mathrm{Var}_{(a,b)}(|f_j|) - \mathrm{Var}(|f_j|, \mathcal{P}) &\geq \mathrm{Var}\left(|f_j|, \mathcal{P} \cup \bigcup_{i=1}^{K} \widetilde{\mathcal{P}}_i^*\right) - \mathrm{Var}\left(|f_j|, \mathcal{P} \cup \bigcup_{i=1}^{K} \{a_{i,1}^*, a_{i,n_i-1}^*\}\right) \\
&= \sum_{i=1}^{K} \mathrm{Var}\left(|f_j|, \widetilde{\mathcal{P}}_i^*\right) - \mathrm{Var}\left(|f_j|, \{a_{i,1}^*, a_{i,n_i-1}^*\}\right) \\
&\geq \sum_{i=1}^{K} \mathrm{Var}\left(\widetilde{M} f_j, \widetilde{\mathcal{P}}_i\right) - \mathrm{Var}\left(\widetilde{M} f_j, \{a_{i,1}, a_{i,n_i-1}\}\right) - \frac{\varepsilon}{K} \\
&\geq \mathrm{Var}\left(\widetilde{M} f_j, \widetilde{\mathcal{P}}\right) - \mathrm{Var}\left(\widetilde{M} f_j, \widetilde{\widetilde{\mathcal{P}}}\right) - \varepsilon,
\end{aligned}$$

where $\widetilde{\widetilde{\mathcal{P}}} := \{a_{i,k}; i \in \{1, \ldots, K\}, k \in \{0, 1, n_i - 1, n_i\}\}$. In particular, we note that $\widetilde{\widetilde{\mathcal{P}}}$ consists of at most $3K + 1$ elements and thus

$$\mathrm{Var}\left(\widetilde{M} f_j, \widetilde{\widetilde{\mathcal{P}}}\right) < \mathrm{Var}\left(\widetilde{M} f, \widetilde{\widetilde{\mathcal{P}}}\right) + 12K \|f_j - f\|_{BV} < \mathrm{Var}_{(a,b)}\big(\widetilde{M} f\big) + \varepsilon$$

follows by Lemma 5.2.2 for j big enough. Combining the above inequalities with (5.3), we arrive at

$$\mathrm{Var}_{(a,b)}(|f_j|) - \mathrm{Var}(|f_j|, \mathcal{P}) > \mathrm{Var}_{(a,b)}\big(\widetilde{M} f_j\big) - \mathrm{Var}_{(a,b)}\big(\widetilde{M} f\big) - 4\varepsilon,$$

which, in view of (5.2), gives

$$\operatorname{Var}_{(a,b)}\big(\widetilde{M}f_j\big) < \operatorname{Var}_{(a,b)}\big(\widetilde{M}f\big) + 6\varepsilon,$$

provided that j is big enough. Consequently, $\lim_{j\to\infty} \operatorname{Var}_{(a,b)}\big(\widetilde{M}f_j\big) = \operatorname{Var}_{(a,b)}\big(\widetilde{M}f\big)$. ■

Having obtained Proposition 5.3.1 we continue with the remaining tools required. Our general purpose in the next few lemmas is to get more information about the derivative of the maximal function. In particular, we are interested in studying the behavior of the sequence $\big\{\big(\widetilde{M}f_j\big)'(x)\big\}_{j\in\mathbb{N}}$ for a given point x.

Lemma 5.3.1. *Fix $f \in BV(\mathbb{R})$ and let $\{f_j; j \in \mathbb{N}\} \subset BV(\mathbb{R})$ be such that $\lim_{j\to\infty} \|f_j - f\|_{BV} = 0$. If $⨍_{I_{x,j}} |f_j| = \widetilde{M}f_j(x)$ with $I_{x,j} \ni x$, and $\chi_{I_{x,j}} \to \chi_I$ a.e. with $0 < |I| < \infty$, then we have $⨍_I |f| = \widetilde{M}f(x)$.*

Proof Follows a slight modification in [CMP17, Lemma 12]. ■

Let us now define $D := \big\{x \in \mathbb{R}; \widetilde{M}f(x) > \overline{|f|}(x)\big\}$. This is a slight modification of the disconnecting set used in [CMP17] and in our proofs the role of D is very similar to the role of its prototype in [CMP17]. Since $\widetilde{M}f$ is absolutely continuous and $\overline{|f|}$ is upper semicontinuous, we have that D is open. We notice that if $x \in D \setminus \big\{x; \widetilde{M}f(x) = \widetilde{M}f(\infty)\big\}$, then there exists a finite interval $I \ni x$ such that $⨍_I |f|(x) = \widetilde{M}f(x)$. Indeed, there exists a sequence $\{(a_k, b_k)\}_{k\in\mathbb{N}}$ such that $x \in (a_k, b_k)$ and $⨍_{(a_k,b_k)} |f| \to \widetilde{M}f(x)$. Since $\widetilde{M}f(x) > \big\{\overline{|f|}(x), \widetilde{M}f(\infty)\big\}$, we have $\{b_k - a_k; k \in \mathbb{N}\} \subset (\epsilon, \epsilon^{-1})$ for some $\epsilon > 0$. Thus, by taking a subsequence (if required), we get $a_k \to a$ and $b_k \to b$, with $b-a \in (0, \infty)$. By the boundedness of f we conclude that $⨍_{[a,b]} |f| = \widetilde{M}f(x)$. Also, let us observe that $\max\big\{⨍_{[a,x]} |f|, ⨍_{[x,b]} |f|\big\} \geq ⨍_{[a,b]} |f|$, therefore $\widetilde{M}f(x) = ⨍_{[a,x]} |f|$ or $\widetilde{M}f(x) = ⨍_{[x,b]} |f|$.

The next result states that for a.e. $x \in D$ the derivative of the maximal function $\widetilde{M}f$ can be described by an explicit formula.

Lemma 5.3.2. *Let $f \in BV(\mathbb{R})$. Assume that $\widetilde{M}f$ is differentiable and $|f|$ is continuous at x (that happens a.e. because $\widetilde{M}f$ and $|f|$ have bounded variation). Let us suppose that $x \in D$ is such that there exists an interval $I_x \ni x$ with $|I_x| < \infty$ such that $⨍_{I_x} |f| = \widetilde{M}f(x)$ and $I_x \subset [x, \infty)$ or $I_x \subset (-\infty, x]$. Then*

$$\big(\widetilde{M}f\big)'(x) = \begin{cases} \frac{\int_{I_x}|f|}{|I_x|^2} - \frac{|f|(x)}{|I_x|} = \frac{1}{|I_x|}\Big(\widetilde{M}f(x) - |f|(x)\Big) & \text{if } I_x \subset [x,\infty), \\ \frac{|f|(x)}{|I_x|} - \frac{\int_{I_x}|f|}{|I_x|^2} = \frac{1}{|I_x|}\Big(|f|(x) - \widetilde{M}f(x)\Big) & \text{otherwise.} \end{cases}$$

Also, if $\widetilde{M}f(x) = \widetilde{M}f(\infty)$, then we have $\big(\widetilde{M}f\big)'(x) = 0$.

Proof The last claim follows because x is a local minimum of $\widetilde{M}f$. Assume without loss of generality that $I_x = (x, a_x)$, $a_x > x$ (the other case is similar). We have, for $h > 0$, that

$$\frac{\widetilde{M}f(x) - \widetilde{M}f(x-h)}{h} \leq \frac{\fint_{(x,a_x)} |f| - \fint_{(x-h,a_x)} |f|}{h}$$
$$= \frac{\frac{\int_x^{a_x} |f|}{a_x - x} - \frac{\int_x^{a_x} |f| + \int_{x-h}^{x} |f|}{a_x - x + h}}{h} \to \frac{\int_{I_x} |f|}{(a_x - x)^2} - \frac{|f|(x)}{a_x - x}$$

as $h \to 0$. Therefore $\big(\widetilde{M}f\big)'(x) \leq \frac{\int_{I_x} |f|}{(a_x - x)^2} - \frac{|f(x)|}{a_x - x}$. Also, for $h > 0$ we have

$$\frac{\widetilde{M}f(x+h) - \widetilde{M}f(x)}{h} \geq \frac{\frac{\int_{x+h}^{a_x} |f|}{a_x - x - h} - \frac{\int_x^{a_x} |f|}{a_x - x}}{h} \to \frac{\int_{I_x} |f|}{(a_x - x)^2} - \frac{|f|(x)}{a_x - x}$$

as $h \to 0$. This concludes the proof. ■

Now, we can use the obtained formula to prove the following result regarding pointwise convergence.

Lemma 5.3.3. *Fix $f \in BV(\mathbb{R})$ and let $\{f_j; j \in \mathbb{N}\} \subset BV(\mathbb{R})$ be such that $\lim_{j\to\infty} \|f_j - f\|_{BV} = 0$. Then*

$$\big(\widetilde{M}f_j\big)' \to \big(\widetilde{M}f\big)'$$

a.e. in D.

Proof The claim is trivial if D has measure zero, so assume this is not the case. We define D_j as the analogue of D for f_j. Let us take $x \in D$ such that $\widetilde{M}f_j$ and $\widetilde{M}f$ are differentiable at x for every j and f and f_j is continuous at x. By the uniform convergence we have that $x \in D_j$ for j big enough. We also make the following observation. If there are intervals $I_{x,j} \ni x$ such that $\fint_{I_{x,j}} |f_j| = \widetilde{M}f_j(x)$, then the quantities $|I_{x,j}|$ are bounded below uniformly. Indeed, if for a sequence $\{j_k\}_{k\in\mathbb{N}}$ we have $|I_{x,j_k}| \to 0$, then we would have $\fint_{I_{x,j_k}} |f_{j_k}| \to |f|(x) < \widetilde{M}f(x)$ by the uniform convergence and continuity of f at x, contradicting the pointwise convergence of the maximal functions.

Assume first that $x \in D \setminus \big\{y; \widetilde{M}f(y) = \widetilde{M}f(\infty)\big\}$ and take $\varepsilon > 0$ such that $\widetilde{M}f(x) > \widetilde{M}f(\infty) + 2\varepsilon$. Then for j big enough we have $\widetilde{M}f_j(x) > \widetilde{M}f_j(\infty) + \varepsilon$. Also, there exists $N > |x|$ such that for j big enough and each $y \in \mathbb{R} \setminus [-N, N]$ we have $|f_j|(y) < \widetilde{M}f(\infty) + \varepsilon < \widetilde{M}f_j(x)$. We can observe then that $I_{x,j} \subset [-N, N]$ for j big enough. Let us assume that we have $\delta > 0$ and a sequence $\{j_k\}_{k\in\mathbb{N}}$ such that

$$\Big|\big(\widetilde{M}f_{j_k}\big)'(x) - \big(\widetilde{M}f\big)'(x)\Big| > \delta. \tag{5.4}$$

Without loss of generality assume that $I_{j_k} = (x, a_{j_k})$ (the other case is treated analogously). Since $x < a_{j_k} < N$, there exists a subsequence (that we also denote by j_k) such that $a_{j_k} \to a \in [x, N]$. Moreover, in view of the previous observation, we have $a \neq x$. Thus, Lemma 5.3.1 gives $\fint_{(x,a)} |f| = \widetilde{M}f(x)$ and consequently, in view of Lemma 5.3.2, we obtain $\big(\widetilde{M}f\big)'(x) = \frac{\int_{(x,a)}|f|}{(a-x)^2} - \frac{|f|(x)}{a-x}$. Also, $\big(\widetilde{M}f_{j_k}\big)'(x) = \frac{\int_{I_{x,j}}|f_{j_k}|}{(a_{j_k}-x)^2} - \frac{|f_{j_k}|(x)}{a_{j_k}-x}$ holds. However, by the uniform convergence we have

$$\frac{\int_{I_{x,j}}|f_{j_k}|}{(a_{j_k}-x)^2} - \frac{|f_{j_k}|(x)}{a_{j_k}-x} \to \frac{\int_{(x,a)}|f|}{(a-x)^2} - \frac{|f|(x)}{a-x},$$

reaching a contradiction with (5.4). Thus, we conclude this case.

Now, if $\widetilde{M}f(x) = \widetilde{M}f(\infty)$, then by Lemma 5.3.2 we have $\big(\widetilde{M}f\big)'(x) = 0$. Also, if for a subsequence j_k we have $\widetilde{M}f_{j_k}(x) = \widetilde{M}f_{j_k}(\infty)$, then $\big(\widetilde{M}f_{j_k}\big)'(x) = 0$. Therefore, this subcase follows and we can assume that $x \in D_j \setminus \big\{\widetilde{M}f_j(x) = \widetilde{M}f_j(\infty)\big\}$. It is now enough to prove that $\frac{\int_{I_{x,j}}|f_j|}{(a_j-x)^2} - \frac{|f_j|(x)}{a_j-x} \to 0$. Let us suppose that for some $\delta > 0$ and a subsequence j_k we have $\big|\big(\widetilde{M}f_{j_k}\big)'\big| > \delta$. As before, we assume the case $I_{j_k} = (x, a_{j_k})$. We claim that there exists $C(\delta, f) > 0$ such that for j_k big enough we have $|I_{x,j_k}| < C(\delta, f) < \infty$. Indeed, in view of

$$\frac{2\|f_{j_k}\|_\infty}{|I_{x,j_k}|} > \left|\frac{\int_{I_{x,j_k}}|f_{j_k}|}{(a_{j_k}-x)^2} - \frac{|f_{j_k}|(x)}{a_{j_k}-x}\right| > \delta,$$

we have $\frac{2\|f_{j_k}\|_\infty}{\delta} > |I_{x,j_k}|$ and thus $\|f_{j_k}\|_\infty \to \|f\|_\infty$ gives our claim. Now, since $|I_{x,j_k}| < C(\delta, f)$, we have that $a_{j_k} \in (x, x + C(\delta, f))$ for j big enough. Consequently, there exists a subsequence (that we also denote by j_k) such that $a_{j_k} \to a$ for some $a \in (x, x + C(\delta, f)]$. Then by Lemma 5.3.1 we have that $\fint_{(x,a)} |f| = \widetilde{M}f(x)$. Therefore, Lemma 5.3.2 gives

$$\frac{\int_{(x,a)}|f|}{(a-x)^2} - \frac{|f|(x)}{a-x} = \big(\widetilde{M}f\big)'(x)$$

and the left-hand side must be equal to 0. Since we have

$$\big(\widetilde{M}f_{j_k}\big)'(x) = \frac{\int_{I_{x,j_k}}|f_{j_k}|}{(a_{j_k}-x)^2} - \frac{|f_{j_k}|(x)}{a_{j_k}-x} \to \frac{\int_{(x,a)}|f|}{(a-x)^2} - \frac{|f|(x)}{a-x} = 0$$

by the uniform convergence, we reach a contradiction. This concludes the proof. ■

It remains to take a look at the set $C := \mathbb{R} \setminus D$. This set plays the role of the connecting set in [CMP17].

Lemma 5.3.4. *Let $f \in BV(\mathbb{R})$. Then for a.e. $x \in C$ we have $\big(\widetilde{M}f\big)'(x) = 0$.*

Proof Assume that $\overline{|f|}$ and $\widetilde{M}f$ are differentiable at x (this happens a.e. because $\overline{|f|}$ and $\widetilde{M}f$ have bounded variation). Then, since $\widetilde{M}f(x) = \overline{|f|}(x)$ and $\widetilde{M}f \geq \overline{|f|}$, we have $\big(\widetilde{M}f\big)'(x) = \overline{|f|}'(x)$. Now, assume, in order to get a contradiction, that $\overline{|f|}'(x) > 0$ (the other case is analogous). Then there exist $h_0, L > 0$ such that $\overline{|f|}(x+h) \geq \overline{|f|}(x) + Lh$ for every $0 < h < h_0$. Thus, for a.e. $0 < h < h_0$ we have $|f|(x+h) \geq \overline{|f|}(x) + Lh$, which implies $\widetilde{M}f(x) \geq \frac{\int_0^{h_0} \overline{|f|}(x)+Lh}{h_0} = \overline{|f|}(x) + \frac{Lh_0}{2} > \overline{|f|}(x)$, a contradiction. ■

Combining the previous results we obtain the following.

Lemma 5.3.5. *Fix $f \in BV(\mathbb{R})$ and let $\{f_j; j \in \mathbb{N}\} \subset BV(\mathbb{R})$ be such that $\lim_{j\to\infty} \|f_j - f\|_{BV} = 0$. Then*

$$\left\| \left((\widetilde{M}f_j)' - (\widetilde{M}f)' \right) \chi_D \right\|_1 \to 0.$$

Proof By the classic Brezis–Lieb lemma [BL83], the boundedness of the map $f \mapsto \widetilde{M}f$ from $BV(\mathbb{R})$ to itself and Lemma 5.3.3, we just need to prove the following,

$$\left\| (\widetilde{M}f_j)'\chi_D \right\|_1 \to \left\| (\widetilde{M}f)'\chi_D \right\|_1. \tag{5.5}$$

By Fatou's Lemma, Proposition 5.3.1 and Lemma 5.3.4, we have

$$\begin{aligned} \int_D \left| (\widetilde{M}f)' \right| &\leq \liminf_{j\to\infty} \int_D \left| (\widetilde{M}f_j)' \right| \leq \limsup_{j\to\infty} \int_D \left| (\widetilde{M}f_j)' \right| \\ &\leq \lim_{j\to\infty} \int_{\mathbb{R}} \left| (\widetilde{M}f_j)' \right| = \int_{\mathbb{R}} \left| (\widetilde{M}f)' \right| = \int_D \left| (\widetilde{M}f)' \right|, \end{aligned}$$

from where (5.5) follows. ■

5.4 Proof of Theorem 5.1.1

Finally, we are ready to prove the main result. In what follows C_j denotes the set analogous to C defined for f_j instead of f.

Proof Since by Lemmas 5.2.1 and 5.2.2 we have $\widetilde{M}f_j(-\infty) \to \widetilde{M}f(-\infty)$, it remains to prove that

$$(\widetilde{M}f_j)' \to (\widetilde{M}f)'$$

in $L^1(\mathbb{R})$. We make the following claim

$$\int_{C\cap D_j} \left| (\widetilde{M}f_j)' \right| \to 0. \tag{5.6}$$

Indeed, by Proposition 5.3.1, Lemma 5.3.4 and Lemma 5.3.5 we have

$$\begin{aligned} \lim_{j\to\infty} \int_D \left| (\widetilde{M}f_j)' \right| = \int_D \left| (\widetilde{M}f)' \right| = \int_{\mathbb{R}} \left| (\widetilde{M}f)' \right| &\geq \limsup_{j\to\infty} \left(\int_{D_j\cap C} \left| (\widetilde{M}f_j)' \right| + \int_D \left| (\widetilde{M}f_j)' \right| \right) \\ &= \limsup_{j\to\infty} \int_{D_j\cap C} \left| (\widetilde{M}f_j)' \right| + \lim_{j\to\infty} \int_D \left| (\widetilde{M}f_j)' \right| \end{aligned}$$

and the claim follows. Consequently, by (5.6) and Lemma 5.3.5 we get

$$\begin{aligned}\int_{\mathbb{R}} \left|(\widetilde{M}f_j)' - (\widetilde{M}f)'\right| &= \int_{C\cap D_j} \left|(\widetilde{M}f_j)' - (\widetilde{M}f)'\right| + \int_{C\cap C_j} \left|(\widetilde{M}f_j)' - (\widetilde{M}f)'\right| \\ &\quad + \int_D \left|(\widetilde{M}f_j)' - (\widetilde{M}f)'\right| \\ &= \int_{C\cap D_j} \left|(\widetilde{M}f_j)'\right| + \int_D \left|(\widetilde{M}f_j)' - (\widetilde{M}f)'\right| \to 0\end{aligned}$$

as $j \to \infty$, from where we conclude our result. ■

5.4.1 Concluding remarks

We end our discussion by showing that the assumptions $f, f_j \in BV(\mathbb{R})$ are important, not only $f - f_j \in BV(\mathbb{R})$.

Example 5.4.1. Let $A = \cup_{k=1}^{\infty}(4k-2, 4k)$ and take

$$f = \chi_{(-\infty,0)\cup A}, \quad \text{and} \quad f_n = f + \frac{1}{n}\chi_{(0,4n+2)}.$$

Then we have $\|f_n - f\|_{BV} \to 0$, while $\|\widetilde{M}f_n - \widetilde{M}f\|_{BV} \not\to 0$.

Indeed, the first claim is obvious and for the second one we argue as follows. We observe that $\widetilde{M}f \equiv 1$ and $\widetilde{M}f_n(x) = 1 + \frac{1}{n}$ for $x \in \{3, 7, \ldots, 4n-1\}$. Moreover, if $n \geq 3$, then for any $x \in \{1, 5, \ldots, 4n+1\}$ we have

$$\widetilde{M}f_n(x) \leq \max\left\{1, \frac{2}{3} + \frac{1}{n}\right\} = 1,$$

which is due to the fact that for any interval $I \ni x$ we have $|I \cap A \cap (0, 4n+2)| \leq \frac{2}{3}|I \cap (0, 4n+2)|$. Thus, for $\mathcal{P}_n = \{1, 3, \ldots, 4n+1\}$ we have $\mathrm{Var}\left(\widetilde{M}f_n - \widetilde{M}f, \mathcal{P}_n\right) \geq 2n \cdot \frac{1}{n} \not\to 0$.

Chapter 6

On the continuity of convolution type maximal operators at the derivative level

6.1 Introduction

In this chapter we study continuity at the endpoint of the derivative of centered convolution type maximal operators. The following kernels are of major relevance for our purposes:

$$\varphi_1(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}} \, \frac{1}{(|x|^2+1)^{(d+1)/2}} \qquad \text{(Poisson kernel)}$$

$$\varphi_2(x) = \frac{1}{(4\pi)^{d/2}} \, e^{-|x|^2} \qquad \text{(Heat kernel)}$$

$$\varphi_3^\alpha(x) = C_d^\alpha \frac{1}{(|x|^2+1)^{(d+1-\alpha)/2}} \qquad \text{(Fractional Poisson kernel),}$$

where $0 < \alpha < 1$ and C_d^α is such that $\|\varphi_3^\alpha\|_1 = 1$. This last kernel was studied by Caffarelli and Silvestre in [CS07] where its relation to the fractional Laplacian was investigated. In this chapter, for a given kernel ϕ we write $\tilde{u}(x,t) = u * \phi_t(x)$ and $u^*(x) = \sup_{t>0} \tilde{u}(x,t)$.

In our main theorem we make use of the essential *subharmonicity property* that these kernels have to conclude the one-dimensional continuity of these maximal operators at the derivative level, solving a question suggested by Carneiro[1].

Theorem 6.1.1. *Let $\phi \in \{\varphi_1, \varphi_2, \varphi_3^\alpha\}$. Then the map*

$$u \mapsto (u^*)'$$

is continuous from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$.

[1]Personal communication.

In the case of φ_3^α we first have to prove that the aforementioned map is well defined and bounded. This is obtained by similar methods than the ones developed in [CFS18] and [CS13]. This is explained in §6.2.

The methods developed in the aforementioned works [CMP17] and in Chapter 4 and 5 are not enough to conclude Theorem 6.1.1. In those works it is relevant that the maximal operators considered there have the flatness property; that is, the maximal functions have zero derivative a.e. at the points where they coincide with the original functions. This property does not typically hold when dealing with centered maximal operators, so a new approach is required in this case. In order to overcome this difficulty, our strategy is strongly tied with the *subharmonicity property* that these kernels satisfy. We use this property in order to obtain a *local boundedness* that is stable under linear perturbations. This allows us to discretize some important aspects of the proof. Complementing this with some previous methods developed in Chapter 5 we obtain our result. These new tools are explained in §6.3.

6.2 Preliminaries

Here we develop the preliminaries for the proof of our theorem. Given $u \in W^{1,1}(\mathbb{R})$ we write its disconnecting set as

$$D := \{x \in \mathbb{R}; u^*(x) > u(x)\}.$$

We say that $\phi \in L^1(\mathbb{R})$ has the *subharmonicity property* when for any $u \in W^{1,1}(\mathbb{R})$ the associated maximal operator u^* is subharmonic in D. We notice that, given that we are in the one-dimensional setting, this property implies that u^* is convex in D. By [CS13, Lemmas 8 and 12] we know that property holds for both φ_1 and φ_2. In the next proposition we establish the same for φ_3^α.

Proposition 6.2.1. *For $\phi = \varphi_3^\alpha \in L^1(\mathbb{R}^d)$, $\alpha \in (0,1)$, we have that u^* is continuous in $\mathbb{R}^d$ and subharmonic in the set $\{x \in \mathbb{R}^d; u^*(x) > u(x)\}$ for any $u \in W^{1,1}(\mathbb{R}^d) \cap C(\mathbb{R}^d)$. Moreover, the map $u \mapsto (u^*)'$ is well defined and bounded from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$.*

Proof Following [CS13, Lemma 7(i)] we can conclude that u^* is continuous for $u \in W^{1,1}(\mathbb{R}^d) \cap C(\mathbb{R}^d)$. Therefore, following the proof of [CS13, Theorem 2(ii)], we need to prove the fact that u^* is subharmonic in the set $\{u^* > u\}$ to conclude the last assertion of our proposition. In order to conclude this subharmonicity, let us observe that, according to [CS07, §2.4], the function $\tilde{u}(\cdot, t) := u * (\varphi_3^\alpha)_t(\cdot)$ solves the Cauchy problem

$$\begin{aligned} \triangle_x \tilde{u} + \frac{\alpha}{t}\tilde{u}_t + \tilde{u}_{tt} = 0 \qquad & \text{for } (x,t) \in \mathbb{R}^d \times (0,\infty) \\ \tilde{u}(x,0) = u(x) \qquad & \text{for } x \in \mathbb{R}^d. \end{aligned}$$

Therefore, by combining [GT01, Theorem 3.1] and the remark thereafter with the proof of [CFS18, Lemma 7], we just need to prove the following: for any compact ball $B_r(x_0)$ and

$\varepsilon > 0$, there exists t_ε big enough such that for any $z \in B_r(x_0)$ we have $\tilde{u}(z,t) < \varepsilon$ for any $t > t_\varepsilon$. This claim follows from

$$|\tilde{u}(z,t)| \leq \|(\varphi_3^\alpha)_t\|_\infty \|u\|_1,$$

and the fact that $\|(\varphi_3^\alpha)_t\|_\infty \to 0$ when $t \to \infty$. ■

Let us recall the following result.

Lemma 6.2.1 ([CMP17, Lemma 14])**.** *Let $u \in W^{1,1}(\mathbb{R})$ and $\{u_j\}_{j\geq 1} \subset W^{1,1}(\mathbb{R})$ such that $\|u_j - u\|_{1,1} \to 0$. Then $\||u_j| - |u|\|_{1,1} \to 0$.*

The previous result allow us to always assume that the u_j and u are nonnegative, a simplification that we adopt henceforth. Now we prove a general statement about the uniform convergence of maximal functions.

Proposition 6.2.2. *Let $u_j \to u$ in $W^{1,1}(\mathbb{R})$. Then*

$$u_j^* \to u^*$$

uniformly.

Proof Let $x \in \mathbb{R}$, and let $t_x, t_{x,j} \geq 0$ such that $\tilde{u}(x, t_x) = u^*(x)$ and $\tilde{u}_j(x, t_{x,j}) = u_j^*(x)$, where we use the notation $\tilde{u}(x, 0) := u(x)$. Then

$$\begin{aligned}
|u_j^*(x) - u^*(x)| &= \max\{u_j^*(x) - u^*(x), u^*(x) - u_j^*(x)\} \\
&\leq \max\{\tilde{u}_j(x, t_{x,j}) - \tilde{u}(x, t_{x,j}), \tilde{u}(x, t_x) - \tilde{u}_j(x, t_x)\} \\
&\leq \|u_j - u\|_\infty \\
&\leq \|u_j - u\|_{1,1}.
\end{aligned}$$

■

In the following we assume $\phi \in \{\varphi_1, \varphi_2, \varphi_3^\alpha\}$ and that $d = 1$. Recall that in that case u^* is weakly differentiable and continuous. In the next lemma we reduce our analysis to a bounded interval.

Lemma 6.2.2. *If $u_j \to u$ in $W^{1,1}(\mathbb{R})$, for every $\varepsilon > 0$ there exist $R > 0$ and there exists j big enough such that we have*

$$\int_{[-R,R]^c} \left|(u_j^*)'\right| + |(u^*)'| < \varepsilon$$

Proof We prove that for any function $w \in W^{1,1}(\mathbb{R})$ and $R > 0$ we have

$$\int_{[R,\infty)} |(w^*)'| \leq |w^*(R) - w(R)| + \int_{[R,\infty)} |w'|,$$

the other required estimate follows by symmetry. If we write $\{x \in (R,\infty); w^*(x) > w(x)\} = \bigcup_{i=1}^\infty (a_i, b_i)$ we have $w^*(a_i) = w(a_i)$ and $w^*(b_i) = w(b_i)$ unless $a_i = R$. If $a_i \neq R$ we have

$\int_{(a_i,b_i)} |(u^*)'| \le \int_{(a_i,b_i)} |(u)'|$ by the *subharmonicity property.* By the same property we have that, if $a_i = R$ and w^* attains its minimum for the interval (a_i, b_i) at the point c_i, we have (with possibly $b_i = \infty$)

$$\int_{(R,b_i)} |(w^*)'| = w^*(R) - 2w^*(c_i) + w^*(b_i)$$
$$\le |w^*(R) - w(R)| + w(R) - 2w(c_i) + w(b_i) \le |w^*(R) - w(R)| + \int_{(R,b_i)} |w'|,$$

from where we conclude our claim. Now, in order to conclude our lemma we take R such that $\int_{[R,R]^c} |w'| \le \frac{\varepsilon}{4}$, $w^*(R) - w(R) + w^*(-R) - w(-R) < \frac{\varepsilon}{4}$ and j such that $\|w_j' - w'\|_1 < \frac{\varepsilon}{4}$ and $w_j^*(R) - w_j(R) + w_j^*(-R) - w_j(-R) < \frac{\varepsilon}{4}$, where in this last choosing we use Proposition 6.2.2. ■

Another important ingredient in our strategy is presented in the next proposition. For a partition $\mathcal{P} := \{a_1 < \cdots < a_m\}$ and $w : \mathbb{R} \to \mathbb{R}$ we define

$$\mathrm{Var}\,(w, \mathcal{P}) := \sum_{i=1}^{m-1} |w(a_{i+1}) - w(a_i)|.$$

Proposition 6.2.3. *Let $u_j \to u \in W^{1,1}(\mathbb{R})$. Then*

$$\|(u_j^*)'\|_1 \to \|(u^*)'\|_1$$

Proof By Lemma 6.2.2 is enough to prove that, for any (a, b), we have

$$\lim_{j\to\infty} \int_{[a,b]} \left|(u_j^*)'\right| = \int_{[a,b]} |(u^*)'|\,.$$

Since fo any $w \in W^{1,1}(\mathbb{R})$ we have

$$\int_{[a,b]} |w'| = \sup_{\mathcal{P}\subset[a,b]} \mathrm{Var}\,(w, \mathcal{P}),$$

by Fatou's lemma we obtain

$$\liminf_{j\to\infty} \int_{[a,b]} \left|(u_j^*)'\right| \ge \int_{[a,b]} |(u^*)'|\,.$$

Now, given $\varepsilon > 0$, we prove that

$$\limsup_{j\to\infty} \int_{\mathbb{R}} \left|(u_j^*)'\right| \le \int_{[a,b]} |(u^*)'| + 3\varepsilon.$$

Let us take a partition $\mathcal{P} = \{a = a_0 < a_1 \cdots < a_K = b\}$ such that

$$\mathrm{Var}\ (u, \mathcal{P}) > \int_{[a,b]} |u'| - \varepsilon$$

and

$$\text{Var}\ (u^*, \mathcal{P}) > \int_{[a,b]} |(u^*)'| - \varepsilon.$$

By uniform convergence we have

$$\text{Var}\ (u_j, \mathcal{P}) > \int_{[a,b]} |(u_j)'| - 2\varepsilon \tag{6.1}$$

and

$$\text{Var}\ \left(u_j^*, \mathcal{P}\right) > \int_{[a,b]} \left|(u_j^*)'\right| - 2\varepsilon \tag{6.2}$$

for j big enough. Now, let us consider $\widetilde{\mathcal{P}} = \widetilde{\mathcal{P}}(j) \supset \mathcal{P}$ with $\widetilde{\mathcal{P}} \subset [a,b]$ such that

$$\text{Var}\ \left(u_j^*, \widetilde{\mathcal{P}}\right) > \int_{[a,b]} \left|(u_j^*)'\right| - \varepsilon.$$

Without loss of generality we can assume that $\widetilde{\mathcal{P}}$ is such that $[a_i, a_{i+1}] \cap \widetilde{\mathcal{P}} = \{a_{i-1} = a_{i,0} < \cdots < a_{i,n_i} = a_i\}$ satisfies that $\text{sign}(u_{i,k}^* - u_{i,k+1}^*) = -\text{sign}(u_{i,k+1}^* - u_{i,k+2}^*)$ for every $k = 0, \ldots, n_i - 2$. For each such i we denote $\widetilde{\mathcal{P}}_i = \{a_{i,1}, \ldots, a_{i,n_i-1}\}$ and claim that it is possible to find another partition $\widetilde{\mathcal{P}}_i^* = \{a_{i,1}^*, \ldots, a_{i,n_i-1}^*\} \subset (a_{i-1}, a_i)$ such that

$$\text{Var}\ \left(u_j, \widetilde{\mathcal{P}}_i^*\right) - \text{Var}\ \left(u_j, \{a_{i,1}^*, a_{i,n_i-1}^*\}\right) = \text{Var}\ \left(u_j^*, \widetilde{\mathcal{P}}_i\right) - \text{Var}\ \left(u_j^*, \{a_{i,1}.a_{i,n_i-1}\}\right) \tag{6.3}$$

For $n_i \leq 2$ it follows by convention. For $n_i \geq 3$, by the *subharmonicity property* if $k \in \{i, \ldots, n_i - 1\}$ is such that $u_j^*(a_{i,k}) > \max\{u_j^*(a_{i,k-1}), u_j^*(a_{i,k+1})\}$, there exists $y \in (a_{i,k-1}, a_{i,k+1})$ such that $u_j(y) = u_j^*(a_{i,k})$. We choose $a_{i,k}^* = y$ in that case. Now, if $u_j^*(a_{i,k}) < \min\{u_j^*(a_{i,k+1}^*), u_j^*(a_{i,k-1}^*)\}$ (where $a_{i,0}^* = a_{i,0}$ and $a_{i,n_i}^* = a_{i,n_i}$) and $k < n_i - 1$, since $u_j(a_{i,k}) \leq u_j^*(a_{i,k}) < u_j(a_{i,k+1}^*)$ by continuity there exists $y \in (a_{i,k}, a_{i,k+1})$ such that $u_j(y) = u_j^*(a_{i,k})$. We choose $y = a_{i,k}^*$. The case $k = n_i - 1$ is done analogously, but instead choosing $y \in (a_{i,n_i-2}, a_{i,n_i-1})$ with the same property. From here (6.3) follows. Now, we apply (6.3) in order to obtain the following inequality

$$\begin{aligned}
\int_{[a,b]} |(u_j)'| - \text{Var}\,(u_j, \mathcal{P}) &\geq \text{Var}\left(u_j, \mathcal{P} \cup \bigcup_{i=1}^{K} \widetilde{\mathcal{P}}_i^*\right) - \text{Var}\left(u_j, \mathcal{P} \cup \bigcup_{i=1}^{K} \{a_{i,1}^*, a_{i,n_i-1}^*\}\right) \\
&= \sum_{i=1}^{K} \text{Var}\left(u_j, \widetilde{\mathcal{P}}_i^*\right) - \text{Var}\left(u_j, \{a_{i,1}^*, a_{i,n_i-1}^*\}\right) \\
&\geq \sum_{i=1}^{K} \text{Var}\ \left(u_j^*, \widetilde{\mathcal{P}}_i\right) - \text{Var}\ \left(u_j^*, \{a_{i,1}, a_{i,n_i-1}\}\right) \\
&\geq \text{Var}\left(u_j^*, \widetilde{\mathcal{P}}\right) - \text{Var}\left(u_j^*, \widetilde{\widetilde{\mathcal{P}}}\right),
\end{aligned}$$

where $\widetilde{\widetilde{\mathcal{P}}} := \{a_{i,k}; i \in \{1, \ldots, K\}, k \in \{0, 1, n_i - 1, n_i\}\}$. Notice that $|\widetilde{\widetilde{\mathcal{P}}}| \leq 3K+1$, therefore:

$$\text{Var}\left(u_j^*, \widetilde{\widetilde{\mathcal{P}}}\right) < \text{Var}\left(u^*, \widetilde{\widetilde{\mathcal{P}}}\right) + 12K\|u_j - u\|_\infty < \int_{(a,b)} |(u^*)'| + \varepsilon$$

for j big enough. Combining these estimates with (6.2), we get

$$\int_{[a,b]} |(u_j)'| - \text{Var}\,(u_j, \mathcal{P}) > \int_{[a,b]} \left|(u_j^*)'\right| - \int_{[a,b]} |(u^*)'| - \varepsilon.$$

Then, we have by (6.1) that

$$\int_{[a,b]} |(u_j^*)'| \leq \int_{[a,b]} |(u^*)'| + 3\varepsilon,$$

from where we conclude. ■

Now we state a classical property about convergence of convex functions.

Lemma 6.2.3. *Let $\{w_j\}_{j\in\mathbb{N}}$ and $(l_j, r_j) \subset \mathbb{R}$ with $w_j : \mathbb{R} \to \mathbb{R}$ such that w_j is convex in (l_j, r_j) for each $j \in \mathbb{N}$. Assume that $\lim_{j\to\infty} l_j = l$ and $\lim_{j\to\infty} r_j = r$ and that $w_j \to w$ uniformly. Then w is convex in (l, r) and*

$$\lim_{j\to\infty} w_j'(x) = w'(x),$$

for a.e. $x \in (l, r)$.

Proof For $a, b \in (l, r)$, then for j big enough $a, b \in [l_j, r_j]$ and therefore $w_j(\frac{a+b}{2}) \leq \frac{w_j(a)+w_j(b)}{2}$. Then by the pointwise convergence we have $w(\frac{a+b}{2}) \leq \frac{w(a)+w(b)}{2}$ from where the convexity follows. Then, the absolutely continuity of w in (l, r) follows. The last claim follows as in [Roc70, Theorem 25.7]. ■
The last preliminary lemma is the following.

Lemma 6.2.4. *Let $\phi \in \{\varphi_1, \varphi_2, \varphi_3^\alpha\}$. If $u_j \to u$ in $W^{1,1}(\mathbb{R})$, then*

$$(u_j^*)'(x) \to (u^*)'(x)$$

for a.e. $x \in D$.

Proof Follows by an adaptation of [CMP17, Lemmas 5 and 13]. A simpler proof using the *subharmonicity property* follows by Lemma 6.2.3 above. ■

6.3 Novel tools

In this section we develop new additional tools to address the continuity problem. Let us take $\varepsilon > 0$, consider $v_\varepsilon = \sum_{i=1}^{N} \alpha_i \chi_{(a_i,a_{i+1})}$, such that $\|u' - v_\varepsilon\|_1 < \varepsilon$. That is, we approximate the derivative of our limit function by a simple function. We define, as usual $D_j := \{x \in \mathbb{R}; u_j^*(x) > u_j(x)\}$. We know that D_j is an open set, we write it as an union of intervals in a convenient way that depends on our approximation v_ε. That is

$$D_j = D_j^1 \cup \bigcup_{i=0}^{N+1} D_j^{2,i},$$

where D_j^1 is the union of the intervals contained in D_j that contain at least one element of the set $\{a_1, \ldots, a_{N+1}\}$. The sets $D_j^{2,i}$, for $i = 1, \ldots, N$ are the union of intervals contained in (a_i, a_{i+1}), and $D_j^{2,0}$ and $D_j^{2,N+1}$ are the union of the intervals contained in D_j that are contained in $(-\infty, a_1)$ and (a_{N+1}, ∞), respectively. We write

$$D_j^1 = \bigcup_{r=1}^{N+1} (c_r(j), d_r(j)),$$

where $(c_r(j), d_r(j)) \ni a_r$ (possibly some intervals are empty or the same) and

$$D_j^{2,i} = \bigcup_{k=1}^{\infty} (c_k^i(j), d_k^i(j)).$$

The heart of our proof is the following lemma, where we prove that in the sets $D_j^{2,i}$ the function u_j^* is close to u_j at the derivative level. In the proof the *subharmonicity property* plays a major role.

Lemma 6.3.1. *If $\|u' - u_j'\|_1 < \varepsilon$ we have that*

$$\int_{\bigcup_{i=0}^{N+1} D_j^{2,i}} \left|(u_j^*)' - u_j'\right| < 4\varepsilon.$$

Proof Let us define $a_0 := -\infty$, $a_{N+2} := \infty$ and $\alpha_0 := 0 =: \alpha_{N+1}$. Let us see that, for $i = 0, \ldots, N+1$,

$$\int_{\cup_{k=1}^{\infty}(c_k^i(j), d_k^i(j))} \left|(u_j^*)' - u_j'\right| < 2 \int_{(a_i,a_{i+1})} \left|u_j' - \alpha_i\right|,$$

from where the result follows since

$$\sum_{i=0}^{N+1} \int_{(a_i,a_{i+1})} \left|u_j' - \alpha_i\right| < \varepsilon + \sum_{i=0}^{N+1} \int_{(a_i,a_{i+1})} \left|u' - \alpha_i\right| < 2\varepsilon.$$

Indeed, consider $L_i : (a_i, a_{i+1}) \to \mathbb{R}$ a line with $L_i'(x) = \alpha_i$ for all x and $i = 0, \ldots, N+1$. Then we observe that

$$\begin{aligned}\int_{(c_k^i(j), d_k^i(j))} \left|(u_j^*)' - u_j'\right| &= \int_{(c_k^i(j), d_k^i(j))} \left|(u_j^* - L_i)' - (u_j - L_i)'\right| \\ &\leq \int_{(c_k^i(j), d_k^i(j))} \left|(u_j^* - L_i)'\right| + \left|(u_j - L_i)'\right|.\end{aligned}$$

At this point, note that

$$\int_{(c_k^i(j), d_k^i(j))} \left|(u_j^* - L_i)'\right| \leq \int_{(c_k^i(j), d_k^i(j))} \left|(u_j - L_i)'\right|. \tag{6.4}$$

In fact, since u_j^* is convex in (c_k^i, d_k^i) we have that $u_j^* - L_i$ is also convex in that interval, therefore $u_j^* - L_i$ has no local maxima in that interval, considering that $u_j^* - L_i \geq u_j - L_i$ and that they coincide at the endpoints of the interval we conclude the claim. Now since $|(u_j - L_i)'| = \left|u_j' - \alpha_i\right|$ we conclude our lemma. ∎

Now, we need to control the (finitely many) remaining intervals in D_j.

Lemma 6.3.2. *We have that*

$$\int_{\cup_{r=1}^{N+1}(c_r(j), d_r(j))} \left|(u_j^*)' - (u^*)'\right| \to 0.$$

Proof Assume that there exists, for some r, an $\varepsilon_2 > 0$ such that $\int_{(c_r(j), d_r(j))} \left|(u_j^*)' - (u^*)'\right| > \varepsilon_2$ for a subsequence of j (that we also index by j). Let us take a subsequence such that $c_r(j) \to c_r$, $d_r(j) \to d_r$ when $j \to \infty$ (possibly $c_r = -\infty$ or $d_r = +\infty$). Then, Lemma 6.2.3 implies that u^* is convex in (c_r, d_r) and that $(u_j^*)' \to (u^*)'$ a.e in (c_r, d_r). Therefore, by the Brezis-Lieb lemma we just need to prove that

$$\lim_{j \to \infty} \int_{(c_r(j), d_r(j))} \left|(u_j^*)'\right| = \int_{(c_r, d_r)} \left|(u^*)'\right|.$$

If we write $m_r(j) = \min\limits_{x \in (c_r(j), d_r(j))} u_j^*(x)$ and $m_r = \min\limits_{x \in (c_r, d_r)} u^*(x)$, we have that

$$\int_{(c_r(j), d_r(j))} \left|(u_j^*)'\right| = u_j^*(c_r(j)) - 2m_r(j) + u_j^*(d_r(j))$$

and

$$\int_{(c_r, d_r)} \left|(u^*)'\right| = u^*(c_r) - 2m_r + u^*(d_r).$$

Therefore the desired convergence is a consequence of the uniform convergence and the continuity of u^*. This concludes the proof of the lemma. ∎

6.4 Proof of Theorem 6.1.1

With the tools developed in the last section we are in position to prove our theorem. First, we claim the following:

$$\lim_{j\to\infty}\int_{D_j\cap C}\left|(u_j^*)'-(u^*)'\right|\to 0. \tag{6.5}$$

Noticing that $u'=(u^*)'$ a.e. in the set of integration we have

$$\begin{aligned}\int_{D_j\cap C}\left|(u_j^*)'-(u^*)'\right| &= \int_{D_j\cap C}\left|(u_j^*)'-u'\right| \\ &\leq \int_{D_j}\left|(u_j^*)'-u'\right| \\ &\leq \int_{D_j^1}\left|(u_j^*)'-u'\right| + \int_{\cup_{i=0}^{N+1}D_j^{2,i}}\left|(u_j^*)'-u'\right| \\ &\leq \int_{D_j^1}\left|(u_j^*)'-u'\right| + \int_{\cup_{i=0}^{N+1}D_j^{2,i}}\left|(u_j^*)'-u_j'\right| + \|u'-u_j'\|_1 \\ &\leq \int_{D_j^1}\left|(u_j^*)'-u'\right| + 5\varepsilon,\end{aligned}$$

for j big enough, where we use Lemma 6.3.1 in the last line. Since

$$\int_{D_j^1}\left|(u_j^*)'-u'\right| < \varepsilon$$

for j big enough by the Lemma 6.3.2, we have

$$\limsup_{j\to\infty}\int_{D_j\cap C}\left|(u_j^*)'-(u^*)'\right| \leq 6\varepsilon.$$

Since $\varepsilon>0$ is arbitrary we conclude the proof of our claim (6.5).

From (6.5) and since

$$\int_{C_j\cap C}\left|(u_j^*)'-(u^*)'\right| = \int_{C_j\cap C}\left|u_j'-u'\right|\to 0,$$

we conclude that

$$\int_{C}\left|(u_j^*)'-(u^*)'\right|\to 0. \tag{6.6}$$

Now, in order to prove our Theorem 6.1.1 we need to conclude that

$$\int_{D}\left|(u_j^*)'-(u^*)'\right|\to 0.$$

Indeed, in light of Lemma 6.2.4, by the Brezis-Lieb lemma we only require that

$$\int_D \left|(u_j^*)'\right| \to \int_D |(u^*)'| \,,$$

and this is a consequence of (6.6) and Proposition 6.2.3. This concludes the proof of Theorem 6.1.1.

6.4.1 Concluding remarks

The same scheme of proof presented here allows one to establish the analogous of Theorem 6.1.1 for a more general class of maximal operators of convolution type. The key properties that we require are that the maximal function u^* is continuous and has the *subharmonicity property*, and one has to then deal with minor technicalities that might appear (and for simplicity we do not enter in all such variations). For instance, one could consider the operators defined in [CFS18, §1.2], in which the approximation of the identity are slightly different.

Chapter 7

Continuity for the one-dimensional centered Hardy-Littlewood maximal operator at the derivative level

7.1 Introduction

In the present chapter, we establish the continuity for the centered Hardy-Littlewood maximal operator, solving a question posed by Carneiro, Madrid and Pierce in [CMP17, Question A] and establishing, in the one-dimensional case, the endpoint version of [HO04, Question 3] at the derivative level.

Theorem 7.1.1. *We have that the map*

$$f \mapsto (Mf)'$$

is continuous from $W^{1,1}(\mathbb{R})$ to $L^1(\mathbb{R})$.

We notice that the map considered here is well defined and bounded (see Lemma 7.2.1). We highlight that the methods developed in Chapter 4, 5, 6 and [CMP17] are not enough to conclude our result. For instance, in Chapter 4, 5 and [CMP17] it is important that the operator $\widetilde{M}$ has the flatness property; this is, that the maximal functions have a.e. zero derivative at the points where they coincide with the original function. In Chapter 6, the subharmonicity property, which the maximal functions considered there satisfy, plays a crucial role in the proof of the continuity. The centered Hardy-Littlewood maximal operator does not satisfy either of these properties, therefore, new insights are required in order to achieve our result. Our method is based on a decomposition of M as a maximum of two operators M_1 and M_2, both of them depending on f and on a simple function g_ε that approximates f' in $L^1(\mathbb{R})$. The operator M_1, the local one, is restricted to balls that are contained in the support of an interval determined by g_ε. On the other hand, the operator M_2, the global one, is restricted to balls that are not contained in any of these lines. The idea is that, since the operator M_1 is well behaved with respect to some lines, it is possible to conclude that $M_1 f_j$ is close to f_j at the derivative level, for any j big enough. A different approach is needed in order to deal with the contribution of the operator M_2, for this we shall take advantage of the fact that the radii considered in M_2 are generally bounded by below. In essence, this yields a smoother nature to this operator that is helpful for our purposes.

7.2 Preliminaries

In this section we discuss some preliminary results for our purposes. Let us consider $f_j \to f$ in $W^{1,1}(\mathbb{R})$. In order to prove Theorem 7.1.1, by [CMP17, Lemma 14] we may assume henceforth that $f_j, f \geq 0$. Also, since the case $f = 0$ of Theorem 7.1.1 follows by the boundedness, we assume that $f \neq 0$. We start with the well known Luiro's formula.

Proposition 7.2.1 (Case $p = 1$ of [Lui07, Theorem 3.1])**.** *Let us take $g \in W^{1,1}(\mathbb{R})$. Assume that Mg is differentiable at the point x, if $Mg(x) = \fint_{[x-r,x+r]} |g|$ with $r > 0$, we have that*

$$(Mg)'(x) = \fint_{[x-r,x+r]} |g|'.$$

Proof This follows from [BGRMW21, Proposition 2.4] and the remark thereafter. ■

The next result provide us with a local control for the variation of M. For any interval I (not necessarily finite) and $g \in L^1(\mathbb{R})$ we define

$$M_I g(x) := \sup_{[x-r,x+r] \subset I} \fint_{[x-r,x+r]} |g|.$$

Lemma 7.2.1. *If $f \in W^{1,1}(I)$, we have that $M_I f$ is absolutely continuous and that there exists an universal constant C, such that*

$$\int_I |(M_I f)'| \leq C \int_I |f'|.$$

Proof The absolutely continuity of $M_I f$ can be concluded by following the reasoning in [Kur15, Corollary 1.3]. The boundedness follows from [Kur15, Remark 6.4]. ■

Also, we need the following uniform control near a finite number of points.

Lemma 7.2.2. *Let $f_j \to f$ in $W^{1,1}(\mathbb{R})$. Let $\{p_1, \ldots, p_s\}$ be a finite set. For any $\varepsilon > 0$ there exists $\delta > 0$ such that, for j big enough, we have*

$$\sum_{i=1}^{s} \int_{[p_i-\delta, p_i+\delta]} |(Mf_j)'| < \varepsilon$$

.

Proof This proof follows a similar path than the one presented originally in [GR20, Proposition 19]. It is enough to prove that there exists $\delta > 0$ such that

$$\int_{[p_i-\delta,p_i+\delta]} |(Mf_j)'| < \frac{\varepsilon}{s}$$

for any fixed i and j big enough. Let us take $\delta_i > 0$ such that

$$\int_{(a_i-\delta_i,a_i+\delta_i)} |f'| < \frac{\varepsilon}{2Cs},$$

where C is the universal constant that appears in Lemma 7.2.1. For j big enough we have

$$\int_{(a_i-\delta_i,a_i+\delta_i)} |f_j'| < \frac{\varepsilon}{2Cs}.$$

For any given $\ell \in \mathbb{Z}_{>0}$ let us define

$$A^1_{\ell,j} := \left\{ x \in \left(a_i - \frac{\delta_i}{\ell}, a_i + \frac{\delta_i}{\ell}\right); Mf_j(x) = M_{(a_i-\delta_i,a_i+\delta_i)}f_j(x) \right\}$$

and

$$A^2_{\ell,j} = \left\{ x \in \left(a_i - \frac{\delta_i}{\ell}, a_i + \frac{\delta_i}{\ell}\right); Mf_j(x) > M_{(a_i-\delta_i,a_i+\delta_i)}f_j(x) \right\}.$$

Since $Mf_j \geq M_{(a_i-\delta_i,a_i+\delta_i)}f_j$ know that $(Mf_j)' = (M_{(a_i-\delta_i,a_i+\delta_i)}f_j)'$ a.e. in $A^1_{\ell,j}$. Therefore

$$\int_{A^1_{\ell,j}} |(Mf_j)'| \leq \int_{(a_i-\delta_i,a_i+\delta_i)} |(M_{(a_i-\delta_i,a_i+\delta_i)}f_j)'| \leq C \int_{(a_i-\delta_i,a_i+\delta_i)} |f_j'| \leq \frac{\varepsilon}{2s}.$$

Also, for a.e. $x \in A^2_{\ell,j}$, there exists $r_x \geq \delta_i - \frac{\delta_i}{\ell} = \frac{\delta_i(\ell-1)}{\ell}$ such that $\fint_{[x-r_{j,x},x+r_{j,x}]} f_j = Mf_j(x)$. Then, by Luiro's formula (Proposition 7.2.1), we have $(Mf_j)'(x) = \fint_{[x-r_{j,x},x+r_{j,x}]} f_j'$, and therefore $|(Mf_j)'(x)| \leq \frac{1}{2r_{j,x}}\|f_j'\|_1$. Thus, for $x \in A^2_{\ell,j}$ we have

$$|(Mf_j)'(x)| \leq \frac{\delta_i \ell}{2(\ell-1)}\|f_j'\|_1.$$

In consequence, we have

$$\int_{A^2_{\ell,j}} |(Mf_j)'| \leq \int_{(a_i-\frac{\delta_i}{\ell},a_i+\frac{\delta_i}{\ell})} \frac{\delta_i \ell}{2(\ell-1)}\|f_j'\|_1 \leq \frac{\delta_i^2}{(\ell-1)}\|f_j'\|_1.$$

From here, we conclude our lemma by choosing ℓ such that $\frac{\delta_i^2}{\ell-1} < \frac{\varepsilon}{4s}$, $\delta := \frac{\delta_i}{\ell}$ and by taking j big enough such that $\frac{\|f'\|_1}{2} \leq \|f_j'\|_1 \leq \frac{3\|f'\|_1}{2}$. ■

Also, we need the following uniform control near infinity.

Lemma 7.2.3 ([BM19, Proposition 4.11])**.** *Let $f_j \to f$ in $W^{1,1}(\mathbb{R})$ and $\varepsilon > 0$ be given. There exists $K > 0$ such that, for j big enough, we have*

$$\int_{(-K,K)^c} |(Mf_j)'| < \varepsilon.$$

7.3 The auxiliary maximal operators

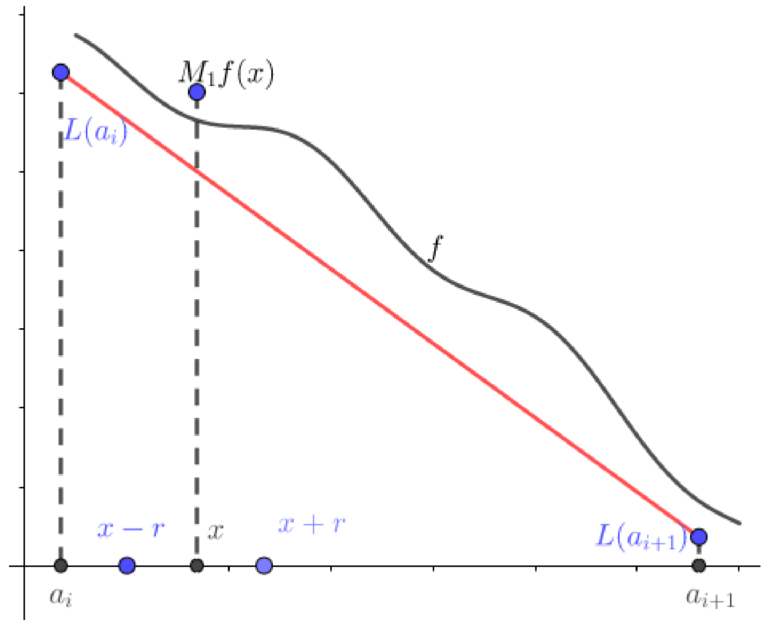

Figure 7.1: In the figure the scope of L is α_i and $[x-r, x+r]$ is an admissible interval for x and M_1.

In this section we define the main objects of our work. Let us take $\varepsilon > 0$ and consider $g_\varepsilon = \sum_{i=0}^{N} \alpha_i \chi_{(a_i, a_{i+1})}$ such that $\|f' - g_\varepsilon\|_1 < \varepsilon$. That is, we approximate the derivative of our limit function by a simple function. We write $a_0 = -\infty$, $a_{N+1} = \infty$ and $\mathcal{P} := \{a_1, \ldots, a_N\}$. We assume that $\mathcal{P}$ is non-empty. We observe that $\alpha_0 = \alpha_n = 0$. Now, we define our auxiliary maximal operators M_1, M_2 as follows: for any $h \in L^1(\mathbb{R})$ and $x \in \mathbb{R}$ we set

$$M_1 h(x) := \sup_{r < d(x,\mathcal{P})} \fint_{[x-r,x+r]} |h|,$$

and

$$M_2 h(x) := \sup_{r \geq d(x,\mathcal{P})} \fint_{[x-r,x+r]} |h|.$$

We now state some basic results about our operators M_1, M_2.

Lemma 7.3.1. *Let $f_j \to f$ in $W^{1,1}(\mathbb{R})$. We have*

$$M_i f_j \to M_i f$$

uniformly, for $i = 1, 2$.

Proof It follows from the fact that $|M_i f_j - M_i f| \leq |M_i(f_j - f)| \leq \|f_j - f\|_\infty$. ∎

7.3.1 Properties of M_2

For any $K,\delta>0$ such that the intervals $(a_i-\delta,a_i+\delta)$ are pairwise disjoint, let us define $U_{\delta,K}=(-K,K)\setminus\cup_{i=1}^n(a_i-\delta,a_i+\delta)$. We observe that for any $x\in U_{\delta,K}$ and any $g\in W^{1,1}(\mathbb{R})$ there exists a radius $r_x\geq\delta$ such that $\fint_{[x-r_x,x+r_x]}|g|=M_2\,g(x)$. We have then the following.

Lemma 7.3.2. *For any $g\in W^{1,1}(\mathbb{R})$ we have that $M_2\,g$ is weakly differentiable in $U_{\delta,K}$.*

Proof For any $x,y\in U_{\delta,K}$ with $M_2\,g(x)>M_2\,g(y)$, we have

$$M_2\,g(x)-M_2\,g(y)=\fint_{[x-r_x,x+r_x]}|g|-\fint_{[y-r_y,y+r_y]}|g|\leq\fint_{[x-r_x,x+r_x]}|g|-\fint_{[y+|x-y|-r_x,y+|x-y|+r_x]}|g|$$
$$\leq\|g\|_1\left(\frac{1}{2r_x}-\frac{1}{2r_x+2|x-y|}\right)\leq\|g\|_1C(\delta)|x-y|,$$

where $C(\delta)$ is the Lipschitz constant of the function $\frac{1}{2x}$ in the set $[\delta,\infty)$. Therefore, we have that $M_2\,g$ is Lipschitz in this set, from where we conclude our lemma. ■

In the next result we present a formula for the derivative of $M_2\,g$ that has similarities with the one presented in [GRK21, Lemma 10]. We use the notation $x\pm\infty=\pm\infty$.

Lemma 7.3.3. *Let $g\in W^{1,1}(\mathbb{R})$. Let $x\in U_{\delta,K}$ be such that $M_2\,g$ is differentiable at x, and let r_x such that $M_2\,g(x)=\fint_{[x-r_x,x+r_x]}|g|$ with $r_x\geq d(x,\mathcal{P})$. Assume that $\frac{a_i+a_{i+1}}{2}<x<a_{i+1}$. Then, we have*

$$(M_2\,g)'(x)=\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x^2}-\frac{|g|(x-r_x)}{r_x}.$$

Proof Observe that, for $h>0$, we have

$$\frac{M_2\,g(x)-M_2\,g(x-h)}{h}\leq\frac{\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x}-\frac{\int_{[x-r_x-2h,x+r_x]}|g|}{2r_x+2h}}{h}$$
$$=\frac{\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x}-\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x+2h}-\frac{\int_{[x-r_x-2h,x-r_x]}|g|}{2r_x+2h}}{h}$$
$$\to\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x^2}-\frac{|g|(x-r_x)}{r_x}$$

when $h\to0$, where we use the continuity of g. Therefore $(M_2\,g)'(x)\leq\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x^2}-\frac{|g|(x-r_x)}{2r_x}$. Also, for $h>0$, since $x<a_i\leq x+r_x$ (and hence the interval $[x-r_x+2h,x+r_x]$ is admissible for $x+h$ for the operator M_2), we have

$$\frac{M_2\,g(x+h)-M_2\,g(x)}{h}\geq\frac{\frac{\int_{[x-r_x+2h,x+r_x]}|g|}{2r_x-2h}-\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x}}{h}$$
$$=\int_{[x-r_x,x+r_x]}|g|\left(\frac{\frac{1}{2r_x-2h}-\frac{1}{2r_x}}{h}\right)-\frac{\int_{[x-r_x,x-r_x+2h]}|g|}{(2r_x-2h)h}$$
$$\to\frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x^2}-\frac{|g|(x-r_x)}{r_x}$$

when $h \to 0$, and therefore $(M_2\, g)'(x) \geq \frac{\int_{[x-r_x,x+r_x]}|g|}{2r_x^2} - \frac{|g|(x-r_x)}{r_x}$, from where we conclude our lemma. ■

Lemma 7.3.4. *Let $f_j \to f$ in $W^{1,1}(\mathbb{R})$. Let $x \notin \mathcal{P}$. Assume that $M_2 f_j(x) = \fint_{[x-r_{j,x},x+r_{j,x}]}|f_j|$ for some $r_{x,j} \geq d(x,\mathcal{P})$. If $r_{j,x} \to r$ then*

$$M_2 f(x) = \fint_{[x-r,x+r]}|f|.$$

Proof This follows as [CMP17, Lemma 12]. ■

Now we can conclude the pointwise a.e convergence at the derivative level.

Lemma 7.3.5. *Let $f_j \to f$ in $W^{1,1}(\mathbb{R})$. Then, for a.e $x \in U_{\delta,K}$, we have*

$$(M_2 f_j)'(x) \to (M_2 f)'(x).$$

Proof Let us assume that x is such that $M_2 f_j$, for every j, and $M_2 f$ are differentiable at the point x and $x \in \left(\frac{a_i+a_{i+1}}{2}, a_{i+1}\right)$ for some i. The other case follows analogously. Now, for every j, let us take $r_{x,j} \geq d(x,\mathcal{P})$ such that $M_2 f_j(x) = \fint_{[x-r_{j,x},x+r_{j,x}]}|f_j|$. By Lemma 7.3.3 we have that

$$(M_2 f_j)'(x) = \frac{\int_{[x-r_{x,j},x+r_{x,j}]}|f_j|}{2r_{x,j}^2} - \frac{|f_j|(x-r_{x,j})}{r_{x,j}}.$$

Assume that there exists a subsequence $\{j_k\}_{k\in\mathbb{N}}$ such that $|(M_2 f_{j_k})'(x) - (M_2 f)'(x)| > \rho > 0$. Let us take $R > 0$ such that $\int_{[x-R,x+R]}|f| > \frac{\|f\|_1}{2}$. For j big enough we have that $\int_{[x-R,x+R]}|f_j| > \frac{\|f_j\|_1}{2}$. Since

$$\frac{\|f_j\|_1}{4R} < \frac{\int_{[x-R,x+R]}|f_j|}{2R} \leq \fint_{[x-r_{x,j},x+r_{x,j}]}|f_j| \leq \frac{\|f_j\|_1}{2r_{x,j}},$$

we note that $r_{x,j} \leq 2R$. Therefore, there exists a subsequence of $\{j_k\}_{k\in\mathbb{N}}$ (that we keep calling $\{j_k\}_{k\in\mathbb{N}}$ with a harmless abuse of notation) such that $r_{x,j_k} \to r > 0$. Thus, by Lemma 7.3.4, we have

$$\begin{aligned}(M_2 f_j)'(x) &= \frac{\int_{[x-r_{x,j},x+r_{x,j}]}|f_j|}{2r_{x,j}^2} - \frac{|f_j|(x-r_{x,j})}{r_{x,j}}\\ &\to \frac{\int_{[x-r,x+r]}|f|}{2r^2} - \frac{|f|(x-r)}{r} = (M_2 f)'(x).\end{aligned}$$

From this we conclude our lemma. ■

We are now in position to conclude our desired $L^1(U_{\delta,K})$ convergence.

Proposition 7.3.1. *We have $(M_2 f_j)' \to (M_2 f)'$ in $L^1(U_{\delta,K})$.*

Proof Let us take $x \in U_{\delta,K}$ with $x \in \left(\frac{a_i+a_{i+1}}{2}, a_{i+1}\right)$ and such that $M_2 f_j$, for every j, and $M_2 f$ are all differentiable at the point x. The symmetric case follows similarly. By Lemma 7.3.3 we have that (using the notation of the previous lemma)

$$\begin{aligned}|(M_2 f_j)'(x)| &= \left| \frac{\int_{[x-r_{x,j},x+r_{x,j}]} |f_j|}{2r_{x,j}^2} - \frac{|f_j|(x-r_{x,j})}{r_{x,j}} \right| \\ &\leq \frac{\|f_j\|_1}{2\delta^2} + \frac{\|f_j\|_\infty}{\delta} \leq 2\|f\|_{1,1}\left(\frac{1}{2\delta^2}+\frac{1}{\delta}\right),\end{aligned}$$

for j big enough. Therefore, by combining the dominated convergence theorem with Lemma 7.3.5, we conclude our proposition. ■

7.3.2 Properties of M_1

About our local operator M_1, by Lemma 7.2.1 we have that M_1 is weakly differentiable in $\mathbb{R} \setminus \mathcal{P}$. We now prove the following.

Proposition 7.3.2. *Let $f_j \to f$ in $W^{1,1}(\mathbb{R})$ (recall that we assume $f_j, f \geq 0$). We have that, for j big enough*

$$\|(M_1 f_j)' - f_j'\|_1 \leq 2(C+1)\varepsilon,$$

where C is the universal constant appearing in Lemma 7.2.1.

Proof Let $L_i : (a_i, a_{i+1}) \to \mathbb{R}$ be a line such that $L_i' = \alpha_i$ and $L_i \leq 0$ (since $\alpha_0 = \alpha_n = 0$, L_0 and L_n are constant). We observe that

$$\int_{(a_i,a_{i+1})} |(M_1 f_j)' - (f_j)'| \leq \int_{(a_i,a_{i+1})} |(M_1 f_j)' - L_i'| + |L_i' - (f_j)'|. \tag{7.1}$$

Let us notice that, for every $x \in (a_i, a_{i+1})$, we have

$$\begin{aligned}M_1 f_j - L_i &= \left(\sup_{r<d(x,\{a_i,a_{i+1}\})} \fint_{[x-r,x+r]} f_j\right) - L_i \\ &= \sup_{r<d(x,\{a_i,a_{i+1}\})} \fint_{[x-r,x+r]} (f_j - L_i) = M_1(f_j - L_i).\end{aligned}$$

Therefore, we have

$$\begin{aligned}\int_{(a_i,a_{i+1})} |(M_1 f_j)' - L_i'| &= \int_{(a_i,a_{i+1})} |(M_1 f_j - L_i)'| \\ &= \int_{(a_i,a_{i+1})} |(M_1(f_j - L_i))'| \\ &= C\int_{(a_i,a_{i+1})} |(f_j - L_i)'|.\end{aligned}$$

Combining this with (7.1) we have that

$$\int_{(a_i,a_{i+1})} |(M_1 f_j)' - (f_j)'| \le (C+1) \int_{(a_i,a_{i+1})} |f_j' - \alpha_i|$$
$$\le (C+1) \left(\int_{(a_i,a_{i+1})} |f' - \alpha_i| + |f' - f_j'| \right).$$

Therefore, we have

$$\|(M_1 f_j)' - f_j'\|_1 \le (C+1) \left(\varepsilon + \|f' - f_j'\|_1 \right).$$

Since $\|f' - f_j'\|_1 < \varepsilon$ for j big enough, we conclude our proposition. ■
Analogously, we conclude that $\|(M_1 f)' - (f)'\|_1 \le 2(C+1)\varepsilon$, and therefore $\|(M_1 f_j)' - (M_1 f)'\|_1 \le (4C+5)\varepsilon$, for j big enough.

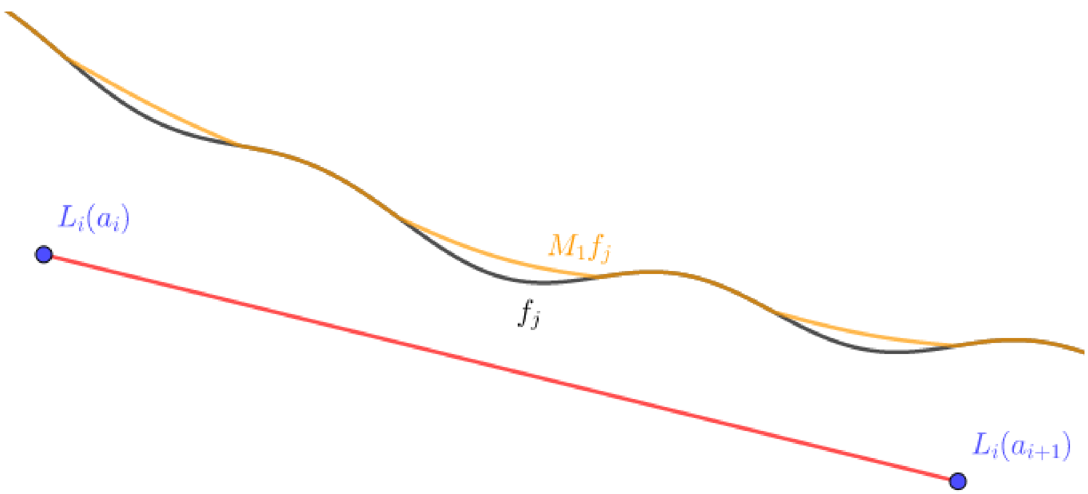

Figure 7.2: f_j and $M_1 f_j$ are close at the derivative level to L_i when j is big enough.

7.4 Proof of Theorem 7.1.1

Now we are able to conclude our result.
Proof By choosing K big enough and δ small enough such that Lemmas 7.2.2 and 7.2.3 hold, we have that

$$\int_{\mathbb{R}\setminus U_{\delta,K}} |(Mf_j)' - (Mf)'| < 2\varepsilon, \tag{7.2}$$

for j big enough. Now we focus on $U_{\delta,K}$. We follow a similar strategy than in [CMP17, Lemma 11]. We observe that $M = \max\{M_1, M_2\}$. Let us write $X_j := \{x \in U_{\delta,K}; M_1 f_j(x) > M_2 f_j(x)\}$, $Y_j := \{x \in U_{\delta,K}; M_1 f_j(x) = M_2 f_j(x)\}$ and $Z_j := \{x \in U_{\delta,K}; M_1 f_j(x) < M_2 f_j(x)\}$. We define X, Y and Z analogously, but this time with respect to f instead of f_j. We observe that $(Mf_j)' = (M_1 f_j)'$ a.e. in X_j, $(Mf_j)' = (M_1 f_j)' = (M_2 f_j)'$ a.e in Y_j and $(M_2 f_j)'(x) = (Mf_j)'(x)$ in Z_j. Analogous properties hold for f in X, Y and Z. Let us observe that

$$\begin{aligned}
&\int_X |(Mf_j)' - (Mf)'| \\
&\qquad = \int_{X\cap X_j} |(Mf_j)' - (Mf)'| + \int_{X\cap Y_j} |(Mf_j)' - (Mf)'| + \int_{X\cap Z_j} |(Mf_j)' - (Mf)'| \\
&\qquad \leq \int_{X\cap X_j} |(M_1 f_j)' - (M_1 f)'| + \int_{X\cap Y_j} |(M_1 f_j)' - (M_1 f)'| + \int_{X\cap Z_j} |(M_2 f_j)' - (M_1 f)'| \\
&\qquad \leq \int_{U_{\delta,K}} |(M_1 f_j)' - (M_1 f)'| + \int_{X\cap Z_j} |(M_2 f_j)' - (M_2 f)'| + \int_{X\cap Z_j} |(M_2 f)' - (M_1 f)'|.
\end{aligned}$$

By Lemma 7.3.1 we have that $\chi_{X\cap Z_j} \to 0$ a.e., therefore by the dominated convergence theorem we have $\int_{X\cap Z_j} |(M_2 f)' - (M_1 f)'| < \varepsilon$ for j big enough. Then, by combining Propositions 7.3.1 and 7.3.2 with this we have that there exists and universal constant $\tilde{C}$ such that

$$\int_X |(Mf_j)' - (Mf)'| < \tilde{C}\varepsilon,$$

for j big enough. Similarly, we conclude an analogous statement about Y and Z. Therefore, considering (7.2), we have that there exist an universal constant $\tilde{\tilde{C}}$ such that

$$\|(Mf)' - (Mf_j)'\|_1 < \tilde{\tilde{C}}\varepsilon,$$

for j big enough. From this we conclude our result. ■

Chapter 8

Sharp inequalities for maximal operators on finite graphs I

8.1 Introduction

An interesting framework of study of maximal operators is the following. Let $G = (V, E)$ be a graph and $f : V \to \mathbb{R}$ a real valued function. We define the Hardy-Littlewood maximal function of f along G at the point $v \in V$ by

$$M_G f(v) := \max_{r\geq 0} \frac{1}{|B(v,r)|} \sum_{m\in B(v,r)} |f(m)|, \tag{8.1}$$

where $B(v,r) = \{m \in V; d_G(v,m) \leq r\}$, where d_G is the metric induced by the edges of G (that is, the distance between two vertices is the number of edges in a shortest path connecting them). A more general version of this, is the so called fractional maximal function defined by

$$M_{\alpha,G} f(v) := \max_{r\geq 0} \frac{1}{|B(v,r)|^{1-\alpha}} \sum_{m\in B(v,r)} |f(m)|$$

for all $\alpha \in (0,1]$. Both operators have uncentered versions defined by

$$\widetilde{M}_{\alpha,G} f(v) = \max_{B(w,r)\ni v} \frac{1}{|B(w,r)|^{1-\alpha}} \sum_{m\in B(w,r)} |f(m)|$$

for the fractional one, and $\widetilde{M}_G = \widetilde{M}_{0,G}$ for the classical one. In this chapter we study the regularity properties of these objects acting on l^p-spaces and bounded $p-$variation spaces. We focus on the classical maximal function defined in (8.1).

Given $p \in (0,\infty)$ we define the p-variation of a function $f : V \to \mathbb{R}$ as follows

$$\operatorname{Var}_p f := \left(\frac{1}{2} \sum_n \sum_{\substack{m \\ d_G(n,m)=1}} |f(n) - f(m)|^p \right)^{1/p}.$$

8.1.1 Conjectures and results for the p-variation in finite graphs

For a given graph $G = (V, E)$ and $0 < p < \infty$, we define

$$\mathbf{C}_{G,p} := \sup_{f:V\to\mathbb{R};\mathrm{Var}_p f>0} \frac{\mathrm{Var}_p M_G f}{\mathrm{Var}_p f}.$$

Liu and Xue ([LX20]) obtained optimal results for $n = 3$ and for the general case $n > 3$ they found some bounds and posed some interesting conjectures. More precisely, they proved that if G is the complete graph with n vertices K_n or the star graph with n vertices S_n, then

$$1 - \frac{1}{n} \leq \mathbf{C}_{G,p} \leq 1$$

for $0 < p < \infty$, and for $n = 3$ the lower bound becomes an equality. Moreover, Liu and Xue posed the following conjecture [LX20, Conjecture 1(i)].

Conjecture A (for the complete graph K_n): For every $n \geq 2$ and $p \in (0, \infty)$ we have

$$\mathbf{C}_{K_n,p} = 1 - \frac{1}{n}.$$

In this chapter we give a positive answer to this conjecture for all $p \geq \frac{\log 4}{\log 6} \approx 0.77$. This range is certainly not optimal and is an interesting problem to try to extend it. Also, we prove the conjecture for every $0 < p < 1$ when $n = 4$. That is the content of our Theorem 8.1.1.

Theorem 8.1.1 (Complete graph)**.** *Let $0 < p \leq \infty$ and $K_n = (V, E)$ be the complete graph with n vertices $(a_1, a_2, \ldots, a_n)$. Then*

(i) If $p > 1$, then

$$\mathbf{C}_{K_n,p} = 1 - \frac{1}{n}.$$

(ii) If $0 < p \leq 1$ and $n = 4$,

$$\mathbf{C}_{K_n,p} = 1 - \frac{1}{n}.$$

(iii) If $n \geq 3$ and $1 \geq p \geq \frac{\log 4}{\log 6} \approx 0.77$, then

$$\mathbf{C}_{K_n,p} = 1 - \frac{1}{n}.$$

Moreover, in all the cases any function that vanishes everywhere but in one vertex is an extremizer (we call this kind of function Dirac's delta).

We notice that given the different behavior of the function $x \mapsto x^p$ when $p > 1$ and $p \le 1$ very contrasting techniques are needed in each case.
The second conjecture that they posed is the following [LX20, Conjecture 1(ii)].

Conjecture B (for the star graph S_n): For any $n \ge 2$ and $p \in (0,1]$ we have

$$\mathbf{C}_{S_n,p} = 1 - \frac{1}{n}.$$

In this case we prove that, in fact, this equality is not true for $p > 1$. In fact, for $n = 3$, we find values for $\mathbf{C}_{S_3,p}$ different to the ones conjectured in that case. However, we give a positive answer to this conjecture when $1/2 \le p \le 1$ for all $n \ge 2$. Moreover, we give a positive answer to the conjecture when $0 < p < 1/2$ if n is sufficiently large, this is the content of our Theorem 8.1.2.

Theorem 8.1.2 (Star graph)**.** *Let $S_n = (V, E)$ be a start graph with n vertices $(a_1, a_2, \ldots, a_n)$, with center at a_1. Then, the following hold.*

(i) For all $1 < p < \infty$ we have that

$$\mathbf{C}_{S_3,p} = \frac{(1 + 2^{p/(p-1)})^{(p-1)/p}}{3} < 1. \tag{8.2}$$

(ii) If $p = 1$, then

$$\mathbf{C}_{S_n,p} = 1 - \frac{1}{n}. \tag{8.3}$$

(iii) If $n = 4$ and $0 < p < 1$, or $n \ge 5$ and $\frac{1}{2} \le p \le 1$, then

$$\mathbf{C}_{S_n,p} = 1 - \frac{1}{n}. \tag{8.4}$$

Moreover, (8.4) holds for every $\frac{1}{2} > p > 0$ when $n \ge C(p)$, for some finite constant $C(p)$ depending only on p.

The range $(\frac{1}{2}, 1)$ in (iii) is certainly not optimal, to find improvements on this range is an interesting problem.
In the following we discuss the third conjecture proposed by Liu and Xue [LX20, Conjecture 1(iii) and (iv)].

Conjecture C (boundedness and continuity): Let $0 < p, q \le \infty$ and $0 \le \alpha < 1$. The operator $M_{\alpha,G}$ is bounded and continuous from $BV_p(G)$ to $BV_q(G)$, where $BV_p(G) := \{f : V \to \mathbb{R}; \operatorname{Var}_p f < \infty\}$ is endowed with $\|f\|_{\widetilde{BV_p(V)}} := \operatorname{Var}_p f$, note that $\|\cdot\|_{\widetilde{BV_p(V)}}$ depends strongly on G not only on the set of vertices V.

We prove that the boundedness holds as conjectured. Moreover, we prove that with a slight modification the continuity affirmation is true. That is the content of our next theorem. We also prove that a modification is strictly required. This is related with the fact that $\|\cdot\|_{\widetilde{BV_p(G)}}$ is not a norm (think about constant functions for example), on the other hand, taking $a_0 \in V$ we have that $\|f\|_{BV_p(G)} := \|f\|_{\widetilde{BV_p(G)}} + |f(a_0)|$ is a norm.

Theorem 8.1.3. *Let $G_n = (V, E)$ be a graph with n vertices $(a_1, a_2, \dots, a_n)$. The following statements hold.*

(i) [Boundedness] Let $\alpha \in [0,1)$. For all $0 < p, q \leq \infty$ there exists a constant $C(n,p,q) > 0$ such that

$$\mathrm{Var}_q M_{\alpha,G_n} f \leq C(n,p,q)\mathrm{Var}_p f. \tag{8.5}$$

for all functions $f : V \to \mathbb{R}$.

(ii) [Continuity] Let $0 < p, q \leq \infty$. Consider a sequence of functions $f_j : V \to \mathbb{R}$ such that $\|f_j - f_0\|_{BV_p(G)} \to 0$ as $j \to \infty$.

1. *Assuming that $\lim_{j\to\infty} \min_{x\in V} |f(x) - f_j(x)| = 0$. Then*

$$\mathrm{Var}_q(M_{\alpha,G_n} f - M_{\alpha,G_n} f_j) \to 0 \text{ as } j \to \infty. \tag{8.6}$$

2. *(8.6) could fail to be true without the extra assumption that $\lim_{j\to\infty} \min_{x\in V} |f(x) - f_j(x)| = 0$.*

(iii) M_{α,G_n} is bounded and continuous from $(BV_p(G_n), \|\cdot\|_{BV_p(G_n)})$ to $(BV_p(G_n), \|\cdot\|_{BV_p(G_n)})$.

8.1.2 Optimal l^2 bounds for maximal operators on finite graphs

We are also interested in the l^p norm of M_G when acting on finite graphs. That is, to find the exact value of the expression

$$\sup_{f:V\to\mathbb{R}, f\neq 0} \frac{\|M_G f\|_p}{\|f\|_p} =: \|M_G\|_p,$$

where $\|g\|_p := \left(\sum_{v\in V} |g(v)|^p\right)^{\frac{1}{p}}$, for $g : V \to \mathbb{R}$.

These norms were first treated by Soria and Tradacete, who found $\|M_G\|_p$ when $G = S_n$ and $G = K_n$, where $p \in (0,1)$ (see [ST16, Proposition 2.7] and [ST16, Theorem 3.1]). Their results rely strongly in Jensen's inequality for the function $x \mapsto x^p$ where $p \leq 1$, so those methods are not available when $p > 1$. In fact, they claimed that this problem was difficult when $p > 1$ (see [ST16, Remark 2.8]). The following inequality was proved by Soria and Tradacete [See [ST16], Proposition 2.7]

$$\left(1 + \frac{n-1}{n^2}\right)^{1/2} \leq \|M_{K_n}\|_2 \leq \left(1 + \frac{n-1}{n}\right)^{1/2}.$$

Our next result is a formula for the precise value of $\|M_{K_n}\|_2$ for $n \geq 2$. We also find extremizers for all $n \geq 2$. Moreover, we prove that $\|M_{K_{3n}}\|_2 = \|M_{K_3}\|_2$, for all $n \geq 2$. We list these results as follows.

Theorem 8.1.4. *Let $K_n = (V, E)$ be the complete graph with n vertices $V = \{a_1, a_2, \ldots, a_n\}$. Then we have*

$$\|M_{K_n}\|_2 = \max_{k \in \{\lfloor \frac{n}{3} \rfloor, \lceil \frac{n}{3} \rceil\}} \left(1 - \frac{k}{2n} + \frac{(4kn - 3k^2)^{1/2}}{2n}\right)^{1/2},$$

where $\lfloor x \rfloor := \max\{k \in \mathbb{Z}; k \leq x\}$ (it is the integer part of x) and $\lceil x \rceil := \min\{k \in \mathbb{Z}; k \geq x\}$.

In particular, we have.

Corollary 8.1.1. *If $n = 3m$ for some $m \in \mathbb{N}$, then*

$$\|M_{K_{3m}}\|_2 = \left(\frac{4}{3}\right)^{1/2}.$$

For $n = 2$ we have $\|M_{K_2}\|_2 = \frac{(3+5^{1/2})^{1/2}}{2}$.

Similarly, the following inequality was also proved by Soria and Tradacete [See [ST16], Proposition 3.4]

$$\left(1 + \frac{n-1}{4}\right)^{1/2} \leq \|M_{S_n}\|_2 \leq \left(\frac{n+5}{2}\right)^{1/2}.$$

Our next result is a formula for the precise value of $\|M_{S_n}\|_2$. Moreover, we find some extremizers.

Theorem 8.1.5. *Let $n \geq 4$ and $S_n = (V, E)$ be the star graph with n vertices $V = \{a_1, a_2, a_3, \ldots, a_n\}$ and center at a_1. Then, the following holds.*

$$\|M_{S_n}\|_2 = \left(1 + \frac{n-4}{8} + \frac{(n^2 + 8n)^{1/2}}{8}\right)^{1/2}. \tag{8.7}$$

Remark 8.1.1. It was observed by Soria and Tradecete that in the case $n = 2$ the optimal constant is $\frac{[3+5^{1/2}]^{1/2}}{2}$ (See remark 2.8 in [ST16]), this coincides with our formula (8.7).

8.2 Proof of optimal bounds for the p-variation of maximal functions

We start by proving our results on K_n.

8.2.1 Optimal bounds for the p-variation on K_n: proof of Theorem 8.1.1

For every result listed in Theorem 8.1.1 we can see that, taking $f = \delta_{a_1}$ in the definition of $\mathbf{C}_{K_n,p}$, we have the following.

$$\mathbf{C}_{K_n,p} \geq 1 - \frac{1}{n}.$$

In the following we prove, in each case, that

$$\mathbf{C}_{K_n,p} \leq 1 - \frac{1}{n}. \tag{8.8}$$

A very important tool in the case $p \leq 1$ will be Karamata's inequality, we include the precise statement of this for completeness:

Lemma 8.2.1 (Karamata's Inequality)**.** *Let I be an interval of the real line and let f denote a real valued, convex function defined on I. If $x_1, \dots, x_n$ and $y_1, \dots, y_n$ are numbers in I such that $(x_1, \dots, x_n)$ majorizes $(y_1, \dots, y_n)$, then*

$$f(x_1) + \cdots + f(x_n) \geq f(y_1) + \cdots + f(y_n).$$

Here majorization means that $x_1, \dots, x_n$ and $y_1, \dots, y_n$ satisfies

$$x_1 \geq \cdots \geq x_n \quad \text{and} \quad y_1 \geq \cdots \geq y_n,$$

and we have the inequalities

$$x_1 + x_2 + \cdots + x_i \geq y_1 + y_2 + \cdots + y_i \text{ for all } \; i \in \{1, \dots, n\},$$

and the equality

$$x_1 + x_2 + \cdots + x_n = y_1 + y_2 + \cdots + y_n.$$

Remark 8.2.1. In the case $0 < p \leq 1$, in the proof of our Theorems 8.1.1 and 8.1.2, we will use Karamata's inequality several times in the particular case when $f(x) = -x^p$.

Proof [Proof of Theorem 8.1.1 (i)]

Since by the triangular inequality we have that $\operatorname{Var}_p|f| \leq \operatorname{Var}_p f$ for any function $f : V \to \mathbb{R}$, we can assume without loss of generality that f is nonnegative. Let

$$m := m_n := \frac{\sum_{i=1}^{n} f(a_i)}{n},$$

and for all $k \in \{1, 2, \dots, n-1\}$ we define

$$m_k = \frac{\sum_{i=1}^{k} f(a_i)}{k}.$$

Reordering if necessary, we can assume without loss of generality that

$$f(a_n) \geq f(a_{n-1}) \geq \cdots \geq f(a_r) \geq m > f(a_{r-1}) \geq \cdots \geq f(a_1),$$

thus we have that

$$M_{K_n} f(a_i) = f(a_i) \ \forall \ i \geq r \quad \text{and} \quad M_{K_n} f(a_i) = m \ \forall \ i < r.$$

Let us keep in mind in the following that

$$(\mathrm{Var}_{\,p} M_{K_n} f)^p = \sum_{i,j \in \{r,\ldots,n\}} |f(a_i) - f(a_j)|^p + (r-1) \sum_{i=r}^{n} |f(a_i) - m|^p.$$

Observe that $m_1 \leq m_2 \leq m_3 \leq \cdots \leq m_{n-1} \leq m$. Therefore

$$\begin{aligned}
(\mathrm{Var}_{\,p} M_{K_n} f)^p &\leq (n-1)(f(a_n) - m)^p + (n-2)(f(a_{n-1}) - m)^p \\
&\quad \cdots + (r-1)(f(a_r) - m)^p \\
&\leq (n-1)(f(a_n) - m)^p + (n-2)(f(a_{n-1}) - m_{n-1})^p \\
&\quad \cdots + (r-1)(f(a_r) - m_r)^p. \qquad (8.9)
\end{aligned}$$

Then, we note that by Hölder's inequality

$$f(a_i) - m_i \leq \frac{\sum_{t=1}^{i-1} |f(a_i) - f(a_t)|}{i} \leq \frac{\left(\sum_{t=1}^{i} |f(a_i) - f(a_t)|^p\right)^{1/p} (i-1)^{1/p'}}{i}.$$

where $p' = \frac{p}{p-1}$ denotes the conjugate of p as usual (remind that $p > 1$). Combining the two previous estimatives we obtain

$$\begin{aligned}
(\mathrm{Var}_{\,p} M_{K_n} f)^p &\leq (n-1)(f(a_n) - m)^p + (n-2)(f(a_{n-1}) - m_{n-1})^p \\
&\quad \cdots + (r-1)(f(a_r) - m_r)^p \\
&\leq (n-1)\frac{\left(\sum_{t=1}^{n-1} |f(a_n) - f(a_t)|^p\right)^{p/p} (n-1)^{p/p'}}{n^p} \\
&\quad + (n-2)\frac{\left(\sum_{t=1}^{n-2} |f(a_{n-1}) - f(a_t)|^p\right)^{p/p} (n-2)^{p/p'}}{(n-1)^p} \\
&\quad \cdots + (r-1)\frac{\left(\sum_{t=1}^{r-1} |f(a_r) - f(a_t)|^p\right)^{p/p} (r-1)^{p/p'}}{r^p} \\
&\leq \left(\frac{n-1}{n}\right)^p \sum_{t=1}^{n-1} |f(a_n) - f(a_t)|^p \\
&\quad + \left(\frac{n-2}{n-1}\right)^p \sum_{t=1}^{n-2} |f(a_{n-1}) - f(a_t)|^p \\
&\quad + \cdots + \left(\frac{r-1}{r}\right)^p \sum_{t=1}^{r-1} |f(a_r) - f(a_t)|^p \\
&\leq \left(\frac{n-1}{n}\right)^p (\mathrm{Var}_{\,p} f)^p.
\end{aligned}$$

From where we conclude (8.8) in this case. Concluding the proof of this assertion of Theorem 8.1.1.

■

Case $p \leq 1$: proof of assertion (ii) and (iii) in Theorem 8.1.1.

We keep the notation of the previous proof and the assumption that

$$f(a_n) \geq \dots f(a_r) \geq m > f(a_{r-1}) \geq \cdots \geq f(a_1).$$

For $0 < p \leq 1$, the simplest case of the theorem is when $r = n$.

Lemma 8.2.2. *For every $0 < p \leq 1$ and $n \geq 2$, if $r = n$, we have $\mathbf{C}_{K_n,p} = 1 - \frac{1}{n}$.*

Proof This can be proved directly by

$$\begin{aligned}(n-1)|f(a_n) - m|^p \leq &(n-1)\left|\frac{n-1}{n}(f(a_n) - f(a_1))\right|^p \\ &\leq \left(\frac{n-1}{n}\right)^p \Big(|f(a_n) - f(a_1)|^p \\ &\quad + \sum_{i=2}^{n-1} |f(a_n) - f(a_i)|^p + |f(a_i) - f(a_1)|^p\Big) \\ &\leq \left(\frac{n-1}{n}\right)^p (\mathrm{Var}_p f)^p,\end{aligned}$$

where, in the second inequality, we used that if $a, b \geq 0$, then $(a+b)^p \leq a^p + b^p$. ■

Therefore, in the following we assume that $r < n$.

Proof [Proof of Theorem 8.1.1 (ii)] Now we prove the assertion for $n = 4$. Since the case $r = 4$ was already solved, we have two cases left. First we treat the case $r = 3$.

Case $r = 3$. We have the following inequality.

$$\begin{aligned}\left(\frac{3}{4}\right)^p \big(|f(a_4) - f(a_3)|^p + |f(a_3) - f(a_2)|^p + &|f(a_2) - f(a_1)|^p\big) \\ &\geq |f(a_4) - f(a_3)|^p + |f(a_3) - m|^p. \qquad (8.10)\end{aligned}$$

Step 1: Proving (8.10). In order to prove this, we write $f(a_3) - f(a_2) = x$ and $f(a_4) - f(a_3) = y$, then $m = \frac{f(a_1)+3f(a_2)+2x+y}{4}$ and

$$\frac{f(a_1) + 3f(a_2) + 2x + y}{4} \leq f(a_2) + x \implies f(a_1) + y \leq f(a_2) + 2x, \qquad (8.11)$$

also

$$m \geq f(a_2) \implies f(a_2) \leq f(a_1) + 2x + y. \tag{8.12}$$

Then

$$\left(\frac{3}{4}\right)^p (|f(a_4) - f(a_3)|^p + |f(a_3) - f(a_2)|^p + |f(a_2) - f(a_1)|^p)$$
$$= \left(\frac{3}{4}\right)^p (y^p + x^p + (f(a_2) - f(a_1))^p),$$

Consider first the case where $f(a_2) - f(a_1) + 2x \leq 4y$. Here, we observe that $f(a_2) - f(a_1) + 2x \in [f(a_2) - f(a_1) + 2x - y, 4y]$, and since $(f(a_2) - f(a_1) + 2x) + 3y = (f(a_2) - f(a_1) + 2x - y) + 4y$ we have that $3y \in [f(a_2) - f(a_1) + 2x - y, 4y]$, and then, by Karamata's inequality, we have

$$(3y)^p + (f(a_2) - f(a_1) + 2x)^p \geq (4y)^p + (f(a_2) - f(a_1) + 2x - y)^p.$$

Now, since $(3x)^p + (3(f(a_2) - f(a_1)))^p \geq (f(a_2) - f(a_1) + 2x)^p$, we obtain

$$(3y)^p + (3x)^p + (3(f(a_2) - f(a_1)))^p \geq (4y)^p + (f(a_2) - f(a_1) + 2x - y)^p, \tag{8.13}$$

from where (8.10) follows by observing that $4(f(a_3) - m) = 4(f(a_3) - \frac{f(a_1)+3f(a_2)+2x+y}{4}) = f(a_2) - f(a_1) - 2x - y \leq f(a_2) - f(a_1) + 2x - y$.

Now we consider the other case, where $f(a_2) - f(a_1) + 2x \geq 4y$. We do some previous considerations. First, we have that $3^p \geq \frac{4^p - 3^p}{4^p} + 1$ and $3^p \geq \frac{4^p - 3^p}{2^p} + 2^p$, both consequences of the following application of the AM-GM inequality

$$12^p + 3^p > 6^p + 3^p \geq 2(18)^{p/2} > 2(4)^p.$$

Also, let us observe that

$$\begin{aligned}(4^p - 3^p)\left(\frac{f(a_2) - f(a_1)}{4} + \frac{x}{2}\right)^p &\leq (4^p - 3^p)\left(\frac{f(a_2) - f(a_1)}{4}\right)^p + (4^p - 3^p)\left(\frac{x}{2}\right)^p \\ &= \frac{4^p - 3^p}{4^p}(f(a_2) - f(a_1))^p + \frac{4^p - 3^p}{2^p}x^p\end{aligned} \tag{8.14}$$

and

$$(f(a_2) - f(a_1) + 2x)^p \leq (f(a_2) - f(a_1))^p + 2^p x^p. \tag{8.15}$$

Now, by considering that (here we use $f(a_2) - f(a_1) + 2x \geq 4y$)

$$(4^p - 3^p)\left(\frac{f(a_2) - f(a_1)}{4} + \frac{x}{2}\right)^p + (f(a_2) - f(a_1) + 2x)^p \geq (4^p - 3^p)(y)^p + (f(a_2) - f(a_1) + 2x - y)^p,$$

we have that by (8.14) and (8.15) (and the already mentioned inequalities for 3^p):

$$\begin{aligned}(3(f(a_2) - f(a_1)))^p + (3x)^p &\geq \left(\frac{4^p - 3^p}{4^p} + 1\right)(f(a_2) - f(a_1))^p + \left(\frac{4^p - 3^p}{2^p} + 2^p\right)x^p \\ &\geq (4^p - 3^p)\left(\frac{f(a_2) - f(a_1)}{4} + \frac{x}{2}\right)^p + (f(a_2) - f(a_1) + 2x)^p \\ &\geq (4^p - 3^p)y^p + (f(a_2) - f(a_1) + 2x - y)^p.\end{aligned}$$

From this (8.13) follows, and therefore we conclude Step 1.
In the following, we also need the inequality

$$\left(\frac{3}{4}\right)^p (|f(a_4) - f(a_2)|^p + |f(a_3) - f(a_1)|^p) \geq |f(a_4) - m|^p + |f(a_3) - m|^p. \tag{8.16}$$

Step 2: Proving (8.16). We have that (8.16) is equivalent to

$$\begin{aligned}(3x + 3y)^p + (3x + 3(f(a_2) - f(a_1)))^p \geq (f(a_2) - &f(a_1) + 2x + 3y)^p \\ &+ (f(a_2) - f(a_1) + 2x - y)^p,\end{aligned}$$

since $4(f(a_4) - m) = 4(f(a_2) + x + y - \frac{f(a_1)+3f(a_2)+2x+y}{4}) = f(a_2) - f(a_1) + 2x + 3y$ and $4(f(a_3) - m) = 4(f(a_4) - m) - 4y = f(a_2) - f(a_1) + 2x - y$. Here we distinguish among two cases, the first when $x + 4y \geq f(a_2) - f(a_1)$. Here, by the concavity of the function $x \mapsto x^p$, since

$$4x + 2(f(a_2) - f(a_1)) + 4y \geq 3x + 3y \geq 2x + f(a_2) - f(a_1) - y,$$

and

$$4x + 2(f(a_2) - f(a_1)) + 4y \geq 3x + 3(f(a_2) - f(a_1)) \geq 2x + f(a_2) - f(a_1) - y,$$

by Karamata's inequality for $-x^p$ we have

$$\begin{aligned}(3x + 3y)^p + (3x + 3(f(a_2) - f(a_1)))^p &\geq (4x + 2(f(a_2) - f(a_1)) + 4y)^p \\ &+ (2x + (f(a_2) - f(a_1)) - y)^p \\ &\geq ((f(a_2) - f(a_1)) + 2x + 3y)^p \\ &+ (f(a_2) - f(a_1) + 2x - y)^p,\end{aligned}$$

from where (8.16) follows.

Now we deal with the the other case, where $x + 4y \leq f(a_2) - f(a_1)$. We can prove that (this is independent to $x + 4y \leq f(a_2) - f(a_1)$):

$$\begin{aligned}(f(a_2) - f(a_1) + 2x + 2y)^p + (f(a_2) - f(a_1) + 2x)^p \geq (&f(a_2) - f(a_1) + 2x + 3y)^p \\ &+ (f(a_2) - f(a_1) + 2x - y)^p,\end{aligned} \tag{8.17}$$

by Karamata's inequality for the function $-x^p$, considering that $f(a_2) - f(a_1) + 2x - y \leq f(a_2) - f(a_1) + 2x \leq f(a_2) - f(a_1) + 2x + 2y \leq f(a_2) - f(a_1) + 2x + 3y$. Also, since $y \leq \frac{f(a_2)-f(a_1)}{4}$ and $2x + y \geq f(a_2) - f(a_1)$ (recall (8.12)), we have $x \geq \frac{3(f(a_2)-f(a_1))}{8}$, therefore we obtain (by just expanding)

$$(3x)(3x + 3(f(a_2) - f(a_1))) \geq \left(\frac{3}{2}(f(a_2) - f(a_1)) + 2x\right)(f(a_2) - f(a_1) + 2x)$$

and, as a consequence,

$$\log(3x) + \log(3x + 3(f(a_2) - f(a_1))) \geq \log\left(\frac{3}{2}(f(a_2) - f(a_1)) + 2x\right) + \log(f(a_2) - f(a_1) + 2x). \tag{8.18}$$

Let us observe that, since $x \leq x + 4y \leq f(a_2) - f(a_1)$, we have

$$\log(3x) \leq \log(f(a_2) - f(a_1) + 2x) \leq \log\left(\frac{3}{2}(f(a_2) - f(a_1) + 2x\right), \tag{8.19}$$

let us take then $v := \log(f(a_2)-f(a_1)+2x)+\log(\frac{3}{2}(f(a_2)-f(a_1)+2x)-\log(3x)$, by (8.18) we have $v \leq \log(3x + 3(f(a_2) - f(a_1)))$ and by (8.19) we have $v \geq \log\left(\frac{3}{2}(f(a_2) - f(a_1)) + 2x\right)$. By Karamata's inequality, now applied to the convex function $x \mapsto e^{px}$, by considering (8.19) we have

$$e^{p\log(\frac{3}{2}(f(a_2)-f(a_1))+2x)} + e^{p\log(f(a_2)-f(a_1)+2x)} \leq e^{p\log(3x)} + e^{pv} \leq e^{p\log(3x)} + e^{p\log(3x+3(f(a_2)-f(a_1)))},$$

and therefore:

$$\begin{aligned}(3x + 3y)^p + (3x + 3(f(a_2) - f(a_1)))^p &\geq (3x)^p + (3x + 3(f(a_2) - f(a_1)))^p \\ &= e^{p\log(3x)} + e^{p\log(3x+3(f(a_2)-f(a_1)))} \\ &\geq e^{p\log(\frac{3}{2}(f(a_2)-f(a_1))+2x)} + e^{p\log(f(a_2)-f(a_1)+2x)} \\ &= (\frac{3}{2}(f(a_2) - f(a_1)) + 2x)^p + (f(a_2) - f(a_1) + 2x)^p \\ &\geq (f(a_2) - f(a_1) + 2x + 2y)^p + (f(a_2) - f(a_1) + 2x)^p,\end{aligned}$$

where we used $f(a_2) - f(a_1) \geq 4y$ in the last inequality. Therefore we obtain, by combining this with (8.17), the desired inequality (8.16).

Step 3: Conclusion of case $r = 3$. The case $r = 3$ then follows by combining (8.16), (8.10) and the inequality $\left(\frac{3}{4}\right)^p (f(a_4) - f(a_1))^p = \left(f(a_4) - \frac{f(a_4)+3f(a_1)}{4}\right)^p \geq (f(a_4) - m)^p$. In fact, adding these three inequalities we obtain that

$$\begin{aligned}\left(\frac{3}{4}\right)^p (\mathrm{Var}_p f)^p = &\left(\frac{3}{4}\right)^p (|f(a_4) - f(a_3)|^p + |f(a_3) - f(a_2)|^p + |f(a_2) - f(a_1)|^p) \\ &+ \left(\frac{3}{4}\right)^p (|f(a_4) - f(a_2)|^p + |f(a_3) - f(a_1)|^p) + \left(\frac{3}{4}\right)^p |f(a_4) - f(a_1)|^p \\ \geq &\, (|f(a_4) - f(a_3)|^p + |f(a_3) - m|^p) \\ &+ (|f(a_4) - m|^p + |f(a_3) - m|^p) + |f(a_4) - m|^p \\ = &\, (\mathrm{Var}_p M_{K_4} f)^p.\end{aligned}$$

Case $r = 2$. Here, we have that $m \leq f(a_2) \implies 2x + y \leq f(a_2) - f(a_1)$ since $f(a_2) - m = f(a_2) - \frac{f(a_1)+3f(a_2)+2x+y}{4} = \frac{f(a_2)-f(a_1)-2x-y}{4}$. We prove first the inequality

$$|f(a_4) - f(a_3)|^p + |f(a_3) - f(a_2)|^p + |f(a_2) - m|^p \leq \left(\frac{3}{4}\right)^p (|f(a_4) - f(a_3)|^p + |f(a_3) - f(a_2)|^p + |f(a_2) - f(a_1)|^p), \tag{8.20}$$

Step 1: Proving (8.20). Our desired inequality (8.20) is equivalent (since $4(f(a_2) - m) = 4(f(a_2) - \frac{f(a_1)+3f(a_2)+2x+y}{4}) = f(a_2) - f(a_1) - 2x - y$) to

$$(4y)^p + (4x)^p + (f(a_2) - f(a_1) - 2x - y)^p \leq (3y)^p + (3x)^p + (3(f(a_2) - f(a_1)))^p.$$

We observe that

$$(4y)^p + (4x)^p + (f(a_2) - f(a_1) - 2x - y)^p \leq (4y)^p + (4x)^p + (f(a_2) - f(a_1))^p,$$

also, since $x \leq \frac{f(a_2)-f(a_1)}{2}$ and $y \leq f(a_2) - f(a_1)$, we have

$$(4^p - 3^p)(x^p + y^p) \leq (4^p - 3^p)\left[\left(\frac{1}{2}\right)^p + 1\right](f(a_2) - f(a_1))^p \leq (3^p - 1)(f(a_2) - f(a_1))^p,$$

because $4^p + 8^p + 2^p \leq 2(6)^p + 3^p$ by Jensen inequality. Therefore

$$\begin{aligned}(4y)^p + (4x)^p + (f(a_2) - f(a_1) - 2x - y)^p &\leq 4^p x^p + 4^p y^p + (f(a_2) - f(a_1))^p \\ &\leq (3y)^p + (3x)^p + (3(f(a_2) - f(a_1)))^p,\end{aligned}$$

from where it follows (8.20). This conclude Step 1.

Also, we have that

$$(|f(a_4) - f(a_2)|^p + |f(a_3) - f(a_1)|^p)\left(\frac{3}{4}\right)^p \geq (|f(a_4) - f(a_2)|^p + |f(a_3) - m|^p), \tag{8.21}$$

Step 2: Proving (8.21). We have that (8.21) is equivalent to

$$(4x + 4y)^p + (f(a_2) - f(a_1) + 2x - y)^p \leq (3x + 3y)^p + (3(x + f(a_2) - f(a_1)))^p,$$

this happens since $f(a_4) - f(a_2) = x + y$, $f(a_3) - f(a_1) = x + f(a_2) - f(a_1)$ and $f(a_3) - m = f(a_2) + x - \frac{f(a_1)+3f(a_2)+2x+y}{4} = \frac{f(a_2)-f(a_1)+2x-y}{4}$. Also, since by Jensen $4^p + 2^p \leq 2(3)^p$ and because of $m \leq f(a_2) \implies 2x + y \leq f(a_2) - f(a_1) \implies x + y \leq f(a_2) - f(a_1)$, we have

$$\begin{aligned}(4^p - 3^p)(x + y)^p + (2(f(a_2) - f(a_1)))^p \leq (3(f(a_2) - f(a_1)))^p &\leq (3(f(a_3) - f(a_1)))^p \\ &= (3(x + f(a_2) - f(a_1)))^p.\end{aligned}$$

By observing that $2x \leq f(a_2) - f(a_1)$ and then $f(a_2) - f(a_1) + 2x - y \leq 2(f(a_2) - f(a_1)$ we have

$$\begin{aligned}(4^p - 3^p)(x+y)^p + (f(a_2) - f(a_1) + 2x - y)^p &\leq (4^p - 3^p)(x+y)^p + (2(f(a_2) - f(a_1)))^p \\ &\leq (3(x + f(a_2) - f(a_1)))^p,\end{aligned}$$

from where we obtain (8.21) and therefore we conclude Step 2.

Step 3: Conclusion of $r = 2$ and $n = 4$. The case $r = 2$, and thus our result in $n = 4$ follows by combining (8.20), (8.21) and the inequality $f(a_4) - m \leq (\frac{3}{4})(f(a_4) - f(a_1))$. We conclude this part and thus the assertion. ∎

Now, we prove our assertion for general n and $p \in [\frac{\log 4}{\log 6}, 1]$.

Proof [Proof of Theorem 8.1.1 (iii)] The strategy that we follow in order to prove this assertion is the following inductive argument. Proving (8.8) is equivalent to proving that for each $f : V \to \mathbb{R}_{\geq 0}$, we have

$$\sum_{i,j \in \{r,\ldots,n\}} |f(a_i) - f(a_j)|^p + \sum_{i=r}^{n} |f(a_i) - m|^p \leq \left(1 - \frac{1}{n}\right)^p \left(\sum_{i,j \in \{1,\ldots,n\}} |f(a_i) - f(a_j)|^p\right). \tag{8.22}$$

In order to prove the inequality above for $\frac{\log 4}{\log 6} \leq p \leq 1$, we establish in this range a control over the contribution of the vertex a_r to (8.22), that is, we prove that

$$\sum_{i=r+1}^{n} |f(a_i) - f(a_r)|^p + (n-r)|f(a_r) - m|^p \leq \left(1 - \frac{1}{n}\right)^p \left(\sum_{i=1}^{n} |f(a_r) - f(a_i)|^p\right). \tag{8.23}$$

Then, we observe that the analogous of (8.22) for the graph K_{n-1}, obtained by deleting the vertex a_r (and the respective edges) from K_n, and a good choice of $\widetilde{f} : V \setminus \{a_r\}$ would give us a proper control over the edges of K_n that were not considered in (8.23). By combining this control with (8.23) we would obtain (8.22) for our initial f. Thus, the induction concludes our result (since we know that the result holds for $n = 3$).

We proceed to our proof by proving first (8.23). We write $x_i := f(a_i) - f(a_r)$ for $i = n, \ldots, r+1$, $u = f(a_r) - m$, $y_i = f(a_r) - f(a_i)$ for $i = r-1, \ldots, 1$. We have then, since

$$\begin{aligned} m = \frac{\sum_{i=1}^{n} f(a_i)}{n} &= \frac{\sum_{i=1}^{r-1} f(a_r) - y_i + f(a_r) + \sum_{i=r+1}^{n} f(a_r) + x_i}{n} \\ &= f(a_r) + \frac{\sum_{i=r+1}^{n} x_i - \sum_{i=1}^{r} y_i}{n}, \end{aligned}$$

that $\displaystyle\sum_{i=r+1}^{n} x_i + nu = \sum_{i=1}^{r-1} y_i$. Then, (8.23) is equivalent to:

$$\sum_{i=r+1}^{n} x_i^p + (r-1)u^p \leq \left(1-\frac{1}{n}\right)^p \left(\sum_{i=r+1}^{n} x_i^p + \sum_{i=1}^{r-1} y_i^p\right). \tag{8.24}$$

In order to prove that, since $y_i \geq u$ for every $i = 1, \ldots, r-1$ and $\sum_{i=1}^{r-1} y_i = (r-2)u + \left(\sum_{i=1}^{r-1} y_i - (r-2)u\right)$, we have $\sum_{i=1}^{r-1} y_i - (r-2)u \geq y_j \geq u$ for every $j = 1, \ldots, r-1$. Moreover, $y_1 \geq y_2 \geq \cdots \geq y_{r-1}$ and $\left(\sum_{i=1}^{r-1} y_i - (r-2)u\right) + ku \geq \sum_{i=1}^{k+1} y_i$ for each $k = 0, \ldots r-2$ since $\sum_{i=k+2}^{r-1} y_i \geq (r-2-k)u$, then by Karamata's inequality, we have (we also use here that $\sum_{i=r+1}^{n} x_i + nu = \sum_{i=1}^{r-1} y_i$):

$$\sum_{i=1}^{r-1} y_i^p \geq (r-2)u^p + \left(\sum_{i=1}^{r-1} y_i - (r-2)u\right)^p = (r-2)u^p + \left((n-r+2)u + \sum_{i=r+1}^{n} x_i\right)^p,$$

also, by Jensen's inequality we have

$$\left((n-r+2)u + \sum_{i=r+1}^{n} x_i\right)^p \geq 2^{p-1}\left[((n-r+2)u)^p + \left(\sum_{i=r+1}^{n} x_i\right)^p\right].$$

Therefore, combining the two previous inequalities we obtain

$$\begin{aligned}\left(1-\frac{1}{n}\right)^p &\left(\sum_{i=r+1}^{n} x_i^p + \sum_{i=1}^{r-1} y_i^p\right)\\ &\geq \left(1-\frac{1}{n}\right)^p \left[\sum_{i=r+1}^{n} x_i^p + (r-2)u^p + 2^{p-1}((n-r+2)u)^p + 2^{p-1}\left(\sum_{i=r+1}^{n} x_i\right)^p\right].\end{aligned}$$

Then, since $\left(\sum_{i=r+1}^{n} x_i\right)^p \geq (n-r)^{p-1}\left(\sum_{i=r+1}^{n} x_i^p\right)$ (by Hölder's Inequality) we get

$$\begin{aligned}\left(1-\frac{1}{n}\right)^p &\left(\sum_{i=r+1}^{n} x_i^p + \sum_{i=1}^{r-1} y_i^p\right)\\ &\geq \left(1-\frac{1}{n}\right)^p \left(\sum_{i=r+1}^{n} x_i^p(1+2^{p-1}(n-r)^{p-1}) + u^p(r-2+2^{p-1}(n-r+2)^p)\right).\end{aligned}$$

Then, in order to obtain (8.24) it is enough to see that

$$1 \leq \left(1 - \frac{1}{n}\right)^p \left(1 + 2^{p-1}(n-r)^{p-1}\right) \tag{8.25}$$

and

$$r - 1 \leq \left(1 - \frac{1}{n}\right)^p \left(r - 2 + 2^{p-1}(n-r+2)^p\right). \tag{8.26}$$

To prove (8.25) it is enough to see that (since $r \leq n-1 \implies n-r \geq 1$ and $n \geq 3 \implies 1 - \frac{1}{n} \geq \frac{2}{3}$)

$$1 \leq \left(\frac{2}{3}\right)^p (1 + 2^{p-1})$$

and that is equivalent to $\left(\frac{3}{2}\right)^p \leq 1 + 2^{p-1}$, which follows since $p \in (0,1)$ and then $(\frac{3}{2})^p \leq \frac{3}{2} = 1 + \frac{1}{2} \leq 1 + \frac{2^p}{2}$. Now, to prove (8.26), we observe that (since $p \geq \frac{\log 4}{\log 6}$, $n - r + 2 \geq 3$),

$$2^{p-1}(n-r+2)^p \geq 2^{p-1}(3)^p = \frac{6^p}{2} \geq 2,$$

and therefore

$$\left(1 - \frac{1}{n}\right)^p (r - 2 + 2^{p-1}(n-r+2)^p) \geq \left(1 - \frac{1}{n}\right)^p (r) \geq \left(1 - \frac{1}{n}\right) r \geq (r-1),$$

where in the last inequality we use $1 - \frac{1}{n} \geq 1 - \frac{1}{r}$ since $r \leq n$. From this we conclude (8.26) and thus (8.24) follows.

Now we follow with the remaining steps of our proof. Assume that inequality $\mathrm{Var}_p M_{K_{n-1}} f \leq (1 - \frac{1}{n-1})\mathrm{Var}_p f$, holds for every $\widetilde{f} : V(K_{n-1}) \to \mathbb{R}_{\geq 0}$ in K_{n-1}, whenever $p \geq \frac{\log 4}{\log 6}$ (it holds for $n = 3, 4$). Then, if $b_1, \ldots, b_{n-1}$ are the vertices of the K_{n-1} graph, we define $\widetilde{f}$ as $\widetilde{f}(b_i) = f(a_{i+1})$ for $i = r, \ldots, n-1$ and $\widetilde{f}(b_i) = f(a_i)$ for $i = 1, \ldots, r-1$. We write $\widetilde{m} = \frac{\sum_{i=1}^{n-1} \widetilde{f}(b_i)}{n-1}$. Since $f(a_r) \geq m$, we observe that $\widetilde{m} = \frac{nm - f(a_r)}{n-1} \leq m$. Then, we write $\widetilde{f}(b_s) \geq \widetilde{m} > \widetilde{f}(b_{s-1})$, where we observe that $s \leq r$. By the inductive hypothesis, we have

$$\begin{aligned} \sum_{i,j \in \{s,\ldots,n-1\}} |\widetilde{f}(a_i) - \widetilde{f}(a_j)|^p + (s-1) \sum_{i=s}^{n-1} |\widetilde{f}(a_i) - \widetilde{m}|^p \\ \leq \left(1 - \frac{1}{n-1}\right)^p \left(\sum_{i,j \in \{1,\ldots,n-1\}} |\widetilde{f}(a_i) - \widetilde{f}(a_j)|^p \right). \end{aligned} \tag{8.27}$$

By noticing that

$$\sum_{i=s}^{n-1} |\widetilde{f}(a_i) - \widetilde{f}(a_j)|^p = \sum_{i,j\in\{r+1,\dots,n\}} |f(a_i) - f(a_j)|^p + \sum_{i=r+1}^{n}\sum_{j=s}^{r-1} |f(a_i) - f(a_j)|^p$$
$$+ \sum_{i,j\in\{s,\dots,r-1\}} |f(a_i) - f(a_j)|^p \tag{8.28}$$
$$\geq \sum_{i,j\in\{r+1,\dots,n\}} |f(a_i) - f(a_j)|^p + (r-s)\sum_{i=r+1}^{n} |f(a_i) - m|^p$$

where in the last inequality we used that $f(a_i) \geq m > f(a_j)$ for $i \geq r+1$ and $j \leq r-1$. Then, by combining (8.27), (8.28) and using that $(s-1)\sum_{i=s}^{n-1} |\widetilde{f}(a_i) - \widetilde{m}|^p \geq (s-1)\sum_{i=r+1}^{n} |f(a_i) - m|^p$ (since $\widetilde{m} \leq m$), we have

$$\left(1 - \frac{1}{n-1}\right)^p \left(\sum_{i,j\in\{1,..r-1,r+1,..,n\}} |f(a_i) - f(a_j)|^p\right)$$
$$\geq \sum_{i,j\in\{r+1,\dots,n\}} |f(a_i) - f(a_j)|^p + (r-s)\sum_{i=r+1}^{n} |f(a_i) - m|^p$$
$$+ (s-1)\sum_{i=r+1}^{n} |f(a_i) - m|^p$$
$$\geq \sum_{i,j\in\{r+1,\dots,n\}} |f(a_i) - f(a_j)|^p + (r-1)\sum_{i=r+1}^{n} |f(a_i) - m|^p.$$

Combining this with (8.23) we conclude

$$\sum_{i,j\in\{r,..,n\}} |f(a_i) - f(a_j)|^p + (r-1)\sum_{i=r}^{n} |f(a_i) - m|^p \leq \left(1 - \frac{1}{n}\right)^p \left(\sum_{i,j\in\{1..,n\}} |f(a_i) - f(a_j)|^p\right),$$

that is equivalent to (8.8) in this case. This concludes the proof of our theorem. ■

Remark 8.2.2. We observe that proving (8.24) in a larger range implies a proof of Theorem 8.1.1 (iii) in the same range. This is the case because the remaining of the proof is independent of the condition $p \geq \frac{\log 4}{\log 6}$.

8.2.2 Optimal bounds for the p-variation on S_n: proof of Theorem 8.1.2

Now we deal with the problems related to the p-variation of the maximal operator in S_n.

Proof [Proof of Theorem 8.1.2 (i)] We assume without loss of generality that f is nonnegative. We analyse three different cases. *Case 1:* $f(a_1) \geq \max\{f(a_2), f(a_3)\}$.
In this case we have that $M_{S_3} f(a_1) = f(a_1)$, then

$$\begin{aligned}
(\mathrm{Var}_{\,p} M_{S_3} f)^p &\leq \left(f(a_1) - \frac{f(a_1)+f(a_2)}{2}\right)^p + \left(f(a_1) - \frac{f(a_1)+f(a_3)}{2}\right)^p \\
&\leq \frac{1}{2^p}(\mathrm{Var}_{\,p} f)^p.
\end{aligned}$$

Case 2: $f(a_1) \leq \min\{f(a_2), f(a_3)\}$. We assume without loss of generality that $f(a_1) \leq f(a_3) \leq f(a_2)$. Then, we have that

$$\begin{aligned}
(\mathrm{Var}_{\,p} M_{S_3} f)^p &= \left(f(a_2) - \frac{f(a_1)+f(a_2)+f(a_3)}{3}\right)^p + \left(\left[f(a_3) - \frac{f(a_1)+f(a_2)+f(a_3)}{3}\right]_+\right)^p \\
&= \left(\frac{f(a_2)-f(a_1)+f(a_2)-f(a_3)}{3}\right)^p + \left(\left[\frac{f(a_3)-f(a_1)-(f(a_2)-f(a_3))}{3}\right]_+\right)^p \\
&= \left(\frac{2(f(a_2)-f(a_1))-(f(a_3)-f(a_1))}{3}\right)^p + \left(\left[\frac{2(f(a_3)-f(a_1))-(f(a_2)-f(a_1))}{3}\right]_+\right)^p \\
&\leq \left(\frac{2(f(a_2)-f(a_1))-(f(a_3)-f(a_1))}{3} + \left[\frac{2(f(a_3)-f(a_1))-(f(a_2)-f(a_1))}{3}\right]_+\right)^p \\
&\leq \frac{(1+2^{p'})^{p/p'}}{3^p}(\mathrm{Var}_{\,p} f)^p.
\end{aligned}$$

Where we have used the fact that $p > 1$ in the fourth line and the final step follows by Hölder's inequality.

Case 3: $\min\{f(a_2), f(a_3)\} < f(a_1) < \max\{f(a_2), f(a_3)\}$. We assume without loss of generality that $f(a_3) < f(a_1) < f(a_2)$. Then, we have

$$\begin{aligned}
(\mathrm{Var}_{\,p} M_{S_3} f)^p &= (f(a_2) - M_{S_3} f(a_1))^p + \left(M_{S_3} f(a_1) - \frac{f(a_1)+f(a_2)+f(a_3)}{3}\right)^p \\
&\leq \left(f(a_2) - \frac{f(a_1)+f(a_2)+f(a_3)}{3}\right)^p \\
&= \left(\frac{2(f(a_2)-f(a_1))+(f(a_1)-f(a_3))}{3}\right)^p \\
&\leq \frac{(1+2^{p'})^{p/p'}}{3^p}(\mathrm{Var}_{\,p} f)^p.
\end{aligned}$$

In the second line we used the fact that $p > 1$. In the final inequality (in the last two cases) we have used that $c_1 d_1 + c_2 d_2 \leq (c_1^{p'} + c_2^{p'})^{1/p'}(d_1^p + d_2^p)^{1/p}$ for all $c_1, c_2, d_1, d_2 \geq 0$ by Hölder's inequality. This conclude the proof of

$$\mathbf{C}_{S_3,p} \leq \frac{(1+2^{p/(p-1)})^{(p-1)/p}}{3}$$

in (8.2). Finally, we observe that

$$\mathbf{C}_{S_3,p} \geq \frac{(1+2^{p/(p-1)})^{(p-1)/p}}{3}. \tag{8.29}$$

For that we consider the function $f : V \to \mathbb{R}$ defined by

$$f(a_3) = 2, f(a_1) = 3 \text{ and } f(a_2) = 3 + 2^{\frac{1}{p-1}}.$$

Then, $\text{Var}_p f = (1 + 2^{\frac{p}{p-1}})^{\frac{1}{p}}$. Moreover,

$$M_{S_3} f(a_2) = f(a_2) = 3 + 2^{\frac{1}{p-1}} \text{ and } M_{S_3}(a_3) = M_{S_3} f(a_1) = \frac{2+3+3+2^{\frac{1}{p-1}}}{3}.$$

Thus

$$\text{Var}_p M_{S_3} f = M_{S_3} f(a_2) - M_{S_3} f(a_1) = \frac{1+2^{\frac{p}{p-1}}}{3}.$$

Therefore

$$\frac{\text{Var}_p M_{S_3} f}{\text{Var}_p f} = \frac{(1+2^{p'})^{\frac{1}{p'}}}{3}.$$

So, we obtain (8.29) and thus (8.2). ■

The proof of the previous result provides an example where the value

$$\sup_{f:V\to\mathbb{R};\text{Var}_p f>0} \frac{\text{Var}_p M_G f}{\text{Var}_p f}$$

is not attained by any *Dirac's delta.* This is a sign of the complexity of this problem when $p > 1$, since is not clear how the extremizers should behave for $n > 3$.

In the case $p = 2$, an interesting example is the following: let $S_n = (V, E)$ as in the Theorem 8.1.2 where a_1 is the vertex of degree $n-1$, consider the function $f : V \to \mathbb{R}$ defined by

$$f(a_1) = n,\ f(a_2) = n + (n-1), \text{ and } f(a_i) = n-1 \text{ for all } 3 \leq i \leq n.$$

In this case

$$M_{S_n} f(a_2) = n + (n-1) \text{ and } M_{S_n} f(a_i) = n + \frac{1}{n} \text{ for all } i \neq 2.$$

Then

$$\frac{\text{Var}_2 M_{S_n} f}{\text{Var}_2 f} = \frac{n-1-\frac{1}{n}}{[(n-1)^2+(n-2)]^{1/2}} = \frac{[(n-1)^2+(n-2)]^{1/2}}{n} > \frac{n-1}{n}.$$

This provides further evidence to the fact that in general the extremizers on S_n are different when $p > 1$ than when $p \leq 1$.
Now we deal with the next assertion of our theorem. Taking $f = \delta_{a_2}$ on the definition of $\mathbf{C}_{S_n,p}$ we have that

$$\mathbf{C}_{S_n,p} \geq 1 - \frac{1}{n}.$$

In the following we prove the inequality

$$\mathbf{C}_{S_n,p} \leq 1 - \frac{1}{n}, \tag{8.30}$$

from where both assertion follow. This inequality is equivalent to

$$\mathrm{Var}_{\,p} M_{S_n} f \leq (1 - \frac{1}{n}) \mathrm{Var}_{\,p} f, \tag{8.31}$$

for all functions $f : V \to \mathbb{R}$.
Proof [Proof of Theorem 8.1.2 (ii)] We assume without loss of generality that f is nonnegative. Let

$$m = \frac{1}{n} \sum_{i=1}^{n} f(a_i).$$

Then

$$\begin{aligned}
&\operatorname{Var} M_{S_n} f \\
&= \sum_{i=2}^{n} |M_{S_n} f(a_i) - M_{S_n} f(a_1)| \\
&= \sum_{M_{S_n} f(a_i) > M_{S_n} f(a_1)} M_{S_n} f(a_i) - M_{S_n} f(a_1) + \sum_{M_{S_n} f(a_1) > M_{S_n} f(a_i)} M_{S_n} f(a_1) - M_{S_n} f(a_i) \\
&= \sum_{M_{S_n} f(a_i) > M_{S_n} f(a_1)} f(a_i) - M_{S_n} f(a_1) + \sum_{M_{S_n} f(a_1) > M_{S_n} f(a_i)} f(a_1) - M_{S_n} f(a_i) \\
&\le \sum_{M_{S_n} f(a_i) > M_{S_n} f(a_1)} f(a_i) - m + \sum_{M_{S_n} f(a_1) > M_{S_n} f(a_i)} f(a_1) - m \\
&= \sum_{M_{S_n} f(a_i) > M_{S_n} f(a_1)} \left[\frac{n-1}{n} (f(a_i) - f(a_1)) + \sum_{j \ne i} \frac{f(a_1) - f(a_j)}{n} \right] \\
&\qquad + \sum_{M_{S_n} f(a_1) > M_{S_n} f(a_i)} \sum_{k=2}^{n} \frac{f(a_1) - f(a_k)}{n} \\
&= \sum_{M_{S_n} f(a_i) > M_{S_n} f(a_1)} (f(a_i) - f(a_1)) \left[\frac{n-1}{n} - \frac{(|\{i; M_{S_n} f(a_i) > M_{S_n} f(a_1)\}| - 1)}{n} \right. \\
&\qquad\qquad \left. - \frac{|\{i; M_{S_n} f(a_1) > M_{S_n} f(a_i)\}|}{n} \right] \\
&\qquad + \sum_{M_{S_n} f(a_1) > M_{S_n} f(a_i)} (f(a_1) - f(a_k)) \left[\frac{|\{i; M_{S_n} f(a_i) > M_{S_n} f(a_1)\}|}{n} \right. \\
&\qquad\qquad \left. + \frac{|\{i; M_{S_n} f(a_1) > M_{S_n} f(a_i)\}|}{n} \right] \\
&\le \frac{n-1}{n} \operatorname{Var} f,
\end{aligned}$$

from where (8.30) follows and therefore our result. ■

Proof [Proof of Theorem 8.1.2 (iii)] We write $f(a_2) \ge \cdots \ge f(a_r) \ge m > f(a_{r+1}) \ge \cdots \ge f(a_n)$. We distinguish two cases, the first being $f(a_1) \le m$.

Case 1: $f(a_1) \le m$. Let us keep in mind in the following that in this case

$$(\operatorname{Var}_p M_{S_n} f)^p = \sum_{i=2}^{r} |f(a_i) - m|^p.$$

In this case it is enough to prove inequality (8.31) when $f(a_i) < f(a_1)$ for $i > r$. In fact, if (8.31) fails for some f with $f(a_i) > f(a_1)$ and $i > r$, it also fails for the function $\widetilde{f}$ defined by

$\widetilde{f}(e) = f(e)$ for every $e \notin \{a_2, a_i\}$, $\widetilde{f}(a_i) = 2f(a_1) - f(a_i)$ and $\widetilde{f}(a_2) = f(a_2) + f(a_i) - \widetilde{f}(a_i)$ (notice that $\widetilde{f}(a_i) < \widetilde{f}(a_1)$ by construction). This holds because $\widetilde{m} = \frac{\sum_{j=1}^n \widetilde{f}(a_j)}{n} = \frac{\sum_{i=j}^n f(a_j)}{n} = m$, by definition, and

$$(1 - \frac{1}{n})^p (f(a_2) - f(a_1))^p - (f(a_2) - m)^p \geq (1 - \frac{1}{n})^p (\widetilde{f}(a_2) - \widetilde{f}(a_1))^p - (\widetilde{f}(a_2) - \widetilde{m})^p. \tag{8.32}$$

This is the case because (8.32) is equivalent to

$$\left(1 - \frac{1}{n}\right)^p (\mathrm{Var}_p f)^p - (\mathrm{Var}_p M_{S_n} f)^p \geq \left(1 - \frac{1}{n}\right)^p (\mathrm{Var}_p \widetilde{f})^p - (\mathrm{Var}_p M_{S_n} \widetilde{f})^p,$$

since the other terms in this inequality remain unchanged when we do the transformation $f \to \widetilde{f}$ (notice that, by construction, $|f(a_1) - f(a_i)| = |\widetilde{f}(a_1) - \widetilde{f}(a_i)|$.) We have that (8.32) holds because

$$|\widetilde{f}(a_2) - \widetilde{f}(a_1)|^p - |f(a_2) - f(a_1)|^p \leq |\widetilde{f}(a_2) - \widetilde{m}|^p - |f(a_2) - m|^p,$$

inequality that follows because of $\widetilde{f}(a_1) = f(a_1), m = \widetilde{m}$, the concavity of the function $x \mapsto x^p$ (and thus the function $x \to (x+c)^p - x^p$ is decreasing for $x, c > 0$, here considering $c = \widetilde{f}(a_2) - f(a_2)$) and the fact that $f(a_2) - f(a_1) \geq f(a_2) - m$. By iterating the previous argument we get the desired reduction.

We write $f(a_i) - m = x_i$ for $i = 2, ...r$; $m - f(a_1) = u$ and $y_i = f(a_1) - f(a_i)$ for $i = r+1, ..., n$. Observe that given our reduction we have $y_i \geq 0$. We observe that since

$$m = \frac{\sum_{i=1}^n f(a_i)}{n} = \frac{\sum_{i=2}^r (m + x_i) + f(a_1) + \sum_{i=r+1}^n (f(a_1) - y_i)}{n} = \frac{\sum_{i=2}^r (m + x_i) + m - u + \sum_{i=r+1}^n (m - u - y_i)}{n},$$

we have

$$\sum_{i=2}^r x_i = u + \sum_{i=r+1}^n (u + y_i),$$

from where we obtain $u \leq \frac{\sum_{i=2}^r x_i}{n-r+1}$. Also, let us observe that (8.31) is equivalent in this case to

$$\sum_{i=2}^r |x_i|^p \leq \left(1 - \frac{1}{n}\right)^p \left(\sum_{i=2}^r |x_i + u|^p + \sum_{i=r+1}^n |y_i|^p\right). \tag{8.33}$$

Observe that $\sum_{i=r+1}^n |y_i|^p \geq \left|\sum_{i=r+1}^n y_i\right|^p = \left|\sum_{i=2}^r x_i - (n - r + 1)u\right|^p$. Then, for $x_2, \ldots, x_r, n, r$ and p fixed, we define the function

$$g(z) := \sum_{i=2}^r |x_i + z|^p + \left|\sum_{i=2}^r x_i - (n - r + 1)z\right|^p,$$

we observe that for $z \in \left[0, \frac{\sum_{i=2}^r x_i}{n-r+1}\right]$ this function is concave (sum of concave functions), therefore $g(z) \geq \min\left\{g\left(\frac{\sum_{i=2}^r x_i}{n-r+1}\right), g(0)\right\}$ in that interval. Then, we have

$$\begin{aligned}\left(1-\frac{1}{n}\right)^p\left(\sum_{i=2}^r|x_i+u|^p+\sum_{i=r+1}^n|y_i|^p\right) &\geq \left(1-\frac{1}{n}\right)^p\left(\sum_{i=2}^r|x_i+u|^p+\left|\sum_{i=2}^r x_i-(n-r+1)u\right|^p\right)\\ &= \left(1-\frac{1}{n}\right)^p g(u)\\ &\geq \left(1-\frac{1}{n}\right)^p \min\left\{g\left(\frac{\sum_{i=2}^r x_i}{n-r+1}\right), g(0)\right\}\\ &\geq \left(1-\frac{1}{n}\right)^p \min\left\{\left(\sum_{i=2}^r|x_i|^p+\left|\sum_{i=2}^r x_i\right|^p\right), \left(\sum_{i=2}^r\left|x_i+\frac{\sum_{i=2}^r x_i}{n-r+1}\right|^p\right)\right\}.\end{aligned}$$

Therefore, in order to prove (8.33) it is enough to prove that

$$\sum_{i=2}^r|x_i|^p \leq \left(1-\frac{1}{n}\right)^p\left(\sum_{i=2}^r|x_i|^p+\left|\sum_{i=2}^r x_i\right|^p\right), \tag{8.34}$$

and

$$\sum_{i=2}^r|x_i|^p \leq \left(1-\frac{1}{n}\right)^p\left(\sum_{i=2}^r\left|x_i+\frac{\sum_{i=2}^r x_i}{n-r+1}\right|^p\right), \tag{8.35}$$

for (8.34) we observe that $\left|\sum_{i=2}^r x_i\right|^p \geq \max_{i=2,..r}|x_i|^p \geq \frac{\sum_{i=2}^r|x_i|^p}{r-1}$, so

$$\left(1-\frac{1}{n}\right)^p\left(\sum_{i=2}^r|x_i|^p+\left|\sum_{i=2}^r x_i\right|^p\right) \geq \left(\sum_{i=2}^r|x_i|^p\right)\left(1-\frac{1}{n}\right)^p\left(1+\frac{1}{r-1}\right) \geq \left(\sum_{i=2}^r|x_i|^p\right),$$

where we use that for $r \leq n-1$ we have

$$\left(1-\frac{1}{n}\right)^p\left(1+\frac{1}{r-1}\right) \geq \left(1-\frac{1}{n}\right)\left(1+\frac{1}{r-1}\right) \geq \left(1-\frac{1}{n}\right)\left(1+\frac{1}{n-1}\right) = 1.$$

From this we conclude this inequality.

For (8.35), we notice that $x_i + \frac{\sum_{i=2}^r x_i}{n-r+1} \geq x_i\left(1+\frac{1}{n-r+1}\right)$. Thus, since (for $n \geq r \geq 2$,) we have

$$\left(1-\frac{1}{n}\right)^p\left(1+\frac{1}{n-r+1}\right)^p \geq \left(1-\frac{1}{n}\right)\left(1+\frac{1}{n-r+1}\right) \geq \left(1-\frac{1}{n}\right)\left(1+\frac{1}{n-1}\right) = 1$$

we conclude this inequality, and therefore this case. Notice that this argument holds for every $p \in (0,1)$.

Case 2: $f(a_1) > m$. Here, we observe that if $f(a_2) \leq f(a_1)$ then $|M_{S_n}f(a_1) - M_{S_n}f(a_i)| \leq \frac{|f(a_1)-f(a_i)|}{2}$ for all $i \geq 2$ and thus (8.31) follows in this case (since $\mathrm{Var}_p M_{S_n} f \leq \frac{1}{2}\mathrm{Var}_p f$). So we can assume that $f(a_2) > f(a_1)$. Let us take k such that $f(a_2) \geq f(a_3) \geq \dots f(a_k) \geq f(a_1) > f(a_{k+1})$, and s is the minimum such that $f(a_1) + f(a_s) \geq 2m$. Let us keep in mind that, in this case, we have

$$(\mathrm{Var}_p M_{S_n} f)^p = \sum_{i=2}^{k} |f(a_i) - f(a_1)|^p + \sum_{j=k+1}^{s} \left|\frac{f(a_1) - f(a_i)}{2}\right|^p + \sum_{i=s+1}^{n} |f(a_1) - m|^p .$$

Let us write $u = f(a_1) - m$, $f(a_i) - f(a_1) = x_i$ for $i = 2, \dots, k$ and $y_i = f(a_1) - f(a_i)$ for $i = k+1, \dots, n$. We observe that, since

$$m = \frac{\sum_{i=1}^{n} f(a_i)}{n} = \frac{\sum_{i=2}^{k}(m + u + x_i) + m + u + \sum_{i=k+1}^{n}(m + u - y_i)}{n},$$

we have $\displaystyle\sum_{i=2}^{k} x_i + nu = \sum_{k+1}^{n} y_i$. Then (8.31) is equivalent to

$$\sum_{i=2}^{k} x_i^p + \sum_{i=k+1}^{s} \left(\frac{y_i}{2}\right)^p + \sum_{s+1}^{n} u^p \leq \left(1 - \frac{1}{n}\right)^p \left(\sum_{i=2}^{k} x_i^p + \sum_{i=k+1}^{n} y_i^p\right). \tag{8.36}$$

It is useful to solve first the case $k = n-1$ (observe, that then $s = n-1$). In this case we observe that $y_n = \displaystyle\sum_{i=2}^{k} x_i + nu$. Then, we need to prove

$$\sum_{i=2}^{k} x_i^p + u^p \leq \left(1 - \frac{1}{n}\right)^p \left[\sum_{i=2}^{k} x_i^p + \left(\sum_{i=2}^{k} x_i + nu\right)^p\right]. \tag{8.37}$$

We notice first that, by Jensen's inequality we have

$$\frac{(n-2)n^p + 1}{n-1} \leq \left(\frac{(n-2)n + 1}{n-1}\right)^p = (n-1)^p,$$

then $(n-2)^{1-p}(n^p - (n-1)^p) \leq (n-2)(n^p - (n-1)^p) \leq (n-1)^p - 1$, where this last inequality is just another way of writing the previous claim. Therefore, by Jensen's inequality (in the second inequality), we have

$$\begin{aligned}
((n-1)^p - 1)\left(\sum_{i=2}^{k} x_i + nu\right)^p &\geq ((n-1)^p - 1)\left(\sum_{i=2}^{k} x_i\right)^p \\
&\geq ((n-1)^p - 1)(n-2)^{p-1} \sum_{i=2}^{k} x_i^p \\
&\geq (n^p - (n-1)^p) \sum_{i=2}^{k} x_i^p,
\end{aligned}$$

where in the last inequality we use what we obtained before. Then,

$$(n-1)^p \left[\left(\sum_{i=2}^{k} x_i + nu \right)^p + \sum_{i=2}^{k} x_i^p \right] \geq n^p \sum_{i=2}^{k} x_i^p + \left(\sum_{i=2}^{k} x_i + nu \right)^p,$$

thus, we have

$$\begin{aligned}\sum_{i=2}^{k}(nx_i)^p + (nu)^p &\leq \sum_{i=2}^{k}(nx_i)^p + \left(\sum_{i=2}^{k} x_i + nu \right)^p \\ &\leq (n-1)^p \left[\sum_{i=2}^{k} x_i^p + \left(\sum_{i=2}^{k} x_i + nu \right)^p \right],\end{aligned}$$

concluding the inequality (8.37). So, we assume in the following that $k \leq n-2$.

We observe that $u \leq \frac{y_i}{2}$ for $i = s+1, ..n$, and thus

$$\sum_{i=2}^{k} x_i^p + \sum_{i=k+1}^{s} \left(\frac{y_i}{2}\right)^p + \sum_{s+1}^{n} u^p \leq \sum_{i=2}^{k} x_i^p + \sum_{i=k+1}^{n} \left(\frac{y_i}{2}\right)^p,$$

therefore (8.36) would follow if

$$\sum_{i=2}^{k} x_i^p \left(1 - \left(1 - \frac{1}{n}\right)^p\right) \leq \left(\left(1 - \frac{1}{n}\right)^p - \frac{1}{2^p}\right) \left(\sum_{i=k+1}^{n} y_i^p\right).$$

Indeed, by Jensen's inequality $\sum_{i=k+1}^{n} y_i^p \geq \left(\sum_{i=k+1}^{n} y_i\right)^p \geq \left(\sum_{i=2}^{k} x_i\right)^p \geq (k-1)^{p-1} \left(\sum_{i=2}^{k} x_i^p\right)$. So, we need $(k-1)^{1-p}\left[1 - \left(1 - \frac{1}{n}\right)^p\right] \leq \left(1 - \frac{1}{n}\right)^p - \frac{1}{2^p}$. Since $k-1 \leq n-3$ is enough

$$(n-3)^{1-p} \left(1 - \left(1 - \frac{1}{n}\right)^p\right) \leq \left(1 - \frac{1}{n}\right)^p - \frac{1}{2^p}, \tag{8.38}$$

but that is equivalent to

$$(n-3)^{1-p}(n^p - (n-1)^p) \leq (n-1)^p - \left(\frac{n}{2}\right)^p,$$

therefore, it is enough to prove (we use here $n^p - (n-1)^p \leq p(n-1)^{p-1}$ by the fundamental theorem of calculus)

$$(n-3)^{1-p} p(n-1)^{p-1} \leq (n-1)^p - \left(\frac{n}{2}\right)^p, \tag{8.39}$$

or, the stronger bound (since $\left(\frac{n-3}{n-1}\right)^{1-p} \leq 1$), $p \leq (n-1)^p - (\frac{n}{2})^p$. Fixed p, it is possible to observe that this last inequality holds for n big enough. Therefore, we conclude the last statement of Theorem 8.1.2 (iii). Now we assume that $1 > p \geq \frac{1}{2}$. First observe that for $n \geq 6$ we have that $p \leq (n-1)^p - \left(\frac{n}{2}\right)^p$, in fact $g(n) = (n-1)^p - \left(\frac{n}{2}\right)^p$ is increasing for $n \geq 2$ because its derivative is $p(n-1)^{p-1} - \frac{p}{2}\left(\frac{n}{2}\right)^{p-1} \geq 0$ since $2 \geq 2^{1-p} \geq \left(\frac{2(n-1)}{n}\right)^{1-p}$ So, we need to prove $p \leq 5^p - 3^p$, indeed $g(p) = 5^p - 3^p - p$ is convex for $p \geq 0$ (its second derivative is $\log(5)^2 5^p - \log(3)^2 3^p \geq 0$) thus since $g(0) = 0$ and $g\left(\frac{1}{2}\right) = \sqrt{5} - \sqrt{3} - \frac{1}{2} \geq 0$ for every $p \geq \frac{1}{2}$ we get $\alpha g(p) = \alpha g(p) + \beta g(0) \geq g\left(\frac{1}{2}\right) > 0$, for some $\alpha, \beta \geq 0$. From where we conclude this inequality. Therefore, considering [LX20, Theorem 1.4], the only cases left are $n = 4$ and $n = 5$. For $n = 4$, considering (8.39), we just need

$$\left(\frac{1}{3}\right)^{1-p} p \leq 3^p - 2^p,$$

or, equivalently, $p \leq 3 - 3(\frac{2}{3})^p$, but $g(p) = 3 - 3(\frac{2}{3})^p - p$ is concave in $(0,1)$, so, since $g(0) = 0 = g(1)$, we conclude in this case. Notice that this argument holds for every $1 > p > 0$, and therefore the case $n = 4$ is completed.

Finally, for $n = 5$, we just need (considering (8.39))

$$\left(\frac{1}{2}\right)^{1-p} p \leq 4^p - \left(\frac{5}{2}\right)^p,$$

or equivalently

$$\frac{p}{2} \leq 2^p - \left(\frac{5}{4}\right)^p,$$

but $g(p) = 2^p - \left(\frac{5}{4}\right)^p - \frac{p}{2}$ is convex for $p \geq 0$ (because its second derivative is $\log(2)^2 2^p - \log\left(\frac{5}{4}\right)^2 \left(\frac{5}{4}\right)^p \geq 0$) then since $\sqrt{2} - \sqrt{\frac{5}{4}} - \frac{1}{4} \geq 0$ and $g(0) = 0$ we conclude this case similarly as for $n \geq 6$. Since we finish the analysis of cases, we conclude the proof of the theorem. ■

Remark 8.2.3. It is possible, in fact, to prove (8.38) for every $0 < p < 1$ when $n = 5$, thus proving Theorem 8.1.2(iii) for every $0 < p < 1$ in this case. We omit the details for the sake of simplicity.

8.2.3 Qualitative results: proof of Theorem 8.1.3

In the last part of this section we prove our versions of the qualitative results conjectured in Conjecture C.

Proof [Proof of Theorem 8.1.3 (i)] We assume without loss of generality that f is nonnegative. Also, in the following we assume that G_n is connected, since the general case follows from there. Given $u, v \in G_n := \{a_1, a_2, \ldots, a_n\}$, such that $M_{\alpha,G_n} f(u) > M_{\alpha,G_n} f(v)$, we observe that there exists $k \leq n-1$ such that

$$M_{\alpha,G_n} f(u) = \frac{|B(u,k)|^\alpha}{|B(u,k)|} \sum_{a_i \in B(u,k)} f(a_i),$$

then

$$\begin{aligned}
M_{\alpha,G_n}f(u) - M_{\alpha,G_n}f(v) &\le \frac{|B(u,k)|^\alpha}{|B(u,k)|}\sum_{a_i\in B(u,k)} f(a_i) - \frac{n^\alpha}{n}\sum_{i=1}^{n} f(a_i)\\
&\le n^\alpha\left[\frac{1}{|B(u,k)|}\sum_{a_i\in B(u,k)} f(a_i) - \frac{1}{n}\sum_{i=1}^{n} f(a_i)\right]\\
&\le n^\alpha(f(x)-f(y))\\
&\le n^\alpha(n-1)^{\max\{1-\frac{1}{p},0\}}\mathrm{Var}_p f.
\end{aligned}$$

Where, in the third line $x \in G_n$ is chosen such that $f(x) := \max\{f(a_i); a_i \in B(u,k)\}$ and $y \in G_n$ is chosen such that $f(y) := \min\{f(a_i); a_i \in G_n\}$. In the fourth line we used Hölder's inequality.

Therefore

$$\begin{aligned}
\mathrm{Var}_q M_{\alpha,G_n} &= \left(\frac{1}{2}\sum_{u\in G_n}\sum_{v\in N_{G_n}(u)} |M_{\alpha,G_n}f(u) - M_{\alpha,G_n}f(v)|^q\right)^{1/q}\\
&\le \left(\frac{n(n-1)}{2}\right)^{1/q} n^\alpha(n-1)^{\max\{\frac{p-1}{p},0\}}\mathrm{Var}_p f\\
&= C(n,p,q)\mathrm{Var}_p f.
\end{aligned}$$

■

Proof [Proof of Theorem 8.1.3 (ii)] We start observing that for all $j \ge 1$

$$\begin{aligned}
\|f-f_j\|_{l^\infty(G_n)} &= \max_{y\in V}|f(y)-f_j(y)| - \min_{x\in V}|f(x)-f_j(x)| + \min_{x\in V}|f(x)-f_j(x)|\\
&\le \mathrm{Var}(f-f_j) + \min_{x\in V}|f(x)-f_j(x)|\\
&\le n^{\max\{1-1/p,0\}}\mathrm{Var}_p(f-f_j) + \min_{x\in V}|f(x)-f_j(x)|. \qquad (8.40)
\end{aligned}$$

Then, assuming that $\lim_{j\to\infty}\min_{x\in V}|f(x)-f_j(x)| = 0$, we have that

$$\|f-f_j\|_{l^\infty(G_n)} \to 0 \quad \text{as} \quad j\to\infty.$$

Moreover, for any $u, v \in G_n$ we have that

$$\begin{aligned}
M_{\alpha,G_n}f(u) - M_{\alpha,G_n}f_j(u) - [M_{\alpha,G_n}f(v) - M_{\alpha,G_n}f_j(v)] &\le M_{\alpha,G_n}(f-f_j)(u) + M_{\alpha,G_n}(f-f_j)(v)\\
&\le 2\|f-f_j\|_{l^1(G_n)}\\
&\le 2n\|f-f_j\|_{l^\infty(G_n)} \to 0 \quad \text{as} \quad j\to\infty.
\end{aligned}$$

Therefore

$$\mathrm{Var}_q(M_{\alpha,G_n}f - M_{\alpha,G_n}f_j) \le \left(\frac{n(n-1)}{2}\right)^{1/q} 2n\|f-f_j\|_{l^\infty(G_n)} \to 0 \quad \text{as} \quad j\to\infty.$$

Finally, we observe that without the assumption that $\lim_{j\to\infty}\min_{x\in V}|f(x)-f_j(x)|=0$ the continuity property could fail, with this purpose in mind consider the following situation: Let $G_n = S_n$ the star graph with n vertices $V=\{a_1,a_2,\ldots,a_n\}$ and center at a_1, for simplicity we take $\alpha=0$ and $p=q=1$. We define the function f by $f(a_1)=2$ and $f(a_i)=1$ for all $i\neq 1$ thus $M_{S_n}f(a_1)=2$ and $M_{S_n}f(a_i)=3/2$ for all $i\neq 1$. Then, we consider the sequence of functions $(f_j)_{j\in\mathbb{N}}$ defined by $f_j(a_i)=f(a_i)-3$ for all $a_i\in V$ and for all $j\in\mathbb{N}$. Then $\mathrm{Var}\,(f-f_j)=0$ for all $j\in\mathbb{N}$, moreover $M_{S_n}f_j(a_1)=\frac{1+2(n-1)}{n}$ and $M_{S_n}f_j(a_i)=2$ for all $i\neq 1$. Therefore

$$\begin{aligned}\mathrm{Var}\,(M_{S_n}f-M_{S_n}f_j) &\geq M_{S_n}f(a_1)-M_{S_n}f_j(a_1)-[M_{S_n}f(a_2)-M_{S_n}f_j(a_2)]\\ &=2-\frac{1+2(n-1)}{n}-[3/2-2]\\ &=\frac{1}{n}+\frac{1}{2}\quad\text{for all}\quad j\in\mathbb{N}.\end{aligned}$$

Then $\mathrm{Var}\,(M_{S_n}f-M_{S_n}f_j)\nrightarrow 0$ as $j\to\infty$.

■

Proof [Proof of Theorem 8.1.3 (iii)] The boundedness follows using part (i) and the following inequality which is true for some $k\leq n-1$

$$\begin{aligned}M_{\alpha,G_n}f(a_0)&=\frac{1}{|B(a_0,k)|^{1-\alpha}}\sum_{m\in B(a_0,k)}|f(m)|\\ &=\frac{1}{|B(a_0,k)|^{1-\alpha}}\sum_{m\in B(a_0,k)}(|f(m)|-|f(a_0)|)+|B(a_0,k)|^{\alpha}|f(a_0)|\\ &\leq |B(a_0,k)|^{\alpha}(\max_{m\in B(a_0,k)}|f(m)-f(a_0)|+|f(a_0)|)\\ &\leq |B(a_0,k)|^{\alpha}(\mathrm{Var}\,f+|f(a_0)|)\\ &\leq |B(a_0,k)|^{\alpha}n^{\max\{1-1/p,0\}}(\mathrm{Var}_{\,p}f+|f(a_0)|)\\ &\leq n^{\alpha+\max\{1-1/p,0\}}\|f\|_{BV_p(G_n)}.\end{aligned}$$

The continuity follows using part (ii) and the following observations

$$0\leq \mathrm{Var}_{\,p}(f-f_j)+\min_{x\in V}|f(x)-f_j(x)|\leq \mathrm{Var}_{\,p}(f-f_j)+|(f-f_j)(a_0)|=\|f-f_j\|_{BV_p(G_n)},$$

and

$$\begin{aligned}|M_{\alpha,G_n}f(a_0)-M_{\alpha,G_n}f_j(a_0)|&\leq M_{\alpha,G_n}(f-f_j)(a_0)\\ &\leq \|f-f_j\|_{l^1(G_n)}\\ &\leq n\|f-f_j\|_{l^\infty(G_n)}\\ &\leq n^{1+\max\{1-1/p,0\}}\mathrm{Var}_{\,p}(f-f_j)+n\min_{x\in V}|f(x)-f_j(x)|\\ &\leq n^{1+\max\{1-1/p,0\}}\|f-f_j\|_{BV_p(G_n)},\end{aligned}$$

which is a consequence of (8.40). ■

8.3 Proof of optimal bounds for the 2-norm of maximal functions

In this subsection we prove our results concerning the values $\|M_G\|_2$ for our graphs of interest.

8.3.1 2-norm of the maximal operator in K_n: proof of Theorem 8.1.4 and Corollary 8.1.1

We start by proving that Corollary 8.1.1 follows by Theorem 8.1.4.

Proof [Proof of Corollary 8.1.1] The inequality

$$\|M_{K_n}f\|_2 \leq \left(\frac{4}{3}\right)^{1/2}\|f\|_2$$

follows from the Theorem 8.1.4, since $k = n/3$ in the right hand side. On the other hand, we consider the following example: we define $f : V \to \mathbb{R}$ by

$$f(a_i) = 4 \text{ for all } 1 \leq i \leq \frac{n}{3} \text{ and } f(a_i) = 1 \text{ for all } \frac{n}{3} + 1 \leq i \leq n.$$

Then, in this case we have

$$M_{K_n}f(a_i) = 4 \text{ for all } 1 \leq i \leq \frac{n}{3} \text{ and } M_{K_n}f(a_i) = 2 \text{ for all } \frac{n}{3} + 1 \leq i \leq n.$$

Therefore

$$\|M_{K_n}f\|_2 = \left(\frac{\frac{16n}{3} + \frac{4(2n)}{3}}{\frac{16n}{3} + \frac{2n}{3}}\right)^{1/2}\|f\|_2 = \left(\frac{4}{3}\right)^{1/2}\|f\|_2.$$

■

Now we prove our bound that holds for K_n for every $n \geq 2$.

Proof [Proof of Theorem 8.1.4] We assume without loss of generality that f is nonnegative. Consider the case

$$f(a_1) \geq f(a_2) \geq \cdots \geq f(a_k) \geq m \geq f(a_{k+1}) \geq \cdots \geq f(a_n).$$

Then, in this case

$$M_{K_n}f(a_i) = f(a_i) \text{ for all } 1 \leq i \leq k, \text{ and } M_{K_n}f(a_i) = m \text{ for all } k+1 \leq i \leq n.$$

Therefore by AM-GM inequality we have that

$$\begin{aligned}
\|M_{K_n}f\|_2^2 &= \sum_{i=1}^{k} f(a_i)^2 + (n-k)m^2 \\
&= \left(1+\frac{n-k}{n^2}\right)\sum_{i=1}^{k} f(a_i)^2 + \frac{n-k}{n^2}\sum_{i=k+1}^{n} f(a_i)^2 \\
&\quad + \frac{2(n-k)}{n^2}\sum_{1\le i<j\le k} f(a_i)f(a_j) + \frac{2(n-k)}{n^2}\sum_{k+1\le i<j\le n} f(a_i)f(a_j) \\
&\quad + \frac{2(n-k)}{n^2}\sum_{\substack{1\le i\le k\\ k+1\le j\le n}} f(a_i)f(a_j) \\
&\le \left(1+\frac{n-k}{n^2}\right)\sum_{i=1}^{k} f(a_i)^2 + \frac{n-k}{n^2}\sum_{i=k+1}^{n} f(a_i)^2 \qquad (8.41) \\
&\quad + \frac{(n-k)(k-1)}{n^2}\sum_{i=1}^{k} f(a_i)^2 + \frac{(n-k)(n-k-1)}{n^2}\sum_{i=k+1}^{n} f(a_i)^2 \\
&\quad + \frac{2(n-k)}{n^2}\sum_{\substack{1\le i\le k\\ k+1\le j\le n}} f(a_i)f(a_j) \\
&= A_k\sum_{i=1}^{k} f(a_i)^2 + B_K\sum_{i=k+1}^{n} f(a_i)^2 + \frac{2(n-k)}{n^2}\sum_{\substack{1\le i\le k\\ k+1\le j\le n}} f(a_i)f(a_j),
\end{aligned}$$

where $A_k := 1 + \frac{(n-k)k}{n^2}$ and $B_k := \frac{(n-k)^2}{n^2}$. Observe that $A_k - B_k = \frac{3nk-2k^2}{n^2}$ and by the AM-GM inequality

$$\begin{aligned}
\|M_{K_n}f\|_2^2 &\le A_k\sum_{i=1}^{k} f(a_i)^2 + B_k\sum_{i=k+1}^{n} f(a_i)^2 + \frac{2(n-k)}{n^2}\sum_{\substack{1\le i\le k\\ k+1\le j\le n}} f(a_i)f(a_j) \\
&\le A_k\sum_{i=1}^{k} f(a_i)^2 + B_k\sum_{i=k+1}^{n} f(a_i)^2 + \frac{1}{n^2}\sum_{\substack{1\le i\le k\\ k+1\le j\le n}} (xf(a_i)^2 + yf(a_j)^2) \qquad (8.42) \\
&= \left(A_k + \frac{(n-k)x}{n^2}\right)\sum_{i=1}^{k} f(a_i)^2 + \left(B_k + \frac{ky}{n^2}\right)\sum_{i=k+1}^{n} f(a_i)^2
\end{aligned}$$

for all $0 < x, y$ such that $xy = (n-k)^2$. Then, we choose x, y such that

$$A_k + \frac{(n-k)x}{n^2} = B_k + \frac{ky}{n^2}.$$

So, x is the positive solution for the equation

$$(3nk-2k^2)x+(n-k)x^2=k(n-k)^2.$$

More precisely

$$x:=\frac{-(3nk-2k^2)+(4kn^3-3n^2k^2)^{1/2}}{2(n-k)}.$$

Therefore, combining (8.41) and (8.42) we obtain

$$\begin{aligned}\|M_{K_n}f\|_2^2 &\leq \max_{k\in[1,n-1]}\left(A_k+\frac{(n-k)x}{n^2}\right)\sum_{i=1}^n f(a_i)^2\\ &= \max_{k\in[1,n-1]}\left(1+\frac{(n-k)k}{n^2}+\frac{(4kn^3-3n^2k^2)^{1/2}-(3nk-2k^2)}{2n^2}\right)\sum_{i=1}^n f(a_i)^2\\ &= \max_{k\in[1,n-1]}\left(1-\frac{k}{2n}+\frac{(4kn-3k^2)^{1/2}}{2n}\right)\sum_{i=1}^n f(a_i)^2.\end{aligned}$$

Then, we consider the function $g:[1,n-1]\to\mathbb{R}$ defined by $g(t):=-t+(4tn-3t^2)^{1/2}$.

Observe that

$$\max_{t\in[1,n-1]} g(t)=g\left(\frac{n}{3}\right).$$

Moreover, g is increasing in $[1,n/3]$ and decreasing in $[n/3,n-1]$. Therefore

$$\|M_{K_n}f\|_2^2\leq \max_{k\in\{\lfloor\frac{n}{3}\rfloor,\lceil\frac{n}{3}\rceil\}}\left(1-\frac{k}{2n}+\frac{(4kn-3k^2)^{1/2}}{2n}\right)\|f\|_2^2. \tag{8.43}$$

Finally, observe that in order to have an equality in (8.43) it is enough to have equality in (8.41) and (8.42). Moreover, the equality in (8.41) is attained if and only if $f(a_i)=f(a_1)=\gamma$ for all $1\leq i\leq k$, and $f(a_j)=f(a_{k+1})=\eta$ for all $k+1\leq j\leq n$, for some $0<\eta<\gamma$. We can assume without loss of generality that $\eta=1$. On the other hand, the equality in (8.42) is attained if and only if $y^{1/2}=x^{1/2}\gamma=(n-k)^{1/2}\gamma^{1/2}$, or equivalently $\gamma=\frac{n-k}{x}$. Therefore, in order to obtain an equality in (8.43) for $k\in\{\lfloor\frac{n}{3}\rfloor,\lceil\frac{n}{3}\rceil\}$ we consider the function $g_k:V\to\mathbb{R}$ defined by

$$g_k(a_i)=\gamma:=\frac{2(n-k)^2}{(4kn^3-3n^2k^2)^{1/2}-(3nk-2k^2)}\quad \text{for all}\ \ 1\leq i\leq k,$$

and $g_k(a_j)=1$ for all $k+1\leq j\leq n$. Then, by construction

$$\|M_{K_n}\|_2=\max_{k\in\{\lfloor\frac{n}{3}\rfloor,\lceil\frac{n}{3}\rceil\}}\frac{\|M_{K_n}g_k\|_2}{\|g_k\|_2}.$$

this shows that our bound is optimal, moreover we have found extremizers. Observe that, in the particular case when $n=3k$, we obtain $\gamma=4$ as in the Corollary 8.1.1.

■

8.3.2 2-norm of the maximal operator in S_n: proof of Theorem 8.1.5

Now, we prove our result concerning the 2-norm of our maximal operator on S_n.
Proof [Proof of Theorem 8.1.5] As usual we assume without loss of generality that f is no negative and we denote by m the average of f along V *i.e.* $m = \frac{\sum_{i=1}^n f(a_i)}{n}$. We observe that $M_{S_n}f(a_1) = f(a_1)$ or $M_{S_n}f(a_1) = m$. We study this two cases separately.

Case 1: $M_{S_n}f(a_1) = f(a_1)$. Assume without loss of generality that $M_{S_n}f(a_i) = f(a_i)$ for all $1 \leq i \leq k$, $M_{S_n}f(a_i) = \frac{f(a_i)+f(a_1)}{2}$ for all $k+1 \leq i \leq k+r$, and $M_{S_n}f(a_i) = m$ for all $k+r+1 \leq i \leq n$. By Cauchy-Schwarz inequality we have

$$m^2 \leq \frac{\sum_{i=1}^n f(a_i)^2}{n}.$$

Using this inequality, we get

$$\begin{aligned}
\|M_{S_n}f\|_2^2 &\leq \left(1+\frac{r}{4}\right) f(a_1)^2 + \sum_{i=2}^{k} f(a_i)^2 + \frac{1}{4}\sum_{i=k+1}^{k+r} f(a_i)^2 + \frac{2}{4}\sum_{i=k+1}^{k+r} f(a_i)f(a_1) + \frac{s}{n}\sum_{i=1}^{n} f(a_i)^2 \\
&= \left(1+\frac{r}{4}+\frac{s}{n}\right) f(a_1)^2 + \left(1+\frac{s}{n}\right)\sum_{i=2}^{k} f(a_i)^2 + \left(\frac{1}{4}+\frac{s}{n}\right)\sum_{i=k+1}^{k+r} f(a_i)^2 \\
&\quad + \frac{2}{4}\sum_{i=k+1}^{k+r} f(a_i)f(a_1) + \frac{s}{n}\sum_{i=k+r+1}^{n} f(a_i)^2.
\end{aligned}$$

where $s := n-k-r$. Moreover, for all $k+1 \leq i \leq k+r$, we have that

$$\frac{2}{4} f(a_i)f(a_1) \leq x f(a_1)^2 + y f(a_i)^2$$

for all $x, y > 0$ such that $xy \geq \frac{1}{16}$. We can choose x and y such that

$$y - rx = 1 + \frac{r-1}{4} \text{ and } xy = \frac{1}{16}.$$

or equivalently

$$x := \frac{[(r+9)(r+1)]^{1/2} - (r+3)}{8r}.$$

Therefore, for all $n \geq 4$ we have

$$\begin{aligned}
\|M_{S_n}f\|_2^2 &\leq \max_{\{k,r\in\mathbb{N};1\leq k+r\leq n\}} \left(1+\frac{n-k-r}{n}+\frac{r}{4}+\frac{[(r+9)(r+1)]^{1/2}-(r+3)}{8}\right)\|f\|_2^2 \\
&\leq \left(1+\frac{n-1}{4}+\frac{(n^2+8n)^{1/2}-(n+2)}{8}\right)\|f\|_2^2.
\end{aligned}$$

Case 2: $M_{S_n}f(a_1) = m$. In this case $k \geq 2$. Following the same strategy (and notation), for all $n \geq 4$ we obtain that

$$\|M_{S_n}f\|_2^2 \leq \left(\frac{r}{4} + \frac{s+1}{n}\right) f(a_1)^2 + \left(1 + \frac{s+1}{n}\right) \sum_{i=2}^{k} f(a_i)^2 + \left(\frac{1}{4} + \frac{s+1}{n}\right) \sum_{i=k+1}^{k+r} f(a_i)^2$$

$$+ \frac{2}{4} \sum_{i=k+1}^{k+r} f(a_i)f(a_1) + \frac{s+1}{n} \sum_{i=k+r+1}^{n} f(a_i)^2.$$

$$\leq \max_{\{k,r\in\mathbb{N};1\leq k+r\leq n\}} \left\{\frac{n-k-r+1}{n} + \frac{r+1}{4}, \frac{n-k-r+1}{n} + 1\right\} \|f\|_2^2$$

$$= \max_{\{k,r\in\mathbb{N};1\leq k+r\leq n\}} \left\{\frac{n-k-r+1}{n} + \frac{r+1}{4}, \frac{n-1}{n} + 1\right\} \|f\|_2^2.$$

The inequality

$$\|M_{S_n}\|_2 \leq \left(1 + \frac{n-1}{4} + \frac{(n^2+8n)^{1/2} - (n+2)}{8}\right)^{1/2} := C_n$$

follows from these two estimates.

Finally, we observe that $\|M_{S_n}\|_2 = C_n$. Consider the function $g : V \to \mathbb{R}$ defined by $g(a_i) = 1$ for all $1 \leq i \leq n-1$ and $g(a_0) = \gamma$, where we choose γ to be a positive real number larger than 1, such that γ is a solution for the quadratic equation

$$aX^2 + bX + c := \left(C_n^2 - 1 - \frac{(n-1)}{4}\right) x^2 - \frac{n-1}{2} x + C_n^2(n-1) - \frac{n-1}{4} = 0.$$

The existence of γ follows from the definition of C_n, since we can see that $b^2 - 4ac = 0$ and $\frac{-b}{2a} > 1$. More precisely

$$\gamma = -\frac{b}{2a} = \frac{2(n-1)}{(n^2+8n)^{1/2} - (n+2)}.$$

For this particular function we have

$$\frac{\|M_{S_n}g\|_2}{\|g\|_2} = \left(\frac{\gamma^2 + (n-1)\left(\frac{\gamma+1}{2}\right)^2}{\gamma^2 + (n-1)}\right)^{1/2} = C_n.$$

This concludes the proof of our theorem. ∎

Chapter 9

Sharp inequalities for maximal operators on finite graphs II

9.1 Introduction

In this chapter, as in the previous one, we are interested in optimal constants for maximal operators defined over finite graphs. Recall the definition (8.1) for the centered Hardy-Littlewood operator acting on the graph G. The $p-$norm (quasi-norm in the range $0 < p < 1$) of these operators is defined as

$$\|M_G\|_p := \sup_{\substack{f:V\to\mathbb{R} \\ f\neq 0}} \frac{\|M_G f\|_p}{\|f\|_p},$$

where $\|g\|_p = \left(\sum_{v\in V} |g(v)|^p\right)^{\frac{1}{p}}$, for any $g : V \to \mathbb{R}$. In §9.2 we address this problem, we fully characterize the extremizers for $\|M_{S_n}\|_p$ for all $p \in (1, 2]$. Moreover, we obtain some partial characterization for this objects for all $p > 2$, and we obtain a similar result for the extremizers of $\|M_{K_n}\|_p$ in the range $p > 1$.

For a given $p > 1$ and G, it could be difficult to determine the value of $\|M_G\|_p$. That happens even in the model cases $G = K_n$ and $G = S_n$. However, it was proved by Soria and Tradacete (see Proposition 3.4 in [ST16]) that

$$\left(1 + \frac{n-1}{2^p}\right) \le \|M_{S_n}\|_p^p \le \left(\frac{n+5}{2}\right).$$

They also presented similar bounds for $\|M_{K_n}\|_p^p$. We notice that both lower bounds go to 1 when $p \to \infty$.

In §9.3 we discuss the behavior of both $\|M_{S_n}\|_p^p$ and $\|M_{K_n}\|_p^p$. In particular, we prove that

$$\inf_{p>0} \|M_{G_n}\|_p^p > 1,$$

for any graph G of n vertices. This improves qualitatively the aforementioned estimates of Soria and Tradacete. Also, in §9.3 we prove that for all $n \geq 25$ we have

$$\lim_{p\to\infty} \|M_{S_n}\|_p^p = \frac{1+\sqrt{n}}{2}.$$

Moreover, we obtain a similar result for M_{K_n}.

9.1.1 The p-variation of maximal functions

As in the previous chapter, for a function $g : V \to \mathbb{R}_+$ we write

$$\mathrm{Var}_p\, g = \left(\sum_{(v_1,v_2)=e\in E} |g(v_1) - g(v_2)|^p \right)^{\frac{1}{p}},$$

and we define

$$\mathbf{C}_{G,p} = \sup_{\substack{f:V\to\mathbb{R} \\ \mathrm{Var}\, f\neq 0}} \frac{\mathrm{Var}_p M_G f}{\mathrm{Var}_p f}.$$

In the previous chapter we proved some optimal inequalities for the p-variation of maximal operators on finite graphs. Moreover, in those previous situations the extremizers were delta functions. However, In the case $p > 1$, delta functions are not extremizers for the $p-$variation of M_{S_n}. In §9.4 we find the precise value of $\mathbf{C}_{S_n,2}$. Moreover, we fully describe the extremizers in this case. In §9.5, we obtain some complementary results extending [Mad17, Theorem 1] to the range $p \in [\frac{1}{2}, 1)$. In [Mad17, Theorem 1] it is proved that

$$\mathrm{Var}_1 M_{\mathbb{Z}} f \leq 2\|f\|_1,$$

when considering $\mathbb{Z}$ as a graph where consecutive numbers are joined by an edge. This inequality is sharp. The motivation behind this inequality is to try to get an intuition about which is the optimal constant C in the estimate

$$\mathrm{Var}_1 M_{\mathbb{Z}} f \leq C \mathrm{Var}_1 f,$$

that was proved to be true for $C = (2 \cdot 120 \cdot 2^{12} \cdot 300 + 4)$ in [Tem13]. Since $2\|f\|_1 \geq \mathrm{Var}_1 f$ it is believed that $C = 1$ is the optimal constant, but this remains an open problem. In §9.5 we find the best constant C_p such that

$$\mathrm{Var}_p M_{\mathbb{Z}} f \leq C_p \|f\|_p$$

for $p \in [\frac{1}{2}, 1]$. This motivates us to make some conjectures. We also establish the analogous optimal result for $p = \infty$.

9.2 Extremizers for the p-norm of maximal operators on graphs

In this section we prove the existence of extremizers for the p-norm and provide some further properties about these functions.

Proposition 9.2.1. *Let $G = (V, E)$ be a connected finite graph and $p > 0$. We have that there exists $f : V \to \mathbb{R}_{\geq 0}$ such that*

$$\frac{\|M_G f\|_p}{\|f\|_p} = \|M_G\|_p$$

Proof We write $|V| = n$ and $V =: \{a_1, \dots, a_n\}$. Given $y := (y_1, \dots, y_n) \in [0,1]^n \cap \{\max_{i=1,\dots,n} y_i = 1\} =: A$ we define $f_y : V \to \mathbb{R}_{\geq 0}$ by $f_y(a_i) = y_i$. We observe that $M_G f_y(a_i)$ is continuous with respect to y in A (since is the maximum of continuous functions). Then, the function $\frac{\|M_G f_y\|_p}{\|f_y\|_p}$ is continuous with respect to y in A. Thus it achieves its maximum at a point $y_0 \in A$. We claim that

$$\frac{\|M_G f_{y_0}\|}{\|f_{y_0}\|_p} = \|M_G\|_p.$$

In fact, for every $g : V \to \mathbb{R}_{\geq 0}$ we have that the quantity

$$\frac{\|M_G g\|_p}{\|g\|_p}$$

remains unchanged by applying the transformation

$$g \mapsto \frac{g}{\max_{i=1,\dots,n} g(a_i)}.$$

This last function is equal to f_y for some $y \in A$, from where we conclude the result.

■

Our next results intend to characterize the extremizers when $G = K_n$ and $G = S_n$.

Proposition 9.2.2. *Let $K_n = (V, E)$ be the complete graph with $n > 2$ vertices where $V = \{a_1, a_2 \dots, a_n\}$ and let $p > 1$. If*

$$\frac{\|M_{K_n} f\|_p}{\|f\|_p} = \|M_{K_n}\|_p,$$

then $|f|$ only takes two values.

Proof First, by taking a Dirac's delta it is easy to see that $\|M_{K_n}\|_p > 1$. Now, assume that $f \geq 0$ satisfies

$$\frac{\|M_{K_n} f\|_p}{\|f\|_p} = \|M_{K_n}\|_p$$

. We have then that

$$\sum_{i=1}^{r} f(a_i)^p + (n-r)m^p = \|M_{K_n}\|^p \left(\sum_{i=1}^{n} f(a_i)^p\right).$$

Therefore, by Hölder's inequality we have that

$$\begin{aligned}(n-r)m^p &= \left(\|M_{K_n}\|_p^p - 1\right)\left(\sum_{i=1}^{r} f(a_i)^p\right) + \|M_{K_n}\|_p^p \left(\sum_{i=r+1}^{n} f(a_i)^p\right) \\ &\geq r(\|M_{K_n}\|_p^p - 1)\left(\frac{\sum_{i=1}^{r} f(a_i)}{r}\right)^p + (n-r)\|M_{K_n}\|_p^p \left(\frac{\sum_{i=r+1}^{n} f(a_i)}{n-r}\right)^p.\end{aligned}$$

Then, if we take the function $\widetilde{f}(a_j) = \frac{\sum_{i=1}^{r} f(a_i)}{r}$ for $j = 1, \dots, r$, and $\widetilde{f}(a_j) = \frac{\sum_{i=r+1}^{n} f(a_i)}{n-r}$ for $j = r+1, \dots, n$, we have

$$\frac{\|M_{K_n}\widetilde{f}\|_p}{\|\widetilde{f}\|_p} \geq \|M_{K_n}\|_p,$$

with equality if and only if $f(a_i) = \widetilde{f}(a_i)$ for every $i = 1, \dots, n$. So, we conclude the result. ■

We also get the following result.

Proposition 9.2.3. *Let $S_n = (V, E)$ be the star graph with n vertices $V = \{a_1, a_2, \dots, a_n\}$ with center at a_1 and let $p \geq 1$. There exists $f : V \to \mathbb{R}$ with*

$$\frac{\|M_{S_n} f\|_p}{\|f\|_p} = \|M_{S_n}\|_p,$$

such that $f(a_1) = \max f$ and $f_{|V \setminus a_1}$ takes (at most) two values.

Proof By Proposition 9.2.1 there exists $g \geq 0$ such that

$$\frac{\|M_{S_n} g\|_p}{\|g\|_p} = \|M_{S_n}\|_p.$$

Now, we proceed in three steps.
Step 1: We can assume that $g(a_1) \geq g(a_j)$ for all $j \in \{2, 3, \dots, n\}$. We assume without loss of generality that $g(a_2) \geq \dots \geq g(a_r) \geq g(a_1) \geq \dots \geq g(a_n)$, consider $\widetilde{g}(x) := g(x)$ for $x \in V \setminus \{a_2, a_1\}$, $\widetilde{g}(a_2) := g(a_1)$ and $\widetilde{g}(a_1) := g(a_2)$. We observe that

$$M_{S_n}\widetilde{g}(a_1)^p + M_{S_n}\widetilde{g}(a_2)^p = g(a_2) + \max\left\{m^p, \left(\frac{g(a_2) + g(a_1)}{2}\right)^p\right\} \geq M_{S_n}g(a_2)^p + M_{S_n}g(a_1)^p.$$

Also, for $x \in V \setminus \{a_1, a_2\}$, we have that $M_{S_n}\widetilde{g}(x) \geq M_{S_n}g(x)$ since

$$\max\left\{m, \frac{g(x) + \widetilde{g}(a_1)}{2}, g(x)\right\} \geq \max\left\{m, \frac{g(x) + g(a_1)}{2}, g(x)\right\}.$$

Therefore, we get

$$\sum_{i=1}^{n} M_{S_n}\widetilde{g}(a_1)^p \geq \sum_{i=1}^{n} M_{S_n} g(a_i)^p,$$

since clearly we have that $\|\widetilde{g}\|_p = \|g\|_p$. We conclude that

$$\frac{\|M_{S_n}\widetilde{g}\|_p}{\|\widetilde{g}\|_p} = \|M_{S_n}\|_p.$$

So, we can assume that $g(a_1) \geq g(a_j)$ for every j.

Then, we assume without loss of generality that $g(a_1) \geq \cdots \geq g(a_r) \geq 2m - g(a_1) > g(a_{r+1}) \geq \ldots g(a_n)$.

Step 2: We can assume that $g(a_{r+1}) = g(a_{r+2}) = \cdots = g(a_n)$. We consider the function $\widetilde{g} : V \to \mathbb{R}$ defined by $\widetilde{g}(a_i) = \frac{\sum_{i=r+1}^{n} g(a_i)}{n-r}$ for every $i = r+1 \ldots n$ and $\widetilde{g} = g$ otherwise. We have (similarly as in the previous proposition) that

$$\frac{\|M_{S_n}\widetilde{g}\|_p}{\|\widetilde{g}\|_p} \geq \|M_{S_n}\|_p.$$

Therefore, we can assume that $g(a_i) = g(a_n)$ for every $i \geq r+1$.

Step 3: We can assume that $g(a_2) = g(a_3) = \cdots = g(a_r)$. Now consider

$$\widetilde{g}(a_i) = \left(\frac{\sum_{j=2}^{r} g(a_j)^p}{r-1}\right)^{\frac{1}{p}}$$

for $i = 2, \ldots, r$ and $\widetilde{g} = g$ elsewhere. Since $\sum_{i=1}^{n} |\widetilde{g}(a_i)|^p = \sum_{i=1}^{n} |g(a_i)|^p$ it is enough to prove that

$$\sum_{i=1}^{n} |M_{S_n}\widetilde{g}(a_i)|^p \geq \sum_{i=1}^{n} |M_{S_n} g(a_i)|^p.$$

Let us observe first that

$$\widetilde{m} := \frac{\sum_{i=1}^{n} \widetilde{g}(a_i)}{n} \geq \frac{\sum_{i=1}^{n} g(a_i)}{n} = m,$$

since

$$(r-1)\left(\frac{\sum_{i=2}^{r} g(a_i)^p}{r-1}\right)^{1/p} \geq \sum_{i=2}^{r} g(a_i)$$

by Hölder's inequality. Thus, for $i = r+1, \ldots, n$ we have that $M_{S_n}\widetilde{g}(a_i) \geq \widetilde{m} \geq m = M_{S_n} g(a_i)$. Also, we observe that for all $i \in \{2, \ldots, r\}$ we have

$$M_{S_n}\widetilde{g}(a_i) \geq \frac{g(a_1) + \widetilde{g}(a_i)}{2}.$$

So, it is enough to prove that

$$(r-1)\left(\frac{g(a_1)+\left(\frac{\sum_{j=2}^r g(a_j)^p}{r-1}\right)^{1/p}}{2}\right)^p = \sum_{i=2}^r \left(\frac{g(a_1)+\widetilde{g}(a_i)}{2}\right)^p$$
$$\geq \sum_{i=2}^r M_{S_n} g(a_i)^p = \sum_{i=2}^r \left(\frac{g(a_1)+g(a_i)}{2}\right)^p,$$

but that it is equivalent to

$$g(a_1)+\left(\frac{\sum_{j=2}^r g(a_j)^p}{r-1}\right)^{1/p} \geq \left(\frac{\sum_{i=2}^r (g(a_1)+g(a_i))^p}{r-1}\right)^{1/p},$$

which is a consequence of Minkowsky's inequality. From where we conclude our required result.

■

Theorem 9.2.1. *For all $n \geq 3$, let $S_n = (V,E)$ be the star graph with n vertices $V = \{a_1, a_2, \ldots, a_n\}$ with center at a_1. For all $p \in (1,2]$ we have that*

$$\|M_{S_n}\|_p = \left(\sup_{x\in[0,1)} \frac{1+(n-1)(\frac{x+1}{2})^p}{1+(n-1)x^p}\right)^{\frac{1}{p}}.$$

Proof First, let us assume that $n > 3$. Let $f : V \to \mathbb{R}$ be a function such that $\frac{\|M_{S_n} f\|_p}{\|f\|_p} = \|M_{S_n}\|_p$ as in Proposition 9.2.3. After a normalization (if necessary) we can assume that $f(a_1) = 1$. By Proposition 9.2.3 we have that $f_{|V\setminus a_1}$ only takes two values, let us say $x \leq y \leq 1$, x s-times and y t-times. We will prove that $x = y$. We observe that if both x and y satisfy $x, y \geq 2m_f - 1$ by the same argument as in Proposition 9.2.3 we conclude that $x = y$. The same happens if $x, y \leq 2m_f - 1$. So, the only case remaining is when $x < 2m_f - 1 < y$. Then, we observe that by taking a Dirac's delta in a_1 we have that $\|M_{S_n}\|_p^p \geq 1 + \frac{n-1}{2^p} \geq \frac{n+3}{4}$. Let us first assume that $y < 1$, given ε such that

$$1 > y+\varepsilon > 2\left(\frac{1+t(y+\varepsilon)+sx}{n}\right)-1,$$

we consider $f_\varepsilon : V \to \mathbb{R}$ defined by $f_\varepsilon(a_i) = f(a_i)+\varepsilon$ for all a_i such that $f(a_i) = y$ and $f_\varepsilon = f$ elsewhere. If we consider the function (defined in a neighborhood of 0)

$$L(\varepsilon) := \|M_{S_n} f_\varepsilon\|_p^p - \|M_{S_n}\|_p^p \|f_\varepsilon\|_p^p,$$

we have $L(0) = 0$ and $L(\varepsilon) \leq 0$ in a neighborhood of 0. Therefore $L'(0) = 0$, that is

$$0 = \left(1 + t\left(\frac{1+y+\varepsilon}{2}\right)^p + s\left(\frac{1+sx+t(y+\varepsilon)}{n}\right)^p - \|M_{S_n}\|_p^p(1+t(y+\varepsilon)^p + sx^p)\right)'$$
$$= \frac{tp(\frac{1+y}{2})^{p-1}}{2} + s\frac{tp}{n}\left(\frac{1+sx+ty}{n}\right)^{p-1} - tp\|M_{S_n}\|_p^p(y^{p-1}). \tag{9.1}$$

We observe that in fact $y \geq \frac{1+x}{2}$, if not we would have $m_f \leq \frac{1+x}{2}$, a contradiction. However, that implies that $y \geq m_f$ since it is equivalent to $(s+1)y = (n-t)y \geq sx+1$, which is true because $(s+1)\left(\frac{1+x}{2}\right) \geq sx+1$. Then

$$\frac{n+3}{4}y^{p-1} \leq \|M_{S_n}\|_p^p y^{p-1} \leq \frac{(\frac{1+y}{2})^{p-1}}{2} + \frac{n-2}{n}\left(\frac{1+sx+ty}{n}\right)^{p-1} < \frac{(\frac{1+y}{2})^{p-1}}{2} + \frac{n-2}{n}y^{p-1}.$$

Then

$$\frac{n-1}{2} + \frac{4}{n} < \left(\frac{1+y}{2y}\right)^{p-1} \leq \frac{1+y}{2y}. \tag{9.2}$$

Also, we observe that since $2m_f - 1 > x > 0$, we have $m_f > \frac{1}{2}$, so if $x < y \leq \frac{1}{4}$, we have

$$\frac{1}{2} < m_f = \frac{1+sx+ty}{n} < \frac{1+\frac{s+t}{4}}{n} = \frac{1+\frac{n-1}{4}}{n}.$$

Therefore, $\frac{n}{4} < \frac{3}{4}$, a contradiction. So $y > \frac{1}{4}$, then

$$\frac{1+y}{2y} < \frac{10}{4} = \frac{5}{2}.$$

Then, by (9.2), we obtain $\frac{n-1}{2} + \frac{4}{n} < \frac{5}{2}$, that is false for $n \geq 4$. We conclude this case. The only remaining case is when $y = 1$. In this case, we have that (where L is defined in an interval $(\delta, 0]$, with δ close to 0)

$$\frac{L(0) - L(-\varepsilon)}{\varepsilon} \geq 0.$$

Therefore taking $-\varepsilon \to 0^-$, similarly as we obtained (9.1), we have that $0 \leq \frac{tp}{2} + s\frac{tp}{n}m^{p-1} - tp\|M_{S_n}\|_p^p$ and that implies $\frac{n+3}{4} \leq \frac{1}{2} + 1$, which is false for $n > 3$. Therefore we conclude this case. The remaining case $n = 3$ is treated as follows. By the same argument (and notation) above, if $x < 2m_f - 1 < y$, we have $\|M_{S_3}\|_p^p \geq 1 + \frac{2}{2^p} \geq \frac{3}{2}$ and $y \geq \frac{1}{2}$. Proceeding as before, similarly as we obtained (9.1), we have that

$$\|M_{S_3}\|_p^p y^{p-1} \leq \frac{(\frac{1+y}{2})^{p-1}}{2} + \frac{1}{3}(m_f)^{p-1} \leq \frac{(\frac{1+y}{2})^{p-1}}{2} + \frac{1}{3}y^{p-1},$$

therefore $\frac{7}{3} \leq (\frac{1+y}{2y})^{p-1} \leq (\frac{3}{2})^{p-1} \leq \frac{3}{2}$, a contradiction. So, we conclude the result. ■

Remark 9.2.1. An adaptation of the proof above also shows that for any $p > 1$ there exists a positive constant $N(p)$ such that for any $n > N(p)$ we have

$$\mathbf{C}_{S_n,p} = \left(\sup_{x\in[0,1)} \frac{1+(n-1)(\frac{x+1}{2})^p}{1+(n-1)x^p}\right)^{\frac{1}{p}}.$$

9.3 Asymptotic behavior of $\|M_G\|_p$

In the next propositions we study the behavior of $\|M_{K_n}\|_p$ and $\|M_{S_n}\|_p$ as $p \to \infty$. We start with a useful elementary lemma.

Lemma 9.3.1. *Assume that for $\{p_k\}_{k\in\mathbb{N}} \subset [1,\infty)$ such that $p_k \to \infty$ we have $x_{1,p_k}, \dots, x_{n,p_k} \geq 0$ such that $\lim_{k\to\infty} x_{i,p_k}^{p_k} \to x_i < \infty$, for every $i = 1, \dots, n$. Then we have that*

$$\lim_{k\to\infty} \left(\frac{\sum_{i=1}^n x_{i,p_k}}{n} \right)^{p_k} = (x_1 x_2 \dots x_n)^{\frac{1}{n}}.$$

Proof By AM-GM inequality we have

$$\left(\frac{\sum_{i=1}^n x_{i,p_k}}{n} \right)^{p_k} \geq \left(x_{1,p_k}^{p_k} x_{2,p_k}^{p_k} \dots x_{n,p_k}^{p_k} \right)^{\frac{1}{n}} \to (x_1 x_2 \dots x_n)^{\frac{1}{n}}.$$

So, we just need to prove that

$$\limsup_{k\to\infty} \left(\frac{\sum_{i=1}^n x_{i,p_k}}{n} \right)^{p_k} \leq (x_1 x_2 \dots x_n)^{\frac{1}{n}}.$$

Given $\varepsilon > 0$, for k big enough we have that $x_{i,p_k} \leq (x_i + \varepsilon)^{\frac{1}{p_k}}$ for every $i = 1, \dots, n$. Then, we observe that

$$\lim_{k\to\infty} \left(\frac{\sum_{i=1}^n (x_i + \varepsilon)^{\frac{1}{p_k}}}{n} \right)^{p_k} = ((x_1 + \varepsilon)(x_2 + \varepsilon) \dots (x_n + \varepsilon))^{\frac{1}{n}}$$

by the L'Hospital rule after applying log in both sides. Therefore, for every given $\varepsilon > 0$ we have

$$\limsup_{k\to\infty} \left(\frac{\sum_{i=1}^n x_{i,p_i}}{n} \right)^{p_k} \leq ((x_1 + \varepsilon)(x_2 + \varepsilon) \dots (x_n + \varepsilon))^{\frac{1}{n}},$$

from where we conclude the result. ∎

Now we continue by analyzing the behavior of $\|M_{K_n}\|_p^p$ when p goes to ∞. In the following lemma we construct an example that helps us to achieve that goal.

Lemma 9.3.2. *Let $K_n = (V, E)$ be the complete graph with n vertices $V = \{a_1, a_2 \dots, a_n\}$. Then,*

$$\liminf_{p\to\infty} \|M_{K_n}\|_p^p \geq \sup_{\alpha>1, k\in\{1,\dots,n\}} \frac{k\alpha^{\frac{n}{k}} + \alpha(n-k)}{k\alpha^{\frac{n}{k}} + n - k}.$$

Proof For fixed k and $\alpha > 1$ we define the function $f : V \to \mathbb{R}_{\geq 0}$ given by $f_p(a_i) = \frac{n\alpha^{\frac{1}{p}} - (n-k)}{k}$ for $i \leq k$, and $f_p(a_i) = 1$ elsewhere. Thus we have $m_p := \frac{\sum_{i=1}^n f_p(a_i)}{n} = \alpha^{\frac{1}{p}}$. Moreover, we observe that

$$\lim_{p\to\infty} \left(\frac{n\alpha^{\frac{1}{p}} - (n-k)}{k} \right)^p = \alpha^{\frac{n}{k}},$$

therefore

$$\liminf_{p\to\infty} \|M_{K_n}\|_p^p \geq \lim_{p\to\infty} \frac{k\left(\frac{n\alpha^{\frac{1}{p}}-(n-k)}{k}\right)^p + (n-k)m_p^p}{k\left(\frac{n\alpha^{\frac{1}{p}}-(n-k)}{k}\right)^p + (n-k)}$$
$$= \frac{k\alpha^{\frac{n}{k}} + (n-k)\alpha}{k\alpha^{\frac{n}{k}} + (n-k)},$$

from where we conclude the result. ■

We observe that the previous proof gives us the lower bound

$$\|M_{K_n}\|_p^p \geq \sup_{\alpha>1, k\in\{1,\ldots,n\}} \frac{k\left(\frac{n\alpha^{\frac{1}{p}}-(n-k)}{k}\right)^p + (n-k)\alpha}{k\left(\frac{n\alpha^{\frac{1}{p}}-(n-k)}{k}\right)^p + (n-k)},$$

for every $p \geq 1$. Now we claim that this lower bound gives essentially the behavior when $p \to \infty$ for $\|M_{K_n}\|_p^p$. This is the content of the following theorem.

Theorem 9.3.1. *Let $n \geq 3$ and let $K_n = (V, E)$ be the complete graph with n vertices $V = \{a_1, a_2 \ldots, a_n\}$. Then,*

$$\lim_{p\to\infty} \|M_{K_n}\|_p^p = \sup_{\alpha>1, k\in\{1,\ldots,n\}} \frac{k\alpha^{\frac{n}{k}} + \alpha(n-k)}{k\alpha^{\frac{n}{k}} + n - k}.$$

Proof By the previous lemma we just need to prove that

$$\limsup_{p\to\infty} \|M_{K_n}\|_p^p \leq \sup_{\alpha>1, k\in\{1,\ldots,n\}} \frac{k\alpha^{\frac{n}{k}} + \alpha(n-k)}{k\alpha^{\frac{n}{k}} + n - k} := C_n.$$

Observe that $C_n > 1$ since $\alpha > 1$. Moreover, by Proposition 9.2.1 for all $p > 1$ there exists a function $f_p : V \to \mathbb{R}$ such that $\|M_{K_n}\|_p = \frac{\|M_{K_n} f_p\|_p}{\|f_p\|_p}$. Let us assume that there exists a sequence $p_i \to \infty$, such that:

$$\frac{\|M_{K_n} f_{p_i}\|_{p_i}^{p_i}}{\|f_{p_i}\|_{p_i}^{p_i}} > c, \tag{9.3}$$

for a fixed constant $c > C_n$. We assume without loss of generality that $f(a_1) \geq f(a_2) \cdots \geq f(a_n)$. By Proposition 9.2.2, we know that f_{p_i} only takes two values, if the minimum of these two values is 0, after a normalization (if necessary) we could assume $f_{p_i}(a_j) = 1$ for $j \leq k_0 < n$ and $f_{p_i} = 0$ elsewhere, then

$$\frac{\|M_{K_n} f_{p_i}\|_{p_i}^{p_i}}{\|f_{p_i}\|_{p_i}^{p_i}} = \frac{k_0 + (n-k_0)(\frac{k_0}{n})^{p_i}}{k_0} \leq 1 + (n-1)\left(\frac{n-1}{n}\right)^{p_i} \to 1,$$

a contradiction for p_i big enough. So we can assume without loss of generality that f_{p_i} takes two different positive values, and after a normalization, we can assume that the minimum value of f_{p_i} is 1. Let us call the other value by $y_{p_i} > 1$. Let us take a subsequence of p_i (that we also call p_i) such that $f_{p_i}(a_r) = y_{p_i}$ for $r \leq k$ (for some fixed $k \in \{1, \dots, n\}$) and $f_{p_i}(a_r) = 1$ elsewhere. We claim that $y_{p_i} \to 1$. In fact, if there exist a subsequence (that we also call p_i) such that $y_{p_i} \geq \rho >, 1$ we have

$$\frac{m_{p_i}}{y_{p_i}} = \frac{k + (n-k)\frac{1}{y_{p_i}}}{n} \leq \frac{k + (n-k)\frac{1}{\rho}}{n} < 1.$$

Therefore

$$\begin{aligned}\frac{\|M_{K_n} f_{p_i}\|_{p_i}^{p_i}}{\|f_{p_i}\|_{p_i}^{p_i}} = \frac{k y_{p_i}^{p_i} + (n-k) m_{p_i}^{p_i}}{k y_{p_i}^{p_i} + (n-k)} &\leq 1 + (n-k)\left(\frac{m_{p_i}}{y_{p_i}}\right)^{p_i} \\ &\leq 1 + (n-k)\left(\frac{k + (n-k)\frac{1}{\rho}}{n}\right)^{p_i} \to 1,\end{aligned}$$

a contradiction. Now we claim that the $y_{p_i}^{p_i}$ are uniformly bounded. Assume that for a subsequence (that we also call y_{p_i}) we have $y_{p_i}^{p_i} \to \infty$. We consider the function $g(x) = nx^{\frac{n-\frac{1}{2}}{n}} - kx - (n-k) = 0$, we observe that $g(x) \geq 0$ for $x \in \left[1, \left(\frac{(n-\frac{1}{2})}{k}\right)^{2n}\right]$. In fact $g(1) = 0$ and g is increasing in $\left[1, \left(\frac{(n-\frac{1}{2})}{k}\right)^{2n}\right]$ since $g'(x) = \left(n - \frac{1}{2}\right) x^{\frac{-1}{2n}} - k \geq 0$. Now, for p_i big enough we have $y_{p_i} \in \left[1, \left(\frac{(n-\frac{1}{2})}{k}\right)^{2n}\right]$. Thus $n y_{p_i}^{\frac{n-\frac{1}{2}}{n}} - k y_{p_i} - (n-k) \geq 0$ and then

$$m_{p_i} = \frac{k y_{p_i} + n - k}{n} \leq y_{p_i}^{\frac{n-\frac{1}{2}}{n}}.$$

Therefore

$$\begin{aligned}\frac{\|M_{K_n} f_{p_i}\|_{p_i}^{p_i}}{\|f_{p_i}\|_{p_i}^{p_i}} = \frac{k y_{p_i}^{p_i} + (n-k) m_{p_i}^{p_i}}{k y_{p_i}^{p_i} + (n-k)} &\leq 1 + (n-k)\left(\frac{m_{p_i}}{y_{p_i}}\right)^{p_i} \\ &\leq 1 + (n-k)\left(y_{p_i}^{-\frac{1}{2n}}\right)^{p_i} \to 1,\end{aligned}$$

reaching a contradiction. So, we have that $y_{p_i}^{p_i}$ are uniformly bounded. Let us take a subsequence of p_i (that we also denote p_i for simplicity) such that $y_{p_i}^{p_i}$ and $m_{p_i}^{p_i}$ converge. Let us write $\lim_{p_i \to \infty} y_{p_i}^{p_i} = \alpha_1$ and $\lim_{p_i \to \infty} m_{p_k}^{p_k} = \alpha_2$. Then, by Lemma 9.3.1 we have (taking $x_{s,p_k} = y_{p_k}$ for $s \leq k$ and $x_{s,p_k} = 1$ for $s > 1$) $\alpha_2 = \alpha_1^{\frac{k}{n}}$.

This implies

$$\lim_{p_i \to \infty} \frac{\|M_{K_n} f_{p_i}\|_{p_i}^{p_i}}{\|f_{p_i}\|_{p_i}^{p_i}} = \lim_{p_i \to \infty} \frac{k y_{p_i}^{p_i} + (n-k) m_{p_i}^{p_i}}{k y_{p_i}^{p_i} + (n-k)} = \frac{k \alpha_2^{\frac{n}{k}} + (n-k)\alpha_2}{k \alpha_2^{\frac{n}{k}} + (n-k)} \leq C_n.$$

Then, it is not possible to have a sequence like in (9.3), therefore

$$\limsup_{p\to\infty} \|M_{K_n}\|_p^p \leq C_n$$

as desired. ■

Now, we start analyzing the behavior of $\|M_{S_n}\|_p^p$ when p goes to ∞. In the following lemma we construct an example that helps us to achieve this goal.

Lemma 9.3.3. *Let $n \geq 3$ and let $S_n = (V, E)$ be the star graph with n vertices $V = \{a_1, a_2, \ldots, a_n\}$ with center at a_1. Then,*

$$\liminf_{p\to\infty} \|M_{S_n}\|_p^p \geq \frac{1+\sqrt{n}}{2}.$$

Proof For fixed $k > 1$ we define $y_{k,p} = 2k^{\frac{1}{p}} - 1$, we observe that $\left(\frac{1+y_{k,p}}{2}\right)^p = k$. Let us consider the function $f_{k,p} : V \to \mathbb{R}_{\geq 0}$ by $f_{k,p}(a_1) = y_{k,p}$ and $f_{k,p}(a_i) = 1$ for $i > 1$. Then, we have

$$\|M_{S_n}\|_p^p \geq \left(\frac{\|M_{S_n} f_{k,p}\|_p}{\|f_{k,p}\|_p}\right)^p = \frac{\left(2k^{\frac{1}{p}} - 1\right)^p + (n-1)k}{\left(2k^{\frac{1}{p}} - 1\right)^p + (n-1)}.$$

We observe that by L'Hospital $\lim_{p\to\infty} \left(2k^{\frac{1}{p}} - 1\right)^p = k^2$, therefore we have

$$\liminf_{p\to\infty} \|M_{S_n}\|_p^p \geq \frac{k^2 + (n-1)k}{k^2 + (n-1)}.$$

By taking $k = \sqrt{n} + 1$ we have $\frac{k^2+(n-1)k}{k^2+(n-1)} = \frac{\sqrt{n}+1}{2}$, from where we conclude our proposition. ■

We observe that the proof above gives us the estimate

$$\|M_{S_n}\|_p^p \geq \frac{\left(2(1+\sqrt{n})^{\frac{1}{p}} - 1\right)^p + (n-1)(1+\sqrt{n})}{\left(2(1+\sqrt{n})^{\frac{1}{p}} - 1\right)^p + (n-1)}$$

for every $p \geq 1$. Moreover, we observe that $\left(2(1+\sqrt{n})^{\frac{1}{p}} - 1\right)^p$ is an increasing function on p. This is the case because the derivative of $p \log\left(2(1+\sqrt{n})^{\frac{1}{p}} - 1\right)$ is

$$\log\left(2(1+\sqrt{n})^{\frac{1}{p}} - 1\right) - \frac{(1+\sqrt{n})^{\frac{1}{p}}\log(1+\sqrt{n})}{(2(1+\sqrt{n})^{\frac{1}{p}} - 1)p} \geq \log(2(1+\sqrt{n})^{\frac{1}{p}} - 1) - \frac{\log(1+\sqrt{n})}{p} \geq 0.$$

Thus, we have that

$$\frac{\left(2(1+\sqrt{n})^{\frac{1}{p}}-1\right)^p+(n-1)(1+\sqrt{n})}{\left(2(1+\sqrt{n})^{\frac{1}{p}}-1\right)^p+(n-1)}$$

is decreasing with respect to p. Then

$$\|M_{S_n}\|_p^p \geq \lim_{t\to\infty} \frac{\left(2(1+\sqrt{n})^{\frac{1}{t}}-1\right)^t+(n-1)(1+\sqrt{n})}{\left(2(1+\sqrt{n})^{\frac{1}{t}}-1\right)^t+(n-1)} = \frac{\sqrt{n}+1}{2},$$

for all $p \geq 1$. Note that this lower bound is better than the one observed by Soria and Tradacete $1+\frac{n-1}{2^p} \leq \|M_{S_n}\|_p^p$ (see Proposition 3.4 in [ST16]) whenever $p > \frac{\log(\sqrt{n}+1)}{\log 2}+1$.

Let us define

$$\|M_{S_n}\|_p^* := \sup_{y\geq 1}\left(\frac{y^p+(n-1)(\frac{1+y}{2})^p}{y^p+(n-1)}\right)^{\frac{1}{p}}.$$

Our next goal is to analyze the relation between this object and $\|M_{S_n}\|_p$. We start observing that by definition $\|M_{S_n}\|_p^* \leq \|M_{S_n}\|_p$. Also, we have the following.

Lemma 9.3.4. *Let $n \geq 3$. The following identity holds*

$$\lim_{p\to\infty}\left(\|M_{S_n}\|_p^*\right)^p = \frac{1+\sqrt{n}}{2}.$$

Proof We start observing that the proof of Lemma 9.3.3 also works for $\|M_{S_n}\|_p^*$. Then, it is enough to prove that

$$\limsup_{p\to\infty}\,(\|M_{S_n}\|_p^*)^p \leq \frac{1+\sqrt{n}}{2}.$$

Let us assume that there exists a sequence $p_k \to \infty$ and $y_{p_k} > 1$ such

$$\frac{y_{p_k}^{p_k}+(n-1)\left(\frac{1+y_{p_k}}{2}\right)^{p_k}}{y_{p_k}^{p_k}+(n-1)} \geq c > \frac{1+\sqrt{n}}{2}.$$

We observe that $y_{p_k} \to 1$. In fact if there exists a subsequence k_j such that $y_{p_{k_j}} \geq \rho > 1$, we have $\frac{1+y_{p_{k_j}}}{2y_{p_{k_j}}} \leq \frac{1}{2\rho}+\frac{1}{2} < 1$, then

$$\frac{y_{p_{k_j}}^{p_{k_j}}+(n-1)\left(\frac{1+y_{p_{k_j}}}{2}\right)^{p_{k_j}}}{y_{p_{k_j}}^{p_{k_j}}+(n-1)} \leq 1+(n-1)\left(\frac{1}{2\rho}+\frac{1}{2}\right)^{p_{k_j}} \to 1.$$

Therefore, this cannot be the case. Now we prove that the $y_{p_k}^{p_k}$ are uniformly bounded. In fact, since $y_{p_k} < 2$ for k big enough, we have $y_{p_k}^{\frac{3}{4}} \geq \frac{y_{p_k}+1}{2}$, since $2x^{\frac{3}{4}} - x - 1 \geq 0$ for $x \in [1,2]$. Therefore, if $y_{p_{k_j}}^{p_{k_j}} \to \infty$ we have

$$\frac{y_{p_{k_j}}^{p_{k_j}} + (n-1)\left(\frac{1+y_{p_{k_j}}}{2}\right)^{p_{k_j}}}{y_{p_{k_j}}^{p_{k_j}} + (n-1)} \leq \frac{y_{p_{k_j}}^{p_{k_j}} + (n-1)(y_{p_{k_j}})^{\frac{3p_{k_j}}{4}}}{y_{p_{k_j}}^{p_{k_j}} + (n-1)} \to 1.$$

So, we have that $y_{p_k}^{p_k}$ are uniformly bounded. Let us take a subsequence of $\{p_k\}_{k\in\mathbb{N}}$ (that we also denote $\{p_k\}_{k\in\mathbb{N}}$) such that $y_{p_k}^{p_k}$ and $\left(\frac{1+y_{p_k}}{2}\right)^{p_k}$ converges. Let us write $\lim_{k\to\infty} y_{p_k}^{p_k} = \alpha_1$ and $\lim_{k\to\infty}\left(\frac{1+y_{p_k}}{2}\right)^{p_k} = \alpha_2$. By Lemma 9.3.1 (with $n=2$, $x_{1,p_k} = y_{p_k}$ and $x_{2,p_k} = 1$) we have $\alpha_2 = \sqrt{\alpha_1}$. Then, observe that

$$\lim_{k\to\infty} \frac{y_{p_k}^{p_k} + (n-1)\left(\frac{1+y_{p_k}}{2}\right)^{p_k}}{y_{p_k}^{p_k} + (n-1)\left(\frac{1+y_{p_k}}{2}\right)^{p_k}} = \frac{\alpha_2^2 + (n-1)\alpha_2}{\alpha_2^2 + (n-1)} \leq \frac{\sqrt{n}+1}{2},$$

since this last inequality it is equivalent to $\alpha_2^2(\sqrt{n}-1) - 2(n-1)\alpha_2 + (n-1)(\sqrt{n}+1) \geq 0$, and this is true because $\alpha_2^2(\sqrt{n}-1) - 2(n-1)\alpha_2 + (n-1)(\sqrt{n}+1) = (\sqrt{n}-1)(\alpha_2 - (\sqrt{n}+1))^2$. This concludes the proof. ∎

We conclude this section describing the asymptotic behavior of $\|M_{S_n}\|_p^p$ as $p \to \infty$.

Theorem 9.3.2. *Fix $n \in \mathbb{N}$. Let $S_n = (V,E)$ be the star graph with n vertices $V = \{a_1, a_2, \dots, a_n\}$ with center at a_1. For $n \geq 25$ we have*

$$\lim_{p\to\infty} \|M_{S_n}\|_p^p = \frac{1+\sqrt{n}}{2}.$$

Proof We choose $f_p : V \to \mathbb{R}_{\geq 0}$ such that $\frac{\|M_{S_n} f_p\|_p^p}{\|f_p\|_p^p} = \|M_{S_n}\|_p^p$. First we observe that for all p sufficiently large we have $1 + \frac{n-1}{2^p} < \frac{\sqrt{n}+1}{2}$. Then, by Proposition 9.2.3, we can assume that $f_p(a_1) = y_p \geq 1 = f_p(a_2) = f_p(a_3) = \cdots = f_p(a_{s+1}) \geq x_p = f_p(a_{s+2}) = \cdots = f_p(a_n)$. Moreover, we assume that $\frac{y_p + x_p}{2} < \frac{\sum_{i=1}^n f_p(a_i)}{n} =: m_p$ (otherwise, we would have that $x_p \geq 2m_p - f_p(a_1)$, and we can proceed as in the Step 3 of the proof of Proposition 9.2.3 to conclude that $x_p = 1$). Notice that the case $s = 1$ is not possible since then $\frac{x_p+y_p}{2} \geq \frac{y_p + (n-1)x_p}{n} = m_p$. Let us assume that there exists $1 < s < n-1$ and a sequence $p_k \to \infty$ such that

$$\frac{\|M_{S_n} f_{p_k}\|_{p_k}^{p_k}}{\|f_{p_k}\|_p^p} = \frac{y_{p_k}^{p_k} + s\left(\frac{1+y_{p_k}}{2}\right)^{p_k} + (n-s-1)m_{p_k}^{p_k}}{y_{p_k}^{p_k} + s + (n-s-1)x_{p_k}^{p_k}} > \frac{1+\sqrt{n}}{2}.$$

First we observe that $y_{p_k} \to 1$. If not, there exists a subsequence of $\{p_k\}_{k\in\mathbb{N}}$ (that we also call $\{p_k\}_{k\in\mathbb{N}}$) such that $y_{p_k} > \rho > 1$. Then

$$\frac{m_{p_k}}{y_{p_k}} \leq \frac{y_{p_k}+1}{2y_{p_k}} \leq \frac{1}{2} + \frac{1}{2\rho} < 1$$

and therefore

$$\frac{y_{p_k}^{p_k} + s\left(\frac{1+y_{p_k}}{2}\right)^{p_k} + (n-s-1)m_{p_k}^{p_k}}{y_{p_k}^{p_k} + s + (n-s-1)x_{p_k}^{p_k}} \to 1,$$

a contradiction. Now we claim that the $y_{p_k}^{p_k}$ are uniformly bounded. Assume that for a subsequence (that we also call $\{y_{p_k}\}_{k\in\mathbb{N}}$) we have $y_{p_k}^{p_k} \to \infty$. First observe that for k big enough we have $y_{p_k} > 1$. Then, as in the proof of Lemma 9.3.4, we have that for p_k big enough

$$y_{p_k}^{\frac{3}{4}} \geq \frac{y_{p_k}+1}{2} \geq m_{p_k},$$

from where we have

$$\frac{\|M_{S_n} f_{p_k}\|_{p_k}^{p_k}}{\|f_{p_k}\|_{p_k}^{p_k}} \leq \frac{y_{p_k}^{p_k} + (n-1)y_{p_k}^{\frac{3p_k}{4}}}{y_{p_k}^{p_k} + s + (n-s-1)x_{p_k}^{p_k}} \to 1.$$

Reaching a contradiction. Therefore we can take a subsequence (that we also call $\{p_k\}_{k\in\mathbb{N}}$) such that $\lim_{k\to\infty} y_{p_k}^{p_k} = \alpha_1$, $\lim_{k\to\infty} \left(\frac{1+y_{p_k}}{2}\right)^{p_k} = \alpha_2$, $\lim_{k\to\infty} m_{p_k}^{p_k} = \alpha_3$ and $\lim_{k\to\infty} x_{p_k}^{p_k} = \alpha_4$. By Lemma 9.3.1 we have that $\alpha_2 = \sqrt{\alpha_1}$ and $\alpha_3 = \alpha_1^{\frac{1}{n}}\alpha_4^{\frac{n-s-1}{n}}$. Therefore

$$\lim_{k\to\infty} \frac{\|M_{S_n} f_{p_k}\|_{p_k}^{p_k}}{\|f_{p_k}\|_{p_k}^{p_k}} \leq \frac{\alpha_2^2 + s\alpha_2 + (n-s-1)\alpha_1^{\frac{1}{n}}\alpha_4^{\frac{n-s-1}{n}}}{\alpha_2^2 + s + (n-s-1)\alpha_4},$$

we claim that this last expression is bounded above by $\frac{\sqrt{n}+1}{2}$ in our setting, from where we would conclude. Since $\alpha_4 \leq 1$ it is enough to prove that

$$\alpha_2^2 + s\alpha_2 + (n-s-1)\alpha_2^{\frac{2}{n}} \leq \left(\frac{\sqrt{n}+1}{2}\right)\left(\alpha_2^2 + s + (n-s-1)\alpha_4\right). \tag{9.4}$$

To this end it is sufficient to prove

$$\alpha_2^2 + s\alpha_2 + (n-s-1)\alpha_2^{\frac{2}{n}} \leq \left(\frac{\sqrt{n}+1}{2}\right)\left(\alpha_2^2 + s\right).$$

We observe now that since $m_{p_k} \geq \frac{y_{p_k}+x_{p_k}}{2}$ we have $2(y_{p_k} + 2s + 2(n-s-1)x_{p_k}) \geq ny_{p_k} + nx_{p_k}$ and then $2s + (n-2(s+1))x_{p_k} \geq (n-2)y_{p_k}$. Since $1, x_{p_k} < y_{p_k}$, if we assume $n-2(s+1) \geq 0$ we would get

$$(n-2)y_{p_k} \geq 2sy_{p_k} + (n-2(s+1))y_{p_k} > 2s + (n-2(s+1))x_{p_k} \geq (n-2)y_{p_k},$$

a contradiction. Therefore we have $n-2(s+1)\leq -1$ and then

$$n\leq 2s+1. \tag{9.5}$$

Now we assume that $n\geq 25$. We distinguish among two cases, first when $\alpha_2\geq\left(\frac{6}{5}\right)^n$. Here

$$\alpha_2^2+s\alpha_2+(n-s-1)\alpha_2^{\frac{2}{n}}\leq \alpha_2^2+(n-1)\alpha_2\leq 2\alpha_2^2,$$

since $(n-1)\leq(\frac{6}{5})^n\leq\alpha_2$, where we use that $n\geq 25$. Then since $2\alpha_2^2\leq\left(\frac{\sqrt{25}+1}{2}\right)\alpha_2^2$ we conclude this case. Now we consider the case where $\alpha_2\leq\left(\frac{6}{5}\right)^n$. Here we have $\alpha_2^{\frac{2}{n}}\leq\left(\frac{6}{5}\right)^2$. Therefore we just need to prove that

$$\alpha_2^2+s\alpha_2+(n-s-1)\left(\frac{6}{5}\right)^2\leq\left(\frac{\sqrt{n}+1}{2}\right)\left(\alpha_2^2+s\right),$$

or equivalently

$$0\leq\left(\frac{\sqrt{n}-1}{2}\right)\alpha_2^2-s\alpha_2+s\left(\frac{\sqrt{n}+1}{2}\right)-(n-s-1)\left(\frac{6}{5}\right)^2.$$

We just need to verify then that the discriminant of that equation is less than 0. That is

$$s^2<4\left(\frac{\sqrt{n}-1}{2}\right)\left[s\left(\frac{\sqrt{n}+1}{2}\right)-(n-s-1)\left(\frac{6}{5}\right)^2\right]=(n-1)s-2(\sqrt{n}-1)(n-s-1)\left(\frac{6}{5}\right)^2.$$

Since $(n-1)s=s^2+s(n-s-1)$ we just need $2\left(\frac{6}{5}\right)^2(\sqrt{n}-1)<\frac{n-1}{2}$ (given that $s\geq\frac{n-1}{2}$ by (9.5)) or equivalently $4\left(\frac{6}{5}\right)^2-1<\sqrt{n}$. Since the left hand side is lesser than 5 we conclude this and therefore we conclude (9.4), from where the theorem follows. ■

Remark 9.3.1. From the previous proof it can be deduced that, if we define

$$\mathcal{A}=\left\{(s,\alpha_2,\alpha_4)\in\{1,\dots,n-2\}\times[1,\infty)\times[0,1];\alpha_2^{\frac{2}{n}}\alpha_4^{\frac{n-s-1}{n}}>\alpha_2\sqrt{\alpha_4}\right\},$$

for $n\in[3,24]$ we have

$$\lim_{p\to\infty}\|M_{S_n}\|_p^p=\max\left\{\frac{1+\sqrt{n}}{2},\sup_{(s,\alpha_2,\alpha_4)\in\mathcal{A}}\frac{\alpha_2^2+s\alpha_2+(n-s-1)\alpha_2^{\frac{2}{n}}\alpha_4^{\frac{n-s-1}{n}}}{\alpha_2^2+s+(n-s-1)\alpha_4}\right\}.$$

However, for $n<25$, to compare the inner terms in the right hand side is more difficult.

9.4 The p-variation of maximal operators on graphs

In order to compute $\mathbf{C}_{G,p}$ it is useful to study the functions that attain this supremum. Now we prove that actually these extremizers exist.

Proposition 9.4.1. *Given any connected simple finite graph $G = (V, E)$ and $p \in (0, \infty)$ there exists $f : V \to \mathbb{R}_{\geq 0}$ such that*

$$\frac{\operatorname{Var}_p M_G f}{\operatorname{Var}_p f} = \mathbf{C}_{G,p}$$

Proof We write $|V| = n$ and $V =: \{a_1, a_2, \ldots, a_n\}$. Given $y =: (y_1, \ldots, y_n) \in [0,1]^n \cap \left\{\max_{i=1,\ldots,n} y_i = 1\right\} \cap \left\{\min_{i=1,\ldots,n} y_i = 0\right\} =: A$, we define $f_y : V \to \mathbb{R}_{\geq 0}$ by $f_y(a_i) = y_i$. We observe that $M_G f_y(a_i)$ is continuous in such set for any $i = 1, \ldots, n$. Therefore $\frac{\operatorname{Var}_p M_G f_y}{\operatorname{Var}_p f_y}$ is continuous with respect to y in A since the denominator is never 0. Thus it attains its maximum at a point $y_0 \in A$. We claim that

$$\frac{\operatorname{Var}_p M_G f_{y_0}}{\operatorname{Var}_p f_{y_0}} = \mathbf{C}_{G,p}.$$

In fact, for every $g : V \to \mathbb{R}_{\geq 0}$ we have that the value $\frac{\operatorname{Var}_p M_G g}{\operatorname{Var}_p g}$ remains unchanged by doing the transformation

$$g \mapsto \frac{g - \min_{i=1,\ldots,n} g(a_i)}{\max_{i=1,\ldots,n} g(a_i)}.$$

This last function is equal to f_y for some y, from where we conclude the result. ■

9.4.1 The 2-variation of M_{S_n}

For all $p \geq 1$, it was proved in the previous chapter that $\operatorname{Var}_p M_{K_n} f \leq \left(1 - \frac{1}{n}\right) \operatorname{Var}_p f$, for any real valued function f defined on the vertices of K_n. The equality occurs when f is a delta function. The analogous problem for the star graph S_n is more challenging, it was observed in the previous chapter that in this case delta functions are not extremizers. Our next result solves this problem for $p = 2$.

Theorem 9.4.1. *Let $n \geq 3$ and let $S_n = (V, E)$ be the star graph with n vertices $V = \{a_1, a_2, \ldots, a_n\}$ with center at a_1. The following inequality holds*

$$\operatorname{Var}_2 M_{S_n} f \leq \left(\frac{[(n-1)^2 + n - 2]^{1/2}}{n}\right) \operatorname{Var}_2 f$$

for all $f : V \to \mathbb{R}$. Moreover, this result is optimal.

Proof The proof of this result is divided in two cases, the case 2 is divided in many steps.

Case 1: $f(a_1) \le m_f$. In this case $M_{S_n}f(a_1) = m_f$. If $M_{S_n}f(a_1) \ge M_{S_n}f(a_i) \ge m_f$ for all $2 \le i \le n$ then the result is trivial.

Then, we assume without loss of generality that $M_{S_n}f(a_i) > M_{S_n}f(a_1)$ for all $i \in \{2,3,\dots,k\}$ for some $2 \le k \le n$ and $M_{S_n}f(a_i) = M_{S_n}f(a_1) = m_f$ for all $i \in \{k+1, k+2,\dots,n\}$.

We have that

$$(\mathrm{Var}\,_2 M_{S_n}f)^2 = \sum_{i=2}^{k}(f(a_i)-m_f)^2. \tag{9.6}$$

Moreover, for all $i \in \{2,3,\dots,k\}$ we have that

$$\begin{aligned} 0 < (f(a_i)-m_f) &= \frac{1}{n}\left((n-1)f(a_i) - \sum_{j=1,j\neq i}^{k} f(a_j) - \sum_{j=k+1}^{n} f(a_j)\right) \\ &= \frac{1}{n}\left((n-1)(f(a_i)-f(a_1)) - \sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1)) + \sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right) \end{aligned} \tag{9.7}$$

Let

$$S^+ := \left\{ i \in \{2,3,\dots,k\}; -\sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1)) + \sum_{j=k+1}^{n}(f(a_1)-f(a_j)) > 0 \right\},$$

and $S^- := \{2,3,\dots,k\} \setminus S^+$. Then by (9.6) and (9.7) we have that

$$(\mathrm{Var}\,_2 M_{S_n}f)^2 \le \frac{(n-1)^2}{n^2}\sum_{i\in S^-}(f(a_i)-f(a_1))^2 + \sum_{i\in S^+}(f(a_i)-m_f)^2, \tag{9.8}$$

and

$$\begin{aligned} \sum_{i\in S^+}(f(a_i)-m_f)^2 &= \frac{(n-1)^2}{n^2}\sum_{i\in S^+}(f(a_i)-f(a_1))^2 \\ &+ \frac{2(n-1)}{n^2}\sum_{i\in S^+}(f(a_i)-f(a_1))\left(-\sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1)) + \sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right) \\ &+ \frac{1}{n^2}\sum_{i\in S^+}\left(-\sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1)) + \sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right)^2. \end{aligned} \tag{9.9}$$

Also, we observe that, since $f(a_1) \le m_f$, then

$$\sum_{i=2}^{k}(f(a_i)-f(a_1)) \ge \sum_{i=k+1}^{n} f(a_1)-f(a_i),$$

therefore

$$\begin{aligned}
&\sum_{i\in S^+}\left(-\sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1))+\sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right)^2\\
&\leq\left[\sum_{i\in S^+}\left(-\sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1))+\sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right)\right]^2\\
&\leq\left(|S^+|\sum_{j=k+1}^{n}(f(a_1)-f(a_j))-(|S^+|-1)\sum_{j=2}^{k}(f(a_j)-f(a_1))\right)^2 \qquad (9.10)\\
&\leq\left(\sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right)^2\\
&\leq(n-k)\sum_{j=k+1}^{n}(f(a_1)-f(a_j))^2.
\end{aligned}$$

Moreover, by the AM-GM inequality we have that

$$\begin{aligned}
&2(n-1)\sum_{i\in S^+}(f(a_i)-f(a_1))\left(-\sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1))+\sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right)\\
&\leq\sum_{i\in S^+}(n-2)(f(a_i)-f(a_1))^2 \qquad (9.11)\\
&\quad+\sum_{i\in S^+}\left[\frac{(n-1)^2}{n-2}\left(-\sum_{j=2,j\neq i}^{k}(f(a_j)-f(a_1))+\sum_{j=k+1}^{n}(f(a_1)-f(a_j))\right)^2\right].
\end{aligned}$$

Combining (9.9),(9.10), (9.11) and using that $k\geq 2$ we obtain that

$$\begin{aligned}
\sum_{i\in S^+}(f(a_i)-m_f)^2 \leq&\frac{(n-1)^2+(n-2)}{n^2}\sum_{i\in S^+}(f(a_i)-f(a_1))^2\\
&+\frac{(n-k)}{n^2}\left(1+\frac{(n-1)^2}{n-2}\right)\sum_{j=k+1}^{n}(f(a_1)-f(a_j))^2 \qquad (9.12)\\
&\leq\frac{(n-1)^2+(n-2)}{n^2}\left[\sum_{i\in S^+}(f(a_i)-f(a_1))^2+\sum_{j=k+1}^{n}(f(a_1)-f(a_j))^2\right].
\end{aligned}$$

Finally, combining (9.8) and (9.12) we conclude that

$$(\mathrm{Var}\,_2 M_{S_n}f)^2\leq\frac{(n-1)^2+(n-2)}{n^2}\sum_{i=2}^{n}(f(a_i)-f(a_1))^2.$$

Moreover, we observe that in order to have an equality in (9.12) we need to have $k=2$ (this means that there is only one term larger than $f(a_1)$), in order to have an equality in (9.10) we need to have $f(a_j)=f(a_{k+1})=f(a_3)$ for all $j\geq k+1=3$, and in order to have an equality in (9.11) we need to have $(f(a_2)-f(a_1))=(n-1)(f(a_1)-f(a_3))$. We verify that if $f(a_1)=x>0$, $f(a_j)=x-c$ for all $j\geq 3$ and some $c\in(0,x)$, and $f(a_2)=x+c(n-1)$ then we have an extremizer. In fact, in this case we have that $M_{S_n}f(a_2)=x+c(n-1)$ and $M_{S_n}f(a_j)=x+c/n$ for all $j\neq 2$. Therefore

$$\frac{\mathrm{Var}_2 M_{S_n}f}{\mathrm{Var}_2 f}=\frac{c(n-1-1/n)}{[c^2(n-1)^2+c^2(n-2)]^{1/2}}=\frac{[(n-1)^2+(n-2)]^{1/2}}{n}.$$

Case 2: $f(a_1)>m_f$. For this case we assume without loss of generality that $f(a_2)\geq f(a_3)\geq \dots f(a_s)>f(a_1)\geq f(a_{s+1})\geq\dots f(a_k)>2m_f-f(a_1)\geq f(a_{k+1})\geq\dots f(a_n)$. We observe that $M_{S_n}f(a_i)=f(a_i)$, for $i\leq s$; $M_{S_n}f(a_i)=\frac{f(a_1)+f(a_i)}{2}$ for $s<i\leq k$ and $M_{S_n}f(a_i)=m_f$ for $i>k$. We write $f(a_i)-f(a_1)=x_i$ for $i\leq s$, $f(a_1)-f(a_i)=y_i$ for $i>s$ and $f(a_1)-m_f=u$. Then, our goal is to prove

$$\sum_{i=2}^{s}x_i^2+\sum_{i=s+1}^{k}\left(\frac{y_i}{2}\right)^2+(n-k)u^2\leq\left(1-\frac{n+1}{n^2}\right)\left(\sum_{i=2}^{s}x_i^2+\sum_{i=s+1}^{n}y_i^2\right), \tag{9.13}$$

since $1-\frac{n+1}{n^2}=\frac{(n-1)^2+(n-2)}{n^2}$. Assume that $f:V\to\mathbb{R}_{\geq 0}$ is such that

$$\frac{\mathrm{Var}_2 M_{S_n}f}{\mathrm{Var}_2 f}=\mathbf{C}_{S_n,p}.$$

We prove some properties about f following the ideas of Propositions 9.2.2 and 9.2.3. First, we observe that $s\geq 2$. Otherwise we would have that LHS in (9.13) is less than or equal to

$$\sum_{i=s+1}^{k}\left(\frac{y_i}{2}\right)^2\leq\frac{1}{4}(\mathrm{Var}_2 f)^2<\left(1-\frac{n+1}{n^2}\right)(\mathrm{Var}_2 f)^2.$$

So f could not be an extremizer in that case.

Step 1: $s=2$. We consider $\widetilde{f}:V\to\mathbb{R}$ defined by $\widetilde{f}(a_2)=\sum_{i=2}^{s}f(a_i)-(s-2)f(a_1)$, $\widetilde{f}(a_i)=\widetilde{f}(a_1)$ for $i=3,\dots,s$ and $\widetilde{f}=f$ elsewhere. Clearly $m_{\widetilde{f}}=m_f$ then, defining $\widetilde{x}_i$ and

$\widetilde{y}_i$ and $\widetilde{u}$ analogously to x_i, y_i and u, since $\mathbf{C}^2_{S_n,2} < 1$ we observe that

$$\begin{aligned}
0 &= \mathbf{C}^2_{S_n,2}\left(\sum_{i=2}^{s} x_i^2 + \sum_{i=s+1}^{n} y_i^2\right) - \sum_{i=2}^{s} x_i^2 - \sum_{i=s+1}^{k}\left(\frac{y_i}{2}\right)^2 - (n-k)u^2 \\
&= (\mathbf{C}^2_{S_n,2} - 1)\sum_{i=2}^{s} x_i^2 + (\mathbf{C}^2_{S_n,2} - \frac{1}{4})\left(\sum_{i=s+1}^{k} y_i^2\right) + (\mathbf{C}^2_{S_n,2})\left(\sum_{i=k+1}^{n} y_i^2\right) - (n-k)u^2 \\
&\geq (\mathbf{C}^2_{S_n,2} - 1)\left(\sum_{i=2}^{s} x_i\right)^2 + \left(\mathbf{C}^2_{S_n,2} - \frac{1}{4}\right)\left(\sum_{i=s+1}^{k} y_i^2\right) + (\mathbf{C}^2_{S_n,2})\left(\sum_{i=k+1}^{n} y_i^2\right) - (n-k)u^2 \\
&= (\mathbf{C}^2_{S_n,2} - 1)\sum_{i=2}^{s} \widetilde{x}_i^{\,2} + \left(\mathbf{C}^2_{S_n,2} - \frac{1}{4}\right)\left(\sum_{i=s+1}^{k} \widetilde{y}_i^2\right) + (\mathbf{C}^2_{S_n,2})\left(\sum_{i=k+1}^{n} \widetilde{y}_i^2\right) - (n-k)\widetilde{u}^2.
\end{aligned} \tag{9.14}$$

Therefore

$$\sum_{i=2}^{s} \widetilde{x}_i^2 + \sum_{i=s+1}^{k}\left(\frac{\widetilde{y}_i}{2}\right)^2 + (n-k)\widetilde{u}^2 \geq \mathbf{C}^2_{S_n,2}\left(\sum_{i=2}^{s} \widetilde{x}_i^2 + \sum_{i=s+1}^{n} \widetilde{y}_i^2\right).$$

This implies that

$$\frac{\operatorname{Var}_2 M_{S_n}\widetilde{f}}{\operatorname{Var}_2 \widetilde{f}} \geq \mathbf{C}_{S_n,2},$$

thus (9.14) has to be an equality. Then $\sum_{i=2}^{s} x_i^2 = (\sum_{i=2}^{s} x_i)^2$, therefore, there exists at most one $j \in \{2,\dots,s\}$ such that $x_j \neq 0$. Since we have that $x_j > 0$ for all $j \in \{2,\dots,s\}$ we conclude that $s = 2$.

Step 2: $f(a_j) = f(a_3)$ for all $j \in \{3,4,\dots,k\}$. We define the function $\widetilde{f} : V \to \mathbb{R}_{\geq 0}$ as follows: $\widetilde{f}(a_i) = \frac{\sum_{j=3}^{k} f(a_j)}{k-2}$ for every $i \in \{3,\dots k\}$ and $\widetilde{f} = f$ elsewhere. We define $\widetilde{x}_i$, $\widetilde{y}_i$ and $\widetilde{u}$ analogously to x_i, y_i and u, respectively. We observe that $\sum_{i=3}^{k} \widetilde{y}_i = \sum_{i=3}^{k} y_i$, and by Hölder's inequality we have $\sum_{i=3}^{k} \widetilde{y}_i^2 \leq \sum_{i=3}^{k} y_i^2$. So, similarly as in (9.14), since $\mathbf{C}^2_{S_n,2} > \frac{1}{4}$, we conclude that $\frac{\operatorname{Var}_2 M_{S_n}\widetilde{f}}{\operatorname{Var}_2 \widetilde{f}} \geq \mathbf{C}_{S_n,2}$. Thus $\widetilde{f}$ is also an extremizer, and $\sum_{i=3}^{k} \widetilde{y}_i^2 = \sum_{i=3}^{k} y_i^2$. This implies that $\widetilde{f} = f$.

Step 3: $f(a_j) = f(a_{k+1})$ for all $j \in \{k+1, k+2,\dots,n\}$. Now we define $\widetilde{f} : V \to \mathbb{R}$ as follows: $\widetilde{f}(a_i) = \frac{\sum_{j=k+1}^{n} f(a_j)}{n-k}$ for every $i \geq k+1$, and $f = \widetilde{f}$ elsewhere. Then, we have that

$$\sum_{i=k+1}^{n} \widetilde{y}_i = \sum_{i=k+1}^{n} y_i$$

and

$$\sum_{i=k+1}^{n} \widetilde{y}_i^2 \leq \sum_{i=k+1}^{n} y_i^2.$$

So, by a computation similar to (9.14) we have that

$$\frac{\operatorname{Var}_2 M_{S_n}\widetilde{f}}{\operatorname{Var}_2 \widetilde{f}} \geq \mathbf{C}_{S_n,2}.$$

Then $\widetilde{f} = f$.

Step 4: Conclusion. So, by now we conclude that f takes at most 4 values. In fact, we know that $y_i = y_3$ for $i \leq k$ and $y_i = y_{k+1}$ for $i \geq k+1$. In the following we conclude that $y_3 = y_{k+1}$. We start observing that if $2m_f - f(a_1) = f(a_j)$ for all $j \in \{k+1, \dots, n\}$ then we can conclude as in the Step 2. Moreover, since $f(a_3) \geq f(a_{k+1})$ we have that $y_3 \leq y_{k+1}$.

Let us assume that $y_3 < y_{k+1}$. and there exists $i \in \{k+1, \dots, n\}$ such that $2m_f - f(a_1) > f(a_i)$, We consider now $\widetilde{f}$ defined as follows, $\widetilde{f}(a_k) = f(a_k) - \varepsilon$, $\widetilde{f}(a_i) = f(a_i) + \varepsilon$, and $\widetilde{f} = f$ elsewhere, where ε is small enough such that $f(a_k) - \varepsilon > 2m_f - f(a_1) > f(a_i) + \varepsilon$. We observe that

$$\frac{\operatorname{Var}_2 M_{S_n} f}{\operatorname{Var}_2 f} < \frac{\operatorname{Var}_2 M_{S_n}\widetilde{f}}{\operatorname{Var}_2 \widetilde{f}}.$$

In fact,

$$\operatorname{Var}_2 M_{S_n} f = \sum_{j=2}^{s} x_j^2 + \sum_{j=s+1}^{k} \frac{y_j^2}{4} + (n-k)u^2 < \sum_{j=2}^{s} x_j^2 + \sum_{j=s+1}^{k-1} \frac{y_j^2}{4} + \frac{(y_k+\varepsilon)^2}{4} + (n-k)u^2 = \operatorname{Var}_2 M_{S_n}\widetilde{f}$$

and

$$\operatorname{Var}_2 \widetilde{f} = \sum_{j=2}^{s} x_j^2 + \sum_{j=s+1}^{k-1} y_j^2 + (y_k+\varepsilon)^2 + (y_i-\varepsilon)^2 + \sum_{j=k+1, j\neq i}^{n} y_j^2 < \sum_{j=2}^{s} x_j^2 + \sum_{j=s+1}^{n} y_j^2 = \operatorname{Var}_2 f$$

for ε small enough, since $(y_k+\varepsilon)^2 + (y_{k+1}-\varepsilon)^2 = y_k^2 + y_{k+1}^2 + 2\varepsilon(y_k - y_{k+1}) + 2\varepsilon^2 < y_k^2 + y_{k+1}^2$ given that $y_k - y_{k+1} + \varepsilon = y_3 - y_{k+1} + \varepsilon < 0$ for ε small enough. Therefore $\frac{\operatorname{Var}_2 M_{S_n} f}{\operatorname{Var}_2 f} < \frac{\operatorname{Var}_2 M_{S_n}\widetilde{f}}{\operatorname{Var}_2 \widetilde{f}}$, contradicting the fact that f is an extremizer. Then, $y_3 = y_{k+1}$ or equivalently $f(a_3) = f(a_{k+1})$, therefore f only takes three values. Now we have only two subcases left to analyse:

- *Subcase 1:* $f(a_1) + f(a_n) \geq 2m_f$. In this case $\frac{f(a_1)+f(a_n)}{2} = M_{S_n} f(a_i)$ for $i = 3, \dots, n$ and $y_3 = y_i$ for $i = 3, \dots, n$. Also, we observe that $y_3(n-2) = x_2 + nu$ and $u \geq \frac{y_3}{2}$. Then we need to prove that

$$x_2^2 + (n-2)\frac{y_3^2}{4} \leq \left(1 - \frac{n+1}{n^2}\right)(x_2^2 + (n-2)y_3^2),$$

or, equivalently,

$$\frac{n+1}{n^2} x_2^2 \leq (n-2)\left(3/4 - \frac{n+1}{n^2}\right) y_3^2.$$

Since $y_3(n-2) = x_2 + nu \geq x_2 + \frac{n}{2}y_3$ we have $y_3\left(\frac{n}{2}-2\right) \geq x_2$, therefore it is enough to prove

$$\left(\frac{n}{2}-2\right)^2 \frac{n+1}{n^2} \leq (n-2)\left(3/4 - \frac{n+1}{n^2}\right),$$

and that can be established for $n \geq 3$.

- *Subcase 2:* $f(a_1) + f(a_2) \leq 2m_f$. In this case $M_{S_n}f(a_i) = m_f$ for $i = 3, \ldots, n$. Also, we observe that $u \leq \frac{y_3}{2}$. Thus we need to prove that

$$x_2^2 + (n-2)u^2 \leq \left(1 - \frac{n+1}{n^2}\right)(x_2^2 + (n-2)y_3^2),$$

or equivalently

$$x_2^2\left(\frac{n+1}{n^2}\right) + (n-2)u^2 \leq (n-2)\left(1 - \frac{n+1}{n^2}\right)y_3^2.$$

Indeed, since $y_3(n-2) = x_2 + nu$ we have

$$y_3^2 \geq \frac{x_2^2}{(n-2)^2} + \frac{n^2}{(n-2)^2}u^2.$$

Then it is enough to prove $\frac{n+1}{n^2} \leq \frac{\left(1-\frac{n+1}{n^2}\right)}{n-2}$ and $(n-2)^2 \leq n^2\left(1 - \frac{n+1}{n^2}\right)$. Since both hold for $n \geq 3$, we conclude the result.

■

9.4.2 The p-variation of M_G.

For a finite connected graph $G = (V, E)$ with vertices $V = \{a_1, a_2, \ldots, a_n\}$ we define $d(G) =: \max\{d_G(a_i, a_j); a_i, a_j \in V\}$ and $\Omega_G := \{a_i \in V; \exists a_j \in V$ such that $d(G) = d(a_i, a_j)\}$. For all $H \subset G$ we choose a minimum degree element of H and we denote this by a_H.

Proposition 9.4.2. *Let G be a finite connected graph with n vertices, assume that* $\deg(a_{\Omega_G}) =$ *k and there exists a vertex $x \in V$ such that $d(x, a_{\Omega_G}) \geq d(x, y)$ for all $y \in V$ and there are k disjoint paths from a_{Ω_G} to x. Then*

$$\mathbf{C}_{G,p} \geq 1 - \frac{1}{n}$$

for all $p \in (0, 1]$.

Proof This result follows observing that under these hypothesis we have that

$$\mathbf{C}_{G,p} \geq \frac{\operatorname{Var}_p M_G \delta_{a_{\Omega_G}}}{\operatorname{Var}_p \delta_{a_{\Omega_G}}} = \frac{\operatorname{Var}_p M_G \delta_{a_{\Omega_G}}}{k^{1/p}} \geq 1 - \frac{1}{n}.$$

∎

In particular this results hold for trees (in that case $k = 1$), cycles (in that case $k = 2$), hypercubes Q_n (with 2^n vertices, in that case $k = n$), whenever $k = 1$, etc.

Remark 9.4.1. For all $p \in (0,1)$ we have that $\mathbf{C}_{L_n,p} > 1 - \frac{1}{n}$, here L_n is the line graph. This also happens in many other situations. Moreover, it was proved by the authors in the previous chapter that $\mathbf{C}_{S_n} = 1 - \frac{1}{n}$ for all $p \in [1/2, 1]$ and similarly $\mathbf{C}_{K_n} = 1 - \frac{1}{n}$ for all $p \geq \frac{\log 4}{\log 6}$.

Let Γ_n the family of all connected simple finite graphs with n vertices. Our previous proposition motivates the following question.

Question A: Let $p > 0$. What are the values

$$c_{n,p} = \inf_{G \in \Gamma_n} \mathbf{C}_{G,p} \quad \text{and} \quad C_{n,p} = \sup_{G \in \Gamma_n} \mathbf{C}_{G,p}?$$

Moreover, what are the extremizers? i.e what are the graphs $G \in \Gamma_n$ for which $\mathbf{C}_{G,p} = C_{n,p}$ or $\mathbf{C}_{G,p} = c_{n,p}$?

9.5 Discrete Hardy-Littlewood maximal operator

In this section we write $M := M_{\mathbb{Z}}$ and

$$A_{r,s}f(n) := \frac{1}{r+s+1} \sum_{k=-s}^{r} |f(n+k)|.$$

We write $A_{r,r} =: A_r$ and, as usual, for any function $g : \mathbb{Z} \to \mathbb{R}$ we define the derivative of g at the point n by $g'(n) := g(n+1) - g(n)$ for all $n \in \mathbb{Z}$.

The following result was proved by Madrid for $p = 1$ in [Mad17]. This proof follows a similar strategy, we include some details for completeness.

Theorem 9.5.1. *Let $p \in (0,1]$ and $f : \mathbb{Z} \to \mathbb{R}$ be a function in $\ell^p(\mathbb{Z})$. Then*

$$\operatorname{Var}_p Mf \leq \left(2 \sum_{k=0}^{\infty} \frac{2^p}{(2k+1)^p (2k+3)^p} \right)^{\frac{1}{p}} \|f\|_{\ell^p(\mathbb{Z})} =: \mathbf{C}_p \|f\|_p, \tag{9.15}$$

and the constant $\mathbf{C}_p$ is the best possible. Moreover, the equality for $p \in (\frac{1}{2}, 1]$ is attained if and only if f is a delta function.

Proof We can assume without loss of generality that $f \geq 0$. We observe that for all $n \in \mathbb{Z}$ there exists $r_n \in \mathbb{Z}$ such that $Mf(n) = A_{r_n} f(n)$ (this follows from the fact that $f \in \ell^p(\mathbb{Z})$), then we consider the sets

$$X^- = \{n \in \mathbb{Z}; Mf(n) > Mf(n+1)\} \quad \text{and} \quad X^+ = \{n \in \mathbb{Z}; Mf(n+1) > Mf(n)\}.$$

Then

$$\begin{aligned}(\mathrm{Var}_p Mf)^p &= \sum_{n\in\mathbb{Z}} |Mf(n) - Mf(n+1)|^p \qquad (9.16)\\ &\leq \sum_{n\in X^-} (A_{r_n} f(n) - A_{r_n+1} f(n+1))^P + \sum_{n\in X^+} (A_{r_{n+1}} f(n+1) - A_{r_{n+1}+1} f(n))^p.\end{aligned}$$

Observe that for all $n \in X^-$ we have that

$$\begin{aligned}(A_{r_n} f(n) - A_{r_n+1} f(n+1))^p &\leq \left| \frac{2}{(2r_n+1)(2r_n+3)} \sum_{k=n-r_n}^{n+r_n} f(k) \right|^p \qquad (9.17)\\ &\leq \frac{2^p}{(2r_n+1)^p (2r_n+3)^p} \sum_{k=n-r_n}^{n+r_n} f(k)^p.\end{aligned}$$

Then, for any $m \in \mathbb{Z}$ fixed, we find the maximal contribution of $f(m)^p$ to the right hand side of (9.17).

Case 1: If $n \geq m$.
Since $n \in X^-$. In this case we have that the contribution of $f(m)$ to the right hand side of (9.17) is 0 (if $m < n - r_n$) or $\frac{2^p}{(2r_n+1)^p(2r_n+3)^p}$ (if $n - r_n \leq m$). Thus, the contribution of $f(m)$ to $(A_{r_n} f(n) - A_{r_n+1} f(n+1))^p$ is at most

$$\frac{2^p}{(2r_n+1)^p(2r_n+3)^p} \leq \frac{2^p}{(2(n-m)+1)^p(2(n-m)+3)^p}.$$

Here the equality happen if and only if $r_n = n - m$.

Case 2: If $n < m$.
Since $n \in X^-$. In this case we have that the contribution of $f(m)^p$ to the right hand side of (9.17) is 0 (if $m > n + r_n$) or $\frac{2^p}{(2r_n+1)^p(2r_n+3)^p}$ (if $n + r_n \geq m$). Thus, the contribution of $f(m)^p$ to the right hand side of (9.17) is at most

$$\begin{aligned}\frac{2^p}{(2r_n+1)^p(2r_n+3)^p} &\leq \frac{2^p}{(2(m-n)+1)^p(2(m-n)+3)^p}\\ &< \frac{2^p}{(2(m-n-1)+1)^p(2(m-n-1)+3)^p}.\end{aligned}$$

A similar analysis can be done with the second term of (9.16), in fact, for a fixed $n \in X^+$ we start observing that

$$\begin{aligned}(Mf(n+1)-Mf(n))^p &\leq (A_{r_{n+1}}f(n+1)-A_{r_{n+1}}f(n))^p \\ &\leq \left|\frac{2}{(2r_{n+1}+1)(2r_{n+1}+3)}\sum_{k=n+1-r_{n+1}}^{n+1+r_{n+1}} f(k)\right|^p \\ &\leq \frac{2^p}{(2r_{n+1}+1)^p(2r_{n+1}+3)^p}\sum_{k=n+1-r_{n+1}}^{n+1+r_{n+1}} f(k)^p.\end{aligned}$$

Then, if $n \geq m$, the contribution of $f(m)^p$ to the previous expression is strictly smaller than

$$\frac{2^p}{(2(n-m)+1)^p(2(n-m)+3)^p}.$$

Moreover, if $n < m$, the contribution of $f(m)^p$ is smaller than or equal to

$$\frac{2^p}{(2(m-n-1)+1)^p(2(m-n-1)+3)^p}.$$

Therefore, from (9.16) we conclude that

$$\begin{aligned}(\mathrm{Var}_p Mf)^p &\leq \left[\sum_{m=-\infty}^{n}\frac{2^p}{(2(n-m)+1)^p(2(n-m)+3)^p}\right]\|f\|_p^p \\ &\quad+\left[\sum_{m=n+1}^{+\infty}\frac{2^p}{(2(m-n-1)+1)^p(2(m-n-1)+3)^p}\right]\|f\|_p^p \\ &= 2\sum_{k=0}^{\infty}\frac{2^p}{(2k+1)^p(2k+3)^p}\|f\|_p^p.\end{aligned}$$

We can easily see that if f is a delta function then the previous inequality becomes an equality. On the other hand, for a function $f:\mathbb{Z}\to\mathbb{R}$ such that

$$(\mathrm{Var}_p Mf)^p = 2\sum_{k=0}^{\infty}\frac{2^p}{(2k+1)^p(2k+3)^p}\|f\|_p^p$$

and $f \geq 0$, we consider the set $P := \{s \in \mathbb{Z}; f(s) \neq 0\}$, thus

$$(\mathrm{Var}_p Mf)^p = 2\left(\sum_{k=0}^{\infty}\frac{2^p}{(2k+1)^p(2k+3)^p}\right)\sum_{t\in P} f(t)^p.$$

Then, given $s_1 \in P$, by the previous analysis we note that for all $n \geq s_1$ we must have that $n \in X^-$ and $r_n = n - s_1$. If we take $s_2 \in P$ the same has to be true, this implies that $s_1 = s_2$, therefore $P = \{s_1\}$ which means that f is a delta function.

■

Observe that for all $p \in (0,1]$ the following inequality holds

$$\mathrm{Var}_p f \leq 2^{1/p} \|f\|_p$$

for any function $f : \mathbb{Z} \to \mathbb{R}$. This follows from the fact that $|f(n) - f(n+1)|^p \leq |f(n)|^p + |f(n+1)|^p$ for all $n \in \mathbb{Z}$. Motivated by this trivial bound and our Theorem 9.5.1 we pose the following question:

Conjecture 9.5.1. *Let $p \in (1/2, 1]$ and $f : \mathbb{Z} \to \mathbb{R}$ be a function in $\ell^p(\mathbb{Z})$. Then*

$$\mathrm{Var}_p Mf \leq \left(\sum_{k=0}^{\infty} \frac{2^p}{(2k+1)^p (2k+3)^p} \right)^{\frac{1}{p}} \mathrm{Var}_p f. \tag{9.18}$$

In general, it would be interesting to answer the following question:
Question B: Let $p \in (0, \infty]$. What is the smallest constant C_p such that

$$\|(Mf)'\|_p = \mathrm{Var}_p Mf \leq C_p \mathrm{Var}\, f = C_p \|f'\|_p,$$

for all $f : \mathbb{Z} \to \mathbb{R}$.

We note that for $p = \infty$ we have that $C_\infty = 1$. The upper bound $C_\infty \leq 1$ trivially holds, on the other hand to see that the lower bound $C_\infty \geq 1$ holds it is enough to consider the function $f : \mathbb{Z} \to \mathbb{R}$ defined by $f(n) = \max\{10 - |n|, 0\}$. Moreover, observe that for $p \leq 1/2$ the right hand side of (9.18) is $+\infty$ for any no constant function, so the inequality (9.18) trivially holds in that case. However, this is highly not trivial for $p \in (1/2, 1]$. If true, this results would be stronger than our Theorem 9.5.1. For $p > 1$ even the analogous result to our Theorem 9.5.1 remains open.

Also, complementing our previous results, it would be interesting to answer the following question, this time regarding the uncentered Hardy-Littlewood maximal operator $\widetilde{M}$:
Question C: Let $p \in (0, \infty]$. What is the smallest constant $\widetilde{C}_p$ such that

$$\|(\widetilde{M}f)'\|_p = \mathrm{Var}_p \widetilde{M}f \leq \widetilde{C}_p \mathrm{Var}\, f = \widetilde{C}_p \|f'\|_p,$$

for all $f : \mathbb{Z} \to \mathbb{R}$?.

Our next theorem gives an answer to this question for $p = \infty$. An auxiliary tool is the following lemma.

Lemma 9.5.1. *Let $f : \mathbb{Z} \to \mathbb{R}^+$ be a function such that $\|f'\|_\infty < \infty$ and $\widetilde{M}_f \not\equiv \infty$. Then, we have $\widetilde{M}_f(n) < \infty$ for all $n \in \mathbb{Z}$.*

Proof Assume that there is $n \in \mathbb{Z}$ such that $\widetilde{M}f(n) = \infty$, then, there exists a sequence $\{r_j, s_j\}$ in $\mathbb{Z}^+ \times \mathbb{Z}^+$, with $r_j + s_j \to \infty$ such that $A_{r_j,s_j} f(n) \to \infty$ as $j \to \infty$. For any $m \in \mathbb{Z}$, defining $C = \|f'\|_\infty$ we have

$$A_{r_j,s_j} f(m) \geq A_{r_j,s_j} f(n) - C|m - n|\,, \tag{9.19}$$

therefore $\widetilde{M}_\beta f(m) = \infty$, a contradiction. ■

Theorem 9.5.2. *For all $f : \mathbb{Z} \to \mathbb{R}$ such that $\widetilde{M} f \not\equiv \infty$, we have that*

$$\|(\widetilde{M} f)'\|_\infty \leq \frac{1}{2}\|f'\|_\infty.$$

Moreover, the equality is attained if f is a delta function.

Remark 9.5.1. This theorem is a discrete analogue of the main result obtained in [ACL10] on the continuous setting, in that case the optimal constant is $2^{1/2}-1$. We use an elementary combinatorial argument to establish our result, this technique is completely independent of those in [ACL10].

Proof

We assume without loss of generality that f is nonnegative. Let $n \in \mathbb{Z}$, by Lemma 9.5.1 we have that $\widetilde{M} f(n) < \infty$, then, for all $\varepsilon > 0$ there are $r_{n,\varepsilon}, s_{n,\varepsilon} \geq 0$ such that

$$\widetilde{M} f(n) < \frac{1}{r_{n,\epsilon} + s_{n,\varepsilon} + 1} \sum_{k=-s_{n,\varepsilon}}^{r_{n,\varepsilon}} f(n+k) + \varepsilon. \tag{9.20}$$

We analyze two cases, the argument works similarly for both situations. *Case 1:* $(\widetilde{M} f)'(n) > 0$. In this case we star observing that $r_{n,\varepsilon} = 0$ for all sufficiently small ε (otherwise, from (9.20) we would obtain $\widetilde{M} f(n) \leq \widetilde{M} f(n+1)$). Then, for all sufficiently small ε we have that

$$\begin{aligned}
\widetilde{M} f(n) - \widetilde{M} f(n+1) &\leq \frac{1}{s_{n,\varepsilon}+1} \sum_{k=-s_{n,\varepsilon}}^{0} f(n+k) + \varepsilon - \frac{1}{s_{n,\varepsilon}+2} \sum_{k=-s_{n,\varepsilon}-1}^{0} f(n+1+k) \\
&\leq \left(\frac{1}{s_{n,\varepsilon}+1} - \frac{1}{s_{n,\varepsilon}+2}\right) \sum_{k=-s_{n,\varepsilon}}^{0} f(n+k) - \frac{1}{s_{n,\varepsilon}+2} f(n+1) + \varepsilon \\
&= \frac{1}{(s_{n,\varepsilon}+2)(s_{n,\varepsilon}+1)} \sum_{k=-s_{n,\varepsilon}}^{0} (f(n+k) - f(n+1)) + \varepsilon \\
&\leq \frac{1}{(s_{n,\varepsilon}+2)(s_{n,\varepsilon}+1)} \sum_{k=1}^{s_{n,\varepsilon}+1} k\|f'\|_\infty + \varepsilon \\
&= \frac{1}{(s_{n,\varepsilon}+2)(s_{n,\varepsilon}+1)} \frac{(s_{n,\varepsilon}+1)(s_{n,\varepsilon}+2)}{2} \|f'\|_\infty + \varepsilon \\
&= \frac{1}{2}\|f'\|_\infty + \varepsilon.
\end{aligned}$$

Since this holds for any arbitrary ε, sending ε to 0 we conclude that

$$\widetilde{M} f(n) - \widetilde{M} f(n+1) \leq \frac{1}{2}\|f'\|_\infty.$$

Case 2: $(\widetilde{M}f)'(n) < 0$. This case follows analogously. Since these are the only two possible cases the result follows. ■

Bibliography

[ACL10] J. M. Aldaz, L. Colzani, and J. Pérez Lázaro. Optimal bounds on the modulus of continuity of the uncentered hardy–littlewood maximal function. *J. Geom. Anal.*, 22(1):132–167, November 2010.

[APL07] J. M. Aldaz and J. Pérez Lázaro. Functions of bounded variation, the derivative of the one dimensional maximal function, and applications to inequalities. *Trans. Amer. Math. Soc.*, 359(5):2443–2461, 2007.

[BES19] S. Bortz, M. Egert, and O. Saari. Sobolev contractivity of gradient flow maximal functions, preprint, arxiv.org/abs/1910.13150, 2019.

[BGRMW21] D. Beltran, C. González-Riquelme, J. Madrid, and J. Weigt. Continuity of the gradient of the fractional maximal operator on $W^{1,1}(\mathbb{R}^d)$, preprint, to appear in *Math. Res. Lett.*, arxiv.org/abs/2102.10206, 2021.

[BL83] H. Brezis and E. Lieb. A relation between pointwise convergence of functions and convergence of functionals. *Proc. Amer. Math. Soc.*, 88(3):486–490, 1983.

[BM19] D. Beltran and J. Madrid. Endpoint Sobolev continuity of the fractional maximal function in higher dimensions. *Int. Math. Res. Not.*, 12 2019. rnz281.

[BM20] D. Beltran and J. Madrid. Regularity of the centered fractional maximal function on radial functions. *J. Funct. Anal.*, 279(8):108686, 2020.

[Car] E. Carneiro. Regularity of maximal operators: recent progress and some open problems. to appear in New trends in applied harmonic analysis, volume 2, `doi.org/10.1007/978-3-030-32353-0_3` (2019).

[CFS18] E. Carneiro, R. Finder, and M. Sousa. On the variation of maximal operators of convolution type II. *Rev. Mat. Iberoam.*, 34(2):739–766, 2018.

[CGR21] E. Carneiro and C. González-Riquelme. Gradient bounds for radial maximal functions. *Ann. Fenn. Math.*, 46(1):495–521, 2021.

[CGRM20] E. Carneiro, C. González-Riquelme, and J. Madrid. Sunrise strategy for the continuity of maximal operators, preprint, to appear in *J. Anal. Math.*, arxiv.org/abs/2008.07810, 2020.

[CH12] E. Carneiro and K. Hughes. On the endpoint regularity of discrete maximal operators. *Math. Res. Lett.*, 19(6):1245–1262, 2012.

[CM17] E. Carneiro and J. Madrid. Derivative bounds for fractional maximal functions. *Trans. Amer. Math. Soc.*, 369(6):4063–4092, 2017.

[CMP17] E. Carneiro, J. Madrid, and L. B. Pierce. Endpoint Sobolev and BV continuity for maximal operators. *J. Funct. Anal.*, 273(10):3262–3294, 2017.

[CS07] L. Caffarelli and L. Silvestre. An extension problem related to the fractional laplacian. *Commun. Partial. Differ. Equ.*, 32(8):1245–1260, 2007.

[CS13] E. Carneiro and B. F. Svaiter. On the variation of maximal operators of convolution type. *J. Funct. Anal.*, 265(5):837–865, 2013.

[DX13] F. Dai and Y. Xu. *Approximation Theory and Harmonic Analysis on Spheres and Balls*. Springer New York, 2013.

[EG92] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[GR20] C. González-Riquelme. Sobolev regularity of polar fractional maximal functions. *Nonlinear Anal.*, 198:111889, 2020.

[GR21a] C. González-Riquelme. Continuity for the one-dimensional centered hardy-littlewood maximal operator at the derivative level, preprint, arxiv.org/abs/2109.09691, 2021.

[GR21b] C. González-Riquelme. On the continuity of maximal operators of convolution type at the derivative level, preprint, to appear in *Israel J. Math*, arxiv.org/abs/2105.15020, 2021.

[GRK21] C. González-Riquelme and D. Kosz. BV continuity for the uncentered Hardy–Littlewood maximal operator. *J. Funct. Anal.*, 281(2):109037, 2021.

[GRM21] C. González-Riquelme and J. Madrid. Sharp inequalities for maximal operators on finite graphs. *J. Geom. Anal.*, 31(10):9708–9744, 2021.

[GRM22] C. González-Riquelme and J. Madrid. Sharp inequalities for maximal operators on finite graphs, ii. *J. Math. Anal. Appl.*, 506(2):125647, 2022.

[GT01] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer Berlin Heidelberg, 2001.

[HM10] P. Hajlasz and J. Malý. On approximate differentiability of the maximal function. *Proc. Amer. Math. Soc.*, 138(1):165–174, 2010.

[HO04] P. Hajlasz and J. Onninen. On boundedness of maximal functions in Sobolev spaces. *Ann. Acad. Sci. Fenn. Math.*, 29(1):167–176, 2004.

[Kin97] J. Kinnunen. The Hardy-Littlewood maximal function of a Sobolev function. *Israel J. Math.*, 100:117–124, 1997.

[KS03] J. Kinnunen and E. Saksman. Regularity of the fractional maximal function. *Bull. London Math. Soc.*, 35(4):529–535, 2003.

[Kur15] O. Kurka. On the variation of the Hardy-Littlewood maximal function. *Ann. Acad. Sci. Fenn. Math.*, 40(1):109–133, 2015.

[LM19] H. Luiro and J. Madrid. The variation of the fractional maximal function of a radial function. *Int. Math. Res. Not. IMRN*, (17):5284–5298, 2019.

[Lui07] H. Luiro. Continuity of the maximal operator in Sobolev spaces. *Proc. Amer. Math. Soc.*, 135(1):243–251, 2007.

[Lui18] H. Luiro. The variation of the maximal function of a radial function. *Ark. Mat.*, 56(1):147–161, 2018.

[LX20] F. Liu and Q. Xue. On the variation of the Hardy–Littlewood maximal functions on finite graphs. *Collect. Math.*, 2020.

[Mad17] J. Madrid. Sharp inequalities for the variation of the discrete maximal function. *Bull. Aust. Math. Soc.*, 95(1):94–107, 2017.

[Mad19] J. Madrid. Endpoint Sobolev and BV continuity for maximal operators, II. *Rev. Mat. Iberoam.*, 35(7):2151–2168, 2019.

[Mel03] A. D. Melas. The best constant for the centered Hardy-Littlewood maximal inequality. *Ann. of Math.*, 157(2):647–688, 2003.

[MM79] M. Marcus and V. J. Mizel. Every superposition operator mapping one sobolev space into another is continuous. *J. Funct. Anal.*, 33(2):217–229, 1979.

[MSW02] A. Magyar, E. M. Stein, and S. Wainger. Discrete analogues in harmonic analysis: Spherical averages. *Ann. of Math.*, 155(1):189–208, 2002.

[PPSS18] C. Pérez, T. Picon, O. Saari, and M. Sousa. Regularity of maximal functions on Hardy-Sobolev spaces. *Bull. Lond. Math. Soc.*, 50(6):1007–1015, 2018.

[Ram19] J. P. G. Ramos. Sharp total variation results for maximal functions. *Ann. Acad. Sci. Fenn. Math.*, 44(1):41–64, 2019.

[Roc70] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series. Princeton University Press, Princeton, N. J., 1970.

[ST16] J. Soria and P. Tradacete. Best constants for the Hardy–Littlewood maximal operator on finite graphs. *J. Math. Anal. Appl*, 436(2):661 – 682, 2016.

[Ste70] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions (PMS-30)*. Princeton University Press, 1970.

[SW72] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces (PMS-32)*. Princeton University Press, December 1972.

[SY94] R. Schoen and S.-T. Yau. *Lectures on differential geometry*. Conference Proceedings and Lecture Notes in Geometry and Topology, I. International Press, Cambridge, MA, 1994. Lecture notes prepared by Wei Yue Ding, Kung Ching Chang [Gong Qing Zhang], Jia Qing Zhong and Yi Chao Xu, Translated from the Chinese by Ding and S. Y. Cheng, With a preface translated from the Chinese by Kaising Tso.

[Tan02] H. Tanaka. A remark on the derivative of the one-dimensional Hardy-Littlewood maximal function. *Bull. Austral. Math. Soc.*, 65(2):253–258, 2002.

[Tem13] F. Temur. On regularity of the discrete Hardy-Littlewood maximal function, preprint, arxiv.org/abs/1303.3993, 2013.

[Wei21a] J. Weigt. Endpoint Sobolev bounds for fractional Hardy-Littlewood maximal operators, preprint, arxiv.org/abs/2010.05561, 2021.

[Wei21b] J. Weigt. The variation of the uncentered maximal operator with respect to cubes, preprint, arxiv.org/abs/2109.10747, 2021.