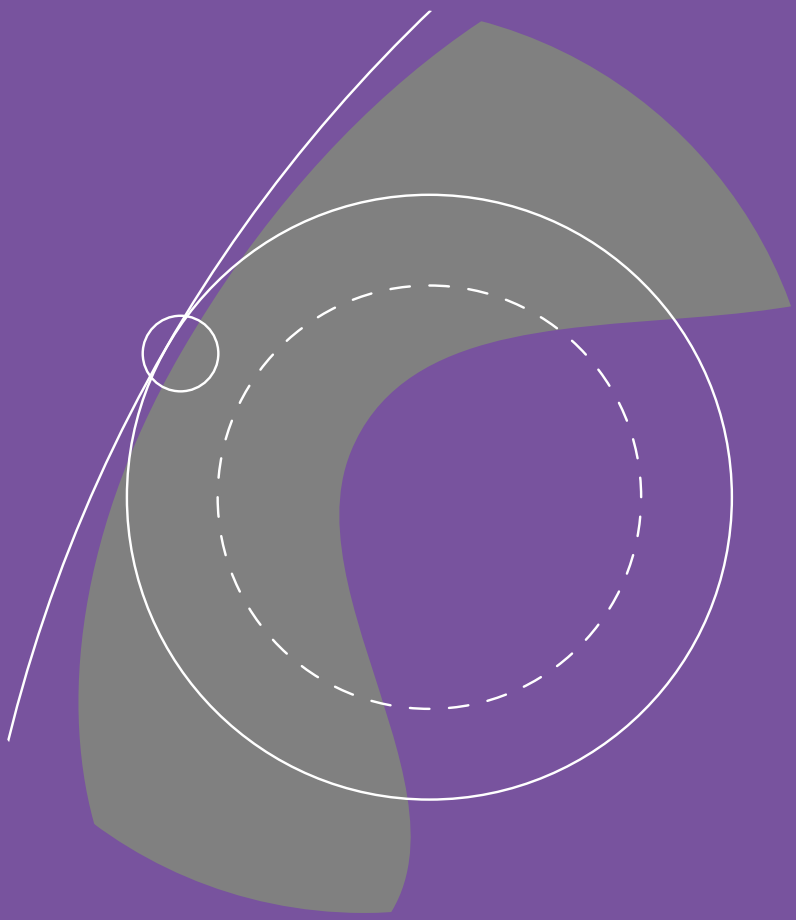
Department of Mathematics and Systems Analysis

Endpoint regularity of maximal functions in higher dimensions

Julian Weigt



Aalto University

DOCTORAL THESES

Aalto University publication series
DOCTORAL THESES 134/2022

Endpoint regularity of maximal functions in higher dimensions

Julian Weigt

A doctoral thesis completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall Otakaari 1, U147 U5 of the school on 21 October 2022 at 12.

Aalto University
School of Science
Department of Mathematics and Systems Analysis
Nonlinear Partial Differential Equations research group

Supervising professor
Professor Juha Kinnunen, Aalto University, Finland

Thesis advisor
Professor Juha Kinnunen, Aalto University, Finland

Preliminary examiners
Professor Jesús Munárriz Aldaz, Universidad Autónoma de Madrid, Spain
Professor David Cruz-Uribe, University of Alabama, United States of America

Opponent
Professor Emanuel Carneiro, Abdus Salam International Centre for Theoretical Physics

Aalto University publication series
DOCTORAL THESES 134/2022

ISBN 978-952-64-0948-1 (printed)
ISBN 978-952-64-0949-8 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-64-0949-8

Unigrafia Oy
Helsinki 2022

Finland

NORDIC SWAN ECOLABEL
Printed matter
4041-0619

Aalto University

Abstract

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Author
Julian Weigt

Name of the doctoral thesis
Endpoint regularity of maximal functions in higher dimensions

Publisher School of Science

Unit Department of Mathematics and Systems Analysis

Series Aalto University publication series DOCTORAL THESES 134/2022

Field of research Mathematical Analysis

Manuscript submitted 13 April 2022 **Date of the defence** 21 October 2022

Permission for public defence granted (date) 17 August 2022 **Language** English

☐ **Monograph** ☒ **Article thesis** ☐ **Essay thesis**

Abstract

It is well known that the Hardy-Littlewood maximal operator is bounded on Lebesgue spaces if the exponent is strictly larger than one, and that this bound fails when the Lebesgue exponent is equal to one. Similarly, the gradient of the Hardy-Littlewood maximal function is bounded by the gradient of the function when the Lebesgue exponent is strictly larger than one, but it has been an open question whether this also holds when the Lebesgue exponent equals one. This endpoint regularity bound has been conjectured to hold, but only proven fully in one imension, using a simple formula for the variation of semi-continuous functions on the real line. In higher-dimensional Euclidean spaces the bound has been proven for the maximal function of radial functions, where again one-dimensional considerations suffice. The only fully known endpoint regularity bounds in higher dimensions concern some fractional maximal operators, which however are not of the same form as in the aforementioned conjecture.

In this thesis we present the first proof of the endpoint boundedness of the gradient of a maximal operator in all dimensions. In the first two papers we prove the endpoint regularity of the uncentered Hardy-Littlewood maximal function of characteristic functions and of the dyadic maximal function of any function. We then generalize and combine the insights which we gained in order to prove further endpoint regularity bounds: We prove the corresponding endpoint bound for the gradient of the centered and of the uncentered fractional Hardy-Littlewood maximal function, and we eventually also prove their endpoint continuity. We conclude this thesis by showing a proof for the endpoint regularity bound for the cube maximal function, answering the long-standing endpoint regularity question for an uncentered maximal operator when averaging over cubes instead of balls. Our results also hold for the local versions of the above maximal operators, excluding fractional maximal operators.

The starting point in our proofs is to view the variation of a function in terms of the coarea formula. We then prove and apply higher-dimensional geometric tools which involve the interplay between volume and perimeter such as the relative isoperimetric inequality, covering lemmas that concern the boundary of a set, dyadic decompositions of functions, and approximation arguments in Sobolev spaces.

Keywords maximal function, variation

ISBN (printed) 978-952-64-0948-1 **ISBN (pdf)** 978-952-64-0949-8

ISSN (printed) 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2022

Pages 172 **urn** http://urn.fi/URN:ISBN:978-952-64-0949-8

Preface

This research project began in autumn 2018 when I started my doctoral studies at Aalto University under the supervision of Juha Kinnunen. First and foremost I want to thank him for his continuous support, both mathematically and otherwise. I am very grateful to have had him as my supervisor.

Olli Saari and Pavel Zorin-Kranich introduced me to the topic of regularity of maximal functions in their homonymous seminar at the University of Bonn during the Winter term 2017/2018. I would like to thank them, because they sparked my initial motivation to study maximal functions.

Next, I want to thank my coauthors David Beltran, Cristian González-Riquelme and José Madrid, and furthermore Constantin Bilz, João Pedro Ramos and Olli Saari whose joint publications do not appear in my narrowly scoped thesis but are closely related.

I thank Jesús Munárriz Aldaz and David Cruz-Uribe for their preliminary examination of my thesis and their helpful comments, and Nidia Obscura Acosta and David Radnell for help with grammar and wording. Thank you to Emanuel Carneiro for agreeing to act as the opponent during my public defence.

I owe thanks to the Vilho, Yrjö and Kalle Väisälä Foundation, the Magnus Ehrnroot Foundation and to the Department of Mathematics and Systems Analysis at Aalto University for their financial support. Moreover, I thank the department for providing a supportive and comfortable working environment.

Finally, I thank Nidia, my family and my friends for their care and support.

Helsinki, September 15, 2022,

Julian Weigt

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I Julian Weigt. Variation of the maximal uncentered characteristic function. Accepted for publication in *Revista Matemática Iberoamericana*, 27 pages, November 2021.

II Julian Weigt. Variation of the dyadic maximal function. Accepted for publication in *International Mathematical Research Notices*, 21 pages, March 2022.

III Julian Weigt. Endpoint sobolev bounds for the uncentered fractional maximal function. Accepted for publication in *Mathematische Zeitschrift*, 21 pages, February 2022.

IV David Beltran, Cristian González-Riquelme, José Madrid and Julian Weigt. Continuity of the gradient of the fractional maximal operator on $W^{1,1}(\mathbb{R}^d)$. Accepted for publication in *Mathematical Research Letters*, 12 pages, November 2021.

V Julian Weigt. The variation of the uncentered maximal operator with respect to cubes. Submitted to *a journal*, September 2021.

Author's Contribution

Publication I: "Variation of the maximal uncentered characteristic function"

The article results from the author's independent research.

Publication II: "Variation of the dyadic maximal function"

The article results from the author's independent research.

Publication III: "Endpoint sobolev bounds for the uncentered fractional maximal function"

The article results from the author's independent research.

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

All authors contributed equally to all parts of the publication.

Publication V: "The variation of the uncentered maximal operator with respect to cubes"

The article results from the author's independent research.

List of Figures

1. Background

For a locally integrable function f on $\mathbb{R}^d$, the function given by

$$\mathrm{M}f(x) = \sup_{r>0} \frac{1}{\mathcal{L}(B(x,r))} \int_{B(x,r)} |f(y)| \, \mathrm{d}y \tag{1.0.1}$$

is called the *Hardy-Littlewood maximal function*. Here, $B(x,r)$ is the open ℓ^2-ball of radius r around the point x,

$$B(x,r) = \{y \in \mathbb{R}^d : (y_1 - x_1)^2 + \ldots + (y_d - x_d)^2 < r^2\},$$

and $\mathcal{L}(B(x,r))$ is its Lebesgue measure. The Hardy-Littlewood maximal operator M, which maps a function to its maximal function, is a classical operator. It is well know that for $p > 1$ it is a bounded operator on the Lebesgue space $L^p(\mathbb{R}^d)$, which consists of all functions $f : \mathbb{R}^d \to \mathbb{R}$ for which the norm

$$\|f\|_{L^p(\mathbb{R}^d)} = \Big(\int_{\mathbb{R}^d} |f(x)|^p \, \mathrm{d}x \Big)^{\frac{1}{p}}$$

is finite. This result is called the *Hardy-Littlewood maximal function theorem*. For $p = 1$ this does not hold. However, it is still a bounded operator from $L^1(\mathbb{R}^d)$ into the weak Lebesgue space $L^{1,\infty}(\mathbb{R}^d)$ consisting of all functions with finite seminorm

$$\|f\|_{L^{1,\infty}(\mathbb{R}^d)} = \sup_{\lambda>0} \lambda \mathcal{L}(\{x \in \mathbb{R}^d : |f(x)| > \lambda\}),$$

where $\mathcal{L}(A)$ denotes the Lebesgue measure or the volume of a set A. The proof of this boundedness is very direct and elementary. We observe that the superlevel set $\mathcal{L}(\{x \in \mathbb{R}^d : |f(x)| > \lambda\})$ is contained in the union of all balls $B(x,r)$ with

$$\frac{1}{\mathcal{L}(B(x,r))} \int_{B(x,r)} |f(y)| \, \mathrm{d}y > \lambda. \tag{1.0.2}$$

Recall the Vitali covering lemma, which states that every set $\mathcal{B}$ of balls with finite measure has a subset $\mathcal{S}$ of disjoint balls such that each ball $B \in \mathcal{B}$ is contained in the ball $B(x,5r)$ for some $B(x,r) \in \mathcal{S}$. We apply the Vitali covering lemma to estimate the measure of the superlevel set of the

maximal function by the measure of a union of disjoint balls, losing only a dimensional constant. If we use that each of these balls satisfies (1.0.2) and use that the balls are disjoint, we recover $\|f\|_{L^1(\mathbb{R}^d)}$ after summing up over all of them. This concludes the proof of $\|\mathrm{M}f\|_{L^{1,\infty}(\mathbb{R}^d)} \leq C\|f\|_{L^1(\mathbb{R}^d)}$. This, together with the easy observation that $\|\mathrm{M}f\|_{L^\infty(\mathbb{R}^d)} \leq \|f\|_{L^\infty(\mathbb{R}^d)}$, where

$$\|f\|_{L^\infty(\mathbb{R}^d)} = \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} |f(x)|,$$

allows us to apply the Marcinkiewicz interpolation theorem and conclude the $L^p(\mathbb{R}^d)$-boundedness of M for all $1 < p \leq \infty$.

For $p = 1$ this boundedness fails, in fact M does not even map $L^1(\mathbb{R}^d)$ into $L^1(\mathbb{R}^d)$. It is straightforward to show that for any non-zero function f its maximal function $\mathrm{M}f(x)$ decays like $\frac{c}{|x|^d}$ for large x and thus is not integrable.

The boundedness of the maximal operator on Lebesgue spaces is a classical tool. For example it plays a central role in the proof of the Lebesgue differentiation theorem, which states that given a locally integrable function f we have for almost every $x \in \mathbb{R}^d$ that

$$\lim_{r\to 0} \frac{1}{\mathcal{L}(B(x,r))} \int_{B(x,r)} f(y)\,\mathrm{d}y = f(x).$$

1.1 Regularity of the Hardy-Littlewood maximal function

The centered Hardy-Littlewood maximal function is defined with the absolute value on the right-hand side of (1.0.1) because of the way the maximal function is usually used. Indeed, when being interested in L^p-norms this is the natural definition because $|f(y)|$ is the relevant quantity, not so much $f(y)$. For investigating the regularity of the maximal function however, it appears more natural to instead define

$$\mathrm{M}f(x) = \sup_{r>0} \frac{1}{\mathcal{L}(B(x,r))} \int_{B(x,r)} f(y)\,\mathrm{d}y \tag{1.1.1}$$

without the absolute value, so this is how we define the maximal function from now on. For example, the function $x \mapsto f(x) + c$ is essentially the same as f in terms of regularity, while the maximal functions with the absolute value of those two functions can be quite different. Since we are interested in regularity we are considering the change of f, which can be very different than the change of $|f|$, and the maximal operator without the absolute value is more practical to deal with here. Note, that we can still recover the original maximal function by considering $\mathrm{M}|f|$. This way we can even conclude the same first order regularity bounds for the original maximal operator, since for the weak gradient we have $|\nabla|f|(x)| \leq |\nabla f(x)|$ for almost every $x \in \mathbb{R}^d$. The function $|f|$ can only be more regular than f,

at least when considering first order regularity. That means considering the maximal function without absolute values also yields slightly more general results.

Without the absolute values, we can see the maximal function is as the pointwise supremum of the convolutions given by

$$x \mapsto \Big(f * \frac{1_{B(0,r)}}{\mathcal{L}(B(0,r))}\Big)(x) = \frac{1}{\mathcal{L}(B(x,r))} \int_{B(x,r)} f(y)\,\mathrm{d}y$$

over all $r > 0$. It is well known that if f is differentiable then for any given r we have

$$\Big\|\nabla\Big(f * \frac{1_{B(0,r)}}{\mathcal{L}(B(0,r))}\Big)\Big\|_{L^1(\mathbb{R}^d)} \leq \|\nabla f\|_{L^1(\mathbb{R}^d)},$$

i.e. a convolution can only increase regularity. This suggests that maybe also the maximal function increases regularity. However, note that it is not possible to deduce a regularity bound for the maximal function from this alone; these convolutions for example can likewise only decrease the L^1-norm of a function, while the maximal function can increase it even to infinity.

One technical detail here is that the maximal function of a differentiable function is not necessarily differentiable everywhere. Even when f is smooth and compactly supported, its maximal function usually has corners. Instead, the appropriate setting is the setting of Sobolev spaces and functions of bounded variation, where we consider the weak gradient instead. A formal construction of these spaces and proofs of most of the fundamental tools we use in these publications can be found in the book [EG15] by Evans and Gariepy. We repeat the definitions and some important properties here. We say that a function $f \in L^p(\mathbb{R}^d)$ belongs to the Sobolev space $W^{1,p}(\mathbb{R}^d)$ if there is a vector valued function $g \in L^p(\mathbb{R}^d;\mathbb{R}^d)$ such that for every compactly supported smooth vector valued function $\varphi \in C_c^\infty(\mathbb{R}^d;\mathbb{R}^d)$ we have

$$\int f(x)\,\mathrm{div}\,\varphi(x)\,\mathrm{d}x = \int \varphi(x)\cdot g(x)\,\mathrm{d}x. \tag{1.1.2}$$

In this case we call $\nabla f = -g$ the *weak gradient* of g. The weak gradient of a function is unique up to changes on sets of measure zero. This is why for the rest of this thesis, all pointwise properties of the weak gradient are implicitly meant to hold only almost everywhere. Note that by integration by parts, if $f \in L^p(\mathbb{R}^d)$ is a continuously differentiable function then its gradient is also its weak gradient. Thus, given a function $f \in W^{1,p}(\mathbb{R}^d)$ we are interested in the weak gradient $\nabla \mathrm{M} f$ of its maximal function.

The regularity of a maximal operator was first studied in [Kin97] in 1997, where Kinnunen proved that for $p > 1$ that for any function f in the Sobolev space $W^{1,p}(\mathbb{R}^d)$ the bound

$$\|\nabla \mathrm{M} f\|_{L^p(\mathbb{R}^d)} \leq C_{d,p}\|\nabla f\|_{L^p(\mathbb{R}^d)} \tag{1.1.3}$$

holds. The proof uses finite differences. For a unit vector $e \in \mathbb{R}^d$ and $h > 0$ it follows from the sublinearity and the homogeneity of the maximal operator M that

$$\begin{aligned}\frac{\mathrm{M}f(x+he) - \mathrm{M}f(x)}{h} &\leq \frac{\mathrm{M}(f(\cdot + he) - f)(x)}{h} \\ &= \mathrm{M}\Big(\frac{f(\cdot + he) - f}{h}\Big)(x).\end{aligned}$$

Using that finite differences converge to the weak gradient in Sobolev spaces, this implies the pointwise inequality

$$|\nabla \mathrm{M}f(x)| \leq \mathrm{M}(|\nabla f|)(x).$$

Then it follows from the Hardy-Littlewood maximal function theorem for $p > 1$ that

$$\|\nabla \mathrm{M}f\|_{L^p(\mathbb{R}^d)} \leq \|\mathrm{M}(|\nabla f|)\|_{L^p(\mathbb{R}^d)} \leq C_{d,p}\|\nabla f\|_{L^p(\mathbb{R}^d)}.$$

Together with the boundedness of the maximal function on $L^p(\mathbb{R}^d)$ this implies that the Hardy-Littlewood maximal operator is bounded on the Sobolev space $W^{1,p}(\mathbb{R}^d)$. Since $\mathrm{M}f$ is not integrable for any non-zero function f, the maximal operator is not bounded on $W^{1,1}(\mathbb{R}^d)$. However it is still unknown if (1.1.3) holds for $p = 1$, i.e. if the bound

$$\|\nabla \mathrm{M}f\|_{L^1(\mathbb{R}^d)} \leq C_d\|\nabla f\|_{L^1(\mathbb{R}^d)} \tag{1.1.4}$$

holds. In 2004 Hajłasz and Onninen asked this and other related questions in their paper [HO04]. By a scaling argument one can see that (1.1.4) holds if and only if the operator $f \mapsto \nabla \mathrm{M}f$ is a bounded operator from $W^{1,1}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$.

In this endpoint $p = 1$ a generalization of the Sobolev space $W^{1,1}(\mathbb{R}^d)$ is useful. For a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, we say that $f \in \mathrm{BV}(\mathbb{R}^d)$, or that f has *bounded variation*, if instead of (1.1.2) we only have

$$\sup\Big\{\int f(x)\operatorname{div}\varphi(x)\,\mathrm{d}x : \varphi \in C_{\mathrm{c}}^\infty(\mathbb{R}^d;\mathbb{R}^d),\ \forall x \in \mathbb{R}^d\ |\varphi(x)| \leq 1\Big\} < \infty. \tag{1.1.5}$$

Every function $f \in \mathrm{BV}(\mathbb{R}^d)$ also has a weak gradient, which may however be a measure, in the sense that there exists a vector valued Borel measure μ on $\mathbb{R}^d$ such that for every $\varphi \in C_{\mathrm{c}}^\infty(\mathbb{R}^d;\mathbb{R}^d)$ we have

$$\int f(x)\operatorname{div}\varphi(x)\,\mathrm{d}x = \int \varphi(x)\,\mathrm{d}\mu(x),$$

and we write

$$\operatorname{var} f = |\mu|(\mathbb{R}^d).$$

If $f \in W^{1,1}(\mathbb{R}^d)$ then $f \in \mathrm{BV}(\mathbb{R}^d)$ and

$$\operatorname{var} f = \|\nabla f\|_{L^1(\mathbb{R}^d)}.$$

This means the bound

$$\operatorname{var} \mathrm{M} f \leq C_d \operatorname{var} f \tag{1.1.6}$$

is very similar to (1.1.4). Lahti showed in [Lah20] that (1.1.4) and (1.1.6) are actually equivalent, meaning that (1.1.4) holds for all functions $f \in W^{1,1}(\mathbb{R}^d)$ if and only if (1.1.6) holds for all functions with bounded variation. In order to show that (1.1.4) implies (1.1.6) one can approximate functions with bounded variation by Sobolev functions. In order to prove the reverse implication, it suffices to show that (1.1.6) implies that for every $f \in W^{1,1}(\mathbb{R}^d)$ we have $\mathrm{M} f \in W^{1,1}(\mathbb{R}^d)$, as then we can conclude

$$\|\nabla \mathrm{M} f\|_{L^1(\mathbb{R}^d)} = \operatorname{var} \mathrm{M} f \leq C_d \operatorname{var} f = \|\nabla f\|_{L^1(\mathbb{R}^d)}.$$

There are many maximal operators under investigation where the averages over balls of the form $B(x,r)$ are replaced by averages over other sets. Many results such as the Hardy-Littlewood maximal function theorem and the boundedness of M on $W^{1,p}(\mathbb{R}^d)$ for $p > 1$ hold for most maximal operators by essentially the same proof as the proof for the centered Hardy-Littlewood maximal operator M. There are no known counterexamples against (1.1.6) for any of the maximal operators. Positive results have so far essentially been restricted to the one-dimensional setting. The first answer to the question by Hajłasz and Onninen was not provided for the centered but for the *uncentered* or *non-centered Hardy-Littlewood maximal function*. It is defined by

$$\widetilde{\mathrm{M}} f(x) = \sup_{B \ni x} \frac{1}{\mathcal{L}(B)} \int_B f(y) \, \mathrm{d}y,$$

where the supremum is taken over all open balls B which contain the point $x \in \mathbb{R}^d$. In 2002 Tanaka proved the variation bound (1.1.4) in [Tan02] for the uncentered maximal function of a function $f : \mathbb{R} \to \mathbb{R}$. His explicit constant in the inequality is 2. In 2007 Aldaz and Pérez Lázaro reduced that constant to the optimal value 1 in [APL07]. They used that in one dimension the variation can be written as

$$\operatorname{var} f = \sup_{n \in \mathbb{N},\ a_1 < \ldots < a_n} \sum_{k=1}^{n-1} |f(a_{k+1}) - f(a_k)|. \tag{1.1.7}$$

To be precise, (1.1.7) holds with equality only for certain representatives such as the lower or upper semi-continuous representative of the function $f \in \mathrm{BV}(\mathbb{R})$, otherwise it only holds with a less-or-equal sign. In particular, it holds for the maximal function $\mathrm{M} f$ of any function $f \in \mathrm{BV}(\mathbb{R}^d)$. Because the maximal function $\mathrm{M} f$ is robust against changes on a set of measure zero, it is the same for each representative of a function $f \in \mathrm{BV}(\mathbb{R}^d)$. We can conclude that in one dimension it is equivalent to prove (1.1.4) with (1.1.7) as the definition for the variation and with (1.1.5).

Furthermore, in (1.1.7) it essentially suffices to consider points $a_1, a_2, \ldots$ such that for even indices a_{2i} is a local maximum of $\widetilde{\mathrm{M}} f$ and for odd indices

a_{2i+1} is a local minimum. As a consequence of the Lebesgue differentiation theorem we have for almost every point $a \in \mathbb{R}$ that $\widetilde{\mathrm{M}}f(a) \geq f(a)$. This holds in particular at local minima of $\widetilde{\mathrm{M}}f$. The key result in [APL07] is that for a local maximum a of $\widetilde{\mathrm{M}}f$ we have $\widetilde{\mathrm{M}}f(a) = f(a)$. They concluded

$$\begin{aligned}
\operatorname{var} \widetilde{\mathrm{M}}f &= \sup_{n\in 2\mathbb{N}+1,\ a_1<\ldots<a_n} \sum_{k=1}^{n-1} |\widetilde{\mathrm{M}}f(a_{k+1}) - \widetilde{\mathrm{M}}f(a_k)| \\
&= \sup_{n\in 2\mathbb{N}+1,\ a_1<\ldots<a_n} \sum_{i=1}^{(n-1)/2} 2\widetilde{\mathrm{M}}f(a_{2i}) - \widetilde{\mathrm{M}}f(a_{2i-1}) - \widetilde{\mathrm{M}}f(a_{2i+1}) \\
&\leq \sup_{n\in 2\mathbb{N}+1,\ a_1<\ldots<a_n} \sum_{i=1}^{(n-1)/2} 2f(a_{2i}) - f(a_{2i-1}) - f(a_{2i+1}) \\
&\leq \operatorname{var} f.
\end{aligned}$$

This argument fails in dimensions $d > 1$. While it is still true that at local maxima a we essentially have $\widetilde{\mathrm{M}}f(a) = f(a)$, in higher dimensions this does not imply much about the variation of the maximal function. It is also known that the constant in (1.1.6) has to be strictly larger than one if $d > 1$. This argument fails for the centered maximal function already in one dimension. It is not true that at a local maximum a of $\mathrm{M}f$ we have $\mathrm{M}f(a) = f(a)$. Nevertheless, Kurka succeeded in 2015 to prove (1.1.6) for the centered maximal operator in one dimension in the very involved paper [Kur15]. He did case distinctions with respect to the shape of triples $a_0 < a_1 < a_2$ with $\mathrm{M}f(a_0) < \mathrm{M}f(a_1) > \mathrm{M}f(a_2)$ and a decomposition into scales. Kurka's constant is much larger than the conjectured constant 1. Together with Constantin Bilz, we managed to prove (1.1.6) for the centered Hardy-Littlewood maximal operator in $d = 1$ with constant 1 in [BW21], however only for characteristic functions.

There are some initial results regarding the higher-dimensional case. In 2017 Luiro proved (1.1.6) in [Lui18] for the uncentered maximal operator, radial functions in $W^{1,1}(\mathbb{R}^d)$, and any dimension d. Note that the maximal function of a radial function is again radial, and that radial functions are also one-dimensional in the sense that they depend only on one one-dimensional parameter. Already in 2009 Aldaz and Pérez Lázaro proved (1.1.6) in [APL09] for block decreasing functions, which are similar to radially decreasing functions.

There is not much hope for interesting results on the second derivative of the maximal function because even for compactly supported smooth functions its maximal function usually has corners, i.e., its second derivative blows up. This has nothing to do with the absolute value in the original definition. The same happens with non-negative functions.

1.2 Regularity of the fractional maximal function

For $0 \leq \alpha \leq d$ the centered *fractional* Hardy-Littlewood maximal function is defined by

$$\mathrm{M}_\alpha f(x) = \sup_{r>0} \frac{r^\alpha}{\mathcal{L}(B(x,r))} \int_{B(x,r)} |f(x)| \, \mathrm{d}x.$$

Note, that for $\alpha = 0$ we recover the original Hardy Littlewood maximal function. In the following we assume $\alpha > 0$. Here it may be more natural to keep the absolute values in the definition, or equivalently only consider non-negative functions. Also note, that if we define the fractional maximal function without absolute values we have $\mathrm{M}_\alpha f(x) \geq 0$ for almost every $x \in \mathbb{R}^d$ by the Lebesgue differentiation theorem. Shifting a function f to $x \mapsto f(x) + c$ does not commute with the maximal operator, in fact it even makes it blow up. If f is non-negative and $c > 0$ then the maximal function $\mathrm{M}_\alpha(f + c)(x)$ is infinite in any point $x \in \mathbb{R}^d$.

In much of the rest of the thesis we will repeatedly use the notation

$$a \lesssim_{c_1, c_2, \ldots} b.$$

By this we mean that there is a constant C depending only on $c_1, c_2, \ldots$ such that $a \leq Cb$ holds for any a, b.

The fractional maximal operator also exhibits a Hardy-Littlewood maximal function theorem, albeit with a different Lebesgue exponent for the target space than for the domain. For $p \leq d/\alpha$ the bound

$$\|\mathrm{M}_\alpha f\|_{L^{pd/(d-\alpha p)}(\mathbb{R}^d)} \lesssim_{d,\alpha,p} \|f\|_{L^p(\mathbb{R}^d)}$$

holds if and only if $p > 1$. The corresponding gradient bound

$$\|\nabla \mathrm{M}_\alpha f\|_{L^{pd/(d-\alpha p)}(\mathbb{R}^d)} \lesssim_{d,\alpha,p} \|\nabla f\|_{L^p(\mathbb{R}^d)} \tag{1.2.1}$$

for $p > 1$ follows from the same proof as for the original Hardy-Littlewood maximal function. Note that $pd/(d - \alpha p) > p$ for $\alpha > 0$, which means that the fractional maximal operator increases the regularity of a function. One way to interpret this is to say that the factor r^α discourages balls with a small radius, and averaging over large balls has a stronger smoothing effect than averaging over small balls. Many results that are known for the Hardy-Littlewood maximal operator have also been proven for the fractional maximal operator. For the uncentered fractional maximal function Carneiro and Madrid proved (1.2.3) for $d = 1$ in [CM17], and Luiro proved (1.2.3) for radial functions in [LM19].

Interestingly, the factor r^α turned out to make the endpoint gradient question (1.2.1) for $p = 1$ easier to investigate, and there has been more progress on the fractional maximal operator than on other maximal operators. In particular, results have already been proven for dimensions larger

than one. Kinnunen and Saksman proved in [KS03] that for all $\alpha \geq 1$ we have

$$|\nabla \mathrm{M}_\alpha f(x)| \leq (d-\alpha)\mathrm{M}_{\alpha-1} f(x) \tag{1.2.2}$$

for almost every $x \in \mathbb{R}^d$. Their proof considers finite differences. Because for almost every $x \in \mathbb{R}^d$ the radii in the supremum of the definition of the fractional maximal operators will not tend to 0 or to ∞, by compactness there is an *optimal radius* r such that

$$\mathrm{M}_\alpha f(x) = \frac{r^\alpha}{\mathcal{L}(B(x,r))} \int_{B(x,r)} |f(y)| \,\mathrm{d}y.$$

Thus for every $y \in \mathbb{R}^d$ we have

$$\begin{aligned}
&\mathrm{M}_\alpha f(x) - \mathrm{M}_\alpha f(y) \\
&\quad \leq \frac{r^\alpha}{\mathcal{L}(B(x,r))} \int_{B(x,r)} |f(z)| \,\mathrm{d}z \\
&\qquad - \frac{(r+|x-y|)^\alpha}{\mathcal{L}(B(y,r+|x-y|))} \int_{B(y,r+|x-y|)} |f(z)| \,\mathrm{d}z \\
&\quad \leq \frac{r^\alpha}{\mathcal{L}(B(x,r))} \int_{B(y,r+|x-y|)} |f(z)| \,\mathrm{d}z \\
&\qquad - \frac{(r+|x-y|)^\alpha}{\mathcal{L}(B(y,r+|x-y|))} \int_{B(y,r+|x-y|)} |f(z)| \,\mathrm{d}z \\
&\quad = \Big(\frac{r^\alpha}{\mathcal{L}(B(x,r))} - \frac{(r+|x-y|)^\alpha}{\mathcal{L}(B(y,r+|x-y|))} \Big) \int_{B(y,r+|x-y|)} |f(z)| \,\mathrm{d}z
\end{aligned}$$

For $|x-y| \to 0$ we have

$$\frac{1}{|x-y|} \Big(\frac{r^\alpha}{\mathcal{L}(B(x,r))} - \frac{(r+|x-y|)^\alpha}{\mathcal{L}(B(y,r+|x-y|))} \Big) \to (d-\alpha) \frac{r^{\alpha-1}}{\mathcal{L}(B(x,r))}.$$

Thus, in the limit $|x-y| \to 0$ we obtain

$$\frac{\mathrm{M}f(x) - \mathrm{M}f(y)}{|x-y|} \leq (d-\alpha) \frac{r^{\alpha-1}}{\mathcal{L}(B(x,r))} \int_{B(y,r)} |f(z)| \,\mathrm{d}z \leq (d-\alpha) \mathrm{M}_{\alpha-1} f(x).$$

We can conclude (1.2.2). In [CM17] Carneiro and Madrid used the Hardy-Littlewood maximal function theorem for $\mathrm{M}_{\alpha-1}$ and Sobolev embedding to show

$$\begin{aligned}
\|\nabla \mathrm{M}_\alpha f\|_{L^{d/(d-\alpha)}(\mathbb{R}^d)} &\lesssim_d \|\mathrm{M}_{\alpha-1} f\|_{L^{d/(d-\alpha)}(\mathbb{R}^d)} \\
&\lesssim_{d,\alpha} \|f\|_{L^{d/(d-1)}(\mathbb{R}^d)} \\
&\lesssim_d \|\nabla f\|_{L^1(\mathbb{R}^d)}.
\end{aligned}$$

This proves the endpoint case $p = 1$ of (1.2.1),

$$\|\nabla \mathrm{M}_\alpha f\|_{L^{d/(d-\alpha)}(\mathbb{R}^d)} \lesssim_d \|\nabla f\|_{L^1(\mathbb{R}^d)}, \tag{1.2.3}$$

however only for $\alpha \geq 1$. For $\alpha \leq 1$ this strategy breaks down because for $\beta < 0$ the fractional maximal function $\mathrm{M}_\beta f$ is not well defined. Some endpoint results for the range $0 < \alpha < 1$ have been obtained in [BRS19]. It is noteworthy that their paper is based on Fourier methods; to the best of my knowledge this the only time where Fourier methods have been useful to obtain endpoint regularity questions of maximal operators.

All the arguments above also work for the uncentered fractional maximal function, where the supremum is taken over all balls that contain x. What is more noteworthy though is that they do work for the centered fractional maximal function. Furthermore, in [BM20] Beltran and Madrid transferred Luiro's result in [LM19] that the uncentered fractional maximal operator satisfies (1.2.3) to the centered fractional function. Compare this to the original Hardy-Littlewood maximal function, where the progress in the uncentered case is significantly greater than in the centered case. All the publications in this thesis are no exception to that trend that for the Hardy-Littlewood maximal operator the centered case is much harder to investigate than the centered case, while for the fractional maximal operator it is not.

1.3 Continuity and other related topics

A slightly more general maximal operator is the *local* maximal operator. It is defined for a domain $\Omega \subset \mathbb{R}^d$ and on functions defined on this domain. Local maximal operators are defined the same way as global maximal operators, except that they only consider averages over balls or cubes which are contained in Ω. Local maximal operators exhibit the same L^p-boundedness, but there can be differences in terms of regularity. The centered and the uncentered local Hardy-Littlewood maximal operator are also believed to satisfy the endpoint regularity bound (1.1.6). This is already known for $p > 1$ due to [KL98] and for $p = 1$ in one dimension due to [APL08]. However, counterexamples show that local fractional maximal operators usually do not satisfy (1.2.3). Positive results and such counterexamples on the regularity of local fractional maximal operators can be found in [HKKT15, RSW20].

Another more general question is if the maximal operator is a continuous operator, that means if $\mathrm{M}f$ converges to $\mathrm{M}g$ if f converges to g. If a maximal operator is continuous in $f = 0$ then it is also bounded. Vice versa, as a map $L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ for $p > 1$, continuity is a direct consequence of the sublinearity of the maximal operator

$$\mathrm{M}(f + g)(x) \leq \mathrm{M}f(x) + \mathrm{M}g(x),$$

and of its $L^p(\mathbb{R}^d)$-boundedness. Similarly, it can be asked if the maximal operator is a continuous operator on the Sobolev space $W^{1,p}(\mathbb{R}^d)$ for $p > 1$.

This question was asked by Iwaniec and first formulated in [HO04]. The above proof does not work on Sobolev spaces because sublinearity does not hold on the gradient level. It not only fails pointwise but even after taking the norm; Example 5.2 in PII provides functions f, g with

$$\operatorname{var} \mathrm{M}^{\mathrm{d}}(f+g) > \operatorname{var} \mathrm{M}^{\mathrm{d}} f + \operatorname{var} \mathrm{M}^{\mathrm{d}} g.$$

In 2007 Luiro nevertheless proved in [Lui07] the continuity of the maximal operator M on $W^{1,p}(\mathbb{R}^d)$ for $p > 1$. A very useful fact to work on continuity is that for almost every $x \in \mathbb{R}^d$ there is a radius $r > 0$ such that

$$\mathrm{M}f(x) = \frac{1}{\mathcal{L}(B)(x,r)} \int_{B(x,r)} f(y)\,\mathrm{d}y$$

or we have $\mathrm{M}f(x) = f(x)$. This follows from the observation that for functions $f \in L^1(\mathbb{R}^d)$ the supremum in the definition of the maximal function is not reached by $r \to \infty$ for any $x \in \mathbb{R}^d$, and if it is approached for $r \to 0$ we can use the Lebesgue differentiation theorem. For a given $x \in \mathbb{R}^d$ we call the corresponding ball $B(x,r)$ the *optimal ball*, or the ball *used by* the maximal operator. As a key tool Luiro showed that for the optimal ball the gradient can be taken inside the integral. In [Lui07] he established the formula

$$\nabla \mathrm{M}f(x) = \frac{1}{\mathcal{L}(B)(x,r)} \int_{B(x,r)} \nabla f(y)\,\mathrm{d}y$$

for the gradient of the maximal function, where $B(x,r)$ is the optimal ball for x. The same is true for the uncentered maximal operator $\widetilde{\mathrm{M}}$. In [Lui10] Luiro extended the continuity to the local maximal operator on $W^{1,p}(\Omega)$.

The corresponding problem in the endpoint $p = 1$ is still open.. Because $\mathrm{M}f \notin L^1(\mathbb{R}^d)$, in this case the question is if the map $f \mapsto \nabla \mathrm{M}f$ is continuous as a map from $W^{1,1}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$. In [CMP17] Carneiro, Madrid and Pierce proved the endpoint continuity for the uncentered maximal operator $\widetilde{\mathrm{M}}$ for $d = 1$ and $W^{1,1}(\mathbb{R})$ and in [GRK21] González-Riquelme and Kosz extended this to $\mathrm{BV}(\mathbb{R})$. In [CGRM22] Carneiro, González-Riquelme and Madrid proved the continuity for radial functions in any dimension. There are no continuity results in the endpoint case $p = 1$ for the centered maximal operator M yet. For the fractional maximal operator, Beltran and Madrid proved the endpoint continuity in [BM21, Mad19] for the uncentered fractional maximal operator in the cases where the boundedness is known.

More regularity results are available for related maximal operators. Instead of averages over certain sets, convolution maximal operators take the supremum over averages against certain kernels. They generalize the maximal operators considered so far, because sets can be represented by their characteristic functions. The study of the regularity of convolution maximal operators has been focused on smooth kernels associated to partial differential equations, [CS13, CFS18, CGR21]. Discrete maximal functions take supreme over a countable set of averages. Regularity

results on discrete maximal operators which average over elements of a partition of unity can be found in [AK10, LN14]. Other discrete maximal functions have $\mathbb{Z}^d$ as a domain instead of $\mathbb{R}^d$ and consider averages over balls or cubes of functions on $\mathbb{Z}^d$. Some results that are known for maximal operators on $\mathbb{R}^d$ have also been proven for discrete maximal operators on $\mathbb{Z}^d$, for example [BCHP12, Mad17, LW19], or on finite graphs, [GRM21]. The discrete and the continuous setting are not the same in every aspect. There are regularity results which have only been proven for the discrete maximal operator on $\mathbb{Z}^d$, [CH12, Tem19]. On the other hand, in dimension $d > 1$ balls in $\mathbb{Z}^d$ are much more complicated objects than balls in $\mathbb{R}^d$. Also multilinear maximal operators of different kinds have been investigated for their regularity properties, for example in [CM08, LW15], as well as maximal operators on different spaces, such as in the metric setting [Buc99, KT07], on Lizorkin-Triebel spaces [Kor02], and Hardy-Sobolev spaces [PPSS18]. Local regularity properties of the maximal function, which are weaker than (1.1.6) have also been studied [HM10, ACPL12]. Carneiro wrote a survey [Car19] with more details on the history of the regularity of maximal operators.

2. Strategy and results

2.1 Overview

In the publications in this thesis we prove endpoint regularity bounds for various maximal operators in higher dimensions. The publications are heavily interlinked, see Figure 2.1. In each publication we develop important tools which also enable later publications. The main results in the publications are mostly independent. The only exception is the main result of PII which is superseded by the main result of PV. The main tool developed in PII continues to be valuable: the proofs in PIII and PV crucially rely on it.

The set of dyadic cubes is

$$\mathfrak{D} = \Big\{ \big[k_1 2^n, (k_1+1)2^n\big) \times \ldots \times \big[k_d 2^n, (k_d+1)2^n\big) : n, k_1, \ldots, k_d \in \mathbb{Z} \Big\}.$$

We define the dyadic maximal function defined by

$$\mathrm{M}^{\mathrm{d}} f(x) = \sup_{Q \in \mathfrak{D},\, Q \ni x} \frac{1}{\mathcal{L}(Q)} \int_Q f(y)\,\mathrm{d}y,$$

where the supremum is taken over all dyadic cubes Q which contain x. It plays the role of the model maximal operator in these publications. In PI we prove (1.1.6) for the dyadic maximal function and the uncentered Hardy-Littlewood maximal function of a characteristic function. The variation of the maximal function is split into two parts, the *high density part* and the *low density part*. We develop tools to control each of them separately. In PII we prove (1.1.6) for the dyadic maximal operator. The tool we develop to control the high density part in PI is used the same way to bound the high density part of the variation of the dyadic maximal function in PII. The main contribution of PII is a tool to estimate the low density part of the variation of the dyadic maximal function. In PIII we prove the corresponding endpoint regularity bound (1.2.3) for the uncentered and the centered fractional maximal operator. We use the extra leg room

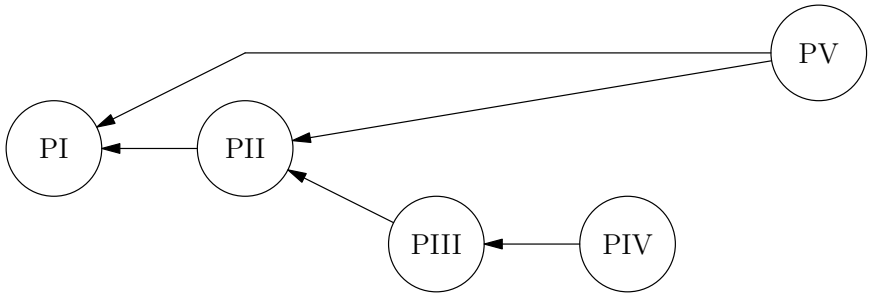


Figure 2.1. The dependency graph of the publications in this thesis.

that comes from the strictly positive fractional exponent to reduce balls to dyadic cubes. Then (1.2.3) follows after we apply the tool from PII. In PIV we prove the endpoint continuity of the gradient of the fractional maximal operator. We make use of many tools developed over the years, more specifically from [BM20, BM21, CMP17, HM10, KS03, Lui07], to reduce to a much simplified setting, and then finally use the main result from PIII to finish the proof. In PV we prove (1.1.6) for the cube maximal operator which averages over all cubes with any orientation, not only dyadic cubes. The high density part of the variation can yet again be bounded using the corresponding tool from PI. For the low density part we prove covering tools for the boundary of a union of cubes. The proofs were inspired by the tools to deal with the low density part of the variation in PI. They also hold for balls verbatim, and overall generalize the tools used for the low density part of the variation for characteristic functions in PI. They facilitate a reduction to an almost dyadic setting, in which subsequently the main tool of PII can be applied to prove the bound. This interdependence between the publications is visualized in Figure 2.1.

Some of the proofs also work for the local versions of the corresponding maximal operator. The strategy initially developed in PI for characteristic functions however relies on blowing up certain optimal balls, which is not admissible for the local maximal operator because these blowups might reach outside of the domain of Ω. Later we found a more refined approach that instead uses contractions, which are admissible for the local maximal function. Interestingly, this local approach turned out to be the one needed to deal with general functions in PV, even for the global maximal operator. The reason is that for general functions we can only rely on local information of the function. Given an optimal ball $B(x,r)$, the function f may have an arbitrarily deep drop just outside of $B(x,r)$, which means that we have no control on the average of f on $B(x,2r)$. Because we may only use local arguments anyways, our results on global maximal functions also work in the local setting. In addition, our proof strategy works vertically in local way. The arguments are invariant under shifting a function f by an additive constant $c \in \mathbb{R}$, and do not use information

of superlevel sets below the level λ we are considering at a time. For characteristic functions on the other hand we have the fixed reference heights $\lambda = 0$ and $\lambda = 1$ and all superlevel sets in between 0 and 1 are equal. The fractional maximal function is an exception, where we again have to use blowups of balls. This is expected, as the endpoint regularity bound (1.2.3) is known to fail for local fractional maximal operators as mentioned earlier. Recall that for the fractional maximal operator we consider only non-negative functions, so we do again have the global reference level $\lambda = 0$. Our arguments in the fractional setting are also not vertically local in the above sense, but usually relate to the level $\lambda = 0$.

The question of Hajłasz and Onninen was originally phrased for the centered and the uncentered Hardy-Littlewood maximal functions. These are the most well known maximal functions, and the paper [Kin97] which initiated the study of regularity of maximal functions is formulated for the centered Hardy-Littlewood maximal operator only. However, from a purely theoretical standpoint there is no clear reason to single out these particular maximal operators. Many of the proofs on the boundedness and regularity of maximal functions, in particular the one in [Kin97] for $p > 1$, work for the uncentered Hardy-Littlewood maximal operator the same way as for maximal operators averaging over other sets, for example cubes. Also our tool to bound the high density part of the variation of a maximal function works for many maximal functions including the uncentered Hardy-Littlewood maximal function. Our proof for the cube maximal operator provides a promising pathway also for the low density part of the uncentered Hardy-Littlewood maximal operator, as most of the strategy works the same way for balls as it works for cubes, if not more easily. However, a key step used to estimate the low density part of the variation remains open for balls. There are ideas how to replace this missing step but it appears to be a hard problem. A priori there was no reason to believe that it would be easier to prove (1.1.4) for the maximal operator with respect to cubes than for balls. For the centered Hardy-Littlewood maximal function the approach we use appears to fail directly from the start; new ideas are probably necessary for that maximal operator.

Some of the tools developed here may also be of interest independent of maximal operators. For example we prove an extension of the relative isoperimetric inequality and a Vitali covering lemma for the boundary. The most promising approach to prove the variation bound for the uncentered Hardy-Littlewood maximal function is to prove a decomposition of a general set according to the scale of its boundary. This decomposition could also be of independent interest because it would have to represent both the volume and the boundary of a set well at the same time.

2.2 General strategy

The base for all proofs is to view the variation of a function in terms of its superlevel sets $\{x \in \mathbb{R}^d : f(x) > \lambda\}$. We write the variation using the *coarea formula*

$$\operatorname{var} f = \int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\partial_* \{x \in \mathbb{R}^d : f(x) > \lambda\}) \, \mathrm{d}\lambda. \tag{2.2.1}$$

This is the derivative version of the layer cake formula for the Lebesgue integral

$$\int f(x) \, \mathrm{d}x = \int_0^{\infty} \mathcal{L}(\{x \in \mathbb{R}^d : f(x) > \lambda\}) \, \mathrm{d}\lambda$$

of a non-negative function $f : \mathbb{R}^d \to [0, \infty)$. Here, $\partial_* A$ is the *measure theoretic boundary* of a measurable set A, which is the set of all points which are neither density points of A nor of its complement $\mathbb{R}^d \setminus A$. For regular enough sets such as balls or cubes the topological and the measure theoretic boundary agree. Their main difference is that the measure theoretic boundary does not take into account measure zero sets. This is in accordance with the idea that Sobolev functions and their important properties, such as their variation, are not affected by changes on sets with measure zero.

We denote by $\mathcal{H}^{d-1}(S)$ the $(d-1)$-dimensional Hausdorff measure of a set S, its surface measure. We say that $\mathcal{H}^{d-1}(\partial_* A)$ is the *perimeter* of a set A. For $\lambda \in \mathbb{R}$ we use the following shorthand notation for the superlevel set

$$\{f > \lambda\} = \{x \in \mathbb{R}^d : f(x) > \lambda\}$$

and we write an integral as

$$\int f = \int f(x) \, \mathrm{d}x.$$

For a ball or cube B we denote the average integral by

$$f_B = \frac{1}{\mathcal{L}(B)} \int_B f.$$

Then the uncentered Hardy-Littlewood maximal function is given by

$$\widetilde{\mathrm{M}} f(x) = \sup_{B \ni x} f_B,$$

where the supremum is taken over all balls B which contain x. To indicate a union of a set $\mathcal{B}$ of balls we use the notation

$$\bigcup \mathcal{B} = \bigcup_{B \in \mathcal{B}} B.$$

For the uncentered maximal function we have

$$\{\widetilde{\mathrm{M}} f > \lambda\} = \bigcup \{B : B \text{ is a ball with } f_B > \lambda\},$$

which was already used in the proof of the Hardy-Littlewood maximal function theorem. In PI we realized that this simple formula for the superlevel set suggests applying the coarea formula (2.2.1) to estimate the variation of the uncentered Hardy-Littlewood maximal function. The same applies to the uncentered fractional Hardy-Littlewood maximal function $\widetilde{\mathrm{M}}_\alpha f$ because its superlevel sets can be written in a similar manner, and to the dyadic maximal function $\mathrm{M}^{\mathrm{d}} f$. That means in order to estimate the variation of uncentered maximal functions, one has to estimate the perimeter of a union of balls or cubes. Some of these estimates are purely geometric, in particular those developed in PI and in PV.

Recall that the Lebesgue differentiation theorem states that for almost every $x \in \mathbb{R}^d$ we have

$$\lim_{r \to 0} f_{B(x,r)} = f(x).$$

One can conclude that for almost every $x \in \mathbb{R}^d$ we have $\mathrm{M}f(x) \geq f(x)$, and the same is true for most maximal functions. That implies that the superlevel set of f is essentially contained in the corresponding superlevel set of $\mathrm{M}f$, which means that $\partial_* \{\widetilde{\mathrm{M}} f > \lambda\}$ cannot intersect the measure theoretic interior of $\{f > \lambda\}$. So we can split $\partial_* \{\widetilde{\mathrm{M}} f > \lambda\}$ into two parts, one part which lies away from the superlevel set of f and the other which is contained in its boundary. For a measurable set A denote by $\overline{A}^*$ its *measure theoretic closure*, the set of all points which are not density points of the complement of A. Then we have

$$\begin{aligned} \partial_* \{\widetilde{\mathrm{M}} f > \lambda\} &= \partial_* \left(\{\widetilde{\mathrm{M}} f > \lambda\} \cup \{f > \lambda\} \right) \\ &\subset \partial_* \{\widetilde{\mathrm{M}} f > \lambda\} \setminus \overline{\{f > \lambda\}}^* \cup \partial_* \{f > \lambda\}. \end{aligned}$$

This means we have

$$\begin{aligned} \operatorname{var} \widetilde{\mathrm{M}} f &= \int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\partial_* \{\widetilde{\mathrm{M}} f > \lambda\}) \, \mathrm{d}\lambda \\ &\leq \int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\partial_* \{\widetilde{\mathrm{M}} f > \lambda\} \setminus \overline{\{f > \lambda\}}^*) \, \mathrm{d}\lambda \\ &\quad + \int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\partial_* \{f > \lambda\}) \, \mathrm{d}\lambda. \end{aligned} \tag{2.2.2}$$

Since the last term equals $\operatorname{var} f$, it suffices to estimate the term on the second to last line, the part of boundary of the superlevel set of the maximal function that lies away from the superlevel set of the function.

2.3 Characteristic functions

For a set E we denote by 1_E the characteristic function of E, i.e. the function with $1_E(x) = 1$ if $x \in E$ and $1_E(x) = 0$ if $x \notin E$. In PI we prove that for every set E the uncentered Hardy-Littlewood maximal function $\widetilde{\mathrm{M}} f$

and the dyadic maximal function $\mathrm{M}^{\mathrm{d}} f$ satisfy (1.1.6) for the characteristic function $f = 1_E$ of any measurable set $E \subset \mathbb{R}^d$. In this case the superlevel sets further simplify to

$$\{\widetilde{\mathrm{M}} 1_E > \lambda\} = \bigcup \{B : \mathcal{L}(E \cap B) > \lambda \mathcal{L}(B)\}.$$

Another simplifying factor with characteristic functions is that we can write their variation as

$$\operatorname{var} 1_E = \mathcal{H}^{d-1}(\partial_* E).$$

For the dyadic maximal function of a characteristic function $\mathrm{M}^{\mathrm{d}} 1_E$ we can simplify further. Because any two dyadic cubes are either disjoint or one is contained in the other, a union of dyadic cubes $\mathcal{Q}$ with finite measure can be written as the disjoint union of all dyadic cubes $Q \in \mathcal{Q}$ such that there is no parent cube $P \in \mathcal{Q}$ with $Q \subsetneq P$. We call such dyadic cubes Q the *maximal dyadic cubes*. For a set E and $\lambda \in \mathbb{R}$ denote by $\mathcal{Q}_\lambda$ the set of maximal dyadic cubes with $\mathcal{L}(E \cap Q) > \lambda \mathcal{L}(Q)$. Then we can write the superlevel set as the disjoint union of all cubes in $\mathcal{Q}_\lambda$,

$$\{\mathrm{M}^{\mathrm{d}} 1_E > \lambda\} = \bigcup \mathcal{Q}_\lambda.$$

Recall the *relative isoperimetric inequality*

$$\min\{\mathcal{L}(B \cap E), \mathcal{L}(B \setminus E)\}^{d-1} \lesssim_d \mathcal{H}^{d-1}(B \cap \partial_* E)^d, \tag{2.3.1}$$

which not only holds for balls B, but also for cubes, and in general for John domains, see [Haj, Theorem 107]. For

$$\lambda \leq \frac{\mathcal{L}(Q \cap E)}{\mathcal{L}(Q)} \leq \frac{1}{2}$$

we can use (2.3.1) to prove

$$\begin{aligned} \mathcal{H}^{d-1}(\partial_* Q) \lesssim_d \mathcal{L}(Q)^{\frac{d-1}{d}} &= \Big(\frac{\mathcal{L}(Q)}{\mathcal{L}(Q \cap E)}\Big)^{\frac{d-1}{d}} \mathcal{L}(Q \cap E)^{\frac{d-1}{d}} \\ &\leq \lambda^{-\frac{d-1}{d}} \mathcal{L}(Q \cap E)^{\frac{d-1}{d}} \\ &\lesssim_d \lambda^{-\frac{d-1}{d}} \mathcal{H}^{d-1}(Q \cap \partial_* E). \end{aligned} \tag{2.3.2}$$

Again, this calculation also works for balls instead of cubes. However for $\lambda \geq \frac{1}{2}$, if E occupies too much of Q, the above proof and also its conclusion fail. But we can repair this by substracting $\overline{E}^*$ from $\partial_* Q$ on the left-hand side. We prove that for any ball or cube Q and any set E with $\mathcal{L}(Q \cap E) \geq \mathcal{L}(Q)/2$ we have

$$\mathcal{H}^{d-1}(\partial_* Q \setminus \overline{E}^*) \lesssim_d \mathcal{H}^{d-1}(\partial_* E \cap \mathring{Q}). \tag{2.3.3}$$

The motivation for the proof of (2.3.3) came from the proof of [KKST08, Theorem 3.1]. If we drop the condition $\mathcal{L}(Q \cap E) \geq \mathcal{L}(Q)/2$ and only assume

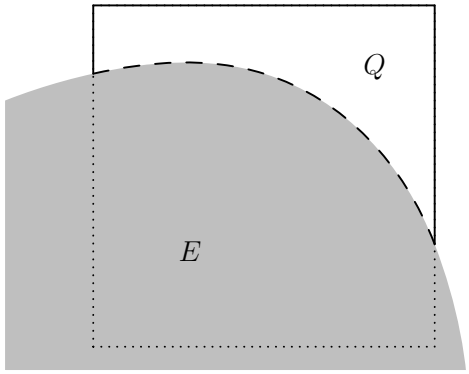


Figure 2.2. A sketch of (2.3.3). We can bound the boundary of a cube Q outside of E by the boundary of E inside Q if E occupies enough of Q.

$\mathcal{L}(Q \cap E) \geq \lambda \mathcal{L}(Q)$ then (2.3.3) fails. But, as the above calculation with the relative isoperimetric inequality shows, we can also save this if we multiply the right-hand side by the factor $\lambda^{-\frac{d-1}{d}}$. So by putting these two arguments together, we can conclude for $\mathcal{L}(Q \cap E) \geq \lambda \mathcal{L}(Q)$ that

$$\mathcal{H}^{d-1}(\partial_* Q \setminus \overline{E}^*) \lesssim_d \lambda^{-\frac{d-1}{d}} \mathcal{H}^{d-1}(\partial_* E \cap \mathring{Q}). \tag{2.3.4}$$

The case distinction between $\lambda \geq \frac{1}{2}$ and $\lambda < \frac{1}{2}$ is what we will refer to as the *high density* and the *low density* case. Although this case distinction is not visible anymore in the statement (2.3.4), it is used in its proof as described above. All the results above hold for balls as they hold for cubes.

Recall that it suffices to estimate the part of the boundary of the superlevel set of the maximal function which lies outside the measure theoretic closure of the superlevel set of the function, (2.2.2). For the dyadic maximal operator we can conclude

$$\begin{aligned}
&\int_0^1 \mathcal{H}^{d-1}(\partial_* \{\mathrm{M}^{\mathfrak{d}} 1_E > \lambda\} \setminus \overline{E}^*) \,\mathrm{d}\lambda \\
&\quad = \int_0^1 \mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{Q}_\lambda \setminus \overline{E}^*\Big) \,\mathrm{d}\lambda \\
&\quad \leq \int_0^1 \sum_{Q \in \mathcal{Q}_\lambda} \mathcal{H}^{d-1}(\partial_* Q \setminus \overline{E}^*) \,\mathrm{d}\lambda \\
&\quad \lesssim_d \int_0^1 \sum_{Q \in \mathcal{Q}} \lambda^{-\frac{d-1}{d}} \mathcal{H}^{d-1}(\partial_* E \cap \mathring{Q}) \,\mathrm{d}\lambda \\
&\quad = \int_0^1 \lambda^{-\frac{d-1}{d}} \mathcal{H}^{d-1}\Big(\partial_* E \cap \bigcup \{\mathring{Q} : Q \in \mathcal{Q}\}\Big) \,\mathrm{d}\lambda \\
&\quad \leq \int_0^1 \lambda^{-\frac{d-1}{d}} \mathcal{H}^{d-1}(\partial_* E) \,\mathrm{d}\lambda \\
&\quad \lesssim_d \mathcal{H}^{d-1}(\partial_* E) \\
&\quad = \operatorname{var} 1_E.
\end{aligned}$$

For the uncentered Hardy-Littlewood maximal operator the problem in the above proof is that the superlevel set of the uncentered Hardy-Littlewood maximal function is not a union of disjoint balls. Instead we have to bound the perimeter of a union of balls which may intersect. First we consider the high density case $\lambda \geq \frac{1}{2}$. In PI we extend (2.3.3) from a single ball to a union $\mathcal{B}$ of balls such that for every $B \in \mathcal{B}$ we have $\mathcal{L}(B \cap E) \geq \mathcal{L}(B)/2$, and obtain

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B} \setminus \overline{E}^*\Big) \lesssim_d \mathcal{H}^{d-1}\Big(\partial_* E \cap \bigcup \mathcal{B}\Big). \tag{2.3.5}$$

Taking the same proof but allowing for a variable condition $\mathcal{L}(B \cap E) \geq \lambda \mathcal{L}(B)$ and tracking the dependence on λ, we obtain

$$\mathcal{H}^{d-1}\Big(\bigcup \{B : \mathcal{L}(E \cap B) > \lambda \mathcal{L}(B)\}\Big) \lesssim_d \lambda^{-\frac{d-1}{d}}(1 - \log \lambda)\mathcal{H}^{d-1}(\partial_* E). \tag{2.3.6}$$

Using (2.3.6) we can prove $\operatorname{var} \widetilde{\mathrm{M}} 1_E \lesssim_d \operatorname{var} 1_E$ for the uncentered Hardy-Littlewood maximal function. We follow the same strategy as for the dyadic maximal function but instead of passing to single cubes we apply (2.3.6), and then use that $\lambda^{-\frac{d-1}{d}}(1 - \log \lambda)$ is integrable from 0 to 1 also with the additional factor $(1 - \log)$.

Note that for a disjoint union of cubes (or balls) (2.3.6) holds without the factor $1 - \log \lambda$ because it follows directly from (2.3.4). In fact, the factor $1 - \log \lambda$ is not necessary even if we consider general sets of balls in (2.3.6). Easy examples show that the rate $\lambda^{-\frac{d-1}{d}}$ is also the best possible rate in λ. In order to get rid of the factor $1 - \log \lambda$ we only need to consider the case $\lambda \leq \frac{1}{2}$, because for $\frac{1}{2} \leq \lambda \leq 1$ the factor $1 - \log \lambda$ is uniformly bounded. It is not surprising that our first proof does not produce the optimal rate in λ because it was initially developed for the high density case $\lambda \geq \frac{1}{2}$ only. Much of PI is devoted to developing a proof of (2.3.6) for the low density case $0 < \lambda \leq \frac{1}{2}$ without the factor $1 - \log \lambda$. And indeed, the discovered proof strategy only works in the low density case. For $0 < \lambda \leq \frac{1}{2}$ we are in the realm where the argument (2.3.2) is valid, and our proof is also based on this. The argument works both for balls and cubes. In fact, the strategy works also for uncentered maximal operators that average over more general sets, for example convex sets with bounded eccentricity.

While this improvement to obtain the optimal the rate in λ is not needed in order to prove the main result in PI, it inspired the strategy to bound the low density part in PV.

2.4 Dyadic maximal function

In PII we prove the variation bound (1.1.6) for the dyadic maximal operator M^{d} applied to any function $f : \mathbb{R}^d \to \mathbb{R}$ with bounded variation. It is not possible to just use the sublinearity of the maximal operator to extend the

variation bound from characteristic to simple and then general functions. The pitfall in that strategy is that while the maximal function is sublinear, this is not true on the gradient level: There are characteristic functions f_1, f_2 such that

$$\operatorname{var} \mathrm{M}(f_1 + f_2) > \operatorname{var} \mathrm{M} f_1 + \operatorname{var} \mathrm{M} f_2,$$

see Example 5.2 in PII. So we need to develop further tools that apply in the case of general functions.

In the previous section on characteristic functions we did a case distinction for levels $\lambda \leq \frac{1}{2}$ and $\lambda > \frac{1}{2}$. For the dyadic maximal function we also do a case distinction into a low density and a high density case, but it will not be a distinction into different levels. For a level $\lambda \in \mathbb{R}$ denote by $\mathcal{Q}_\lambda$ the set of maximal dyadic cubes Q with $f_Q > \lambda$. We sort the cubes $Q \in \mathcal{Q}_\lambda$ into those which intersect the levelset of the function f much, and those which intersect it only little. We say that a cube $Q \in \mathcal{Q}_\lambda$ is a *high density* cube if $\mathcal{L}(Q \cap \{f > \lambda\}) > 2^{-d-1}\mathcal{L}(Q)$ and a *low density* cube if $\mathcal{L}(Q \cap \{f > \lambda\}) \leq 2^{-d-1}\mathcal{L}(Q)2$. Recall that it suffices to estimate the part of the boundary of the superlevel set of the maximal function which lies outside the measure theoretic closure of the superlevel set of the function, see (2.2.2). We split it according to the low density cubes and the high density cubes,

$$\begin{aligned}
&\mathcal{H}^{d-1}(\partial_* \{\mathrm{M}^{\mathrm{d}} f > \lambda\} \setminus \overline{\{f > \lambda\}}^*) \leq \\
&\quad \mathcal{H}^{d-1}\Big(\partial_* \bigcup \{Q \in \mathcal{Q}_\lambda,\ \mathcal{L}(Q \cap \{f > \lambda\}) > 2^{-d-1}\mathcal{L}(Q)\} \setminus \overline{\{f > \lambda\}}^*\Big) \\
&\quad + \mathcal{H}^{d-1}\Big(\partial_* \bigcup \{Q \in \mathcal{Q}_\lambda,\ \mathcal{L}(Q \cap \{f > \lambda\}) \leq 2^{-d-1}\mathcal{L}(Q)\}\Big) \qquad (2.4.1)
\end{aligned}$$

We integrate (2.4.1) over $-\infty < \lambda < \infty$ and call the first summand on the right-hand side the *high density* part of the variation of $\mathrm{M}^{\mathrm{d}} f$ and the second summand the *low density* part.

For the high density part we apply the same argument as for characteristic functions. Using (2.3.4) with $\lambda = 2^{-d-1}$ and the disjointness of the cubes in $\mathcal{Q}_\lambda$, we obtain

$$\begin{aligned}
&\mathcal{H}^{d-1}\Big(\partial_* \bigcup \{Q \in \mathcal{Q}_\lambda : \mathcal{L}(Q \cap \{f > \lambda\}) > 2^{-d-1}\mathcal{L}(Q)\} \setminus \overline{\{f > \lambda\}}^*\Big) \\
&\quad \leq \sum_{\substack{Q \in \mathcal{Q}_\lambda: \\ \mathcal{L}(Q \cap \{f > \lambda\}) > 2^{-d-1}\mathcal{L}(Q)}} \mathcal{H}^{d-1}(\partial_* Q \setminus \overline{\{f > \lambda\}}^*) \\
&\quad \lesssim_d \sum_{\substack{Q \in \mathcal{Q}_\lambda: \\ \mathcal{L}(Q \cap \{f > \lambda\}) > 2^{-d-1}\mathcal{L}(Q)}} \mathcal{H}^{d-1}(\partial_* \{f > \lambda\} \cap \mathring{Q}) \\
&\quad \leq \mathcal{H}^{d-1}(\partial_* \{f > \lambda\}).
\end{aligned}$$

Integrating over λ yields the high density part of the variation of $\mathrm{M}^{\mathrm{d}} f$ on the left-hand side and $\operatorname{var} f$ on the right-hand side.

In the high density case we estimated the perimeter of the superlevel set of the maximal function at level λ by the perimeter of the superlevel set

at the function at that level. In the low density case this can also be done for characteristic functions because all its superlevel sets are the same. For general functions this is not possible, so the low density tools from PI do not suffice. The dependence of the superlevel set of the maximal function at λ on the superlevel sets of f below and above λ adds to the difficulty of the problem. In the low density case we eventually control the perimeter of the superlevel set at of the maximal function at level λ only by the superlevel sets of the function above level λ.

The new idea here is to split the graph of f into small chunks, each of which has perimeter comparable to its volume. Then we bound the contribution of each of those chunks to $\operatorname{var} \mathrm{M}^{\mathrm{d}} f$ by the variation of f within that chunk. More precisely, we decompose the mass

$$\{(x, \lambda) : \lambda \leq f(x)\}$$

of f, which is the area below the graph of f, into pieces of the form

$$Q \times I \cap \{(x, \lambda) : \lambda \leq f(x)\},$$

where Q is a dyadic cube and I is an interval. These pieces $Q \times I$ will satisfy that for every $\lambda \in I$ we have $\mathcal{L}(Q \cap \{f > \lambda\}) \leq \mathcal{L}(Q)/2$. By the relative isoperimetric inequality this implies

$$\mathcal{H}^{d-1}(\partial_* \{f > \lambda\} \cap Q)^d \gtrsim \mathcal{L}(\{f > \lambda\} \cap Q)^{d-1}.$$

The pairs (Q, I) are chosen in such a way that two distinct pieces $Q_1 \times I_1$ and $Q_2 \times I_2$ are disjoint, i.e. $(Q_1 \times I_1) \cap (Q_2 \times I_2) = \emptyset$. Furthermore, we will have for every $\lambda \in I$ that Q is a maximal cube with $f_Q > \lambda$. This will only yield a decomposition of a subset of the mass $\{(x, \lambda) : \lambda \leq f(x)\}$. However, it will contain enough of the mass to account for the variation caused by low density cubes.

Denote by $\mathfrak{D}([0,1)^d)$ the set of dyadic cubes contained in $[0,1)^d$, and by $\mathfrak{D}(Q_0)$ the set of dyadic subcubes of Q_0 is the image of $\mathfrak{D}([0,1)^d)$ under the affine map that maps $[0,1)^d$ into Q_0. That means $Q_0 \in \mathfrak{D}(Q_0)$, and for every cube $Q \in \mathfrak{D}(Q_0)$, all the 2^d disjoint subcubes with half the sidelength of Q which partition Q belong to $\mathfrak{D}(Q_0)$. The main ingredient we prove in PII is that for any $\lambda_0 \in \mathbb{R}$ and any low density cube Q_0 with

$$\mathcal{L}(\{f \geq \lambda_0\} \cap Q_0) \leq 2^{-d-1} \mathcal{L}(Q_0), \tag{2.4.2}$$

we have

$$\mathcal{L}(Q_0)(f_{Q_0} - \lambda_0) \leq 2^{d+1} \int_{f_{Q_0}}^{\infty} \sum_{Q} \mathcal{L}(\{f \geq \lambda\} \cap Q) \, \mathrm{d}\lambda, \tag{2.4.3}$$

where the sum is taken over all low density maximal maximal dyadic subcubes $Q \in \mathfrak{D}(Q_0)$ with $f_Q \geq \lambda$ and $\mathcal{L}(Q \cap \{f \geq \lambda\}) < \mathcal{L}(Q)/2$.

The idea of (2.4.3) is that a very low density cube Q_0 has lots of low density mass above it. The definition of f_Q means that within the rectangle $Q_0 \times [0,\infty)$ there is

$$\int_0^\infty \mathcal{L}(Q_0 \cap \{f > \lambda\})\,\mathrm{d}\lambda = \int_{Q_0} f = \mathcal{L}(Q_0) f_{Q_0}$$

much mass of f. The very low density condition (2.4.2) implies that for all $\lambda \geq \lambda_0$ the majority of Q_0 is not occupied by $\{f > \lambda\}$, so that

$$\begin{aligned}\int_0^{f_{Q_0}} \mathcal{L}(Q_0 \cap \{f > \lambda\})\,\mathrm{d}\lambda &\leq \int_0^{\lambda_0} \mathcal{L}(Q_0)\,\mathrm{d}\lambda + \int_{\lambda_0}^{f_{Q_0}} \mathcal{L}(Q_0 \cap \{f > \lambda\})\,\mathrm{d}\lambda \\ &\leq \lambda_0 \mathcal{L}(Q_0) + (f_{Q_0} - \lambda_0) 2^{-d-1} \mathcal{L}(Q_0) \\ &= f_{Q_0} \mathcal{L}(Q_0) - (1 - 2^{-d})(f_{Q_0} - \lambda_0)\mathcal{L}(Q_0).\end{aligned}$$

That means $(1 - 2^{-d-1})(f_{Q_0} - \lambda_0)\mathcal{L}(Q_0)$ much mass of f must lie within $Q_0 \times [f_{Q_0}, \infty)$. Formula (2.4.3) claims that a certain portion of that mass can even be found within low density cubes. For more details see Section 3.2.

For a dyadic cubes Q_0 denote by $\lambda_{Q_0} \in \mathbb{R}$ the smallest value λ for which Q_0 is a very low density cube, i.e. satisfies (2.4.2). It is straightforward to see that the low density part of the variation of the dyadic maximal function, see (2.4.1), is bounded by

$$\sum_{Q_0} (f_{Q_0} - \lambda_{Q_0}) \mathcal{H}^{d-1}(\partial\, Q_0),$$

where the sum is over all dyadic cubes $Q_0 \in \bigcup_{\lambda \in \mathbb{R}} \mathcal{Q}_\lambda$. Those are the dyadic cubes which the maximal function uses and which have $f_{Q_0} > \lambda_{Q_0}$. Observe, that

$$(f_{Q_0} - \lambda_{Q_0}) \mathcal{H}^{d-1}(\partial\, Q_0) \lesssim_d (f_{Q_0} - \lambda_{Q_0}) \frac{\mathcal{L}(Q_0)}{\mathrm{l}(Q_0)}.$$

Applying (2.4.3) we can bound the low density part of the variation by a double sum

$$\sum_{Q_0} \int_{f_{Q_0}}^\infty \sum_Q \frac{\mathcal{L}(\{f > \lambda\} \cap Q)}{\mathrm{l}(Q_0)}\,\mathrm{d}\lambda$$

over certain dyadic cubes Q_0 and $Q \subset Q_0$. The we apply Fubini's Theorem, observe that for a fixed cube Q the sum of $\mathrm{l}(Q_0)^{-1}$ over all dyadic cubes Q_0 sums to $\mathrm{l}(Q)^{-1}$, and finally apply the relative isoperimetric inequality on the low density cube Q. Because the cubes Q for a given level λ are disjoint, we then recover the variation of f on the right-hand side.

2.5 Cube maximal function

We define the cube maximal function by

$$\mathrm{M}^{\mathrm{c}} f(x) = \sup_{Q \ni x} f_Q,$$

where the supremum is taken over all cubes Q with $x \in Q$ with any orientation. Our results also hold for maximal operators averaging over other sets of cubes, for example only axes parallel cubes, i.e. cubes of the form $Q = [a_1, a_1 + l] \times \ldots \times [a_d, a_d + l]$ with $a_1, \ldots, a_d \in \mathbb{R},\ l > 0$. The idea is to reduce from general cubes to cubes which are in some sense almost dyadic because we already know how to prove (1.1.6) in the case of dyadic cubes. Instead of considering all the uncountably many cubes over which the cube maximal operator averages, we want to select a finite set of cubes, which represent the cube maximal function well. These cubes will depend on the function f. The reduction to such cubes uses strategies similar to the proof of the optimal rate in the low density case for the local maximal function of a characteristic function in PI.

For the moment we will consider balls instead of cubes, but the arguments work the same for cubes. In the simpler case of a set instead of a function, the Vitali covering lemma is a tool to represent an uncountable union of balls by a disjoint union of balls. The balls of the Vitali covering lemma are designed to represent the volume of the original union. However, since we want to estimate the variation, we are actually interested in representing the perimeter of the original union of balls. We discovered a Vitali covering lemma for the boundary which represents the perimeter of a union of balls $\mathcal{B}$ by perimeters of a subset $\mathcal{S}$ of balls. The difference here is that the balls in $\mathcal{S}$ are not disjoint but can still have small overlap, meaning that for any two balls $S, T \in \mathcal{S}$ we only know

$$\mathcal{L}(S \cap T) \leq \varepsilon \min\{\mathcal{L}(S), \mathcal{L}(T)\}.$$

Our Vitali covering lemma for the boundary states that we have

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B}\Big) \lesssim_{d,\varepsilon} \sum_{B \in \mathcal{S}} \mathcal{H}^{d-1}(\partial\, S).$$

Such a subset $\mathcal{S}$ can be found for any $\varepsilon > 0$, but the constant that occurs in the previous estimate grows with ε becoming smaller. Furthermore note that we cannot estimate by the perimeter of the union of the balls in $\mathcal{S}$, but instead we estimate by the sum of the perimeters of each of the balls in $\mathcal{S}$, which is only the same if the balls in $\mathcal{S}$ are actually disjoint. In fact, the balls in $\mathcal{S}$ can even be made disjoint using a different proof, see Section 3.3. It is not clear if this improvement can be of help for the purposes of the proof of (1.1.6) for the cube maximal function.

The previous Vitali covering lemma for the boundary holds also for cubes. However, it is not formalized explicitly in the above form for balls or cubes in any paper of this thesis. Instead, we prove a more specialized version in PV, which deals with the cube maximal function of a function and not just with a union of balls or cubes. It allows us to pass from the low density part of the superlevel set of the cube maximal function to a set of almost disjoint cubes with a dyadic structure. Those cubes are not actually dyadic,

but they satisfy a property that two cubes are either almost disjoint, or one is of a smaller scale than the other. The maximal operator which is allowed to average only over those cubes still essentially retains the low density part of variation of the full cube maximal function. Actually, we cannot guarantee that the variation is retained by this almost dyadic maximal function, but instead we have to consider its variation as if for each levelset all cubes are disjoint.

Note that we can not just apply the previous Vitali covering lemma for the boundary for each levelset of the maximal function to achieve the same, because a priori the Vitali covering lemma for the boundary may produce a very different set of cubes for each levelset. Instead we have to select the almost disjoint cubes in a fashion which works consistently over all superlevel sets, in the sense that for $\lambda_0 < \lambda_1$, the cubes in the set S for level λ_1 can only be dyadic children of the cubes for λ_0 in some way. Unfortunately it is not clear how to generalize the proof for the disjoint version of the Vitali covering lemma for the boundary from a union of cubes to the setting of a maximal function. Thus we really do have to deal with the fact that cubes are only almost disjoint and not completely disjoint.

We would like to use the same strategy to prove the variation bound (1.1.6) for the uncentered fractional Hardy-Littlewood maximal operator which averages over uncentered balls and not cubes. Indeed, the whole proof in PV does work for balls just as for cubes, except for one ingredient: We have no suitable version for (2.4.3) for balls available, and there is no straightforward candidate for that. More thoughts and some attempts on this can be found in Section 3.2.

2.6 Fractional maximal functions

In PIII We prove (1.2.3) for all $0 < \alpha \leq d$ for both the uncentered and the centered Hardy-Littlewood maximal function. The starting point is (1.2.2), which is also the key step in the short proof of (1.2.3) for $1 \leq \alpha \leq d$, discovered by Carneiro and Madrid. For $\alpha < 1$ (1.2.2) does not make sense in this exact form. However, we prove in PIII that it still holds with $\widetilde{\mathrm{M}}_{\alpha-1}f$ replaced by a more refined maximal function $\widetilde{\mathrm{M}}_{\alpha,-1}f$ which continues to make sense for $\alpha < 1$. For a ball B let $r(B)$ be its radius. We define

$$\widetilde{\mathrm{M}}_{\alpha,-1}f(x) = r(B)^{\alpha-1} f_B,$$

where B is the *optimal ball* with $x \in B$ such that $r(B)^\alpha f_B$ is maximal. Such an optimal ball exists for almost every $x \in \mathbb{R}^d$ as pointed out earlier. Note, that for $\alpha \geq 1$ we have $\widetilde{\mathrm{M}}_{\alpha,-1}f(x) \leq \widetilde{\mathrm{M}}_{\alpha-1}f(x)$. In order to prove (1.2.3) it thus suffices to show

$$\|\widetilde{\mathrm{M}}_{\alpha,-1}f\|_{L^{d/(d-\alpha)}(\mathbb{R}^d)} \lesssim_{d,\alpha} \|\nabla f\|_{L^1(\mathbb{R}^d)}.$$

The key observation in the proof is the following. Assume that B_1, B_2 are two balls with $B_1 \subset B_2$ and $r(B_1) \leq r(B_2)/2$ which are both used by the fractional maximal function. Then we must have $r(B_1)^\alpha f_{B_1} > r(B_2)^\alpha f_{B_2}$ which implies $f_{B_1} > 2^\alpha f_{B_2}$. So we get that for $\alpha > 0$ fixed there must be a gap between f_{B_1} and f_{B_2} of length proportional to f_{B_1} itself. This argument can be strengthened in such a way that for any two balls B_1, B_2 of different enough scale either they have a horizontal gap, i.e. a certain minimal distance to another, or a vertical gap, i.e. a certain minimal distance between f_{B_1} and f_{B_2}. The same argument can be used to pass from centered to uncentered balls and even to dyadic cubes. This allows us to use the main argument for the dyadic maximal operator from PII to prove (1.2.3) both for the uncentered and for the centered fractional maximal function.

In their strengthened form, these arguments use blowups of balls. This is why they can not be used for the local fractional maximal operator. As mentioned earlier, this is expected because (1.2.3) does not hold for the local fractional maximal operator. However, the proof still works locally in the sense that we can bound the gradient of the maximal function by the variation of the function within these blown up optimal balls.

This localized result is what we use in PIV to prove the continuity of the gradient of the fractional maximal operator. Consider a sequence of functions $f_1, f_2, \ldots \in W^{1,1}(\mathbb{R}^d)$ converging to another function $f \in W^{1,1}(\mathbb{R}^d)$. The goal is to show that the gradients $\nabla \widetilde{\mathrm{M}}_\alpha f_n$ converge to $\nabla \widetilde{\mathrm{M}}_\alpha f$ in $L^{d/(d-\alpha)}(\mathbb{R}^d)$. We apply various results already established in earlier papers to show that $\nabla \widetilde{\mathrm{M}}_\alpha f_n$ converges on most of the domain. On the part that remains, we estimate $|\nabla \widetilde{\mathrm{M}}_\alpha f_n - \nabla \widetilde{\mathrm{M}}_\alpha f|$ by the triangle inequality and bound the parts $|\nabla \widetilde{\mathrm{M}}_\alpha f_n|$ and $|\nabla \widetilde{\mathrm{M}}_\alpha f|$ separately using the localized version of the result in PIII, finishing the proof of the continuity.

3. Techniques

3.1 The high density case

The first key tool which we prove in PI is the following.

Proposition 3.1.1 (Proposition 4.3 in PI and Proposition 3.5 in PV)**.** Let $\lambda \in (0,1)$. Let $E \subset \mathbb{R}^d$ be a set of locally finite perimeter and let $\mathcal{B}$ be a finite set of balls or cubes such that for each $B \in \mathcal{B}$ we have $\mathcal{L}(E \cap B) > \lambda \mathcal{L}(B)$. Then

$$\mathcal{H}^{d-1}\Big(\partial \bigcup \mathcal{B} \setminus \overline{E}^*\Big) \lesssim_d \lambda^{-\frac{d-1}{d}}(1-\log\lambda)\mathcal{H}^{d-1}\Big(\partial_* E \cap \bigcup \mathcal{B}\Big).$$

Proposition 3.1.1 is local in the sense that it only estimates by E inside the balls in $\mathcal{B}$. This makes it applicable also for the local maximal operator. As we will show later, the term $1-\log\lambda$ can be removed. We initially only wanted to prove in PI that for λ bounded from below the bound in Proposition 3.1.1 holds with a dimensional constant, but we were not interested in the rate in λ.

To summarize the proof idea we consider only $\lambda > \frac{1}{2}$ and $\mathcal{B} = \{B(0,1)\}$. Then we have to show that

$$\mathcal{H}^{d-1}(\partial_* B(0,1) \setminus \overline{E}^*) \lesssim_d \mathcal{H}^{d-1}(\partial_* E \cap B(0,1)).$$

Let $x \in \partial_* B(0,1) \setminus \overline{E}^*$ be a point on the boundary of the unit ball, away from the measure theoretic closure of E. Then we have

$$\lim_{\varepsilon \to 0} \frac{\mathcal{L}(B(x,\varepsilon) \cap E)}{\mathcal{L}(B(x,\varepsilon))} = 0.$$

Furthermore,

$$\mathcal{L}(B(x,2) \cap E) \geq \mathcal{L}(B(0,1) \cap E) \geq \frac{\mathcal{L}(B(0,1))}{2} = \frac{\mathcal{L}(B(x,2))}{2^{d+1}}.$$

By continuity there is an $0 < r \leq 2$ such that

$$\mathcal{L}(B(x,r) \cap E) = \frac{\mathcal{L}(B(x,r))}{2^{d+1}}.$$

Recall (2.3.2), i.e., that by the relative isoperimetric inequality we have

$$\mathcal{H}^{d-1}(\partial\, B(x,r)) \lesssim_d \mathcal{H}^{d-1}(\partial_*\, E \cap B(x,r)),$$

Furthermore, since $r \leq 2$ and since x lies on the boundary of $B(0,1)$ one can show that

$$\mathcal{H}^{d-1}(\partial\, B(0,1) \cap B(x,r)) \lesssim_d \mathcal{H}^{d-1}(\partial\, B(x,r)).$$

In fact, we will later show a generalization of this, Lemma 3.3.1. Chaining these arguments together we obtain

$$\mathcal{H}^{d-1}(\partial\, B(0,1) \cap B(x,r)) \lesssim_d \mathcal{H}^{d-1}(\partial_*\, E \cap B(x,r)).$$

Using more careful considerations we can show this even with the right-hand side replaced by $\mathcal{H}^{d-1}(\partial_*\, E \cap B(x,r) \cap B(0,1))$. Doing this for every point $x \in \partial\, B(0,1) \setminus \overline{E}^*$ we obtain a cover $\mathcal{B}$ of $\partial\, B(0,1) \setminus \overline{E}^*$ with such balls $B(x,r)$. Using the Vitali covering theorem we can essentially consider the balls in $\mathcal{B}$ disjoint and conclude

$$\begin{aligned}
\mathcal{H}^{d-1}(\partial\, B(0,1) \setminus \overline{E}^*) &\lesssim_d \sum_{B(x,r)\in\mathcal{B}} \mathcal{H}^{d-1}(B(x,r) \cap \partial\, B(0,1) \setminus \overline{E}^*) \\
&\leq \sum_{B(x,r)\in\mathcal{B}} \mathcal{H}^{d-1}(B(x,r) \cap \partial\, B(0,1)) \\
&\lesssim_d \sum_{B(x,r)\in\mathcal{B}} \mathcal{H}^{d-1}(\partial_*\, E \cap B(x,r) \cap B(0,1)) \\
&\leq \mathcal{H}^{d-1}(\partial_*\, E \cap B(0,1)).
\end{aligned}$$

If we drop the assumption $\lambda > \frac{1}{2}$, we immediately obtain the rate $\lambda^{-\frac{d-1}{d}}$, even without the factor $1 - \log\lambda$, by observing

$$\mathcal{L}(B(x,2) \cap E) \geq \frac{\lambda}{2^d}\mathcal{L}(B(x,2))$$

and then running the exact same proof described above while tracking the dependence on λ. The main idea still works for the case that we have a general set of balls $\mathcal{B}$ instead of $B(0,1)$, but we have to refine the relative isoperimetric inequality and the Vitali covering argument a bit. Instead of disjoint sets, we can only obtain a collection of sets with bounded overlap at first. To make them disjoint we lose an additional factor $(1 - \log\lambda)$. That means our final rate in λ after this argument is $\lambda^{-\frac{d-1}{d}}(1 - \log\lambda)$.

3.2 Mass estimate for low density maximal dyadic cubes

The main tool we prove in PII is the following.

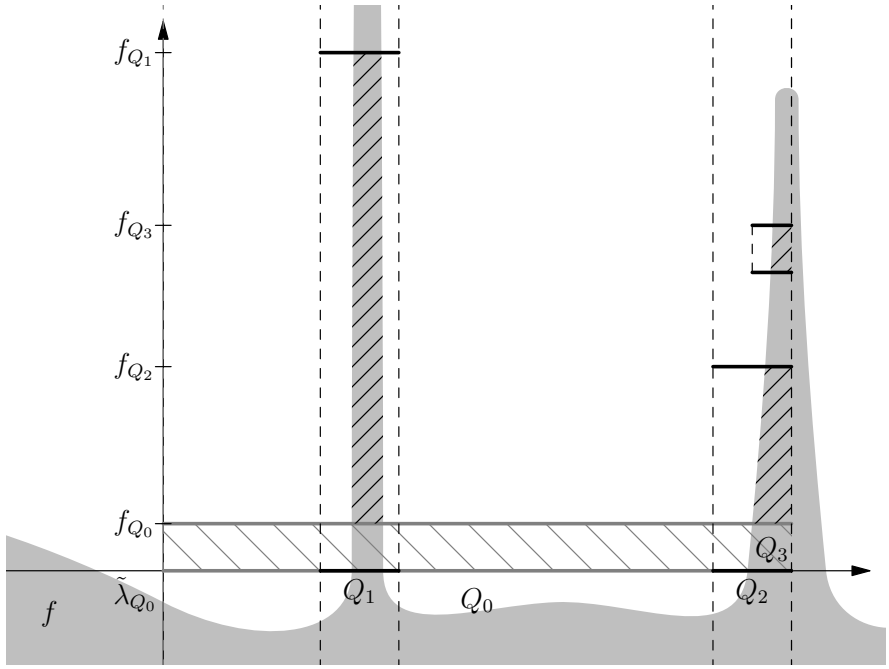


Figure 3.1. Proposition 3.2.1: We can bound the mass from a very low density cube by the mass above of low density cubes.

Proposition 3.2.1 (Corollary 3.3 in PII and Proposition 3.11 in PV)**.** Let Q_0 be a cube, $\lambda_0 \in \mathbb{R}$ and $f \in L^1(Q_0)$ with $\mathcal{L}(\{f \geq \lambda_0\} \cap Q_0) \leq 2^{-d-1}\mathcal{L}(Q_0)$. Then

$$\mathcal{L}(Q_0)(f_{Q_0} - \lambda_0) \leq 2^{d+1} \int_{f_{Q_0}}^{\infty} \mathcal{L}\Big(\{f \geq \lambda\} \cap \bigcup \{Q \in \mathfrak{D}(Q_0) : f_Q \geq \lambda,\ \mathcal{L}(Q \cap \{f \geq \lambda\}) < \mathcal{L}(Q)/2\}\Big)\, \mathrm{d}\lambda.$$

This is the main tool we have to deal with the low density part of a general maximal function. While this lemma only concerns masses, they can readily be converted into contributions to the variation. This is because the masses are low density, for which the relative isoperimetric inequality applies.

To understand the proof idea for Proposition 3.2.1 we consider the case $\lambda_0 = 0$ and $f \geq 0$. Then $\mathcal{L}(Q_0)(f_{Q_0} - \lambda_0) = \int_{Q_0} f$. Let $\mathcal{P}$ be the set of maximal cubes P with $\mathcal{L}(\{f > f_P\} \cap P) \geq 2^{-d-1}\mathcal{L}(P)$. We split the mass of f in Q_0 into three pieces

$$\begin{aligned} \int_{Q_0} f &= \int_0^{\infty} \mathcal{L}(\{f > \lambda\} \cap Q_0)\, \mathrm{d}\lambda \\ &= \underbrace{\int_0^{\infty} \mathcal{L}\Big(\{f > \lambda\} \setminus \bigcup \mathcal{P}\Big)\, \mathrm{d}\lambda}_{M_1} \end{aligned}$$

$$+\underbrace{\int_0^\infty \mathcal{L}\Big(\{f>\lambda\}\cap\bigcup\{P\in\mathcal{P}: f_P>\lambda\}\Big)\,\mathrm{d}\lambda}_{M_2}$$
$$+\underbrace{\int_0^\infty \mathcal{L}\Big(\{f>\lambda\}\cap\bigcup\{P\in\mathcal{P}: f_P\le\lambda\}\Big)\,\mathrm{d}\lambda}_{M_3}.$$

It can be concluded from the Lebesgue differentiation theorem that $M_1 = 0$. By the definition of $\mathcal{P}$ we have for any $P \in \mathcal{P}$ that

$$\begin{aligned}\int_0^\infty \mathcal{L}(\{f>\lambda\}\cap P)\,\mathrm{d}\lambda &= \int_P f = f_P\mathcal{L}(P)\\ &= \int_0^{f_P}\mathcal{L}(P)\,\mathrm{d}\lambda\\ &\le 2^{d+1}\int_0^{f_P}\mathcal{L}(\{f>\lambda\}\cap P)\,\mathrm{d}\lambda.\end{aligned}$$

This implies $M_2 + M_3 \le 2^{d+1}M_2$, and it remains to estimate M_2. We will show that all the mass M_2 is low density mass. Let (x,λ) belong to the mass M_2, i.e., we have $f(x) > \lambda$ and there is a cube $P \in \mathcal{P}$ with $x \in P$ and $f_P > f(x) > \lambda$. Let Q be the maximal cube with $x \in Q$ and $f_Q > \lambda$. We have to show that (x,λ) belongs to the low density mass, i.e. that

$$\mathcal{L}(Q\cap\{f>\lambda\}) \le \mathcal{L}(Q)/2.$$

Let R be the parent cube of Q. Then $f_R \le \lambda$ and since $f_P > \lambda$ we have $P \subset Q \subsetneq R$. By the maximality of P we get

$$\begin{aligned}\mathcal{L}(Q\cap\{f>\lambda\}) &\le \mathcal{L}(Q\cap\{f>f_R\})\\ &\le \mathcal{L}(R\cap\{f>f_R\})\\ &< 2^{-d-1}\mathcal{L}(R) = \mathcal{L}(Q)/2.\end{aligned}$$

This finishes the proof.

By a simple shift, the case of a general λ_0 can be reduced to the case $\lambda_0 = 0$. However, we cannot assume that $f \ge \lambda_0$. The idea is to essentially cut out the region of Q_0 on which $f < \lambda_0$ using the dyadic cubes, and then run the same strategy as above. More precisely, we define $\mathcal{P}$ exactly the same as above, however then only consider those $P \in \mathcal{P}$ with $f_P > 0$; denote them by $\tilde{\mathcal{P}}$. This way, we still have the full mass $\int_{Q_0} f$ supported on $\bigcup\tilde{\mathcal{P}}$, if not more, because those cubes in $\mathcal{P}\setminus\tilde{\mathcal{P}}$ that we cut out carry negative mass. The same argument as above also works for $\tilde{\mathcal{P}}$.

Proposition 3.2.1 is a local statement in that it only concerns f in Q. This means it can also be applied for the local maximal operator. If we assume $\lambda_0 = 0$ and $f \ge 0$ then we can even prove Proposition 3.2.1 for balls instead of cubes with a similar proof as above. Finding the balls corresponding to the set of cubes $\mathcal{P}$ above is a bit more intricate, but also

possible. Analogously to maximal dyadic cubes, by a *maximal ball* B we mean a ball such that for any ball $C \supsetneq B$ we have $f_C < f_B$. They are basically the same as the earlier defined optimal balls. Any maximal ball is an optimal ball, but not necessarily vice versa. The difference however is minuscule, for most arguments the terms can be used interchangeably.

Given a ball $B = B(x, r)$ and a number $c > 0$ we denote the blown up ball by $cB = B(x, cr)$. Assuming $f \geq 0$ we can conclude for a maximal ball B that $f_B > f_{2B} \geq 2^{-d} f_B$. Furthermore, if for a fixed $\varepsilon > 0$ we have

$$\varepsilon < \frac{\mathcal{L}(\{f > \lambda\} \cap B)}{\mathcal{L}(B)} < 1 - \varepsilon,$$

then we also have

$$\varepsilon' < \frac{\mathcal{L}(\{f > \lambda\} \cap 2B)}{\mathcal{L}(2B)} < 1 - \varepsilon'$$

for some fixed $0 < \varepsilon' \leq \varepsilon$. This means that the key properties of balls we care about, f_B and $\mathcal{L}(\{f > \lambda\} \cap B)/\mathcal{L}(B)$, essentially carry over from B to $2B$, in fact to any cB with $c \geq 1$. This means that Vitali covering arguments where we move from balls to their blowups can be used here to make the maximal balls corresponding to the set of cubes $\mathcal{P}$ disjoint.

Using blowups of balls however denies the possibility of applying this argument to the local maximal operator. More importantly it also means that this strategy cannot be applied in the case of general functions. If we drop the assumption $f \geq 0$, then given f_B the value of f_{2B} can become arbitrarily small. It may happen that $2B$ reaches out of our base ball B_0, into a region about which we have no information at all. Moreover, we cannot guarantee that the mass or variation of f outside of B_0 is still available to be used to bound the contribution of B_0 to the variation of the maximal function. We cannot know if we maybe also need this part of f to bound another part of the variation of the maximal function.

Another problem with the proof in the dyadic setting occurs when we cut out the parts of the function where $f(x) < \lambda_0$. Because we have to use blowups when we want to apply this strategy with balls instead of cubes, we lack the necessary precision to cut out pieces using balls. Since initially our balls are not disjoint, we cannot cut out single balls without possibly also affecting other balls by cutting away the mass inside them. And because of the problems mentioned above, we cannot make balls disjoint by applying Vitali covering arguments before having cut out the problematic regions where $f(x) < \lambda_0$.

These problems stand in the way of replacing the cubes in the proof of Proposition 3.2.1 by balls. In fact, it is not even clear if a statement similar to Proposition 3.2.1 holds for balls instead of cubes. Problems might arise when the mass of f is distributed in such a way, that many balls intersect the area where the negative mass $f < 0$ is highly concentrated, but each of these balls use separate reservoirs of positive mass $f > 0$. More precisely,

it may happen for a set of balls $B_1, B_2, \dots$ that for each i we have $f_{B_i} < 0$, while $\int_{\bigcup_i \mathcal{B}_i} f$ is positive and maybe even larger than $\int_{B_0} f$. It is possible that this reduces the amount of low density mass of f within B_0 so much that a ball version of Proposition 3.2.1 fails.

One interpretation of the strategy in the low density case for the dyadic maximal operator is to select dyadic cubes which decompose each superlevel set according to scales. More precisely, for each superlevel set one can extract from the argument a disjoint set of cubes whose volumes and perimeter are comparable to the volume and perimeter of the superlevel set in their vicinities. These cubes cover enough of the superlevel sets to represent all the mass relevant for the low density part of the maximal function. The decision of which scale to assign to which piece is here guided by the size of the optimal cubes of the maximal function. Being able to isolate the mass using dyadic subcubes is imperative for that. It is unclear if the optimal balls for the uncentered Hardy-Littlewood maximal function can also be used for this guidance. Instead, it might be possible to do the decomposition according to scales of each superlevel set more directly, without using any maximal function. This means however that it has to work for whichever way the optimal balls of the maximal function may be arranged. Therefore the decomposition has to be more fine grained and capture much more of the local geometry of the superlevel set than the decomposition using optimal cubes, which poses a challenge. One untapped source of variation of the function which we may use is the variation on levels below the level λ_0. For the dyadic maximal function we only use the levelsets of the function above level λ_0, but for the Hardy-Littlewood maximal function we could also use pieces of the decomposition that lie below that level λ_0. This might even be necessary for a bound like Proposition 3.2.1 to hold for balls. In comparison to the case of a union of balls, a general decomposition would also have to take into account the geometry of holes in the set. The idea is that if we cannot use a blowup of a ball, then it is holes in levelsets below λ which are responsible for this, and we may thus harness their contribution to the variation of f instead of using the blown up ball. There are also examples of functions whose maximal function uses small balls only because larger balls are not allowed due to small and very deep holes in the function. That means we would also consider pieces of negative mass, for example the superlevel set $B(0,10) \setminus B(0,1)$ would probably be decomposed into the piece of *positive* mass $B(0,10)$ and the piece of *negative* mass $B(0,1)$. The disjoint version of the Vitali covering lemma for the boundary is a decomposition which goes in the right direction, because it represents an arbitrary union of balls by a disjoint union of balls in accordance to the scale of the boundary, but only of the positive mass. It is not sufficient here because it does not take into account holes. A decomposition of a general set according to the scale of the boundary and of its positive and negative mass might also be

of independent interest.

3.3 Covering-type lemmas for the boundary

One of the main tasks in PI and PV is to understand the perimeter of arbitrary unions of balls or cubes. The general attempt is to reduce from arbitrary unions of balls or cubes to disjoint unions. There is not one definite theorem, but instead a set of tools that reflect this idea. So, why are we interested in disjoint unions? The most direct example is the low density case for characteristic functions in PI. As described in Section 2.3, to deal with this case it suffices to prove that for any $\lambda \leq \frac{1}{2}$ and any set $\mathcal{B}$ of balls B with $\lambda \leq \frac{\mathcal{L}(B \cap E)}{\mathcal{L}(B)} \leq \frac{1}{2}$ we have

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B}\Big) \lesssim_d \lambda^{-\frac{d-1}{d}} \mathcal{H}^{d-1}\Big(\bigcup \mathcal{B} \cap \partial_* E\Big). \tag{3.3.1}$$

For the case that $\mathcal{B}$ consists only of a single ball this is a direct consequence of the relative isoperimetric inequality, see (2.3.2). This directly extends to the case that $\mathcal{B}$ is a disjoint union, because then we have

$$\begin{aligned}\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B}\Big) &= \sum_{B \in \mathcal{B}} \mathcal{H}^{d-1}(\partial B) \\ &\lesssim_d \lambda^{-\frac{d-1}{d}} \sum_{B \in \mathcal{B}} \mathcal{H}^{d-1}(\partial_* E \cap B) \\ &= \lambda^{-\frac{d-1}{d}} \mathcal{H}^{d-1}\Big(\partial_* E \cap \bigcup \mathcal{B}\Big).\end{aligned}$$

For dyadic cubes all unions can be made disjoint. So as we saw in the case of the dyadic maximal operator, disjoint unions may even allow for a proof of (1.1.6) for general functions.

Our first tool in this spirit is the following. Recall that $r(B)$ is the radius of a ball B.

Lemma 3.3.1. Let $K > 0$ and B_0 be a ball and $\mathcal{B}$ be a set of balls B with $r(B) \geq K r(B_0)$. Then

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B} \cap B_0\Big) \lesssim_d \Big(1 + \frac{1}{K}\Big) \mathcal{H}^{d-1}(\partial B_0).$$

Lemma 3.3.1 reduces the boundary of an arbitrary union of balls to the boundary of the single ball B_0. Note that in PI we formulated Lemma 3.3.1 with the rate K^{-d} instead of the optimal rate K^{-1}. For the proof of Lemma 3.3.1 it suffices to consider $B_0 = B(0,1)$ by translation and scaling. For K large enough we can write $\partial_* \bigcup \mathcal{B} \cap B(0,1)$ as a union of a bounded number of Lipschitz graphs with uniformly bounded domain and Lipschitz constant. We can conclude

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B} \cap B(0,1)\Big) \lesssim 1 \lesssim \mathcal{H}^{d-1}(\partial B(0,1)).$$

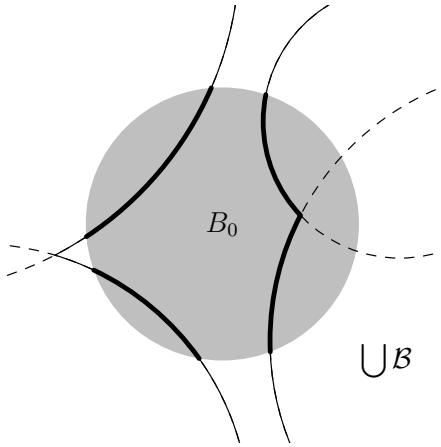


Figure 3.2. The boundary of a union of balls within another ball in Lemma 3.3.1.

For smaller K we cover $B(0,1)$ by $\lesssim K^{-d}$ balls with radius K and apply the previous case to each of these ball, yielding $\lesssim K^{-d}$ contributions of $\lesssim K^{d-1}$ much surface area.

We have recently discovered simplifications of the proof of the bound in the low density case in PI, and even one alternative strategy. In the bigger picture none of the developed strategies have been made obsolete though, because the original proof idea is still being used to prove (1.1.6) for the cube maximal operator in PV.

Strategy 0 As described in Section 2.3, we do not actually need to consider the low density case separately from the high density case in order to prove the gradient bound (1.1.6) for the maximal function of a characteristic function. The proof for the high density case still works well enough in the low density case, as it only introduces the unnecessary although not critical factor $(1 - \log \lambda)$. This strategy also works for the local maximal operator because Proposition 3.1.1 for the high density case is a local estimate.

In order to remove that factor $(1 - \log \lambda)$ we need another strategy, dedicated to the low density case. This strategy is then further developed to deal with the low density part of the maximal function of a general function. These strategies for the low density case centrally involve arguments in the spirit of the following Vitali covering lemma for the boundary.

Proposition 3.3.2. Let $\varepsilon > 0$ and let $\mathcal{B}$ be a bounded set of balls or cubes. Then $\mathcal{B}$ has a subset $\mathcal{S}$ such that for any distinct $S, T \in \mathcal{S}$ we have

$$\mathcal{L}(S \cap T) \leq \varepsilon \min\{\mathcal{L}(S), \mathcal{L}(T)\},$$

and

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B}\Big) \lesssim_d \sum_{B \in \mathcal{S}} \mathcal{H}^{d-1}(\partial B).$$

Strategy 1 Recently we discovered that the balls in Proposition 3.3.2 can actually be made fully disjoint. This improvement can be used to greatly simplify the original proof of (3.3.1) in PI, because we can pass from the boundary of $\bigcup \mathcal{B}$ directly to the boundary of a set of disjoint balls. In that situation we can argue the same way for the dyadic maximal function of a characteristic function for which the proof is very short, as already presented in Section 2.3. It is not clear however if this improved tool can also simplify arguments for the cube maximal function of a general function in PV.

The fully disjoint version of (3.3.1) follows from the proof of the Besicovitch covering theorem, combined with Proposition 3.1.1 in the high density case: i.e., for some dimensional constant $\lambda > 0$. Recall that the Besicovitch covering theorem states that for any bounded set of balls $\mathcal{B}$, there are subsets $\mathcal{B}_1, \dots, \mathcal{B}_N$ of $\mathcal{B}$ with N depending only on the dimension, such that each $\mathcal{B}_i$ consists of disjoint balls and $\bigcup \mathcal{B}_1 \cup \ldots \cup \bigcup \mathcal{B}_N$ contains the centers of all balls in $\mathcal{B}$. This statement alone is not enough for our purposes here. However, when following the proof of the Besicovitch covering theorem, for example in [EG15], one can observe, that for any ball $B(x,r) \in \mathcal{B}$, the ball $B \in \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_N$ which contains x actually has $r(B) \geq \frac{3}{4} r$. This implies

$$\mathcal{L}(B(x,r) \cap B) \geq \varepsilon \mathcal{L}(B)$$

for some $\varepsilon > 0$ depending only on the dimension. Using Proposition 3.1.1 we can conclude

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \Big\{ B(x,r) \in \mathcal{B} : x \in B,\ r(B) \geq \frac{3}{4} r \Big\}\Big) \lesssim_d \mathcal{H}^{d-1}(\partial B),$$

and therefore we can bound

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \mathcal{B}\Big) \lesssim_d \sum_{i=1}^{N} \sum_{B \in \mathcal{B}_i} \mathcal{H}^{d-1}(\partial B).$$

Finally, pick that $i \in \{1, \dots, N\}$ for which $\sum_{B \in \mathcal{B}_i} \mathcal{H}^{d-1}(\partial B)$ is maximal and $\mathcal{S} = \mathcal{B}_i$ is the desired subset of disjoint balls.

Since we are not using blowups, this strategy stays local and thus is also applicable for the local maximal function.

In PI we do not prove Proposition 3.3.2 with fully disjoint balls, instead we develop a more complicated strategy to achieve disjointness. This original approach from PI is still far from obsolete because the fully disjoint version of Proposition 3.3.2 does not work well for general functions. As mentioned in Section 2.5, the reason for this is that in that case we actually use a generalization of Proposition 3.3.2 which deals with all levelsets at the same time, and we do not have a disjoint version of that generalization available. That is why our strategy in PV for the cube maximal function of a general function is still in the same spirit as the one from PI.

Strategy 2 Below we describe the strategy from PI with some small simplifications and avoid Proposition 3.3.2 altogether. The idea is to select disjoint balls $F_1, F_2, \dots$ which represent E well enough in how it shapes the maximal function. That means, each of the balls F_i should contain about as much of E as its own volume, contain at least as much of $\partial_* E$ as its own perimeter, and each ball B of the maximal function has smaller balls F_i in its close vicinity with a combined volume comparable to at least $\lambda \mathcal{L}(B)$. First assume we have obtained such a sequence of balls $F_0, F_1, \dots$. For balls of the same scale we can actually achieve much more than almost disjointness in Proposition 3.3.2: we can even make them widely spaced. That means each ball F_i will have at most one ball $B \in \mathcal{B}$ of each scale with F_i in its vicinity. For each $B \in \mathcal{B}$ we want to estimate the perimeter of B by the perimeter of E inside all the F_i in the vicinity of B. Because we know that the volume of B is comparable to the volume of the F_i in its vicinity, and

$$\mathcal{L}(B) \gtrsim r(B)\mathcal{H}^{d-1}(\partial B)$$

and

$$\mathcal{L}(F_i) \lesssim r(F_i)\mathcal{H}^{d-1}(\partial F_i);$$

we obtain an extra factor $r(F_i)/r(B)$ in the perimeter estimate. We use that this factor converges for every F_i if we sum over all $B \in \mathcal{B}$ with $F_i \subset 2B$, where $2B$ is the ball with the same center as B and twice the radius. That means every F_i may contribute to the mass of infinitely many $B \in \mathcal{B}$ and the same is true for perimeters, however for each ball B it suffices to take a portion of size $r(F_i)/r(B)$ of the perimeter of F_i to account for the perimeter of $\mathcal{H}^{d-1}(\partial B)$. The rate of the parameter λ in this estimate is $\lambda^{-\frac{d-1}{d}}$.

It remains to find the disjoint balls $F_0, F_1, \dots$. Because we only consider the low density balls B of the maximal function, i.e. those with

$$\lambda \leq \frac{\mathcal{L}(B \cap E)}{\mathcal{L}(B)} \leq \frac{1}{2},$$

we can cover almost all of E in B by balls $F \subset B$ with $\mathcal{L}(F \cap E) = \mathcal{L}(F)/2$ due to the Lebesgue differentiation theorem. We do this for every ball B that the maximal function uses. Then we extract a disjoint subsequence of all the balls F using the Vitali covering lemma. What we still have to make sure is that the deletion process in the Vitali covering argument did not erase balls that we might have needed. While each remaining ball F does retain a certain part of all the mass of E within the blown up ball $5F$, it might have happened that we deleted balls $F' \subset 5F$ of much smaller scale than F. This might cause a problem, because we need that for every ball B of the maximal function, $\lambda \mathcal{L}(B)$ much mass in its vicinity is represented by balls F of scale not larger than the scale of $\mathcal{B}$. However, if λ is small enough then this fatal deletion of small balls cannot happen. There cannot be a small ball B in the vicinity of a big ball F, because then B would

be completely contained in the ball $\frac{2}{\lambda}F$, which is an eligible ball for the maximal function because we chose F so that $\mathcal{L}(F \cap E) = \mathcal{L}(F)/2$.

For the local maximal operator we cannot argue like this, because the blown up ball $\frac{2}{\lambda}F$ might reach outside the domain. Instead we pay more attention to how the balls in $\mathcal{B}$ intersect. Let us first consider the case that all balls $\mathcal{B}$ have a similar size. Using the Vitali covering lemma to extract a set $\tilde{\mathcal{B}}$ of disjoint balls and then applying Lemma 3.3.1 to $5B$ for each $B \in \tilde{\mathcal{B}}$, we can bound the perimeter of $\bigcup \mathcal{B}$ by the perimeter of $\bigcup \tilde{\mathcal{B}}$. Then (3.3.1) follows from (2.3.2) applied to each $B \in \tilde{\mathcal{B}}$.

If $\mathcal{B}$ contains balls of arbitrary sizes we group them according to their scale and make the balls in each scale disjoint as above. This means balls of different scales may still intersect. It will suffice to consider intersections of balls of very different scales: We group the scales 2^n into their congruence classes modulo N for some $N \in \mathbb{N}$. We consider each class separately and only lose a factor of N in the final estimate. Different scales in each group have a gap of at least N. That means that when a ball B_1 is of a smaller scale than B_2, then we have $r(B_1) < 2^{2-N} r(B_2)$. Assume that such balls B_1 and B_2 intersect. If B_1 is fully contained in B_2, then we can ignore B_1 because it does not contribute to the perimeter of $\bigcup \mathcal{B}$, so we do not have to consider this case. Otherwise $(1-\delta)B_1$ and $(1-\delta)B_2$ are disjoint for some $\delta > 0$ depending only on N, which means the mass of E in B_2 away from the boundary of B_2 is not touched by B_1. This is the key observation we exploit.

We build the balls $F_1, F_2, \ldots$ that represent the function well the same way as in the global case, except we do not just extract a disjoint subsequence using the Vitali covering argument. The reason is that the Vitali covering argument would for a given ball F_i make unavailable all the mass of E within F_i for other balls F_j. And if F_i lies very near the boundary of its corresponding ball B then we might still need some of the mass to account for a ball B' which intersects B and F_i. Instead we use that we can consider $(1-\delta)F_i$ instead of F_i and still have

$$\varepsilon < \frac{\mathcal{L}((1-\delta)F_i \cap E)}{\mathcal{L}((1-\delta)F_i)} < 1 - \varepsilon.$$

This way we can ensure that either $(1-\delta)F_i$ does not intersect B', or else $r(F_i)$ is bounded by $r(B')$ so that $F_i \subset cB'$, which means that we can use F_i itself to account for the mass in B'.

The benefit of the contraction argument in the previous strategy is not only that it works also for the local maximal operator. We need the same argument in the case of general functions in PV, even when considering the global maximal function. Recall that this connection between the local setting and general functions also occurred in the discussion of Proposition 3.2.1.

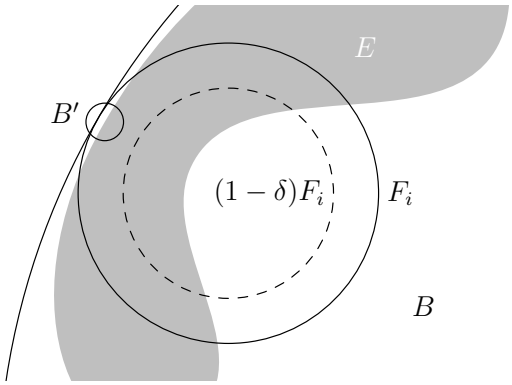


Figure 3.3. Achieving disjointness by contracting F_i to $(1-\delta)F_i$.

Strategy 3 The previous strategy is still a bit simpler than the strategy that we actually use in PI. The original proof implicitly uses Proposition 3.3.2, although we do not state it as a proposition in PI. In order to achieve that $(1-\delta)B_1$ and $(1-\delta)B_2$ are disjoint for some small $\delta > 0$ we used above that we can group the scales into many groups, so that any two balls B_1, B_2 of different scales have very different radii. We do not use this argument in PI, but instead apply Proposition 3.3.2 to the set of balls used by the maximal function, although not explicitly. The resulting balls can be made to intersect so little, that even if B_1 and B_2 have the same radius, the balls $(1-\delta)B_1$ and $(1-\delta)B_2$ are disjoint. While former argument is simpler, in the case of general functions in PV we use the latter argument. There we prove a version of Proposition 3.3.2 adapted to general functions. Proposition 3.3.2 might also be of independent interest.

Below we describe the original proof idea of Proposition 3.3.2. We recall the proof of the original Vitali covering lemma for a set of balls $\mathcal{B}$. First pick the largest ball B_0 in $\mathcal{B}$ and remove all balls from $\mathcal{B}$ which are contained in $5B_0$. Then pick the largest ball B_1 of those leftover balls, and remove all balls from $\mathcal{B}$ contained in $5B_1$. We continue selecting balls inductively in this manner. We obtain a sequence $B_0, B_1, \ldots$ of disjoint balls, and every ball $B \in \mathcal{B}$ is contained in $5B_i$ for some $i \in \mathbb{N}$. We cannot use this sequence of balls for Proposition 3.3.2 though. While clearly we have

$$\mathcal{L}\Big(\bigcup\{B \in \mathcal{B} : B \subset 5B_i\}\Big) \lesssim \mathcal{L}(B_i),$$

we cannot guarantee this inequality to hold if we replace all instances of the Lebesgue measure $\mathcal{L}(\cdot)$ in the previous display by the perimeter $\mathcal{H}^{d-1}(\partial_*(\cdot))$.

In order to prove Proposition 3.3.2 we define a similar sequence, but instead we remove in step i only those balls B from $\mathcal{B}$ with

$$\mathcal{L}(B \cap B_i) \geq \varepsilon \mathcal{L}(B).$$

This way, two distinct balls B_i, B_j satisfy

$$\mathcal{L}(B_i \cap B_j) \leq \varepsilon \min\{\mathcal{L}(B_i), \mathcal{L}(B_j)\}.$$

By Proposition 3.1.1 we have

$$\mathcal{H}^{d-1}\Big(\partial_* \bigcup \{B \in \mathcal{B} : \mathcal{L}(B \cap B_i) > \varepsilon \mathcal{L}(B)\}\Big) \lesssim \mathcal{H}^{d-1}(\partial_* B_i)$$

and we can conclude Proposition 3.3.2.

Because the Vitali covering lemma for the boundary is a local statement, this strategy also works for the local maximal function.

3.4 Gaps created by the fractional exponent

As we have already observed earlier, if the balls that are used by the maximal operator all have comparable size, then we can use Lemma 3.3.1 and a Vitali covering argument to make the balls disjoint. We also observed that for disjoint balls the proof of (1.1.6) is easy. This is true also for the fractional maximal operator. Difficulties occur only when balls of very different scales intersect. This is where we make use of the exponent $\alpha > 0$ of the fractional maximal operator. Let B_1, B_2 be balls with $r(B_2) \geq cr(B_1)$ for some very big $c > 2$. If $2B_1$ and $2B_2$ intersect then $B_1 \subset 4B_2$ which means we must have

$$r(B_1)^\alpha f_{B_1} > r(4B_2)^\alpha f_{4B_2} \geq 4^{\alpha-d} r(B_2)^\alpha f_{B_2}$$

for the ball B_1 to be used by the fractional maximal operator. We can conclude $f_{B_1} > 4^{\alpha-d} c^\alpha f_{B_2}$, so that for c large enough there must be a gap between f_{B_1} and f_{B_2} of a minimal length proportional to f_{B_1} itself. That means for two balls B_1, B_2 of very different scales we can conclude that they either have a gap between them, or there is a gap between f_{B_1} and f_{B_2}. Using Lemma 3.3.1 and a Vitali covering argument we can ensure that the same holds for two balls B_1, B_2 of the same scale.

High density balls are of no use here because the inequality $\mathrm{M}f(x) \geq f(x)$, which was essential to bound the contribution of high density balls, does not hold for for the fractional maximal function $\mathrm{M}_\alpha f$. This is not a problem because for any ball B used by fractional maximal function, at least its blowup $2B$ is a low density ball at level f_B. If B is a high density ball and also

$$\mathcal{L}(\{f > f_B\} \cap 2B\}) \geq (1-\varepsilon)\mathcal{L}(2B),$$

then we must have $f_{2B} \geq (1-\varepsilon) f_B$, which means that

$$r(B)^\alpha f_B < r(2B)^\alpha f_{2B}$$

for ε small enough. That means that the ball B is not actually used by the fractional maximal operator. We can conclude that at least $2B$ is a low

density ball at level f_B. By the argument from the previous paragraph we can make the vertical and horizontal gaps large enough so that each ball B of the fractional maximal function is contained in a low density ball, and between two such low density balls is still a horizontal or vertical gap.

Observe that the blowup of the balls prevents an application to the local fractional maximal operator. This is expected because the endpoint regularity bound does not hold for the local fractional maximal operator.

References

[ACPL12] 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, 2012.

[AK10] Daniel Aalto and Juha Kinnunen. The discrete maximal operator in metric spaces. *J. Anal. Math.*, 111:369–390, 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.

[APL08] J. M. Aldaz and J. Pérez Lázaro. Boundedness and unboundedness results for some maximal operators on functions of bounded variation. *J. Math. Anal. Appl.*, 337(1):130–143, 2008.

[APL09] J. M. Aldaz and F. J. Pérez Lázaro. Regularity of the Hardy-Littlewood maximal operator on block decreasing functions. *Studia Math.*, 194(3):253–277, 2009.

[BCHP12] Jonathan Bober, Emanuel Carneiro, Kevin Hughes, and Lillian B. Pierce. On a discrete version of Tanaka's theorem for maximal functions. *Proc. Amer. Math. Soc.*, 140(5):1669–1680, 2012.

[BM20] David Beltran and José Madrid. Regularity of the centered fractional maximal function on radial functions. *J. Funct. Anal.*, 279(8):108686, 28, 2020.

[BM21] David Beltran and José Madrid. Endpoint Sobolev continuity of the fractional maximal function in higher dimensions. *Int. Math. Res. Not. IMRN*, (22):17316–17342, 2021.

[BRS19] David Beltran, João Pedro Ramos, and Olli Saari. Regularity of fractional maximal functions through Fourier multipliers. *J. Funct. Anal.*, 276(6):1875–1892, 2019.

[Buc99] Stephen M. Buckley. Is the maximal function of a Lipschitz function continuous? *Ann. Acad. Sci. Fenn. Math.*, 24(2):519–528, 1999.

[BW21] Constantin Bilz and Julian Weigt. The one-dimensional centred maximal function diminishes the variation of indicator functions. *arXiv preprint arXiv:2107.12404*, 2021.

[Car19] Emanuel Carneiro. Regularity of maximal operators: recent progress and some open problems. In *New trends in applied harmonic analysis. Vol. 2—harmonic analysis, geometric measure theory, and applications*, Appl. Numer. Harmon. Anal., pages 69–92. Birkhäuser/Springer, Cham, [2019] ©2019.

[CFS18] Emanuel Carneiro, Renan Finder, and Mateus Sousa. On the variation of maximal operators of convolution type II. *Rev. Mat. Iberoam.*, 34(2):739–766, 2018.

[CGR21] Emanuel Carneiro and Cristian González-Riquelme. Gradient bounds for radial maximal functions. *Ann. Fenn. Math.*, 46(1):495–521, 2021.

[CGRM22] Emanuel Carneiro, Cristian González-Riquelme, and José Madrid. Sunrise strategy for the continuity of maximal operators. *Journal d'Analyse Mathématique*, pages 1–48, 2022.

[CH12] Emanuel Carneiro and Kevin Hughes. On the endpoint regularity of discrete maximal operators. *Math. Res. Lett.*, 19(6):1245–1262, 2012.

[CM08] Emanuel Carneiro and Diego Moreira. On the regularity of maximal operators. *Proc. Amer. Math. Soc.*, 136(12):4395–4404, 2008.

[CM17] Emanuel Carneiro and José Madrid. Derivative bounds for fractional maximal functions. *Trans. Amer. Math. Soc.*, 369(6):4063–4092, 2017.

[CMP17] Emanuel Carneiro, José Madrid, and Lillian B. Pierce. Endpoint Sobolev and BV continuity for maximal operators. *J. Funct. Anal.*, 273(10):3262–3294, 2017.

[CS13] Emanuel Carneiro and Benar F. Svaiter. On the variation of maximal operators of convolution type. *J. Funct. Anal.*, 265(5):837–865, 2013.

[EG15] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[GRK21] Cristian González-Riquelme and Dariusz Kosz. BV continuity for the uncentered Hardy-Littlewood maximal operator. *J. Funct. Anal.*, 281(2):Paper No. 109037, 20, 2021.

[GRM21] Cristian González-Riquelme and José Madrid. Sharp inequalities for maximal operators on finite graphs. *J. Geom. Anal.*, 31(10):9708–9744, 2021.

[Haj] Piotr Hajłasz. Jyväskylä lectures on Sobolev spaces. `http://www.pitt.edu/~hajlasz/Notatki/jyv-98b.pdf`.

[HKKT15] Toni Heikkinen, Juha Kinnunen, Janne Korvenpää, and Heli Tuominen. Regularity of the local fractional maximal function. *Ark. Mat.*, 53(1):127–154, 2015.

[HM10] Piotr Hajłasz and Jan Malý. On approximate differentiability of the maximal function. *Proc. Amer. Math. Soc.*, 138(1):165–174, 2010.

[HO04] Piotr Hajłasz and Jani Onninen. On boundedness of maximal functions in Sobolev spaces. *Ann. Acad. Sci. Fenn. Math.*, 29(1):167–176, 2004.

[Kin97] Juha Kinnunen. The Hardy-Littlewood maximal function of a Sobolev function. *Israel J. Math.*, 100:117–124, 1997.

[KKST08] Juha Kinnunen, Riikka Korte, Nageswari Shanmugalingam, and Heli Tuominen. Lebesgue points and capacities via the boxing inequality in metric spaces. *Indiana Univ. Math. J.*, 57(1):401–430, 2008.

[KL98] Juha Kinnunen and Peter Lindqvist. The derivative of the maximal function. *J. Reine Angew. Math.*, 503:161–167, 1998.

[Kor02] Soulaymane Korry. Boundedness of Hardy-Littlewood maximal operator in the framework of Lizorkin-Triebel spaces. *Rev. Mat. Complut.*, 15(2):401–416, 2002.

[KS03] Juha Kinnunen and Eero Saksman. Regularity of the fractional maximal function. *Bull. London Math. Soc.*, 35(4):529–535, 2003.

[KT07] Juha Kinnunen and Heli Tuominen. Pointwise behaviour of $M^{1,1}$ Sobolev functions. *Math. Z.*, 257(3):613–630, 2007.

[Kur15] Ondřej Kurka. On the variation of the Hardy-Littlewood maximal function. *Ann. Acad. Sci. Fenn. Math.*, 40(1):109–133, 2015.

[Lah20] Panu Lahti. On the regularity of the maximal function of a bv function, 2020.

[LM19] Hannes Luiro and José Madrid. The variation of the fractional maximal function of a radial function. *Int. Math. Res. Not. IMRN*, (17):5284–5298, 2019.

[LN14] Hannes Luiro and Juho Nuutinen. On the continuity of discrete maximal operators in Sobolev spaces. *Ann. Acad. Sci. Fenn. Math.*, 39(1):175–185, 2014.

[Lui07] Hannes Luiro. Continuity of the maximal operator in Sobolev spaces. *Proc. Amer. Math. Soc.*, 135(1):243–251, 2007.

[Lui10] Hannes Luiro. On the regularity of the Hardy-Littlewood maximal operator on subdomains of $\mathbb{R}^n$. *Proc. Edinb. Math. Soc. (2)*, 53(1):211–237, 2010.

[Lui18] Hannes Luiro. The variation of the maximal function of a radial function. *Ark. Mat.*, 56(1):147–161, 2018.

[LW15] Feng Liu and Huoxiong Wu. On the regularity of the multisublinear maximal functions. *Canad. Math. Bull.*, 58(4):808–817, 2015.

[LW19] Feng Liu and Huoxiong Wu. A note on the endpoint regularity of the discrete maximal operator. *Proc. Amer. Math. Soc.*, 147(2):583–596, 2019.

[Mad17] José Madrid. Sharp inequalities for the variation of the discrete maximal function. *Bull. Aust. Math. Soc.*, 95(1):94–107, 2017.

[Mad19] José Madrid. Endpoint Sobolev and BV continuity for maximal operators, II. *Rev. Mat. Iberoam.*, 35(7):2151–2168, 2019.

[PPSS18] Carlos Pérez, Tiago Picon, Olli Saari, and Mateus Sousa. Regularity of maximal functions on Hardy-Sobolev spaces. *Bull. Lond. Math. Soc.*, 50(6):1007–1015, 2018.

[RSW20] João P. G. Ramos, Olli Saari, and Julian Weigt. Weak differentiability for fractional maximal functions of general L^p functions on domains. *Adv. Math.*, 368:107144, 25, 2020.

[Tan02] Hitoshi Tanaka. A remark on the derivative of the one-dimensional Hardy-Littlewood maximal function. *Bull. Austral. Math. Soc.*, 65(2):253–258, 2002.

[Tem19] Faruk Temur. Level set estimates for the discrete frequency function. *J. Fourier Anal. Appl.*, 25(3):1008–1025, 2019.

Aalto-DT 134/2022

ISBN 978-952-64-0948-1 (printed)
ISBN 978-952-64-0949-8 (pdf)
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University
School of Science
Department of Mathematics and Systems Analysis
www.aalto.fi

BUSINESS + ECONOMY

ART + DESIGN + ARCHITECTURE

SCIENCE + TECHNOLOGY

CROSSOVER

DOCTORAL THESES