Pointwise convergence, maximal functions and regularity issues in harmonic analysis

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

João Pedro Gonçalves Ramos

aus

Rio de Janeiro, Brasilien.

Bonn 2020

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Christoph Thiele
2. Gutachter: Priv. Doz. Dr. Pavel Zorin-Kranich

Tag der Promotion: 30.01.2020
Erscheinungsjahr: 2020

Dedicated to the loving memory of Maria da Conceição

Abstract

This cumulative thesis is dedicated to the study of different maximal operators related to pointwise convergence in Fourier Analysis and is divided in three main parts.

The first part is dedicated to regularity results for maximal functions. The Hardy–Littlewood maximal function is an essential tool in establishing pointwise convergence in harmonic analysis, and recently more importance has been given to its regularity properties. We make progress in the question of estimating the variation of the maximal function in one dimension, and explore different perspectives of the regularizing properties of fractional maximal functions.

The second part is aimed at maximal versions of classical Fourier restriction theorems. Although the restriction operator has been considered for the past 40 years, it was not until very recently that it was asked whether it can be defined pointwise almost everywhere. We answer this question affirmatively in the two-dimensional case, make progress on the Tomas-Stein exponent case and discuss stronger assertions about Lebesgue points of the Fourier transform.

The third part of this thesis deals with the interplay between Carleson operators and the Hilbert transform along the parabola. An interesting recent conjecture states that the maximally modulated Hilbert transform along the parabola must be bounded in $L^2(\mathbb{R}^2)$. We make partial progress in this question, considering a class of functions essentially constant in directions orthogonal to any fixed line in $\mathbb{R}^2$.

The thesis consists of seven chapters, where Chapters 1 to 6 contain each a scientific article.

In Chapter 0 we develop the historical framework and discuss the motivation for our results, connecting them to the main subject of pointwise convergence and giving a summary of the techniques used.

In Chapter 1 we prove a sharp variation bound for a class of maximal functions interpolating the centered and uncentered maximal functions in one dimension. We also prove a sharp variation bound for Lipschitz truncated uncentered maximal functions. We provide counterexamples showing that our techniques are also sharp.

In Chapter 2 we connect the framework of derivative estimates for fractional maximal functions to Fourier analysis tools. In particular, we prove sharp regularity bounds for certain classes of smooth fractional maximal functions, as well as regularizing bounds for the fractional spherical maximal function.

In Chapter 3 we investigate the regularizing properties of the local fractional maximal function on domains, extending the previous known results to the sharp range in case the domain is smooth enough.

In Chapter 4 we bridge the gap in the recently started line of research of maximal restriction estimates. In particular, we prove that $\mathcal{H}^1-$almost every point in the unit circle is a Lebesgue point of the Fourier transform of an L^p function, $1 \leq p < \frac{4}{3}$.

In Chapter 5 we extend the results in the previous chapter to L^r-norm and spherical Lebesgue points of Fourier transforms of L^p functions. We also devise counterexamples to show sharpness of some of our results and impose restrictions to when the strong maximal function can satisfy full-range maximal restriction estimates.

In Chapter 6 we consider a family of one-dimensional maximally modulated operators arising from the parabolic Carleson operator. We prove uniform bounds in the slope of the line, settling the degenerate case of the conjecture where the Fourier support of the function under consideration collapses into a line.

Acknowledgements

It is indispensable to thank my doctoral advisor Prof. Dr. Christoph Thiele for the priceless guidance throughout the past four years. Since the beginning, Prof. Thiele has been a synonym for a welcoming figure for me, in both personal and professional levels. His patience was of immeasurable help during many turbulent periods researchwise, supporting whenever at the slightest glimpse of lack of motivation. His flexibility and willingness to work were cornerstones to his mentorship, welcoming and sharing excitement at any partial progress, might it be in problems he suggested or in those I started working on my own. On the other hand, his balance of kindness with his pursuit of improvement was always a transforming force within our discussions, aiding to push boundaries of techniques and striving for the sharpest and best way to convey information. His genius never ceased to amaze me, and for all that I am grateful for having had the opportunity to be under his supervision.

Also of immense importance was Prof. Emanuel Carneiro, who from a very early stage served to me as a role model. His brilliant teaching skills and passionate eloquence for talking about mathematics have always been source of inspiration for me and served as guidelines I continuously try to mirror. In addition to all that, his open-mindedness for mathematics has always made it fun to engage in discussions with him. For these and for having awaken in me the taste for Harmonic Analysis, I am extremely thankful.

I would like to thank Felipe Gonçalves not only for the countless hours spent discussing basically any kind of mathematics, but also for being one of my best friends. He and his wife Kate have made my time in Bonn much happier since they arrived, and beyond theorems, the affection and support they gave me throughout the past years are all I could ask for.

My gratitude also goes to Prof. Diogo Oliveira e Silva, who eased the transition when moving to Germany, not only enabling me to speak my mother tongue on a daily basis, but also serving as one of my main sources of mathematical and personal interaction, giving support incessantly, always with a smile on his face.

I am particularly indebted to Olli Saari for being always open to discussions and energetic towards collaborating, what brought about as a consequence Chapters 2 and 3 in this thesis. His friendship has only made my experience in Bonn as a graduate student more pleasant.

I am also very thankful to all my other collaborators, such as David Beltrán, Julian Weigt and Mateus Sousa, for helping me grow and improve professionally and personally. I would need at least another chapter in this thesis to thank all the other people who have contributed meaningfully to the accomplishment of this work, but special thanks go to the Bonn Harmonic Analysis and PDE group for welcoming me and making me feel part of a whole.

Finally, I would like to thank my parents, Roberto and Fátima, my grandparents José and Lúcia, my brother Rafael and the entirety of my family for teaching me the most important things in life. I would not have made it this far without them.

Contents

Chapter 0

Introduction

One of the most fundamental questions in analysis in general is that of *pointwise convergence*. In abstract terms, whenever we are given topological spaces X, Y and a sequence of functions $f_n : X \to Y$, $n \in \mathbb{N}$, we ask when there is another function $f : X \to Y$ such that

$$f_n(x) \to f(x) \text{ for all } x \in X.$$

This question arises naturally in several different areas of mathematics as a tool for understanding underlying properties of mathematical objects. For many analytic purposes, it will suffice to look at a couple of particular instances of pointwise convergence, namely when X is measure space, that is, when it is endowed with a sigma algebra Σ and a measure μ, and the target space $Y = \mathbb{C}$. For that case, pointwise convergence can be generalized to the concept of *almost everywhere convergence.* In that case, we say that a sequence $\{f_n\}_{n\geq 1}$ converges $\mu-$almost everywhere to a function $f : X \to \mathbb{C}$ if there is a set $A \in \Sigma$ such that $\mu(A) = 0$ and

$$f_n(x) \to f(x) \text{ for all } x \in X \setminus A.$$

This notion allows us to explore properties of functions in the measure-theoretic point of view, rather than the more restrictive pointwise one. In concrete terms, we consider the classical example of averages of functions. More specifically, for X a metric measure space, we ask whether the sequence of pointwise averages

$$A_r f(x) := \fint_{B(x,r)} f \, \mathrm{d}\mu \to f(x) \text{ as } r \to 0, \text{ for all } x \in X.$$

The answer to this question, in a general context, demands that the function f possesses a lot in terms of regularity. The asserted convergence is trivial in the case of $f \in C(X)$ continuous, but classical counterexamples evidence that not much more can be said for functions with slightly weaker assumptions.

On the other hand, if we loosen the hypothesis above to demand only that

$$A_r f(x) \to f(x) \text{ for } \mu - \text{almost every } x \in X, \tag{0.1}$$

the answer becomes positive for a much wider class of functions. Indeed, one crucial idea in analysis and related fields is to relate almost everywhere convergence questions as (0.1) to bounds for a suitable *maximal function.* For example, in the case of the problem (0.1), one studies the Hardy–Littlewood maximal function

$$Mf(x) = \sup_{r>0} \fint_{B(x,r)} |f| \, \mathrm{d}\mu \tag{0.2}$$

and its boundedness to conclude almost everywhere convergence. For the case where $X = \mathbb{R}^n$, the classical Hardy–Littlewood–Wiener theorem states that whenever $f \in L^p(\mathbb{R}^n)$ with $p > 1$, then

$$\|Mf\|_p \le C_{p,n}\|f\|_p, \tag{0.3}$$

where the constant C_p depends on p, n but not on the function. Such an inequality automatically implies that the averages converge pointwise to the function. In fact, the convergence stated holds pointwise already for $g \in C(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$, and this class is dense in L^p. If we denote the $n-$dimensional Lebesgue measure by m, it holds that

$$\begin{aligned}
m(\{x \in \mathbb{R}^n \colon |\limsup_{r\to 0} A_r f(x) - &\liminf_{r\to 0} A_r f(x)| > \varepsilon\}) \\
&\le m(\{x \in \mathbb{R}^n \colon |\limsup_{r\to 0}(A_r f - A_r g)(x)| > \varepsilon/3\}) \\
&+ m(\{x \in \mathbb{R}^n \colon |\limsup_{r\to 0} A_r g(x) - \liminf_{r\to 0} A_r g(x)| > \varepsilon/3\}) \\
&+ m(\{x \in \mathbb{R}^n \colon |\liminf_{r\to 0}(A_r f - A_r g)(x)| > \varepsilon/3\}) \\
&\le 2m(\{x \in \mathbb{R}^n \colon M(f-g)(x) > \varepsilon/3\}),
\end{aligned}\tag{0.4}$$

where $g \in C(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$ is arbitrary. But, by Chebyshev's inequality and (0.4), it holds that

$$m(\{x \in \mathbb{R}^n \colon M(f-g) > \varepsilon/3\}) \le \left(\frac{3}{\varepsilon}\right)^p \int_{\mathbb{R}^n} |M(f-g)(x)|^p \, \mathrm{d}x \le \frac{C_p \cdot 3^p}{\varepsilon^p}\|f-g\|_p^p.$$

Taking g such that $\|f-g\|_p \le \varepsilon \cdot \delta$, we obtain that the last display is less than $C_p \cdot 3^p \delta^p$. As $\delta > 0$ was arbitrary, we conclude from (0.4) that

$$m(\{x \in \mathbb{R}^n \colon |\limsup_{r\to 0} A_r f(x) - \liminf_{r\to 0} A_r f(x)| > \varepsilon\}) = 0, \forall \varepsilon > 0,$$

i. e., that $\limsup_{r\to 0} A_r f(x) = \liminf_{r\to 0} A_r f(x)$ almost everywhere. By Minkowski's inequality and continuity of translations in L^p spaces, it holds that

$$A_r f \to f \text{ in } L^p,$$

so that, as the pointwise limit $\lim_{r\to 0} A_r f(x)$ exists, it must be equal to $f(x)$ almost everywhere. Therefore, (0.1) holds for $f \in L^p(\mathbb{R}^n), 1 < p \le \infty$.

Notice that this proof can be modified to the case that the maximal function in (0.2) satisfies only a weak-type inequality of the form

$$\|Mf\|_{p,\infty} = \sup_{\lambda > 0} \lambda m(\{Mf > \lambda\})^{1/p} \le C_{p,n}\|f\|_p.$$

In fact, in the euclidean case, this is the best one can expect in the $p = 1$ case, where it holds that $\|Mf\|_{1,\infty} \le C_n\|f\|_1$, but it can be shown that, whenever $f \not\equiv 0$, then $Mf(x) \ge \frac{C_f}{|x|^n}$, for x sufficiently large and $C_f > 0$ a constant depending on the function f.

The strategy undertaken above for the case of (0.1) is a standard, classical method in analysis. It emphasizes that, whenever we seek pointwise convergence, looking at L^p estimates for maximal functions suffices. Besides the case we discussed about pointwise convergence of averages, this idea has applications in several different subareas of analysis and related fields, among which we mention:

i. the pointwise ergodic theorem [Bir31, AB09], which states that whenever (X, Σ, μ) is a probability space, $T : X \to X$ an ergodic transformation and $f : X \to \mathbb{R}$ is an integrable map, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k x) = \int f \,\mathrm{d}\mu$$

almost everywhere. A particular instance of this theorem is the strong law of large numbers [KW82, Luz18], which states that, given a sequence $\{X_i\}_{i\geq 0}$ of independent, identically distributed random variables with $\mathbb{E}(|X_1|) < +\infty$, it holds that

$$\frac{1}{n}(X_1 + \cdots + X_n) \to \mathbb{E}(X_1)$$

almost surely. The key to proving such an ergodic theorem is a weak-type (1,1) inequality involving the maximal function

$$f^*(x) = \sup_{n\geq 1} \left| \frac{1}{n} \sum_{k=1}^{n} f(T^k x) \right|,$$

whose proof uses strongly the invariance of μ under the action of T;

ii. convergence to the initial datum for partial differential equations, such as the Schrödinger equation

$$\begin{cases} \partial_t u = i\Delta u \text{ in } \mathbb{R}_+ \times \mathbb{R}^n; \\ u(x,0) = u_0(x). \end{cases}$$

The solution to this linear evolution equation will be denoted by $u(x,t) = e^{it\Delta}u_0(x)$, and we seek the smallest value of $s > 0$ so that

$$e^{it\Delta}u_0(x) \to u_0(x) \text{ almost everywhere for each } u_0 \in H^s(\mathbb{R}^n).$$

In this case, there is a plethora of results employing our same underlying principle, such as [CLS19, DGL17, Bou13, Veg88, DZ19] and the references therein. The one-dimensional case, where the relationship between convergence and L^p bounds was first explored, is a result by Carleson [Car80] which states that

$$\| \sup_{t>0} |e^{it\Delta}u_0| \|_{L^4} \leq C \|u_0\|_{H^{1/4}(\mathbb{R})},$$

and we cannot replace $1/4$ by any smaller s;

iii. convergence of Fourier series, as in the celebrated theorem by Carleson [Car66]. For an arbitrary function $f \in L^1(\mathbb{T})$, we define its Fourier coefficients as

$$\widehat{f}(n) = \int_{-1/2}^{1/2} f(x) e^{-2\pi i n x} \,\mathrm{d}x.$$

Classical results state that, if $f \in BV(\mathbb{T})$ is of bounded variation, then the Fourier series

$$S_N f(x) = \sum_{k=-N}^{N} \widehat{f}(n) e^{2\pi i n x} \to f(x)$$

for all x such that f is continuous at x. On the other hand, Kolmogorov's example [Kol23] shows that convergence cannot be expected even at a single point for general $f \in L^1(\mathbb{T})$. The problem of what happens between the case of regular functions $f \in BV$ and of $f \in L^1$ was first stated by Luzin in his doctoral thesis, where he conjectures that, for any $f \in L^2(\mathbb{T})$, it should hold that

$$S_N f(x) \to f(x) \text{ for almost every } x \in \mathbb{T}.$$

This was an open problem for half a century, until Carleson [Car66], considering the maximal function

$$Cf(x) = \sup_{N \in \mathbb{N}} |S_N f(x)|$$

and proving it takes $L^2(\mathbb{T})$ into $L^{2,\infty}(\mathbb{T})$, proved its answer to be affirmative.

Returning to the case of almost everywhere convergence of averages, as in (0.1), we may ask ourselves whether any additional information can be obtained about the exceptional set

$$E_f = \{x \in \mathbb{R}^n \colon A_r f(x) \not\to f(x)\}.$$

For instance, we recall that if $f \in C(\mathbb{R}^n)$, then $E_f = \emptyset$. By Sobolev embedding, any function belonging to the class

$$W^{1,p}(\mathbb{R}^n) = \{g \in L^p(\mathbb{R}^n) \colon \nabla g \in L^p(\mathbb{R}^n)\},$$

with $p > n$, is automatically continuous. On the other hand, if $p < n$, we can only ensure that $W^{1,p} \subset L^{\frac{np}{n-p}}$, and thus the set E_f can be nontrivial. In fact, there are classical examples of functions $h \in W^{1,p}(\mathbb{R}^n)$ whose discontinuity points form a set of Hausdorff dimension $n-p$. The work of Federer and Ziemer [FZ72] shows the converse to this counterexample. That is, for *any* function $f \in W^{1,p}(\mathbb{R}^n)$, the set

$$\tilde{E}_f = \{x \in \mathbb{R}^n \colon A_r f(x) \text{ does not converge }\}$$

has Hausdorff dimension at most $n-p$, and therefore any Sobolev function $f \in W^{1,p}(\mathbb{R}^n)$ can be regularized up to a set of Hausdorff dimension $n-p$. Federer and Ziemer's methods rely heavily on the concept of *Sobolev capacity.* For a set $E \subset \mathbb{R}^n$, we define its Sobolev $p-$capacity as

$$C_p(E) = \inf_{u \in \mathcal{A}(E)} \int_{\mathbb{R}^n} (|u|^p + |\nabla u|^p) \, \mathrm{d}x,$$

where the collection $\mathcal{A}(E)$ is defined to be the set of Sobolev functions $u \in W^{1,p}(\mathbb{R}^n)$ such that $u \geq 1$ is a neighbourhood of the set E. Among many other properties, the $p-$capacity satisfies that, if $C_p(E) = 0$, then $\dim_{\mathcal{H}}(E) \leq n-p$. Therefore, what one needs to prove is that

$$C_p(\tilde{E}_f) = 0 \text{ for } f \in W^{1,p}(\mathbb{R}^n). \tag{0.5}$$

Federer and Ziemer's approach to (0.5) is based upon a thorough decomposition and a geometric-measure theoretic analysis of the exceptional set. We shall follow, however, the insight by Kinnunen [Kin97], as it will lead us to other interesting problems. By repeating the same argument for proving that $A_r f(x) \to f(x)$ for almost every $x \in \mathbb{R}^n$ if $f \in L^p$, we see that proving (0.5) follows from proving a 'weak-type' inequality for the capacity:

$$C_p(\{x \in \mathbb{R}^n \colon Mf(x) > \lambda\}) \leq \frac{C_p}{\lambda^p} \|f\|_{1,p}^p. \tag{0.6}$$

In order to prove (0.6), the idea is simple: if we somehow manage to prove that

$$\text{whenever } u \in W^{1,p}, \text{ then } Mu \in W^{1,p}, \tag{0.7}$$

then Mu belongs automatically to the class $\mathcal{A}(\{x \in \mathbb{R}^n \colon Mf(x) > \lambda\})$. Kinnunen then proceeds not only to prove (0.7), but also the inequality

$$\|Mu\|_{W^{1,p}(\mathbb{R}^n)} \leq C_{p,n}\|u\|_{W^{1,p}(\mathbb{R}^n)}, \text{ whenever } 1 < p \leq +\infty, \tag{0.8}$$

which implies that

$$\begin{aligned} C_p(\{x \in \mathbb{R}^n \colon Mf(x) > \lambda\}) &\leq \frac{1}{\lambda^p}\int_{\mathbb{R}^n}(|Mf|^p + |\nabla Mf|^p)\,\mathrm{d}x \\ &\leq \frac{C_p}{\lambda^p}\int_{\mathbb{R}^n}(|f|^p + |\nabla f|^p)\,\mathrm{d}x = \frac{C_p}{\lambda^p}\|f\|_{1,p}^p. \end{aligned}$$

The first part this thesis is directly related to Kinnunen's inequality (0.8) and its generalizations. In fact, notice that, by the very fact that the maximal function of any non-zero function is not integrable, it holds that $Mf \notin L^1$, so that the endpoint of (0.8) cannot hold as stated. If, however, we dilate the functions in this inequality, we see that (0.8) implies the slightly stronger inequality

$$\|\nabla(Mf)\|_p \leq C_{p,n}\|\nabla f\|_p, \; 1 < p \leq \infty. \tag{0.9}$$

For this inequality, on the other hand, there is no direct obstruction to an endpoint generalization. In a short note, Hajłasz and Onninen [HO04] pose the following question in all dimensions:

Question 0.1 (P.Hajłasz, J. Onninen)**.** *Is the operator $f \to |\nabla Mf|$ bounded from $W^{1,1}(\mathbb{R}^n)$ to $L^1(\mathbb{R}^n)$?*

Besides the scaling argument we gave, Tanaka's result [Tan02] hints at a positive answer to this question. In his manuscript, he proved that the asserted boundedness holds indeed for $n = 1$, with $C = 2$ as an explicit bounding constant:

$$\|(\tilde{M}f)'\|_1 \leq 2\|f'\|_1,$$

where $\tilde{M}f(x) = \sup_{x\in I} \fint_I |f|$ denotes the uncentered Hardy–Littlewood maximal function. In [APL06], the authors proceed to sharpen Tanaka's inequality, proving, again for the uncentered maximal function, that whenever $f \in BV(\mathbb{R})$, then $\tilde{M}f$ is automatically absolutely continuous, and it holds that

$$\|(\tilde{M}f)'\|_1 \leq \|f\|_{BV(\mathbb{R})}.$$

Here, it is interesting to notice that there is an intrinsic difference between the centered and uncentered cases. In the uncentered case, $\tilde{M}f$ has no local maxima in set $\{\tilde{M}f > f\}$, which enables us to easily control the variation. The same behaviour is absent in the centered case, which makes the problem much more complicated. Indeed, the centered case remained open until 2015, when Kurka [Kur15], using an induction-on-scales argument proved that

$$\|Mf\|_{BV(\mathbb{R})} \leq C\|f\|_{BV(\mathbb{R})}, \tag{0.10}$$

where $C = 240,004$. Kurka's argument, although extremely elegant, does not yield the best constant in (0.10). In fact, by thoroughly keeping track and perfecting the steps in his argument, it should be possible to reduce a lot the order of magnitude of C above. This, nonetheless, probably still does not match the conjectured sharp constant:

Conjecture 0.2. *Inequality* (0.10) *holds with* $C = 1$ *for all* $f \in BV(\mathbb{R})$.

The results in Chapter 1 deal with progress towards this conjecture. We define a family of non-tangential operators

$$M^{\alpha} f(x) = \sup_{|x-y| \leq \alpha t} \fint_{y-t}^{y+t} |f(s)| \, \mathrm{d}s$$

which interpolates between the centered and uncentered cases, with $M^0 f(x) = Mf(x)$ and $M^1 f(x) = \tilde{M} f(x)$. Our main result in Chapter 1 is the sharp inequality for certain range of α containing strictly the uncentered case:

Theorem 0.3. *Let* $\alpha \geq \frac{1}{3}$. *It holds that*

$$\|M^{\alpha} f\|_{BV(\mathbb{R})} \leq \|f\|_{BV(\mathbb{R})}, \tag{0.11}$$

for all $f \in BV(\mathbb{R})$, *and this inequality is sharp. Moreover, all extremizers to* (0.11) *are functions* g *such that there exists a point* $x_0 \in \mathbb{R}$ *for which* $g|_{(-\infty, x_0)}$ *is non-decreasing, and* $g|_{(x_0, +\infty)}$ *is non-increasing.*

See Theorem 1.1 for details. The proof of such a result involves proving that $M^{\alpha} f$ does not possess local maxima in the detachment set $\{M^{\alpha} f > f\}$ if $\alpha > \frac{1}{3}$. This, in turn, is implied by proving that a suitable truncated uncentered maximal function coincides with $M^{\alpha} f$ in a neighbourhood of its local maximum, and then resorting back to the fact that the truncated maximal function also fulfills the property of not having local maxima in the detachment set. The condition $\alpha \geq \frac{1}{3}$ comes in a very geometric way when proving that the maximal functions coincide, and Theorem 1.2 shows that this is the furthest one can attain with this method. I.e., for any $\alpha < \frac{1}{3}$, there is a function $f_{\alpha} \in BV(\mathbb{R})$ so that $M^{\alpha}(f_{\alpha})$ has a non-trivial local maximum in the detachment set $\{M^{\alpha}(f_{\alpha}) > f_{\alpha}\}$.

In the subsequent chapters, we continue our investigation of regularity properties of maximal functions. Besides the classical Hardy–Littlewood maximal function, an important maximal function in the literature is its *fractional* variant, given by

$$M_{\alpha} f(x) = \sup_{r>0} r^{\alpha} \fint_{B(x,r)} |f(y)| \, \mathrm{d}y. \tag{0.12}$$

This maximal function is closely related to the Riesz potentials

$$I_{\alpha} f(x) = \int_{\mathbb{R}^n} \frac{|f(y)|}{|x-y|^{n-\alpha}} \, \mathrm{d}y, \tag{0.13}$$

as mentioned in [KS03]. Indeed, it is easy to show that $M_{\alpha} f(x) \leq \frac{1}{|B(0,1)|} I_{\alpha} f(x)$ pointwise, and, although the reverse pointwise inequality does not hold, a result by Muckenhoupt and Wheeden [MW74] proves that the converse holds 'in average': for each $p \in (1, +\infty)$, there is a constant $c_{p,\alpha}$ so that

$$c_{p,\alpha} \|M_{\alpha} f\|_p \leq \|I_{\alpha} f\|_p \leq c_{p,\alpha}^{-1} \|M_{\alpha} f\|_p.$$

It hints at the fact that sometimes working with the fractional maximal function is more convenient than with the Riesz potential. Indeed, Hedberg [Hed72] uses this fractional maximal function to control pointwise the oscillation of Sobolev functions, and in general it is expected that this is, despite the presence of a supremum, easier to work with than the Riesz potential. We shall, however, be interested in a slightly different facet of this maximal function: an inherently regularizing effect. Our motivation is the Riesz potential I_1, for which a calculation shows that

$$\partial_i(I_1 f(x)) = -R_i f(x)$$

almost everywhere, where R_i is the $i-$th Riesz transform of f. As Riesz transforms are bounded in $L^p, 1 < p < +\infty$, we see that I_1 maps L^p into some sort of Sobolev space. As we discussed, the Riesz potential and the fractional maximal function are intimately related to one another, so that one wonders directly whether the same conclusion can be drawn for the latter. The work of Kinnunen and Saksman [KS03] is the first partial answer to this question, where they prove, among other results, that, if $f \in L^p(\mathbb{R}^n)$, $n/p > \alpha \geq 1$, then the pointwise inequality

$$|\nabla(M_\alpha f)| \leq C_{p,n,\alpha}|M_{\alpha-1}f| \tag{0.14}$$

holds. As $M_{\alpha-1} \leq c_n I_{\alpha-1}$ and as Riesz potentials are bounded from L^p to $L^q, q = \frac{np}{n-\alpha p}$, we conclude that

$$\|\nabla(M_\alpha f)\|_{q^*} \leq C_{p,n,\alpha}\|M_{\alpha-1}f\|_{q^*} \leq \tilde{C}_{p,n,\alpha}\|f\|_p, \tag{0.15}$$

with $q^* = \frac{np}{n-(\alpha-1)p}$. On the other hand, the Gagliardo–Nirenberg–Sobolev inequality states that

$$\|f\|_{L^{\frac{n}{n-1}}} \leq C_n\|\nabla f\|_{L^1},$$

so that, if $q' = \frac{n}{n-\alpha}$, $\alpha \geq 1$, inequality (0.15) implies

$$\|\nabla(M_\alpha f)\|_{q'} \leq C_{n,\alpha}\|\nabla f\|_1. \tag{0.16}$$

The remark leading to (0.16) was first made by Carneiro and Madrid [CM17], who additionally considered the one-dimensional *uncentered* fractional maximal function

$$\tilde{M}_\alpha f(x) = \sup_{I \ni x} \frac{1}{|I|^{1-\alpha}} \int_I |f(s)| \, \mathrm{d}s$$

for $0 \leq \alpha < 1$. They proved that, also in this case, for $f \in BV(\mathbb{R})$,

$$\|(\tilde{M}_\alpha f)'\|_{\frac{1}{1-\alpha}} \leq 8^{1-\alpha}\|f\|_{BV(\mathbb{R})}.$$

This raises the question whether inequality (0.16) holds in the higher dimensional setting, for either the centered or uncentered fractional maximal function, with $\alpha \in [0, 1)$. The first result of Chapter 2 is the positive answer to this question for fractional maximal functions with additional smoothness. Namely, if we consider either the *smooth fractional maximal function* given by

$$M_\alpha^\varphi f(x) = \sup_{t>0} t^\alpha |\varphi_t * f(x)| = \sup_{t>0} t^{\alpha-n} \left| \int_{\mathbb{R}^n} \varphi\left(\frac{x-y}{t}\right) f(y) \, \mathrm{d}y \right|,$$

for φ a positive, Schwartz function, or the *lacunary fractional maximal function* given by

$$M_\alpha^{lac} f(x) = \sup_{k \in \mathbb{Z}} 2^{(\alpha-n)k} \left| \int_{B(x,2^k)} f(y) \, \mathrm{d}y \right|,$$

then our first main result in Chapter 2 reads as follows.

Theorem 0.4. *Let $f \in BV(\mathbb{R}^n)$ and suppose that $\alpha \in (0,1)$ and $n \geq 2$. Then for $\mathcal{M}_\alpha \in \{M_\alpha^{lac}, M_\alpha^{\varphi}\}$, there exists a constant C only depending on dimension n, α and φ such that*

$$\|\nabla \mathcal{M}_\alpha f\|_{L^p(\mathbb{R}^n)} \leq C\|f\|_{BV(\mathbb{R}^n)}$$

for $p = n/(n-\alpha)$.

The stretegy of the proof of Theorem 0.4 is to relate the derivative of the fractional maximal functions involved to a suitable maximal Fourier multiplier theorem. Indeed, a main idea is to try to adapt the steps of the proof of (0.16). We need to bypass the domination (0.15), as there is no reasonable way to make sense of M_β when $\beta < 0$. At that point we make use of the connection between maximal functions and Fourier multipliers to pass to a purely Fourier-analytic problem. With classical and modern tools from Fourier analysis at our disposal, such as $g-$function techniques, we prove an inequality bounding the size of $\nabla \mathcal{M}_\alpha f$ by a certain Besov norm of f, which can be dominated, by an argument from [CDDD03], by $\|f\|_{BV}$.

Interestingly, the smoothness conditions on our maximal functions only play a role in the single scale decay for our maximal functions. Smoothness of φ provides additional decay because of smoothness on the Fourier side, while having lacunary sets of radii prevents the scales taken in the maximal functions from interacting with one another, providing us with better bounds.

The second main result in Chapter 2 is a result analogous to (0.15), but for the *spherical* fractional maximal function given by

$$S_\alpha f(x) = \sup_{t>0} t^\alpha \left| \fint_{\partial B(x,t)} f(y)\, \mathrm{d}\sigma(y) \right|.$$

This maximal function has been considered outside the context of derivative bounds previously. In the $\alpha = 0$ case, the main results are due to by Stein [Ste76] and Bourgain [Bou86], proving that these are bounded in $L^p(\mathbb{R}^n)$ for $p > \frac{n}{n-1}$, and this is sharp. For $\alpha > 0$, we mention mainly the works of Schlag [Sch97] and Schlag and Sogge [SS97]. In the context of derivative bounds in the spirit of 0.15, however, no result was known. This was mainly due to the fact that Kinnunen and Saksman's proof from [KS03] uses too heavily the geometric structure of balls to obtain the domination (0.14). In the case of the spherical maximal function, objects are much more singular, and geometric attempts fall apart. The next result represents therefore the first alternative approach to bounds for the derivative of such maximal functions, using Fourier analysis.

Theorem 0.5. *Let $n \geq 5$, $n/(n-2) < p \leq q < \infty$ and*

$$\alpha(p) := \begin{cases} \frac{n^2-2n-1}{n-1} - \frac{2n}{p(n-1)} & \text{if} \quad \frac{n}{n-2} < p \leq \frac{n^2+1}{n^2-2n-1} \\ \frac{n-1}{p} & \text{if} \quad \frac{n^2+1}{n^2-2n-1} < p \leq n-1. \end{cases}$$

Assume that

$$\frac{1}{q} = \frac{1}{p} - \frac{\alpha - 1}{n}, \qquad 1 \leq \alpha < \alpha(p).$$

Then, for any $f \in L^p$, $S_\alpha f$ is weakly differentiable and

$$\|\nabla S_\alpha f\|_{L^q} \leq C_{p,n,\alpha}\|f\|_{L^p}.$$

In order to prove this result, besides passing to the Fourier side to take advantage of the multiplier properties of the Fourier transform of the spherical measure, we need to employ additional techniques in comparison to the proof of Theorem 0.4. In fact, differentiating only makes the regularity of the Fourier multiplier we obtain *worse*, so that we need to study thoroughly single-scale estimates for the spherical maximal function. We achieve better bounds by making use of sharp local smoothing estimates for the wave equation, which, in turn, have only recently become available for $n \geq 5$, thanks to the development of decoupling inequalities (see, e.g., [BD15, GS09, GS10, ŁW02, Wol00] and [BHS, HNS11, LS13, MSS92, Sog91]).

Finally, the last chapter in the first part of this thesis continues to exploit regularity of the fractional maximal function, but this time in the local setting. While, on the one hand, we have inequality 0.14 for the full fractional maximal function M_α, the domain case poses additional difficulties. Namely, if we let the *local* fractional maximal function associated to Ω be

$$M_\alpha^\Omega f(x) = \sup_{0<r<\operatorname{dist}(x,\Omega^c)} r^\alpha \fint_{B(x,r)} |f(y)| \, \mathrm{d}y, \tag{0.17}$$

Heikkinen, Kinnunen, Korvenpää and Tuominen [HKKT15] prove a weaker version of (0.14):

$$|\nabla M_\alpha^\Omega f(x)| \leq c_{\alpha,n} \left(M_{\alpha-1}^\Omega f(x) + S_{\alpha-1} f(x) \right). \tag{0.18}$$

This only enables us to obtain results like (0.15) for $p > n/(n-1)$, as $\alpha \geq 1$ is assumed in their argument, and $\frac{n}{n-1}$ is the least integrabilty condition so that any fractional spherical maximal function $S_{\alpha-1}$ has good enough boundedness properties. The main goal of Chapter 3 is to extend inequality (0.15) to the domain context up until the endpoint $p = 1$.

Theorem 0.6. *Let $\Omega \subset \mathbb{R}^n$ be open, $n \geq 2$, $p > 1$ and $f \in L^p(\Omega)$. Then $M_\alpha^\Omega f$ is weakly differentiable and*

$$\|\nabla M_\alpha^\Omega f\|_{L^p(\Omega)} \leq C \|f\|_{L^p(\Omega)}$$

if any one of the following holds:

- *$\alpha > 1$ and Ω is bounded.*
- *$\alpha = 1$ and Ω^c is convex.*
- *$\alpha = 1$ and Ω is bounded and satisfies a uniform interior ball condition (see Section 3.2.2 for the definition).*
- *$\alpha = 1$ and $p > 1 + \frac{1}{n}$.*

The constant C depends on the dimension, and in the first and third items it also depends on α and the domain.

For the purpose of proving Theorem 0.6, we need a better bound than (0.18). It is crucial to distinguish two cases: when the radius of the ball attaining the sup in the definition of (0.17) is *strictly* less than $\operatorname{dist}(x, \partial\Omega)$, and the case where the maximal function $M_\alpha^\Omega f$ goes all the way until the boundary to attain the supremum. It was shown in [HKKT15, Example 4.1] that the latter is in general not negligible; that is, the set of unconstrained points

$$\left\{ x \in \Omega \colon M_\alpha^\Omega f(x) = \delta(x)^\alpha \fint_{\partial B(x,\delta(x))} |f(y)| \, \mathrm{d}\sigma(y) \right\}$$

for $\delta(x) = \mathrm{dist}(x, \Omega)$, may have positive measure. This set is the one where the authors obtain the $S_{\alpha-1}$−term in (0.18). In order to improve on their method, we study more carefully the derivative of the operator

$$f \mapsto \delta(x)^{\alpha} \fint_{\partial B(x,\delta(x))} |f(y)| \,\mathrm{d}\sigma(y).$$

This turns out to be almost everywhere bounded by $M_{\alpha-1}$ plus a *weighted* spherical average of the form

$$\delta(x)^{\alpha-1} \fint_{\partial B(x,\delta(x))} \frac{|y-b_x|}{\delta(x)} |f(y)| \,\mathrm{d}\sigma(y), \tag{0.19}$$

where b_x is a point on $\partial\Omega$ such that $|x - b_x| = \mathrm{dist}(x, \partial\Omega)$. The term $\frac{|y-b_x|}{\delta(x)}$ is essential here, as it prevents the operator (0.19) from becoming too large near the boundary. The analysis of this operator consists in decomposing the integral defining it into pieces where $\frac{|y-b_x|}{\delta(x)} \sim 2^{-j}$. In order to prove the result for $L^p, p > 1$, we bound each of the pieces in L^∞ and L^1. The L^∞ bounds are the trivial ones, whereas the method for the L^1 bounds can be described as finding the proper substitute for Fubini's theorem. At least morally, the dual to each of the localized pieces is an averaging operator over

$$P_j(y) = \left\{ x \in \Omega \colon \mathrm{dist}(y, x) = \delta(x) \text{ and } \frac{|y-b_x|}{\delta(x)} \sim 2^{-j} \right\}.$$

A crucial observation then becomes that each of the $P_j(y)$, with y fixed, are convex sets. This imposes good bounds on their perimeters, which, on the other hand, implies that each piece in our decomposition possesses good enough L^1 decay to sum up. The main consequence of Theorem 0.6 is an analogue of (0.16) in the domain case:

Corollary 0.7. *Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain, which is either bounded and satisfies the interior ball condition, or such that Ω^c is convex. Then for all $f \in W^{1,1}(\Omega)$ it holds that*

$$\|\nabla M_1^{\Omega} f\|_{L^{n/(n-1)}(\Omega)} \leq C \|f\|_{W^{1,1}(\Omega)}, \tag{0.20}$$

where the constant C only depends on Ω and the dimension.

This inequality was previously out of reach by the methods in [HKKT15]. It would be interesting to replace the $W^{1,1}(\Omega)$ norm on the right hand side of (0.20), as the proof of this corollary only makes use of the full $W^{1,1}$−norm when passing from $\|\nabla E^{\Omega}(f)\|_{L^1(\mathbb{R}^n)}$ to $\|f\|_{W^{1,1}}$, where E^{Ω} denotes the extension operator. This, however, is currently out of the scope of our methods.

The second part of this thesis is dedicated to problems closer in flavour to Example B and partial differential equations. In fact, this is the main motivation behind the so-called Strichartz estimates, initiated by Robert Strichartz in the celebrated paper from 1977 [Str77]. These are the basic setup to relate estimates for PDEs to Fourier restriction theory. Take, for instance, the already discussed case of the Schrödinger equation:

$$\begin{cases} \partial_t u = i\Delta u \text{ in } \mathbb{R}^n \times \mathbb{R}_+; \\ u(x,0) = u_0(x). \end{cases}$$

Its solution is induced by a linear group $u(x,t) = e^{it\Delta}u_0(x)$, where we define the group

$$e^{it\Delta}u_0(x) = \int_{\mathbb{R}^n} e^{2\pi i x\xi} e^{-8i\pi^2 t|\xi|^2} \widehat{u_0}(\xi)\, \mathrm{d}\xi \tag{0.21}$$

via Fourier inversion. By abusing notation and the good will of the reader, we may regard the expression in (0.21) as a formal expression like

$$\int_{\mathbb{R}^n\times\mathbb{R}} e^{2\pi i x\xi} e^{2\pi i t\tau} \widehat{u_0}(\xi)\delta_0(\tau + 4\pi^2|\xi|^2)\, \mathrm{d}\xi\, \mathrm{d}\tau. \tag{0.22}$$

Here, δ_0 is the Dirac delta distribution. This last display resembles too much a $(n+1)-$dimensional Fourier transform, with the exception now that we are no longer manipulating a function, but the "distribution' $\widehat{u_0}(\xi)\delta_0(\tau + 4\pi^2|\xi|^2) = \mu_0(\xi,\tau)$. This distribution admits a rigorous definition: in fact, we simply let

$$\int_{\mathbb{R}^n\times\mathbb{R}} g(\xi,\tau)\, \mathrm{d}\mu_0(\xi,\tau) = \int_{\mathbb{R}^n} g(\xi, -4\pi^2|\xi|^2)\widehat{u_0}(\xi)\, \mathrm{d}\xi,$$

whenever $gin\mathcal{S}(\mathbb{R}^n\times\mathbb{R})$. Among other properties, the most relevant for us now is that the support of μ_0 is contained in the paraboloid $\mathbb{P} = \{(\xi,\tau) \colon \tau = -4\pi^2|\xi|^2\}$. Denoting by $\sigma_0(\xi,\tau) = \delta_0(\tau + 4\pi^2|\xi|^2)$, there is a close relationship between (0.22) and

$$\widehat{g\, \mathrm{d}\sigma_0}(x,t) = \int_{\mathbb{R}^n} g(\xi, -4\pi^2|\xi|^2) e^{2\pi i x\xi - 8i\pi^2 t|\xi|^2}\, \mathrm{d}\xi,$$

for a function g supported on the paraboloid $\mathbb{P}$. In other words, in the context of dispersive partial differential equations, we wish to obtain bounds like

$$\|e^{it\Delta}u_0\|_{L^p(\mathbb{R}^{n+1})} \le C\|u_0\|_2, \tag{0.23}$$

and our considerations have shown that this is tantamount to proving

$$\|\widehat{g\, \mathrm{d}\sigma_0}\|_{L^p(\mathbb{R}^{n+1})} \le C\|g\|_{L^2(\mathbb{P})}. \tag{0.24}$$

Inequality (0.24) is what we call an *extension estimate.* Its dual problem is an instance of the famous *restriction problem* for the Fourier transform:

$$\|\widehat{F}\|_{L^2(\mathbb{P})} \le C\|F\|_{L^{p'}(\mathbb{R}^{n+1})}. \tag{0.25}$$

While inequalities like (0.24) and (0.23) are mainly useful for proving existence results for the underlying PDE, the formulation in (0.25) makes the problem interesting on its own: by the Riemann–Lebesgue lemma, we know that $f \in L^1(\mathbb{R}^n) \Rightarrow \widehat{f} \in C_0(\mathbb{R}^n)$. Therefore, restricting $\widehat{f}$ to any subset of $\mathbb{R}^n$ makes sense. On the other hand, the Fourier transform is an isometric isomorphism in $L^2(\mathbb{R}^n)$, so that it does not make sense to restrict $\widehat{f}$, in general, to any subset of $\mathbb{R}^n$, no matter how small. The question becomes: what happens for $f \in L^p(\mathbb{R}^n), 1 < p < 2$?

We compare two opposite instances. Taking the function $f(x_1,\ldots,x_n) = \frac{\psi(x_2,\ldots,x_n)}{1+|x_1|}$, $\psi : \mathbb{R}^{n-1} \to \mathbb{R}$ smooth, implies $f \in L^p(\mathbb{R}^n), \forall p > 1$, but $\widehat{f} \equiv +\infty$ for $\xi_1 = 0$. Rotating and modulating this function appropriately yields that, for any hyperplane $S \subset \mathbb{R}^n$, there is no meaningful way to restrict $f \in L^p(\mathbb{R}^n), p > 1$. On the other end of the spectrum, if

$S \subset \mathbb{R}^n$ is a set of positive, finite Lebesgue measure, restricton to S can be easily defined. In fact, the Hausdorff–Young inequality implies that

$$\|\widehat{f}\|_{L^q(S)} \le \|\widehat{f}\|_{L^{p'}(\mathbb{R}^n)} |S|^{1/(p'/q)'} \le C_S \|f\|_{L^p},$$

whenever $p' > q$. This inequality holds for $f \in \mathbb{S}(\mathbb{R}^n)$, where $\widehat{f}|_S$ is already defined. We define the operator $f \mapsto \mathcal{R}_S(\widehat{f})$ by density for $f \in L^p(\mathbb{R}^n), 1 < p < 2$, using the definition on Schwartz functions.

A question stemming from this discussion is what happens when S is neither of positive, finite measure, nor contained in a hyperplane. For simplicity, let us look at sets $S \subset \mathbb{R}^n$ with a little more smoothness. That is, let us suppose that S is, in fact, a $(n-1)-$hypersurface, and additionally assume that S does not resemble a hyperplane "too much"; that is, let us assume that it has everywhere non-vanishing curvature. The most basic – and in fact, for us, the almost-unique – example is the $(n-1)-$dimensional sphere $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$. Following the strategy employed in the previous paragraph, we look for inequalities of the form

$$\|\widehat{f}\|_{L^q(\mathbb{S}^{n-1})} \le C \|f\|_{L^p(\mathbb{R}^n)}, \tag{0.26}$$

for some $p, q \ge 1$. The analogue of (0.24) in this case is the inequality

$$\|\widehat{f\,\mathrm{d}\sigma_{n-1}}\|_{L^{p'}(\mathbb{R}^n)} \le C \|f\|_{L^{q'}(\mathbb{S}^{n-1})}, \tag{0.27}$$

with $\mathrm{d}\sigma_{n-1}$ denoting the normalized surface measure of the $(n-1)-$dimensional unit sphere. The main examples to be mentioned here are the following:

i. taking $f \equiv 1$ and using the decay properties of the Fourier transform $\widehat{\sigma_{n-1}}$, we obtain that $p < \frac{2n}{n+1}$;

ii. taking $f = 1_{C_\delta}$, where C_δ denotes a spherical cap of radius $\sim \delta$, and sending $\delta \to 0$, we obtain $p' \ge \frac{n+1}{n-1} q$.

The famous *restriction conjecture* asserts that conditions A and B are also sufficient for inequality (0.27) to hold.

The first instance of an inequality like (0.26) in the literature is in a Lemma by Stein, mentioned in a paper by Fefferman [Fef70], where he proves the $q = 2, 1 \le p < \frac{4n}{3n+1}$ case of such an inequality. Fefferman himself improves on Stein's lemma in two dimensions. One of the consequences of Fefferman's strategy is that (0.26) holds as long as $1 \le p < \frac{4}{3}$, $p' > 3q$, settling the conjecture aside from the $p' = 3q$ case in dimension two. In three dimensions and higher, the next major breakthrough is Tomas's contribution [Tom75], settling the $q = 2$ case of the conjecture completely by extending Stein's range to $1 \le p < \frac{2(n+1)}{n+3}$. Later on, Stein would again contribute to the $q = 2$ case by employing complex interpolation methods to solve the $p = \frac{2(n+1)}{n+3}$ endpoint.

In dimensions $n \ge 3$, the full-range restriction conjecture is still open, with progress on the admissible range of exponents being continuously made; see, for instance, [Tao04, Gut14, Gut18, HR18, Wan18]. The only dimension for which the conjecture has been settled is $n = 2$. After Fefferman's argument [Fef70], several authors worked on the remaining endpoint case, among which we mention Zygmund [Zyg74] and Carleson and Sjölin [CS72]. Later on, Sjölin [Sjö74] would extend the sharp result to a wider class of curves.

Our attention in the second part of the thesis will shift from the possibility of defining a restriction operator to qualitative properties of such a restriction. Inequality (0.26) only enables us to define the restriction $\mathcal{R}_{\mathbb{S}^{n-1}}(\widehat{f})$ as an L^q-limit. The question remains about how to make this definition a *pointwise* one, in the main spirit of this thesis. Inspired by that, Müller, Ricci and Wright [MRW19] were the first ones to obtain positive results in this direction. In a nutshell, the main consequence of their results is that, if $1 \leq p < \frac{8}{7}$, then the regularized values of $\widehat{f}$ agree $\mathcal{H}^1_{\mathbb{S}^1}-$almost everywhere with $\mathcal{R}_{\mathbb{S}^1}(\widehat{f})$. In their methods, however, they analyze the maximal function

$$\mathcal{M}(\widehat{f})(x) = \sup_{r>0} \left| \fint_{B(x,r)} \widehat{f}(y) \, \mathrm{d}y \right|$$

and prove that it satisfies $L^p(\mathbb{R}^2) \to L^q(\mathbb{S}^1)$ bounds in the full $1 \leq p < \frac{4}{3}$, $p' \geq 3q$ two-dimensional restriction range. The only obstacle standing on their way to prove that the restriction operator can be defined almost everywhere in the entire $1 \leq p < \frac{4}{3}$ range rather than $p \in [1, \frac{8}{7})$ is in their linearization trick: letting M denote the centered Hardy–Littlewood maximal function, we have

$$M(\widehat{f}) \leq M(|\widehat{f}|^2)^{1/2} = \mathcal{M}(f * \tilde{f})^{1/2}.$$

Here, $\tilde{f}(x, y) = f(-x, -y)$. In the end, they need $p < \frac{8}{7}$ exactly in order to bound $\|f * \tilde{f}\|_{\tilde{p}} \leq \|f\|_p$, with $\tilde{p} < \frac{4}{3}$. The first and main result of Chapter 4 settles the problem of defining restriction pointwise in the whole range $1 \leq p < \frac{4}{3}$.

Theorem 0.8. *Let $M_R g(x) = \sup_{R \in \mathcal{R}(x)} \fint_R |g(y)| \, \mathrm{d}y$ denote the strong maximal function, where $\mathcal{R}(x)$ denotes the set of axis-parallel rectangles centered at the point x. It holds that*

$$\|M_R(\widehat{f})\|_{L^q(\mathbb{S}^1)} \leq C_p \|f\|_{L^p(\mathbb{R}^2)},$$

where $1 \leq p < \frac{4}{3}$, $p' \geq 3q$. In particular, the regularized values of $\widehat{f}$ coincide $\mathcal{H}^1_{\mathbb{S}^1}-$almost everywhere with $\mathcal{R}_{\mathbb{S}^1}(\widehat{f})$ for $f \in L^p(\mathbb{R}^n), 1 \leq p < \frac{4}{3}$.

As the linearization is where the approach in [MRW19] fails, we develop a new linearization method. We simply write $|\widehat{f}(z)| = \widehat{f}(z) \cdot \frac{\overline{\widehat{f}(z)}}{|\widehat{f}(z)|}$. The function $g(z) = \frac{\overline{\widehat{f}(z)}}{|\widehat{f}(z)|} \in L^\infty(\mathbb{R}^2)$, and thus we are led to consider the weighted maximal operator

$$M_g(\widehat{f})(x) = \sup_{R \in \mathcal{R}(x)} \left| \fint_R \widehat{f}(z) g(z) \, \mathrm{d}z \right|.$$

The point of the proof is then to run the main argument of [Zyg74, CS72] and [MRW19], considering $g \in L^\infty$ to be fixed beforehand. The structure of the exponents in the two-dimensional case allows us to obtain a bound depending only on $\|g\|_\infty$, and the result follows by choosing $g = \frac{\overline{\widehat{f}}}{|\widehat{f}|}$ a posteriori.

Theorem 0.8 is not the only one devoted to generalizing the work of Müller, Ricci and Wright. The results in [MRW19] deal only with the case of curves and, although they can be adapted to higher-dimensional curves, they seem to be tethered to one-dimensional objects. In [Vit17], Vitturi provided a first approach to the problem of maximal restriction in higher dimensions, considering the Tomas-Stein exponent case $1 \le p \le \frac{4}{3}$ in dimension 3, and proving that

$$\|\mathcal{M}(\widehat{f})\|_{L^q(\mathbb{S}^2)} \le C\|f\|_{L^p(\mathbb{R}^3)},$$

whenever $1 \le p \le \frac{4}{3}$, $p' \ge 2q$. The exponent $4/3$ is crucial here, as $(4/3)' = 4$ is an *even* exponent, which allows for the use of additional tools such as the Plancherel identity. Subsequently, Kovač and Oliveira e Silva [KOeS18] generalized Vitturi's results to the variational context, substituting the maximal function $\mathcal{M}$ by a less smooth variation operator, which gives rise to ways of quantifying pointwise convergence.

The most striking result following the work [MRW19] is an abstract implication principle by Kovač [Kov19]. Following an abstract argument connecting classical variational inequalities with a Christ-Kiselev-type argument, Kovač proves that, whenever a restriction estimate of the kind

$$\|\widehat{f}\|_{L^q(S)} \le C\|f\|_{L^p(\mathbb{R}^n)}$$

holds with $p < q$, and we are given a complex measure μ such that

$$|\nabla\widehat{\mu}(\xi)| \le D(1 + |\xi|)^{-1-\eta}, \tag{0.28}$$

for some $D, \eta > 0$, then it holds *automatically* that

$$\|\sup_{t>0} |(\widehat{f} * \mu_t)|\|_{L^q(S)} \le C_{D,\eta,p,q,n}\|f\|_{L^p(\mathbb{R}^n)}. \tag{0.29}$$

Kovač's results are actually even more general than (0.29), enabling one to substitute the supremum by a variation operator on the left hand side. His techniques are surprising, for they depart from the previous line of thought employed for maximal restriction estimates. Until [Kov19], the main strategy for proving inequalities like (0.29) was repeating the proof of the original restriction estimate, dualizing to make use of the maximal function at some point. Kovač, on the other hand, only uses abstract properties of the Fourier transform and of the exponents involved, enabling one to obtain much more striking consequences than the previous ones.

Despite its striking nature, we highlight two minor deficiencies in Kovač's result. The first one is the lack of results about *strong* maximal restriction inequalities of the form

$$\|\sup_{t>0}(|\widehat{f}| * \mu_t)\|_{L^q(S)} \le C_{D,\eta,p,q,n}\|f\|_{L^p(\mathbb{R}^n)}, \tag{0.30}$$

in the spirit of [Ram18] and Chapter 4. This is due to the necessity of writing $\widehat{f} * \psi = \widehat{(f \cdot \widehat{\psi})}$ in the proof of (0.29). The second main deficiency is in the condition (0.28). While it is satisfied by $\mathrm{d}\mu = 1_{B(0,1)}\,\mathrm{d}x$, implying abstract maximal restriction principles involving the maximal function $\mathcal{M}$ above, the maximal *spherical* restriction estimates for

$$\mathcal{S}(\widehat{f})(x) = \sup_{t>0}\left|\fint_{\partial B(x,t)} \widehat{f}(y)\,\mathrm{d}\sigma(y)\right|$$

are only available from Kovač's theorem for $n \ge 4$, as the measure $\mathrm{d}\mu = \mathrm{d}\sigma_{n-1}$ only satisfies (0.28) in that dimension. This is the main motivation for the results in Chapter

5. There, we extend Kovač's maximal restriction results as (0.30) to a wider class of measures in dimension 2, as well as adapt Vitturi's techniques to the three dimensional Tomas-Stein case for a class of measures including the spherical maximal function.

Theorem 0.9.

i. *Let μ be a positive, finite Borel measure defined in $\mathbb{R}^2$, and suppose that the maximal function*

$$M_\mu g(x) := \sup_{t>0} |g| * \mu_t(x).$$

is bounded from $L^r(\mathbb{R}^2) \to L^r(\mathbb{R}^2)$, whenever $r > 2$. Then the following bound holds:

$$\|M_\mu(\widehat{f})\|_{L^q(\mathbb{S}^1)} \le C_{p,\mu}\|f\|_{L^p(\mathbb{R}^2)},$$

where $1 \le p < \frac{4}{3}, p' \ge 3q$.

ii. *Let Let μ be a positive, finite Borel measure defined in $\mathbb{R}^3$, and suppose that the maximal function*

$$M_\mu g(x) := \sup_{t>0} |g| * \mu_t(x).$$

is bounded from $L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$. Then the following bound holds:

$$\|M_\mu(\widehat{f})\|_{L^2(\mathbb{S}^2)} \le C_{p,\mu}\|f\|_{L^p(\mathbb{R}^3)},$$

where $1 \le p \le \frac{4}{3}$.

Another consequence of our techniques is a proof that the part where the methods in [MRW19] fail to prove the full range of the maximal two-dimensional restriction is *only* the use of Young's inequality. Recapitulating the main ideas, the authors use in [MRW19] that

$$M(\widehat{f}) \le M(|\widehat{f}|^2)^{1/2} = \mathcal{M}(f * \tilde{f})^{1/2}$$

in order to prove a pointwise restriction property for the Fourier transform of functions in L^p, $1 \le p < \frac{8}{7}$. The second main result in Chapter 5 shows that also the maximal function

$$M_2(\widehat{f}) := M(|\widehat{f}|^2)^{1/2}$$

satisfies the same restriction inequalities in the two-dimensional setting, and all Tomas-Stein inequalities with exception of the endpoint in the three-dimensional one.

Theorem 0.10.

i. *Let $1 \le r \le 2$. Define the maximal functions $M_r h(x) := (M(|h|^r)(x))^{1/r}$. The following bound holds:*

$$\|M_r(\widehat{f})\|_{L^q(\mathbb{S}^1)} \le C_{p,r}\|f\|_{L^p(\mathbb{R}^2)},$$

where $1 \le p < \frac{4}{3}, p' \ge 3q$.

ii. *Let $1 \le r < 2$. Then the following bound holds:*

$$\|M_r(\widehat{f})\|_{L^2(\mathbb{S}^2)} \le C_{p,r}\|f\|_{L^p(\mathbb{R}^3)},$$

where $1 \le p \le \frac{4}{3}$. Aditionally, the quadratic maximal function M_2 satisfies that

$$\|M_2(\widehat{f})\|_{L^2(\mathbb{S}^2)} \le C_{p,r}\|f\|_{L^p(\mathbb{R}^3)},$$

whenever $1 \le p < \frac{4}{3}$.

The main tool to prove Theorems 0.9 and 0.10 is a combination of the methods of Chapter 4 with a suitable use of a Plancherel trick that allows one to dualize, in order to make use of the boundedness of the operators M_μ and M_r directly. The idea is to devise specific inearizations adapted to $\mu-$ and $r-$averages, extending therefore the main idea of Theorem 0.8.

In the end of Chapter 5 we investigate sharpness of the restrictions in r in Theorem 0.10 by adapting in two different forms the classical Knapp example. Perhaps the most notorious feature is that constructing the Knapp example as a cylinder with varying heights yields different ranges of counterexamples. One special case is that of a hybrid strong $r-$maximal function

$$M_{R,r}(\widehat{f})(x) = \sup_{R \in \mathcal{R}(x)} \left(\fint_R |\widehat{f}(y)|^r \, \mathrm{d}y \right)^{1/r},$$

where we have to adjust the height of our Knapp examples to match that of the sharp circle cap it covers. In the end, this yields in particular that the only dimensions where $M_R(\widehat{f}) = M_{R,1}(\widehat{f})$ can possibly be bounded in the full (conjectured) restriction range are $n = 2, 3$. For $n = 2$, this consists exactly of the results in Theorem 0.8. For $n = 3$, the restriction problem is open, and therefore also this one. An interesting question, however, is whether automatic restriction estimates, in the spirit of Kovač [Kov19], can hold in three dimensions, this time for the strong maximal function M_R. We do not pursue this question in this thesis, despite it being one of high interest.

In the third and final part of this thesis, we investigate questions related to the third and last Example C in our list of instances where maximal functions are useful for proving pointwise almost everywhere convergence. The converse of such an idea, that is, that pointwise convergence of a sequence of operators implies bounds for a maximal function, holds under certain particular circumstances. A celebrated result of Stein [Ste61] proves that, if G is a compact group endowed with the Haar measure and we are given a sequence of operators T_n, each bounded in $L^p(G)$, for some fixed $p \in [1, 2]$, commuting with translations and such that the pointwise limit $Tf(x) = \lim_n T_n f(x)$ exists almost everywhere for each $f \in L^p(G)$, then

$$\| \sup_n |T_n f| \|_{L^{p,\infty}(G)} \leq C_p \|f\|_{L^p(G)}.$$

The most classical application of this result is to justify why Carleson's method above is necessary. Note that the Carleson operator

$$Cf(x) = \sup_{N \in \mathbb{N}} \left| \sum_{k=-N}^{N} \widehat{f}(k) e^{2\pi i k x} \right|$$

fulfills the assumptions of this result, if defined on the one-dimensional torus $\mathbb{T}$. Therefore, in order for Fourier series to converge almost everywhere, C must be bounded from L^2 to $L^{2,\infty}$. Indeed, Carleson's proof involves decompositions of the function and the operator C simultaneously, in order to encompass all symmetries of the operator. Differently than a regular singular integral operator, C commutes not only with translations and dilations, but has an additional *modulation* symmetry. Indeed, instead of bounding C itself, we go about bounding the operator

$$\tilde{C}f(x) = \sup_{N\in\mathbb{Z}} \left| \sum_{k=N}^{+\infty} \widehat{f}(k) e^{2\pi i k x} \right|,$$

whenever this sum makes sense. Modulating f in this setting is the same as translating the starting summing point of k, which, by means of the supremum, does not change the operator $\tilde{C}$. Therefore, as the methods of decomposing a singular integral operator (like, for instance, the Hilbert transform) in order to obtain bounds only encompass information about it being translation and dilation invariant, now there is the need to find an approach that includes modulation symmetries as well.

Carleson's methods [Car66] were the first to propose such a decomposition, but despite being generalized by Hunt [Hun68], they were quite intricate. The original proof was considered to be technical and hard to grasp until Fefferman [Fef73] elucidated the general procedure in a more streamlined way. One curious feature of Fefferman's proof is that they in fact do *not* prove the $L^2 \to L^{2,\infty}$ bound of the Carleson operator directly. This, instead, was only explicitly proven in [LT00], where Lacey and Thiele proved that the *continuous* Carleson operator

$$\mathcal{C}f(x) = \sup_{N\in\mathbb{R}} \left| \int_N^\infty \widehat{f}(\xi) e^{2\pi i x \xi} \, \mathrm{d}\xi \right| \tag{0.31}$$

satisfies $\mathcal{C} : L^2(\mathbb{R}) \to L^{2,\infty}(\mathbb{R})$. Later on, Lacey [Lac04] gave an argument proving the full range of L^p bounds for $\mathcal{C}$. Further improvements and results can be found in [Lie11a, AdR02, AdR] and the references therein.

A fundamental fact used in the proof of the boundedness of $\mathcal{C}$ is that , up to summing a multiple of the identity operator, it equals the *maximally modulated Hilbert transform*

$$\sup_{N\in\mathbb{R}} \left| \int_{\mathbb{R}} f(x-t) e^{iNt} \, \frac{\mathrm{d}t}{t} \right|.$$

Inspired by this, Elias Stein asked in [Ste95] whether the *polynomial Carleson operator*

$$C_d f(x) = \sup_{P\colon \deg P \le d} \left| \int_{\mathbb{R}} f(x-t) e^{iP(t)} \, \frac{\mathrm{d}t}{t} \right| \tag{0.32}$$

is bounded in L^2. The first progress in this question was made by Stein himself, who together with Wainger [MSW01] proved that the operator

$$f \mapsto \sup_{\substack{P\colon \deg P \le d, \\ P(0)=P'(0)=0}} \left| \int_{\mathbb{R}} f(x-t) e^{iP(t)} \, \frac{\mathrm{d}t}{t} \right| \tag{0.33}$$

is L^p bounded, for each $p \in (1, +\infty)$. The absence of linear terms in the polynomials considered in the supremum in (0.33) is crucial in their proof, as it splits the proof into two main parts: the one in which the oscillation given by $iP(t)$ is "low", which makes the operator resemble the usual Hilber transform, and where the oscillation is "high", where the underlying cancelation generated gives us additional decay.

It was not until the paper by Lie [Lie09] where the next major breakthrough in this framework was made. Developing upon ideas from Fefferman's proof, Lie showed that the operator C_2, as in (0.32), is weakly bounded in L^2, extending therefore the result by Lacey and Thiele. Subsequently to that, Lie also considered [Lie11b] the higher-degree polynomial case, and Stein's conjecture was finally settled in its most general version in 2018, when Zorin-Kranich [ZK17] proved that the operators

$$f \mapsto \sup_{R_1 > R_2 > 0} \sup_{Q\colon \deg Q \leq d} \left| \int_{R_1 \leq |x-y| \leq R_2} f(y) e^{iQ(y)} K(x,y) \, \mathrm{d}y \right| \tag{0.34}$$

are $L^p(\mathbb{R}^n)$ bounded for any $p \in (1, +\infty)$, where $K(\cdot,\cdot)$ is a Hölder continuous Calderón-Zygmund operator in $\mathbb{R}^n$; see [ZK17] for the specific definitions needed.

If the Carleson operator can be viewed as a maximally modulated version of a singular integral operator, there are other generalizations of the Hilbert transform that also play an important role in affine areas. Here, we will be interested mainly in the *Hilbert transform along the parabola.* For a function $f \in \mathcal{S}(\mathbb{R}^2)$, we define its Hilbert transform along the parabola to be the operator

$$\mathcal{H}_2 f(x,y) = \text{p.v.} \int_{\mathbb{R}} f(x-t, y-t^2) \, \frac{\mathrm{d}t}{t}. \tag{0.35}$$

This operator has a very intimate connection to partial differential equations. Indeed, let us first consider the initial value problem (IVP)

$$\begin{cases} Lu := \partial_{x_1} u - \partial_{x_2}^2 u = f & \text{for } (x_1, x_2) \in \mathbb{R} \times \mathbb{R}; \\ u(0, x_2) = u_0(x_2) & \text{for } x_2 \in \mathbb{R}. \end{cases} \tag{0.36}$$

This is an example of a *parabolic* partial differential equation; classical results, such as existence and uniqueness of weak solutions for (0.36), can be found for instance in [Eva10, Chapter 7]. We are, however, interested in how regular a solution of (0.36) can be, given certain regularity on u_0, f. Theorems 5,6 and 7 in Section 7.1 of [Eva10] provide us, through classical PDE methods, with some results relating the regularity of the functions involved and of the solution. There is, nonetheless, an alternative approach to the regularity problem for the IVP (0.36). Indeed, it is easy to see that the solution to such initial value problem satisfies

$$Lu = T_1(f) - T_2(f),$$

where the operators T_i are defined by Fourier inversion as $\widehat{T_i(f)} = m_i \widehat{f}$, with $m_1(\xi_1, \xi_2) = \frac{2\pi i \xi_2}{2\pi i \xi_2 + 4\pi^2 \xi_1^2}$, $m_2(\xi_1, \xi_2) = \frac{4\pi^2 \xi_1^2}{2\pi i \xi_2 + 4\pi^2 \xi_1^2}$. One readily notices that the multipliers m_i satisfy an *anisotropic dilation invariance*:

$$m_i(\lambda \xi_1, \lambda^2 \xi_2) = m_i(\xi_1, \xi_2), \, \forall (\xi_1, \xi_2) \in \mathbb{R}^2, \, \lambda > 0. \tag{0.37}$$

A computation shows that the (distributional) Fourier transform of such multipliers are kernels $K_i = \widehat{m_i}$ satisfying another anisotropic homogeneity condition, namely

$$K_i(\lambda x_1, \lambda^2 x_2) = \lambda^{-3} K_i(x_1, x_2). \tag{0.38}$$

After a computation similar to the classical method of rotations (cf. [Gra14a, Section 4.2]), studying the operators T_i becomes basically equivalent to studying operators

$$Tf(x_1, x_2) = \int_0^\pi K(\cos\theta, \sin\theta)\mathcal{H}_\theta f(x_1, x_2)\,(1 + \sin^2\theta)\,\mathrm{d}\theta,$$

where K satisfies the homogeneity condition (0.38), and $\mathcal{H}_\theta$ is just the Hilbert transform along the curve

$$\Gamma_\theta(t) = (t\cos\theta, t^2 \mathrm{sign}(t)\sin\theta).$$

At this point, one sees that studying the operator (0.35) helps understandying the operator T above, and therefore also regularity properties of solutions of parabolic partial differential equations. In this spirit, bounds for the Hilbert transform along the parabola have been studied in [NRW74, NRW76], where the authors prove L^p bounds for $p \in (1, +\infty)$ for a higher-dimensional generalization of this operator. It is worth to notice, however, that unlike the classical Hilbert transform it is not known whether $\mathcal{H}_2 : L^1 \to L^{1,+\infty}$; see [CS87, STW04] and the references therein for developments in this direction.

We will be interested in a hybrid version of (0.31) and (0.35). This has been first considered by Pierce and Yung [BPY15], who proved, in particular, that the operators

$$f(x, y) \mapsto \sup_{P \in \mathcal{P}} \left| \int_{\mathbb{R}^d} f(x - t, y - |t|^2) e^{iP(t)} K(t)\,\mathrm{d}t \right|,$$

where K is a Calderón-Zygmund kernel and $\mathcal{P}$ a subspace of polynomials, are L^p-bounded when $d > 2$ for certain subspaces that avoid linear and some quadratic terms. After that, Guo, Pierce, Roos and Yung [GPRY17] consider the $d = 2$ case for a partial supremum with curves like (t, t^k) and $P(t) = N \cdot t^m$. Of course, the most natural question stemming from [GPRY17] is the most basic hybrid version of (0.31) and (0.35):

Question 0.11. *For $f \in \mathcal{S}(\mathbb{R}^2)$, does the* parabolic Carleson operator

$$\mathcal{C}_2 f(x, y) := \sup_{N, M \in \mathbb{R}} \left| \int_{\mathbb{R}} f(x - t, y - t^2) e^{iNt + iMt^2} \,\frac{\mathrm{d}t}{t} \right| \tag{0.39}$$

map $L^2(\mathbb{R}^2)$ into $L^{2,\infty}(\mathbb{R}^2)$?

The main result in Chapter 6 is partial progress towards Question 0.11. Inspired by the anisotropic invariance (0.37), we consider a simpler family of *one-dimensional* operators associated to the multiplier

$$m(\xi, \eta) = \mathrm{p.v.} \int_{\mathbb{R}} e^{2\pi i \xi t + 2\pi i \eta t^2} \,\frac{\mathrm{d}t}{t}.$$

This multiplier is, indeed, the multiplier associated to the operator $\mathcal{H}_2$, and possesses the invariance property (0.37) as well. If we define the family

$$m_{a,b}(\eta) = m(a\eta + b, \eta), a, b \in \mathbb{R}.$$

This new one-dimensional multiplier is simply the restriction of m along the line $\{(a\eta + b, \eta),\ \eta \in \mathbb{R}\}$. In order to consider horizontal lines, we define additionally

$$m_{\infty,b}(\eta) = m(\eta, b),\ b \in \mathbb{R}.$$

Before stating our main result for this part, we notice that the operator (0.39) is just a *maximally modulated* version of the parabolic Hilbert transform:

$$\mathcal{C}_2 f(x,y) := \sup_{(N,M)\in\mathbb{R}^2} |\mathcal{H}_2(\mathcal{M}_{(N,M)}f)(x,y)|,$$

where $\mathcal{M}_{(N,M)}f(x,y) = e^{iNx+iMy}f(x,y)$. If $T_{a,b}h = (m_{a,b}\widehat{h})^\vee$, then the one-dimensional maximal modulations $\mathcal{C}_{a,b}f(x) = \sup_{N\in\mathbb{R}} |T_{a,b}(\mathcal{M}_N f)(x)|$ can be understood as a degenerate case of the parabolic Carleson operator, where we restrict the supremum in (0.39) to $\{(N,M); M = aN + b\}$. Indeed, think of two-dimensional functions supported on a strip around the line $\{(\eta, a\eta + b)\}$ and let the support 'collapse' to the line. It is immediate then that the operator

$$f \mapsto \sup_{b\in\mathbb{R}} \mathcal{C}_{a,b} f \tag{0.40}$$

is the line-degenerate version of (0.39) *without* constraints on (N,M).

Theorem 0.12. *Let $f \in L^p(\mathbb{R})$. It holds that*

$$\sup_{a\in\mathbb{R}\cup\infty} \|\sup_{b\in\mathbb{R}} \mathcal{C}_{a,b} f\|_{L^p(\mathbb{R})} \le C_p \|f\|_p,$$

for $p \in (1, +\infty)$.

The proof of this result is based on the idea of spotting where the operator (0.40) resembles a certain polynomial Carleson operator and where it possesses useful oscillation. More specifically, the proof is composed of two main parts. The first part is of technical nature, and gets rid of the parameter a and reduces matters to bounding, essentially,

$$\sup_{N,b\in\mathbb{R}} \left| \int_{\mathbb{R}} f(x-t) e^{iNt} e^{ib|t+1|^{1/2}} \frac{\mathrm{d}t}{t} \right| \tag{0.41}$$

in L^p. The second part consists of splitting the operator (0.41) into a part where the phase $b\sqrt{t+1}$ is analytic and therefore is well-approximable by polynomials, and another where this phase is large and causes large oscillations. The latter is treated with standard TT^* methods, whereas the former demands additional work. In the interval $[-1/2, 1/2]$, where $b\sqrt{t+1}$ is analytic, we must consider a part where, comparatively to b, oscillation is still non-trivial, and a part where the best we can expect is for the phase to be close to a polynomial. These two scnearios balance each other due to smoothness and lower bounds on the derivatives of $\sqrt{t+1}$.

We mention that it is also possible to prove Theorem 0.12 as a consequence of a positive answer to Question 0.11. We finish Chapter 6 by proving this fact and commenting on possible extensions and generalizations.

Notation. Throughout this thesis, C will usually denote a positive constant whose value is not important for our purposes and which may change from line to line. Moreover, all proofs of inequalities should be understood as *a priori* proofs, in the sense that every function is assumed to belong to the Schwartz class $\mathcal{S}(\mathbb{R}^n)$ or some other similar class of smooth functions. We omit the classical argument to extend from these dense classes to the entirety of L^p.

We generally use the notation $A \lesssim B$ to denote the existence of a constant $C > 0$ such that $A \leq C \cdot B$. Similarly, we define $A \sim B$ is $A \lesssim B$ and $B \lesssim A$. If the constant C is allowed to depend on a specific parameter τ, we will usually write $A \lesssim_\tau B$. Finally, we often will normalize the Fourier transform as

$$\widehat{f}(\xi) = \mathcal{F}f(\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi ix\xi}\,\mathrm{d}x,$$

so that there are no additional constant emerging when using Plancherel's theorem.

Part I

Maximal functions and their regularity properties

Chapter 1

Sharp total variation results for maximal functions

They say the definition of
madness is doing the same thing
and expecting a different result.

–*T.H.*

This chapter contains the paper [Ram17]. In this article, we prove some total variation inequalities for maximal functions. Our results deal with two possible generalizations of the results contained in Aldaz and Pérez Lázaro's work [APL06], one of whose considers a variable truncation of the maximal function, and the other one interpolates the centered and the uncentered maximal functions. In both contexts, we find sharp constants for the desired inequalities, which can be viewed as progress towards the conjecture that the best constant for the variation inequality in the centered context is one. We also provide counterexamples showing that our methods do not apply outside the stated parameter ranges.

1.1 Introduction

An object of major interest in Harmonic Analysis is the Hardy-Littlewood maximal function, which can be defined as

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

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

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

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

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

where $C = \frac{11+\sqrt{61}}{12}$ is the best constant possible found by A. Melas [Mel03] for M. The same inequality also holds in the case of $\tilde{M}$ above, but this time with $C = 2$ being the best constant, as shown by F. Riesz [Rie32].

In the remarkable paper [Kin97], J. Kinnunen proves, using functional analytic techniques and the aforementioned theorem, that, in fact, M maps the Sobolev spaces $W^{1,p}(\mathbb{R})$ into themselves, for $1 < p \leq \infty$. Kinnunen also proves that this result holds if we replace the standard maximal function by its uncentered version. This opened a new field of studies, and several other properties of this and other related maximal functions were studied. We mention, for example, [CS13, CFS15, HO04, KL98, Lui07].

Since the Hardy-Littlewood maximal function fails to be in L^1 for every nontrivial function f and the tools from functional analysis used are not available either in the case $p = 1$, an important question was whether a bound of the form $\|(Mf)'\|_1 \leq C\|f'\|_1$ could hold for every $f \in W^{1,1}$.

In the uncentered case, H. Tanaka [Tan02] provided us with a positive answer to this question. Explicitly, Tanaka proved that, whenever $f \in W^{1,1}(\mathbb{R})$, then $\tilde{M}f$ is weakly differentiable, and it satisfies that $\|(\tilde{M}f)'\|_1 \leq 2\|f'\|_1$. Here, $W^{1,1}(\mathbb{R})$ stands for the Sobolev space $\{f : \mathbb{R} \to \mathbb{R} \colon \|f\|_1 + \|f'\|_1 < +\infty\}$.

Some years later, Aldaz and Pérez Lázaro [APL06] improved Tanaka's result, showing that, whenever $f \in BV(\mathbb{R})$, then the maximal function $\tilde{M}f$ is in fact *absolutely continuous*, and $\mathcal{V}(\tilde{M}f) = \|(\tilde{M}f)'\|_1 \leq \mathcal{V}(f)$, with $C = 1$ being sharp, where we take the *total variation of a function* to be $\mathcal{V}(f) := \sup_{\{x_1<\cdots<x_N\}=\mathcal{P}} \sum_{i=1}^{N-1} |f(x_{i+1}) - f(x_i)|$, and consequently the space of *bounded variation functions* as the space of functions $f : \mathbb{R} \to \mathbb{R} \colon \exists g;\ f = g$ a.e. and $\mathcal{V}(g) < +\infty$. In this direction, J. Bober, E. Carneiro, K. Hughes and L. Pierce [BCHP12] studied the discrete version of this problem, obtaining similar results.

In the centered case, many questions remain unsolved. Surprisingly, it turned out to be *harder* than the uncentered one, due to the contrast in smoothness of Mf and $\tilde{M}f$. In [Kur15], O. Kurka showed the endpoint question to be true, that is, that $\mathcal{V}(Mf) \leq C\mathcal{V}(f)$, with $C = 240,004$. Unfortunately, his method does not give the best constant possible, with the standing conjecture being that $C = 1$ is the sharp constant.

In [Tem13], F. Temur studied the discrete version of this problem, proving that for every $f \in BV(\mathbb{Z})$ we have $\mathcal{V}(Mf) \leq C'\mathcal{V}(f)$, where $C' > 10^6$ is an absolute constant. The standing conjecture is again that $C' = 1$ in this case, which was in part backed up by J. Madrid's optimal results [MP15]: If $f \in \ell^1(\mathbb{Z})$, then $Mf \in BV(\mathbb{Z})$, and $\mathcal{V}(Mf) \leq 2\|f\|_1$, with 2 being sharp in this inequality.

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

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

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

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

Indeed, from this new definition, we get directly that

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

As in the uncentered case, we can still define 'truncated' versions of these operators, by imposing that $t \leq R$. These operators are far from being a novelty: several references consider those all around mathematics, among those the classical [Ste93, Chapter 2], and the more recent, yet related to our work, [CFS15]. An easy argument (see Section 1.5.1 below) proves that, if $\alpha < \beta$, then

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

This implies already, by the main Theorem in [Kur15], that there exists a constant $A \geq 0$ such that $\mathcal{V}(M^{\alpha} f) \leq A \mathcal{V}(f)$, for all $\alpha > 0$. In the intention of sharpening this result, our first result reads, then, as follows:

Theorem 1.1. *Fix any $f \in BV(\mathbb{R})$. For every $\alpha \in [\frac{1}{3}, +\infty)$, we have that*

$$\mathcal{V}(M^{\alpha} f) \leq \mathcal{V}(f). \tag{1.1}$$

There exists an extremizer f for the inequality (1.1). If $\alpha > \frac{1}{3}$, then any positive extremizer f to inequality (1.1) satisfies:

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

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

Notice that stating that a function $g \in W^{1,1}_{loc}(\mathbb{R})$ is the same as asking it to be locally absolutely continuous. Our ideas to prove this theorem and theorem 1.3 are heavily inspired by the ones in [APL06]. Our aim will always be to prove that, when $f \in BV(\mathbb{R})$, then the maximal function $M^{\alpha} f$ is well-behaved on the detachment set

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

Namely, we seek to obtain that the maximal function does not have any local maxima in the set where it disconnects from the original function. Such an idea, together with the concept of the detachment set E_{α}, are also far from being new, having already appeared at [APL06, CS13, CFS15, Tan02], and recently at [Lui18]. More specific details of this can be found in the next section.

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

$$M^1_{\equiv R} f = \sup_{x \in I\colon |I| \le 2R} \fint_I |f(s)| \mathrm{d}s,$$

where we define

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

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

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

Theorem 1.2. *Let $\alpha < \frac{1}{3}$. Then there exists $f \in BV(\mathbb{R})$ such that $f \ge 0$, $f(x) = \limsup_{y \to x} f(y)$ and a point $x_\alpha \in \mathbb{R}$ such that x_α is a local maximum of $M^\alpha f$, but $M^\alpha f(x_\alpha) > f(x_\alpha)$.*

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

We can inquire ourselves whether we can generalize the results from Aldaz and Pérez Lázaro in yet another direction, though. With this in mind, we notice that Kurka [Kur15] mentions in his paper that his techniques allow one to prove that some Lipschitz truncations of the center maximal function, that is, maximal functions of the form

$$M^0_N f(x) = \sup_{t \le N(x)} \frac{1}{2t}\int_{x-t}^{x+t} |f(s)|\mathrm{d}s,$$

are bounded from $BV(\mathbb{R})$ to $BV(\mathbb{R})$ – with some possibly big constant – if $\mathrm{Lip}(N) \le 1$. Inspired by it, we define the *$N-$truncated uncentered maximal function* as

$$M^1_N f(x) = \sup_{|x-y| \le t \le N(x)} \fint_{y-t}^{y+t} |f(s)|\mathrm{d}s.$$

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

Theorem 1.3. *Let $N : \mathbb{R} \to \mathbb{R}_+$ be a measurable function. If $Lip(N) \le \frac{1}{2}$, we have that, for all $f \in BV(\mathbb{R})$,*

$$\mathcal{V}(M^1_N f) \le \mathcal{V}(f).$$

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

Again, we are also going to use a careful maxima analysis in this case. Actually, we are going to do it both in theorems 1.1 and 1.3 for the non-endpoint cases $\alpha > \frac{1}{3}$ and $\mathrm{Lip}(N) < \frac{1}{2}$, while the endpoints are treated with a limiting argument.

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

Theorem 1.4. *Let $c > \frac{1}{2}$ and*

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

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

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

Acknowledgements. The author would like to thank Christoph Thiele, for the remarks that led him to the full range $\alpha \geq \frac{1}{3}$ at Theorem 1.1, as well as to the proof that this is sharp for this technique, and Olli Saari, for enlightening discussions about the counterexamples in the proof of Theorem 1.4 and their construction. He would also like to thank Emanuel Carneiro and Mateus Sousa for helpful comments and discussions, many of which took place during the author's visit to the International Centre for Theoretical Physics in Trieste, to which the author is grateful for its hospitality, and Diogo Oliveira e Silva, for his thorough inspection and numerous comments on the preliminary versions of this paper. The author would like to thank also the anonymous referees, whose corrections and ideas, among which the use of a new normalization, have simplified and cleaned a lot this manuscript. Finally, the author acknowledges financial support from the Hausdorff Center of Mathematics and the DAAD.

1.2 Basic definitions and properties

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

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

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

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

By considering the arguments and techniques contained in the lemmata from [APL06], we may consider sometimes a function in $BV(\mathbb{R})$ endowed with the normalization $f(x) = \limsup_{y \to x} f(y)$, $\forall x \in \mathbb{R}$. At some other times, however, we might need to work with a normalization a little more friendly to the maximal functions involved. Let, then, for a fixed $\alpha \in (0, 1]$,

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

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

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

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

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

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

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

$$\tilde{\mathcal{N}_\alpha} f(x) = \begin{cases} \mathcal{N}_\alpha f(x), & \text{if } M^\alpha f(x) > \limsup_{y\to x} f(y); \\ M^\alpha f(x), & \text{if } \limsup_{y\to x} f(x) \geq M^\alpha f(x) \geq \mathcal{N}_\alpha f(x). \end{cases} \tag{1.2}$$

Of course, we are only changing the points in which $\liminf_{y\to x} f(y) < \mathcal{N}_\alpha f < \limsup_{y\to x} f(y)$, and thus this normalization does not increase the variation, i. e., $\mathcal{V}(\tilde{\mathcal{N}_\alpha} f) \leq \mathcal{V}(f)$. Again, by adapting the lemmata in [APL06] to this context, one checks that we may assume, without loss of generality, that our function has this normalization. We will, for shortness, say we are using $NORM(\alpha)$ whenever we use this normalization. Notice that $NORM(1)$ is the normalization used by Aldaz and Pérez Lázaro.

We mention also a couple of words about the maxima analysis performed throughout the paper. In the paper [APL06], the authors developed an ingenious way to prove the sharp bounded variation result for the uncentered maximal function. Namely, they proved that, whenever $f \in BV(\mathbb{R})$, then the maximal function $\tilde{M} f$ is actually continuous, and the (open) set

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

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

Property 1.5. *We say that an operator $\mathcal{O}$ defined on the class of bounded variation functions has a good attachment at local maxima if, for every $f \in BV(\mathbb{R})$ and local maximum x_0 of $\mathcal{O} f$ over an interval (a,b), with $\mathcal{O} f(x_0) > \max(\mathcal{O} f(a), \mathcal{O} f(b))$, then either $\mathcal{O} f(x_0) = |f(x_0)|$ or there exists an interval $(a,b) \supset I$ such that $\mathcal{O} f$ is constant on I and there is $y \in I$ such that $\mathcal{O} f(y) = |f(y)|$.*

The intuition behind this principle is that, for such operators, one usually has that $\mathcal{V}(\mathcal{O} f) \leq \mathcal{V}(f)$, as skimming through the proofs in [APL06] suggests. This is, as one should expect, the main tool to prove Theorems 1.1 and 1.3.

1.3 Proof of Theorems 1.1 and 1.2

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

1.3.1 Analysis of maxima for M^α, $\alpha > \frac{1}{3}$

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

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

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

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

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

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

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

$$a \ge y - \alpha t \Rightarrow |a - y| \le \alpha t \Rightarrow M^\alpha f(a) \ge M^\alpha f(y) \ge M^\alpha f(x_0),$$

a contradiction to our assumption of strictness of the maximum. This implies that, as for any $z \in [y - \alpha t, y + \alpha t] \Rightarrow |z - y| \le \alpha t$, the maximal function $M^\alpha f$ is *constant* over the interval $[y - \alpha t, y + \alpha t]$. Moreover, we have that the supremum of

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

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

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

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

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

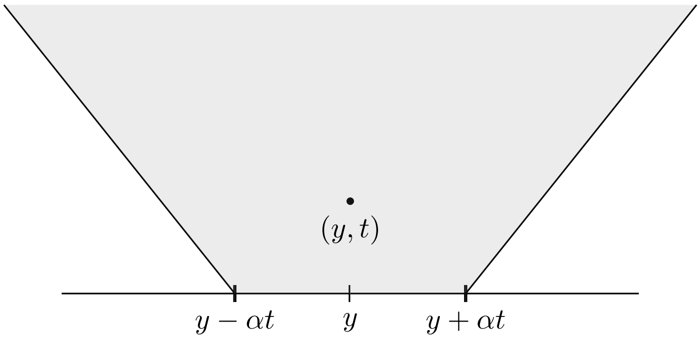

Figure 1.1: The region $C(y, \alpha, t)$.

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

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

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

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

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

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

$\square$

Let $M_{r,A} f(x) = \sup_{0 \leq t \leq 2A} \frac{1}{t} \int_x^{x+t} |f(s)| \, \mathrm{d}s$, and define $M_{l,A} f$ in a similar way, there the subindexes "r" and "l" represent, respectively, "right" and "left". These operators are present out of the context of sharp regularity estimates for maximal functions, just like in [Rie32]. In the realm of regularity of maximal function, though, the first to introduce this notion was Tanaka [Tan02]. As a corollary, we may obtain the following:

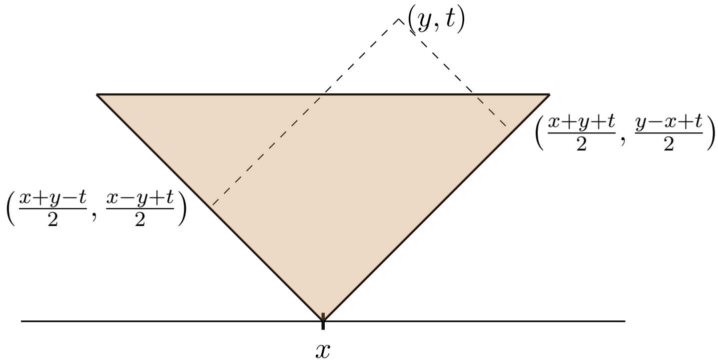

Figure 1.2: Illustration of Lemma 1: the points $\left(\frac{x+y-t}{2}, \frac{x-y+t}{2}\right)$ and $\left(\frac{x+y+t}{2}, \frac{y-x+t}{2}\right)$ are the projections of (y,t) over the lines $t = y - x$ and $t = y + x$, respectively.

Corollary 1.6. *For every $f \in L^1_{loc}(\mathbb{R})$, it holds that*

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

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

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

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

Proof. From Corollary 1.6, we have that

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

$$\le \max\{M_{r,R}f(x), M_{l,R}f(x)\} \le M^1_{\equiv R}f(x).$$

That is exactly what we wanted to prove. □

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

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

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

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

for all $y' \in I$. This is our desired interval for which $M^\alpha f \equiv M^1_{\equiv R}f$. But we already know that, from [APL06, Lemma 3.6], $M^1_{\equiv R}f$ satisfies a *stronger* property of control of maxima. Indeed, in order to fit it into the context of Aldaz and Pérez Lázaro, we note that, by adopting $NORM(1)$, f becomes automatically upper semicontinuous, and also $f \le M^1_{\equiv R}f$ everywhere. In particular, we know that, if $M^1_{\equiv R}f$ is constant in an interval, then it must be *equal* to the function f at *every* point of that interval. But this is exactly our case, as

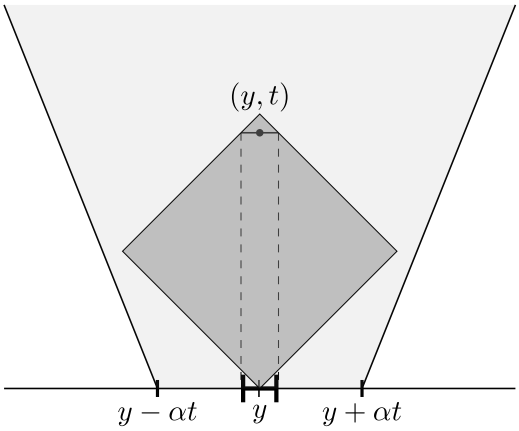

Figure 1.3: In the figure, the **dark gray** area represents the region that our Lemma gives, for some $\frac{1}{2}t < R < \frac{\alpha}{1-\alpha}t$, and the **black** interval is one in which $M^\alpha f = M^1_{\equiv R} f \equiv M^\alpha f(y)$.

we have already noticed that $M^\alpha f$ is constant on $[y - \alpha t, y + \alpha t]$, and therefore also on I. This implies, in particular, that

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

which concludes our analysis of local maxima.

1.3.2 Proof of $\mathcal{V}(M^\alpha f) \leq \mathcal{V}(f)$, for $\alpha \geq \frac{1}{3}$

We remark, before beginning, that this strategy, from now on, is essentially the same as the one contained in [APL06]. We will, therefore, assume that $f \geq 0$ throughout.

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

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

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

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

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

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

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

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

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

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

Thus, it suffices to prove that

$$\limsup_{y\to x} M^\alpha f(y) \ge \limsup_{\varepsilon\to 0} M^\alpha f_\varepsilon(x) \ge \liminf_{\varepsilon\to 0} M^\alpha f_\varepsilon(x) \ge \liminf_{y\to x} M^\alpha f(y),\ \forall x \in \mathbb{R}, \quad (1.3)$$

as then

$$\mathcal{V}(M^\alpha f) = \mathcal{V}(\liminf_{\varepsilon\to 0} M^\alpha f_\varepsilon) = \mathcal{V}(\limsup_{\varepsilon\to 0} M^\alpha f_\varepsilon) = \mathcal{V}(\lim_{j\to\infty} M^\alpha f_{\varepsilon_j}) \le \mathcal{V}(f), \quad (1.4)$$

which follows from the following

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

$$\limsup_{y\to x} g(y) \ge \limsup_{\varepsilon\to 0} g_\varepsilon(x) \ge \liminf_{\varepsilon\to 0} g_\varepsilon(x) \ge \liminf_{y\to x} g(y), \quad (1.5)$$

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

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

Proof of Lemma 4. Let $g_1 = \limsup_{\varepsilon\to 0} g_\varepsilon$, $g_2 = \liminf_{\varepsilon\to 0} g_\varepsilon$. We first prove that $\mathcal{V}(g)$ is less than both $\mathcal{V}(g_1)$, $\mathcal{V}(g_2)$. For that, fix a finite partition of the real line $\mathcal{P} = \{x_1 < \cdots < x_N\}$. In order to estimate the variation $\mathcal{V}_{\mathcal{P}}(g)$, we need to divide into two cases: (i) if g is continuous at every x_i, then we let the partition remain as it is; (ii) if g is *not* continuous at a certain x_i, we then pick two points x_i', x_i'' such that g is continuous at *both* of them, $x_{i-1} < x_i' < x_i < x_i'' < x_{i+1}$ and $g(x_i)$ lies between $g(x_i')$ and $g(x_i'')$. The assumptions on g show that this is *always* possible. Add these new points to the partition $\mathcal{P}$ and call the new one $\mathcal{P}'$. By the way we picked the points x_i', x_i'', we see that the existence of the points x_i in the new partition is superfluous, and therefore we might think of $\mathcal{P}'$ as consisting of the points of $\mathcal{P}$ where g is continuous and the x_i', x_i''. By (1.5), we see that the limit $\lim_{\varepsilon\to 0} g_\varepsilon$ in fact *exists* for each point of this $\mathcal{P}'$ and is equal to g. From that,

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

The desired inequality then follows by taking supremum over all finite partitions of the real line. The reverse inequality consists then of applying the exact same strategy to g_0, g_1, by noticing that, from (1.5), they satisfy the same assumptions as g. $\square$

The reader might notice one still has to prove that $M^\alpha f := g$ satisfies the assumptions in Lemma 4. Indeed, one straightforward way to do so is to use the result in subsection 1.5.1 to conclude that $M^\alpha f$, as a $BV-$function, must fulfill all the properties above. A proof without resorting to this result is however also possible, but we omit it for shortness.

Let us suppose, for the sake of a contradiction, that either the first or the third inequalities in 1.3 are not fulfilled. Therefore, we focus on the first inequality: suppose that there exists a real number x_0, a sequence $\varepsilon_k \to 0$ and a positive real number $\eta > 0$ such that

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

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

$$\fint_{y_k-r_k}^{y_k+r_k} f_{\varepsilon_k}(s)\mathrm{d}s > (1+\eta) \limsup_{y\to x_0} M^\alpha f(y).$$

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

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

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

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

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

Our main claim is then the following:

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

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

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

Proof. The proof goes roughly as the first paragraph of the proof of Lemma 3.9 in [APL06]: let $I_j^\alpha = (l(I_j^\alpha), r(I_j^\alpha)) =: (l_j, r_j)$, and suppose that $M^\alpha f$ is *not* V shaped there. Therefore, there would be a maximal point $x_0 \in I_j^\alpha$ and an interval $J \subset I_j^\alpha$ such that $M^\alpha f$ has a *strict* local maximum at x_0 over J. Then, by the maxima analysis we performed, we see that we have reached a contradiction from this fact alone, as $J \subset E_\alpha$. We omit further details, as they can be found, as already mentioned, at [APL06, Lemma 3.9]. □

We also need the following

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

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

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

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

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

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

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

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

i. if there is an interval $I_{j_0}^\alpha = (-\infty, r_{j_0})$, then it is easy to prove that $\lim_{x\to-\infty} M^\alpha f(x) = L$. Therefore, by 'choosing a point' x' sufficiently large, we see that $|M^\alpha f(x') - f(x')|$ has to be small, and the argument follows;

ii. if there is an interval $I_{j_1}^\alpha = (l_{j_1}, +\infty)$, then a standard argument proves that $\lim_{x\to+\infty} M^\alpha f(x) = \frac{1+\alpha}{2}L + \frac{1-\alpha}{2}M$. But, also by a standard argument, one proves that $M^\alpha f(x) \geq \frac{1+\alpha}{2}L + \frac{1-\alpha}{2}M$ *for each* $x \in \mathbb{R}$. This shows that $M^\alpha f$ cannot be "strictly" V shaped over $I_{j_1}^\alpha$. As we know from Lemma 6 that $M^\alpha f(l_{j_1}) = f(l_{j_1})$ and that $M^\alpha f$ has to be *non-increasing* over $I_{j_1}^\alpha$, choosing a 'sufficiently large' point only helps us, as
$$\mathcal{V}_{I_{j_1}^\alpha}(M^\alpha f) = f(l_{j_1}) - \left(\frac{1+\alpha}{2}L + \frac{1-\alpha}{2}M\right) \leq f(l_{j_1}) - M \leq \mathcal{V}_{I_{j_1}^\alpha}(f).$$

Call the new partition obtained by this procedure $\mathcal{P}'$. By the calculation above and the fact that, if $f \in \mathrm{Lip}(\mathbb{R}) \Rightarrow M^\alpha f \geq f$ everywhere, and in particular $M^\alpha f = f$ at $\mathbb{R}\backslash E_\alpha$, one obtains that
$$\mathcal{V}_{\mathcal{P}}(M^\alpha f) \leq \mathcal{V}_{\mathcal{P}'}(M^\alpha f) \leq \mathcal{V}(f).$$
By taking a supremum over all partitions, we finish the result for $\alpha > \frac{1}{3}$. On the other hand, it is straight from the definition that
$$\beta \leq \alpha \Rightarrow \frac{\beta}{\alpha} M^\alpha f \leq M^\beta f \leq M^\alpha f.$$
This implies that, for a partition $\mathcal{P}$ as above,
$$\sum_{i=1}^{N-1} |M^{\frac{1}{3}} f(x_{i+1}) - M^{\frac{1}{3}} f(x_i)| \leq \lim_{\alpha \searrow \frac{1}{3}} \sum_{i=1}^{N-1} |M^\alpha f(x_{i+1}) - M^\alpha f(x_i)| \leq \mathcal{V}(f).$$
The theorem follows, again, as before.

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

Claim 1.7. *Let $E_\alpha = \{x \in \mathbb{R} : M^\alpha f(x) > f(x)\}$. This set is open for any $f \in BV(\mathbb{R})$ normalized wiht $NORM(\alpha)$ and therefore can be decomposed as*
$$E_\alpha = \cup_{j\in\mathbb{Z}} I_j^\alpha,$$
where each I_j^α is an interval. Furthermore, the restriction of $M^\alpha f$ to each of those intervals is either a monotone function or a V shaped function with a minimum at $c_j \in I_j^\alpha$. Moreover, $M^\alpha f(c_j) < \min\{M^\alpha f(l(I_j^\alpha)), M^\alpha f(r(I_j^\alpha))\}$.

Proof of the claim. The claim seems quite sophisticated, but its proof is simple, once one has done the maxima analysis we have done. The fact that E_α is open is easy to see. In fact, let $x_0 \in E_\alpha$. By the lower semicontinuity of $M^\alpha f$ at x_0 and the fact that we normalized f with $NORM(\alpha)$,
$$\liminf_{z\to x_0} M^\alpha f(z) \geq M^\alpha f(x_0) > \limsup_{z\to x_0} f(z).$$
This shows that, for z close to x_0, the strict inequality should still hold, as desired.

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

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

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

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

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

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

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

still holds, by Lemma 7 and by the fact that $c_j \in I_j^\alpha$. This finishes finally the second proof of Theorem 1.1.

1.3.3 Absolute continuity on the detachment set

We prove briefly the fact that, for $f \in W^{1,1}(\mathbb{R})$, then we have that $M^\alpha f \in W^{1,1}_{loc}(\mathbb{R})$ for *any* $1 > \alpha > 0$, as the case $\alpha = 0$ has been dealt with by Kurka [Kur15], in Corollary 1.4.

Indeed, let

$$E_{\alpha,k} = \{x \in E_\alpha \colon M^\alpha f(x) = \sup_{(y,t)\colon |y-x| \le \alpha t,\, t \ge \frac{1}{2k}} \frac{1}{2t}\int_{y-t}^{y+t} |f(s)|\mathrm{d}s\}.$$

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

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

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

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

i. *g is continuous;*

ii. *g is of bounded variation;*

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

In fact, let S be then a null-measure set on the real line and $f \in W^{1,1}(\mathbb{R})$ – which implies that $M^\alpha f \in C(\mathbb{R})$ and, by the comments in subsection 1.5.1, $\forall \alpha > 0$ and $f \in W^{1,1}(\mathbb{R})$, $M^\alpha f \in BV(\mathbb{R})$ –, and let us invoke [APL06, Lemma 3.1]:

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

$$m^*(f(E)) \le k m^*(E),$$

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

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

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

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

$$\begin{aligned} |M^\alpha f(S)| &\le |M^\alpha f(S \cap E_\alpha^c)| + \sum_{j\in\mathbb{Z}} |M^\alpha f(S \cap I_j^\alpha)| \\ &= |f(S \cap E_\alpha^c)| = 0, \end{aligned}$$

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

1.3.4 Sharpness of the inequality and extremizers

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

Claim 1.8. *Let $f \in BV(\mathbb{R})$ normalized as before satisfy $\mathcal{V}(f) = \mathcal{V}(M^\alpha f)$. If we decompose $E_\alpha = \cup_j I_j^\alpha$, where each of the I_j^α is open and maximal, then*

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

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

$$\begin{cases} \mathcal{V}(M^\alpha f) & \le \mathcal{V}_{\mathcal{P}}(M^\alpha f) + \frac{\varepsilon}{20}, \\ \mathcal{V}(f) & \le \mathcal{V}_{\mathcal{Q}}(f) + \frac{\varepsilon}{20}. \end{cases} \tag{1.6}$$

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

$$\sum_{j\in A} \mathcal{V}_{I_j^\alpha}(f) \ge \sum_{j\in\mathbb{N}\colon I_j^\alpha \neq \emptyset} \mathcal{V}_{I_j^\alpha}(f) - \frac{\varepsilon}{20}, \tag{1.7}$$

then keep the partition as it is before advancing. If not, then add to $\mathcal{S}$ finitely many points, all of them contained in intervals of the form $\overline{I_j^\alpha}$, such that inequality 1.7 holds. Call this new partition $\mathcal{S}$ again, as it still satisfies the inequalities 1.6.

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

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

ii. If $M^\alpha f$ is V shaped over the interval I_j^α, then there is a point c_j such that $M^\alpha f$ is non-increasing on (l_j, c_j) and non-decreasing on (c_j, r_j). Add such a point to our partition.

iii. If $\mathcal{V}_{I_j^\alpha}(f) > \mathcal{V}_{\{x_i \in \mathcal{S}\colon x_i \in I_j^\alpha\}}(f) + \frac{\varepsilon}{2^{20|j|}}$, then add finitely many points to the partition to make the reverse inequality hold (here, $\mathcal{V}_{\{x_i \in \mathcal{S}\colon x_i \in A\}}(g)$ stands for the variation along the finite partition composed solely by elements in the set A).

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

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

which then implies that, by the considerations above,

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \mathcal{V}_{I_j^\alpha}(f) - \frac{\varepsilon}{4} &\le \sum_{j\in A} \mathcal{V}_{I_j^\alpha}(f)\\ &\le \sum_{j\in A} \mathcal{V}_{\{x_i\in\mathcal{S}'\colon x_i\in I_j^\alpha\}}(f) + \varepsilon\\ &\le \sum_{j\in A} \mathcal{V}_{\{x_i\in\mathcal{S}'\colon x_i\in I_j^\alpha\}}(M^\alpha f) + 3\varepsilon\\ &\le \sum_{j\in\mathbb{Z}} \mathcal{V}_{I_j^\alpha}(M^\alpha f) + 3\varepsilon.\end{aligned} \tag{1.8}$$

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

$$\mathcal{V}_{I_j^\alpha}(f) = \mathcal{V}_{I_j^\alpha}(M^\alpha f). \tag{1.9}$$

This finishes the proof of this claim. □

Claim 1.9. *Let f, I_j^α as above. Then f and $M^\alpha f$ are monotone in the closure $\overline{I_j^\alpha}$.*

Proof. Suppose first that $M^\alpha f$ is *not* monotone there. Then it must be V shaped on $\overline{I_j^\alpha}$, and then, by Claim 1.8, we see that the only possibility for that to happen is if $M^\alpha f(c_j) = f(c_j)$, $c_j \in I_j^\alpha$. This is clearly not possible by the definition of I_j^α, and we reach a contradition.

Suppose now that f is *not* monotone over $\overline{I_j^\alpha}$. As $\mathcal{V}_{I_j^\alpha}(f) = \mathcal{V}_{I_j^\alpha}(M^\alpha f)$ by Claim 1.8, and $\mathcal{V}_{I_j^\alpha}(M^\alpha f) = |f(r_j) - f(l_j)|$, then it is easy to see that, no matter what configuration of non-monotonicity we have, it yields a contradiction with the equality for the variations over the interval I_j^α. We skip the details, for they are routinely verified. $\square$

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

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

Claim 1.10. *Such an f is, in fact, non-decreasing on $(-\infty, r(I_j^\alpha)]$.*

Proof. Our proof of this fact will go by contradiction:

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

Subclaim 1.11. *$M^\alpha f(a_j) = f(a_j)$ and $f(a_j-) = f(a_j+)$.*

Proof. If $M^\alpha f(a_j) > f(a_j)$, then there exists a maximal open interval $E_\alpha \supset J_j' \ni a_j$, and, as we proved before, f must be monotone in such an interval. By the definition of a_j, we see that, at least on $(a_j, r(I_j^\alpha)) \cap \overline{J_j'} =: K_j$, the function f has to be non-decreasing. If f is non-constant in K_j, then, by maximality and Claim 1.9, we see that f is non-decreasing in J_j', which contradicts the choice of a_j. Then f has to be constant in K_j and, therefore, non-increasing in J_j'. We wish to show that this cannot happen, so that we conclude the desired equality. We seek to contradict Lemma 7, in the sense that we wish to prove that, actually, $M^\alpha f(r(J_j')) > f(r(J_j'))$. Before we start doing so, we notice that, as f is non-decreasing (and non-constant) on $(a_j, r(I_j^\alpha)]$, and non-increasing on $\overline{J_j'}$, then $r(J_j') \in (l(J_j'), r(I_j^\alpha))$ and f attains a minimum over $[l(J_j'), r(I_j^\alpha)]$ at $r(J_j')$.

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

$$M^\alpha f(r(J_j')) \ge \fint_{l(J_j')}^{2r(J_j')-l(I_j')} f > f(r(J_j')),$$

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

$$M^\alpha f(r(J_j')) \ge \fint_{2r(J_j')-r(I_j^\alpha)-\delta}^{r(I_j^\alpha)+\delta} f > f(r(J_j')),$$

as (i) $[2r(J_j') - r(I_j^\alpha) - \delta, r(I_j^\alpha)] \subset [l(J_j'), r(I_j^\alpha)]$; (ii) f has a local minimum at $r(J_j')$ in $[l(J_j'), r(I_j^\alpha)]$; (iii) if f is constant on $(r(J_j'), r(I_j^\alpha))$, then, for $r(I_j^\alpha)+\delta > p > r(I_j^\alpha)$, it holds that $f(p) > f(r(J_j'))$, as long as we choose $\delta > 0$ to be small enough (this holds because

f is normalized). As we covered the only two possible cases, we arrive at a contradiction, namely, that $M^\alpha f(r(J_j')) > f(J_j')$. This finishes the first part of our subclaim.

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

We must prove yet another fact that will help us:

Subclaim 1.12. *Let*

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

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

Proof. If $a_j \in \mathcal{D}$, then our assertion is proved by Subclaim 1.11. If not, then $\mathcal{D} \subset (b_j, a_j)$. In this case, pick any point d_0 in this intersection.

Case 1: $f(d_0+) = f(d_0-)$. In this case, there is nothing left undone if $f(d_0) = M^\alpha f(d_0)$. Otherwise, we would have that $M^\alpha f(d_0) > f(d_0)$, and then there would be an interval $E_\alpha \supset J_0 \ni d_0$. By the fact that *all* the points in $\mathcal{D}$ must lie in (b_j, a_j), and that f is monotone on J_0, we see automatically that either $f(b_j) \le f(d_0)$, a contradiction, or the right endpoint of J_0 satisfies $f(r(J_0)) \le f(d_0)$. By the definition of d_0, this inequality has to be an equality, and also f must be continuous at $r(J_0)$, by the argument of Remark 1. As an endpoint of a maximal interval $J_0 \subset E_\alpha$, we have then $M^\alpha f(r(J_0)) = f(r(J_0))$.

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

The concluding argument for the proof of the Claim 1.10 goes as follows: let d be the point from Subclaim 1.12. Then we must have that

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

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

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

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

i. If all intervals in the detachment set are *of one single type*, that is, either all *non-increasing* or all *non-decreasing*, our function must then admit a point x_0 such that f is either non-decreasing on $(-\infty, x_0]$ (resp. non-decreasing on $[x_0, +\infty)$,) and $f = M^\alpha f$ on $(x_0, +\infty)$ (resp. on $(-\infty, x_0)$).

ii. If there is at least one interval of each type, then we must have an interval $[R, S]$ such that

- f is non-decreasing on $(-\infty, R]$;
- f is non-increasing on $[S, +\infty)$;
- $f = M^\alpha f$ on (R, S).

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

1.3.5 Proof of Theorem 1.2

We start our discussion by pointing out that the measure $\mathrm{d}\mu = \delta_0 + \delta_1$ satisfies our Theorem.

Proposition 1.13. *Let $0 \le \alpha < \frac{1}{3}$. Then*

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

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

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

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

Proof of Theorem 1.2. Let $f_n(x) = n(\chi_{[0,\frac{1}{n}]} + \chi_{[1-\frac{1}{n},1]})$. It is easy to see that $\int g f_n \mathrm{d}x \to \int g\mathrm{d}\mu(x)$, for each $g \in L^\infty(\mathbb{R})$ that is continuous on $[0, t_0) \cup (t_1, 1]$, for some $t_0 < t_1$.

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

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

To prove convergence, we argue in two steps.

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

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

But then

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

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

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

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

$$M^\alpha f_N(x) = \fint_{y_N - t_N}^{y_N+t_N} f_N(s)\mathrm{d}s, \;\; \Delta(x) \geq t_N \geq \delta(x) > 0.$$

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

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

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

1.4 Proof of Theorems 1.3 and 1.4

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

1.4.1 Analysis of maxima of M_N^1 for Lip$(N) < \frac{1}{2}$

We assume, first of all, that f has $NORM(1)$ normalization. Let (a,b) be an interval on the real line, such that there exists a point x_0, maximum of $M_N^1 f$ over (a,b), with the property that

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

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

Now we want to find a neighbourhood of x_0 such that there is $R = R(x_0) > 0$ such that, for all $x \in I$, $M_{\equiv R}^1 f(x) = M_N^1 f(x_0)$.

By Lemma 1, we can suppose that either $y = x_0 - t$ or $y = x_0 + t$. Without loss of generality, let us assume that $y = x_0 - t$.

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

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

As before, we finish this case with [APL06, Lemma 3.6], as then it guarantees us that $M_{\equiv t+\varepsilon}^1 f(x) = f(x)$ for every point in this interval I.

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

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

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

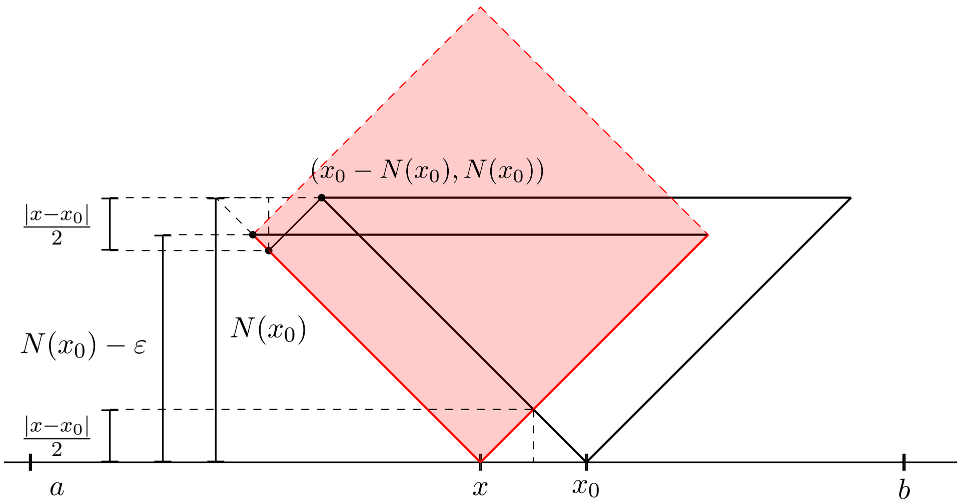

Figure 1.4: Illustration of proof of case (b).

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

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

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

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

But this is equivalent to

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

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

$$|N(x) - N(x_0)| \leq \mathrm{Lip}(N)|x - x_0| \Rightarrow N(x) \geq N(x_0) - \mathrm{Lip}(N)|x - x_0| \geq N(x_0) - \varepsilon \iff$$

$$|x - x_0| \leq \frac{1}{\mathrm{Lip}(N)}\varepsilon.$$

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

$$\{x \in \mathbb{R} : \frac{1}{\mathrm{Lip}(N)}\varepsilon \geq |x - x_0| \geq 2\varepsilon\},$$

it holds that $M_N^1 f(x_0) \geq M_N^1 f(x) \geq M^1_{\equiv N(x_0)-\varepsilon} f(x) \geq M_N^1 f(x_0) \geq M_N^1 f(x)$. By [APL06, Lemma 3.6], $M^1_{\equiv N(x_0)-\varepsilon} f(x) = M_N^1 f(x) = f(x)$. This concludes then that, whenever there is a "strict" local maximum (with respect to the endpoints) of $M_N^1 f$ over an interval (a,b), then there is a $x \in (a,b)$ sucht that $M_N^1 f(a) = f(a)$, as the finishing argument here is then the same as the one used in Theorem 1.1, and we therefore omit it.

1.4.2 The critical case $\mathbf{Lip}(N) = \frac{1}{2}$

The argument is pretty simple: we build explicitly a suitable sequence of approximations of N such that they all have Lipschitz constants less than $\frac{1}{2}$. By our already proved results, this will give us the result also in this case.

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

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

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

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

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

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

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

$$|\tilde{N}(x_i) - N(x_i)| < \delta \Rightarrow |M^1_N f(x_j) - M^1_{\tilde{N}} f(x_j)| < \frac{\varepsilon}{2M}.$$

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

$$\mathcal{V}_{\mathcal{P}}(M^1_N f) \le \mathcal{V}_{\mathcal{P}}(M^1_{\tilde{N}} f) + \varepsilon \le \mathcal{V}(M^1_{\tilde{N}} f) + \varepsilon \le \mathcal{V}(f) + \varepsilon.$$

Taking the supremum over all possible partitions and then taking $\varepsilon \to 0$ finishes also this case, and thus the proof of Theorem 1.3.

1.4.3 Counterexample for $\mathbf{Lip}(N) > \frac{1}{2}$

Finally, we build examples of functions with $\mathrm{Lip}(N) > \frac{1}{2}$ and $f \in BV(\mathbb{R})$ such that

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

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

i. First, let $x_0 = \frac{2}{2\beta+1}$. Let then $N(0) = 1$, $N(x_0) = \frac{x_0}{2}$ and extend it linearly in $(0, x_0)$.

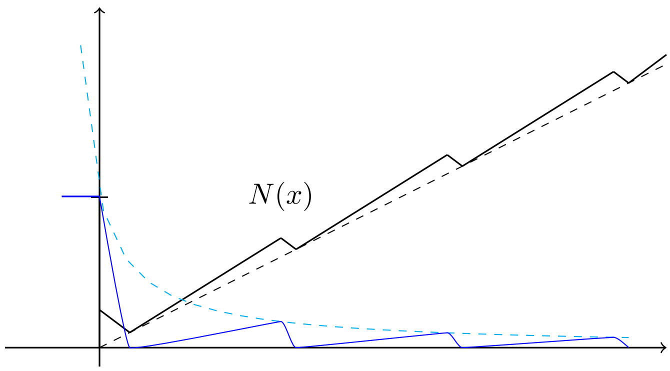

Figure 1.5: A counterexample in the case of $\mathrm{Lip}(N) = \frac{3}{4}$. The dashed lines are the graphs of $\frac{x}{2}$ and $\frac{1}{1+x}$, and the non-dashes ones the graphs of $M_N^1 f$ and N in this case.

ii. Let x'_K be the solution to the equation $\beta x - \beta x_{K-1} + \frac{x_{K-1}}{2} = \frac{x+1}{2} \iff x'_K = x_{K-1} + \frac{1}{\beta - \frac{1}{2}}$.

iii. At last, take $x_K = x'_K + \frac{1}{2\beta+1}$, and define for all $K \geq 1$ $N(x_K) = \frac{x_K}{2}$, $N(x'_K) = \frac{x'_K+1}{2}$, extending it linearly on (x_{K-1}, x'_K) and (x'_K, x_K).

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

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

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

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

In fact, it is not difficult to see that:

i. $M_N^1 f(x_K) = 0$, $\forall K \geq 0$. This is due to the fact that the maximal intervals $(y-t, y+t)$ that satisfy $|x_K - y| \leq t \leq N(x_K)$ are still contained in $[0, +\infty)$, which is of course disjoint from $(-1, 0)$.

ii. $M_N^1 f(x'_K) \geq \frac{1}{x'_K+1}$. This follows from

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

This shows that

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

This construction therefore proves Theorem 1.4.

1.5 Comments and remarks

1.5.1 Monotonicity of maximal $BV-$norms

Theorem 1.1 proves that, if we define

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

then $B(\alpha) = 1$ for all $\alpha \in [\frac{1}{3}, 1]$. We can, however, with the same technique, show that $B(\alpha)$ is non-increasing in $\alpha > 0$, and also that $B(\alpha) \equiv 1\ \forall \alpha \in [\frac{1}{3}, +\infty)$. Indeed, we show that, for $f \in BV(\mathbb{R})$ endowed with $NORM(1)$ normalization and $\beta > \alpha$, then $\mathcal{V}(M^\alpha f) \geq \mathcal{V}(M^\beta f)$. This allows one to conclude, without glancing at Theorem 1.1, that $\mathcal{V}(M^\alpha f) \leq C \cdot \mathcal{V}(f)$, for $C = 240.004$ and all $\alpha \geq 0$, as a consequence of Kurka's [Kur15] result. The argument uses the maximal attachment property and is *independent* of the proof of Theorem 1.1 – which allows us, for instance, to make use of this fact in the proof of Theorem 1.1, as indicated in the subsection 1.3.2: we first assume f to be positive, without loss of generality. Let, as usual, (a, b) be an interval where $M^\beta f$ has a local maximum *inside* it, at, say, x_0, and

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

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

- *If* $M^\beta f(x_0) = f(x_0)$, we do not have absolutely anything to do, as then also $M^\alpha f(x_0) = M^\beta f(x_0)$.
- *If* $M^\beta f(x_0) = u(y, t)$, *for* $t > 0$, we have – as in subsection 1.3.1 – that $(y - \beta t, y + \beta t) \subset (a, b)$. But it is then obvious that

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

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

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

1.5.2 Nontangential maximal functions and classical results

Here, we investigated mostly the regularity aspect of our family M^α of nontangential maximal functions, and looked for the sharp constants in such bounded variation inequalities. One can, however, still ask about the most classical aspect studied by Melas [Mel03]: Let C_α be the least constant such that we have the following inequality:

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

By [Mel03], we have that, for when $\alpha = 0$, then $C_0 = \frac{11+\sqrt{61}}{12}$, and the classical argument of Riesz [Rie32] that $C_1 = 2$. Therefore, $\frac{11+\sqrt{61}}{12} \leq C_\alpha \leq 2,\ \forall \alpha \in (0, 1)$. Nevertheless, the exact values of those constants is, as long as the author knows, still unknown.

1.5.3 Bounded variation results for mixed Lipschitz and nontangential maximal functions

In Theorems 1.3 and 1.4, we proved that, for the *uncentered* Lipschitz maximal function M_N, we have sharp bounded variation results for $\mathrm{Lip}(N) \leq \frac{1}{2}$, and, if $\mathrm{Lip}(N) > \frac{1}{2}$, we cannot even assure *any* sort of bounded variation result.

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

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

then what should be the best constant $L(\alpha)$ such that, for $\mathrm{Lip}(N) \leq L(\alpha)$, then we have some sort of bounded variation result like $\mathcal{V}(M_N^\alpha f) \leq A\mathcal{V}(f)$, and, for each $\beta > L(\alpha)$, there exists a function N_β and a function $f_\beta \in BV(\mathbb{R})$ such that $\mathrm{Lip}(N_\beta) = \beta$ and $\mathcal{V}(M_{N_\beta} f_\beta) = +\infty$? Regarding this question, we cannot state any kind of sharp constant bounded variation result, but the following is still attainable: it is possible to show that the first two lemmas of O. Kurka [Kur15] are adaptable in this context *if* we suppose that

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

and then we obtain our results, with a constant that is even independent of $\alpha \in (0,1)$. On the other hand, our example used above in the proof of Theorem 1.4 is easily adaptable as well, and therefore one might prove the following Theorem:

Theorem 1.14. *Let $\alpha \in [0,1]$ and N be a Lipschitz function with Lip$(N) \leq \frac{1}{\alpha+1}$. Then, for every $f \in BV(\mathbb{R})$, we have that*

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

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

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

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

Chapter 2

Regularity of fractional maximal functions through Fourier multipliers

> A mathematical theorem is in general impossible... Until it is proven, that is when it becomes trivial.
>
> *–M.F.*

This chapter contains the paper [BRS19], a collaboration between the author of this thesis, David Beltrán and Olli Saari. We prove endpoint bounds for derivatives of fractional maximal functions with either smooth convolution kernel or lacunary set of radii in dimensions $n \geq 2$. We also show that the spherical fractional maximal function maps L^p into a first order Sobolev space in dimensions $n \geq 5$.

2.1 Introduction

Define the fractional maximal function as

$$M_\alpha f(x) = \sup_{t>0} \left| \frac{t^\alpha}{|B(x,t)|} \int_{B(x,t)} f \, dy \right|$$

for $\alpha \in [0, n)$. The study of its regularity properties was initiated in [KS03] by Kinnunen and Saksman. They proved the pointwise inequality

$$|\nabla M_\alpha |f|(x)| \leq C M_{\alpha-1} |f|(x), \quad \alpha \geq 1 \tag{2.1}$$

with a constant C only depending on the dimension n and α. This inequality has two interesting consequences. First, M_α maps $L^p(\mathbb{R}^n)$ into a first order Sobolev space. Second, as noted by Carneiro and Madrid [CM17], the pointwise bound together with the Gagliardo–Nirenberg–Sobolev inequality implies

$$\|\nabla M_\alpha f\|_{L^p} \leq C \|M_{\alpha-1} f\|_{L^p} \leq C \|f\|_{L^{n/(n-1)}} \leq C \|\nabla f\|_{L^1} \tag{2.2}$$

for $\alpha \geq 1$ and $p = n/(n - \alpha)$. When $\alpha \in (0, 1)$, inequality (2.1) no longer helps, and the conclusion of (2.2) is an open problem. When M_α is replaced by its non-centred variant, the analogous result is due to Carneiro and Madrid [CM17] for $n = 1$ and Luiro and

Madrid [LM17] for f radial and $n \geq 2$. For other aspects of the regularity of fractional maximal functions, see e.g. [HKKT15, HKNT13] and the references therein.

Our first result is a smooth variant of the inequality (2.2) for $\alpha \in (0,1)$ and $n \geq 2$. Define the lacunary fractional maximal function as

$$M_\alpha^{lac} f(x) := \sup_{k\in\mathbb{Z}} \left| \frac{2^{\alpha k}}{|B(0,2^k)|} \int_{B(x,2^k)} f\,dy \right|.$$

For integrable φ and $t > 0$, let $\varphi_t(x) = t^{-n}\varphi(x/t)$. Assume, for simplicity, that φ is a positive Schwartz function and define the smooth fractional maximal function as

$$M_\alpha^\varphi f(x) = \sup_{t>0} t^\alpha |\varphi_t * f(x)|.$$

The smoothness requirement can be substantially relaxed, see §§2.3.3.

Theorem 2.1. *Let $f \in \dot{\mathrm{BV}}(\mathbb{R}^n)$ and suppose that $\alpha \in (0,1)$ and $n \geq 2$. Then for $\mathcal{M}_\alpha \in \{M_\alpha^{lac}, M_\alpha^\varphi\}$, there exists a constant C only depending on dimension n, α and φ such that*

$$\|\nabla \mathcal{M}_\alpha f\|_{L^p(\mathbb{R}^n)} \leq C|f|_{\mathrm{BV}(\mathbb{R}^n)}$$

for $p = n/(n-\alpha)$.

The proof of this theorem uses the g-function technique familiar from Stein's spherical maximal function theorem. The idea is to follow the scheme behind the short estimation (2.2). The Fourier transform is used to find a substitute for (2.1) at the level of Besov spaces, from which the conclusion then follows by a refined Gagliardo–Nirenberg–Sobolev type embedding theorem [CDDD03]. The last step requires $n > 1$ whereas the smoothness condition on the maximal operator is imposed by Fourier analysis. We stress that the right hand side of the conclusion is BV norm instead of the considerably larger homogeneous Hardy–Sobolev norm one might first expect. The detailed proof is given in §2.3, and all necessary definitions can be found in §2.2. To the best of our knowledge, Fourier transform techniques have not been exploited effectively in the study of endpoint regularity of maximal functions prior to this work.

The background of the question (2.2) goes back to Kinnunen's theorem [Kin97, KL98] asserting that the Hardy–Littlewood maximal function is bounded in $W^{1,p}$ with $p > 1$. His result was later extended to $W^{1,1}$ in the form

$$\|\nabla M f\|_{L^1(\mathbb{R}^n)} \leq C\|\nabla f\|_{L^1(\mathbb{R}^n)} \tag{2.3}$$

by Tanaka [Tan02] when $n = 1$ and Luiro [Lui18] when $n \geq 2$ and f is radial. Here M is the non-centred Hardy–Littlewood maximal function. The same inequality for M_0 (centred maximal function) was established by Kurka [Kur15] when $n = 1$, and the question is open in dimensions $n \geq 2$. Kurka's theorem can be seen as the limiting case $\alpha = 0$ of (2.2).

In connection to (2.3), maximal functions with smooth convolution kernels are better understood than the Hardy–Littlewood maximal function. Inequality (2.3) can be proved with sharp constant for many smooth kernels [CFS15, CS13] whereas the best constant for centred Hardy–Littlewood maximal function is not known (for the non-centred maximal function [APL06] as well as for certain non-tangential maximal functions [Ram17] the constant is one). Similarly, a Hardy–Sobolev bound corresponding to (2.3) is known for smooth maximal functions in all dimensions [PPSS18] whereas the progress for the standard maximal function is limited to the case of radial functions [Lui18]. Finally, there are metric measure spaces where Kinnunen's theorem does not hold but suitable smoother maximal functions satisfy a Sobolev bound [AK10]. Theorem 2.1 can be seen as a part of this line of research attempting to understand (2.2) and (2.3) first in the case of smooth maximal functions.

The second part of the paper studies the regularity of the spherical fractional maximal function

$$S_\alpha f(x) := \sup_{t>0} |t^\alpha \sigma_t * f(x)|, \tag{2.4}$$

where σ_t is the normalized surface measure of the sphere $\partial B(0,t)$. For $\alpha = 0$, one recovers the spherical maximal function of Stein [Ste76] ($n \geq 3$) and Bourgain [Bou86] ($n = 2$). For $\alpha > 0$, $L^p \to L^q$ bounds for this operator follow from the work of Schlag [Sch97] ($n = 2$) and Schlag and Sogge [SS97] ($n \geq 3$). It is natural to ask if the fractional spherical maximal function has regularizing properties similar to (2.1). Our result in this direction is the following.

Theorem 2.2. *Let $n \geq 5$, $n/(n-2) < p \leq q < \infty$ and*

$$\alpha(p) := \begin{cases} \frac{n^2-2n-1}{n-1} - \frac{2n}{p(n-1)} & \text{if } \frac{n}{n-2} < p \leq \frac{n^2+1}{n^2-2n-1} \\ \frac{n-1}{p} & \text{if } \frac{n^2+1}{n^2-2n-1} < p \leq n-1. \end{cases}$$

Assume that

$$\frac{1}{q} = \frac{1}{p} - \frac{\alpha - 1}{n}, \qquad 1 \leq \alpha < \alpha(p).$$

Then, for any $f \in L^p$, $S_\alpha f$ is weakly differentiable and

$$\|\nabla S_\alpha f\|_{L^q} \lesssim \|f\|_{L^p}.$$

The proof of this theorem is also based on the use of the Fourier transform. When $q \geq 2$, we study $L^p \to L^q$ estimates for a maximal multiplier operator in analogy with the estimates in [Sch97, SS97, Lee03] for the spherical maximal function. Since Theorem 2.2 is a statement at the derivative level, the corresponding multiplier enjoys worse Fourier decay than $\widehat{\sigma}$. This forces us to study the behavior in L^p with large p more carefully than what is needed to understand L^p mapping properties of the spherical maximal function. We take advantage of the sharp local smoothing estimate for the wave equation in $L^{n-1}(\mathbb{R}^n)$, which is available whenever $n \geq 5$ thanks to recent advances in decoupling theory (see [BD15, GS09, GS10, ŁW02, Wol00] and [BHS, HNS11, LS13, MSS92, Sog91] for more on decoupling and local smoothing estimates). We remark that results in $n = 4$ could be obtained upon further progress on local smoothing estimates.

Acknowledgements. We would like to thank Juha Kinnunen for his question about regularising properties of the fractional spherical maximal function, which led to this work. We also thank Jonathan Hickman for discussions on the spherical maximal function and local smoothing estimates.

2.2 Notation and Preliminaries

2.2.1 Notation

All function spaces are defined over $\mathbb{R}^n$, and it is written, for instance, L^2 for $L^2(\mathbb{R}^n)$. The letter C denotes a generic constant whose value may vary from line to line. Its dependency on other parameters will be clear from the context. The notation $A \lesssim B$ is used if $A \leq CB$ for such a constant C, and similarly $A \gtrsim B$ and $A \sim B$. The Fourier transform of a tempered distribution $f \in \mathcal{S}'$ is denoted by $\widehat{f}$ or $\mathcal{F}(f)$ and its inverse Fourier transform by $\mathcal{F}^{-1}(f)$ or $f^{\vee}$; in particular for a Schwartz function $f \in \mathcal{S}$,

$$\widehat{f}(\xi) = \mathcal{F}f(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x \cdot \xi} f(x)\, dx.$$

Given any multi-index $\gamma \in \mathbb{N}^n$, ∂^γ denotes

$$\partial^\gamma f = \partial_{x_1}^{\gamma_1} \cdots \partial_{x_n}^{\gamma_n} f.$$

For any $\alpha \in \mathbb{R}$, the notation $(-\Delta)^{\alpha/2}$ is taken to denote the operator associated to the Fourier multiplier $|\xi|^\alpha$.

2.2.2 Besov spaces and Littlewood–Paley pieces

Given a smooth function $\psi \in C_c^\infty$ supported in $\{\xi \in \mathbb{R}^n : 2^{-1} < |\xi| < 2\}$ and such that

$$\sum_{j \in \mathbb{Z}} \psi(2^{-j}\xi) = 1$$

for $\xi \neq 0$, let f_j denote the Littlewood–Paley piece of f at frequency 2^j, given by $\widehat{f_j} = \widehat{f}\psi(2^{-j}\cdot)$. The Besov seminorm for $\dot{B}^s_{p,q}$ for $s \in \mathbb{R}$ and $p, q \in [1, \infty]$ is defined as

$$\|f\|_{\dot{B}^s_{p,q}} = \Big(\sum_{j \in \mathbb{Z}} 2^{qjs} \|f_j\|_{L^p}^q \Big)^{1/q};$$

the seminorms defined through different Littlewood-Paley functions ψ are comparable (see [BL76, Chapter 6] for further details).

2.2.3 BV space

A function f is said to have bounded variation, and denoted by $f \in \dot{\mathrm{B}}\mathrm{V}$, if its variation

$$|f|_{\mathrm{BV}} := \sup \Big\{ \int_{\mathbb{R}^n} f \operatorname{div}(g);\ \ g \in C_c^1(\mathbb{R}^n, \mathbb{R}^n),\ \|g\|_\infty \leq 1 \Big\}$$

is finite, where $g = (g_1, \dots, g_n)$ and the L^∞ norm is defined by

$$\|g\|_\infty := \|(\sum_{i=1}^n g_i^2)^{1/2}\|_{L^\infty}.$$

Note that if f belongs to space $W^{1,1}$, integration by parts allows one to identify

$$|f|_{\mathrm{BV}} = \int_{\mathbb{R}^n} |\nabla f|.$$

See [EG92, Chapter 5] for more.

2.2.4 Finite differences

Denote

$$D^h f(x) = \frac{f(x+h) - f(x)}{|h|}.$$

Recall (see e.g [Eva10, Chapter 5, §5.8, Theorem 3]) that if there is a finite constant A such that

$$\left\| D^h f \right\|_{L^p} \leq A$$

for all $h \in \mathbb{R}^n$, then the weak derivatives of f exist and

$$\|\nabla f\|_{L^p} \leq CA$$

for a constant C only depending on the dimension n. If S is a sublinear operator that commutes with translations, then

$$|D^h S f| \leq |S D^h f|.$$

In particular, if S is a maximal function and f is a positive function, this allows us to reduce the question about differentiability to boundedness of a maximal multiplier for all Schwartz functions f.

2.3 Endpoint results

2.3.1 A model result

It is instructive to start first with a model case for Theorem 2.1. This consists in the study of the single scale version of the (rough) fractional maximal function M_α, defined as

$$M_\alpha^* f = \sup_{1 \leq t \leq 2} \left| \frac{1}{|B(x,t)|} \int_{B(x,t)} f(y)\, dy \right|.$$

Theorem 2.3. *Let $0 < \alpha < 1$, $p = n/(n-\alpha)$ and $n \geq 2$. Then there is a constant C only depending on dimension n and α such that for any $f \in \dot{B}^{1-\alpha}_{p,1}$*

$$\|M_\alpha^* D^h f\|_{L^p} \leq C \|f\|_{\dot{B}^{1-\alpha}_{p,1}}$$

uniformly on $h \in \mathbb{R}^n$.

By the discussion in §§2.2.4, Theorem 2.3 implies an L^p bound for the gradient of M_α^*. It will be shown in §§2.3.2 how the proof of the above estimate gives Theorem 2.1 for sightly smoother versions of the fractional maximal function, such as its lacunary version or maximal functions of convolution type with smooth kernels.

Proof. Write, for $f \in \mathcal{S}$,

$$M_\alpha^*(D^h f)(x) = \sup_{1 \leq t \leq 2} |\mathcal{F}^{-1}((t|\xi|)^\alpha \widehat{1_{B(0,1)}}(t\xi) \mathcal{F}(T_h (-\Delta)^{(1-\alpha)/2} f))|$$

where T_h is the operator defined by

$$\widehat{T_h g}(\xi) = \frac{e^{i\xi \cdot h} - 1}{|\xi||h|} \widehat{g}(\xi) =: a_h(\xi) \widehat{g}(\xi). \tag{2.5}$$

Observe that T_h is a bounded operator on L^p uniformly in $h \in \mathbb{R}^n$ for all $1 < p < \infty$ by the Mikhlin–Hörmander multiplier theorem (see, for instance [Duo01, Theorem 8.10]); it is clear that

$$|\partial^\gamma a_h(\xi)| \lesssim |\xi|^{-|\gamma|} \qquad \text{for all multi-indexes } \gamma \in \mathbb{N}_0^n$$

with implicit constant independent of $h \in \mathbb{R}^n$. Thus, the operator T_h plays no role in determining the range of boundedness for $M_\alpha^* D^h$.

Let $m(\xi) = |\xi|^\alpha \widehat{1_{B(0,1)}}(\xi)$ and $m_t(\xi) := m(t\xi)$ for all $t > 0$. For each $j \in \mathbb{Z}$, let $f_j = \check{\psi}_j * f$ denote the Littlewood-Paley piece of f around the frequency 2^j as in §§2.2.2. Assume momentarily that the following holds.

Proposition 2.4. *Let $g \in \mathcal{S}$. Then for $p = n/(n-\alpha)$ and $0 < \alpha < n/2$,*

$$\| \sup_{1 \le t \le 2} |\mathcal{F}^{-1}(m_t \widehat{g}_j)| \|_{L^p} \lesssim (2^{j\alpha} 1_{\{j \le 0\}} + 1_{\{j > 0\}}) \|g_j\|_{L^p}.$$

Then the proof may be concluded as follows. Decomposing the function f into frequency localised pieces f_j and applying Proposition 2.4 to the function $g = T_h(-\Delta)^{(1-\alpha)/2} f$ one has

$$\begin{aligned}
\| \sup_{1 \le t \le 2} |\mathcal{F}^{-1}(m_t \widehat{g})| \|_{L^p} &\le \sum_{j \in \mathbb{Z}} \| \sup_{1 \le t \le 2} |\mathcal{F}^{-1}(m_t \widehat{g}_j)| \|_{L^p} \\
&\lesssim \sum_{j \in \mathbb{Z}} (2^{j\alpha} 1_{\{j \le 0\}} + 1_{\{j > 0\}}) \|g_j\|_{L^p} \\
&\le \sum_{j \in \mathbb{Z}} 2^{j(1-\alpha)} \|f_j\|_{L^p} \sim \|f\|_{\dot{B}^{1-\alpha}_{p,1}},
\end{aligned} \tag{2.6}$$

where the last step follows from the L^p boundedness of T_h and Young's convolution inequality.

Remark 2.5. By Bernstein's inequality, $2^{j(1-\alpha)} \|f_j\|_{L^p} \lesssim 2^j \|f_j\|_{L^1}$, so one may further bound $\|f\|_{\dot{B}^{1-\alpha}_{p,1}} \lesssim \|f\|_{\dot{B}^1_{1,1}}$ in (2.6).

It remains to prove Proposition 2.4. This is done by interpolating an L^2 bound with an $L^1 - L^{1,\infty}$ bound as in the proof of the spherical maximal function theorem that can be found in the textbooks [Ste93, Chapter XI, §3.3] or [Gra14a, Chapter 5.5]. Writing

$$\mathcal{F}^{-1}(m_t \widehat{g}_j) = t^\alpha \mathcal{F}^{-1}(\widehat{1_{B(0,1)}}(t\xi)(|\xi|^\alpha \widehat{g}_j)),$$

it is clear that

$$\sup_{1 \le t \le 2} |\mathcal{F}^{-1}(m_t \widehat{g})| \lesssim \sup_{1 \le t \le 2} |t^{-n} 1_{B(0,t)} * ((-\Delta)^{\alpha/2} g)| \le M((-\Delta)^{\alpha/2} g)$$

where M is the Hardy–Littlewood maximal function. Bounds on M and Young's convolution inequality then imply

Proposition 2.6. *Let $g \in \mathcal{S}$. Then*

$$\| \sup_{1 \le t \le 2} |\mathcal{F}^{-1}(m_t \widehat{g}_j)| \|_{L^{1,\infty}} \lesssim 2^{j\alpha} \|g_j\|_{L^1}.$$

The L^2 estimate follows by estimating the Fourier decay of m after an application of a Sobolev embedding. This is the part of the proof that allows to take advantage of better symbols m later in §§2.3.3 so we write the proof in detail.

Proposition 2.7. *Let $g \in \mathcal{S}$. Then*

$$\|\sup_{1\le t\le 2} |\mathcal{F}^{-1}(m_t \widehat{g}_j)|\|_{L^2} \lesssim (2^{j\alpha} 1_{\{j\le 0\}} + 2^{j(-\frac{n}{2}+\alpha)} 1_{\{j>0\}}) \|g_j\|_{L^2}.$$

Proof. Let $\tilde{m}(\xi) = \xi \cdot \nabla m(\xi)$ and denote by T_m and $T_{\tilde{m}}$ the operators associated to the multipliers m and $\tilde{m}$. By the fundamental theorem of calculus,

$$\begin{aligned}\sup_{1\le t\le 2} |T_{m_t} g_j| &\le |T_m g_j| + 2\left(\int_1^2 |T_{m_t} g_j||T_{\tilde{m}_t} g_j| \frac{dt}{t}\right)^{1/2} \\ &\le |T_m g_j| + 2\left(\int_1^2 |T_{m_t} g_j|^2 \frac{dt}{t}\right)^{1/4} \left(\int_1^2 |T_{\tilde{m}_t} g_j|^2 \frac{dt}{t}\right)^{1/4}. \end{aligned} \tag{2.7}$$

Taking L^2-norm in the above expression, an application of the Cauchy–Schwarz inequality and Fubini's theorem reduces the problem to compute the L^∞ norm of $m\psi_j$ and $\tilde{m}\psi_j$.

Recall that $\widehat{1_{B(0,1)}}(\xi) = |2\pi\xi|^{-n/2} J_{n/2}(2\pi|\xi|)$, where $J_{n/2}$ denotes the Bessel function of order $n/2$, and

$$J_{n/2}(r) \lesssim r^{n/2} 1_{\{r\le 1\}} + r^{-1/2} 1_{\{r>1\}};$$

see, for instance, [Gra14a, Appendix B] for further details. This immediately yields

$$\|m\psi_j\|_{L^\infty} \lesssim 2^{j\alpha} 1_{\{j\le 0\}} + 2^{j(-\frac{n+1}{2}+\alpha)} 1_{\{j>0\}}. \tag{2.8}$$

Concerning $\tilde{m}$, the relation

$$\frac{d}{dr}[r^{-n/2} J_{n/2}(r)] = -r^{-n/2} J_{n/2+1}(r)$$

and a similar analysis to the one carried above leads to

$$\|\tilde{m}\psi_j\|_{L^\infty} \lesssim 2^{j\alpha} 1_{\{j\le 0\}} + 2^{j(-\frac{n-1}{2}+\alpha)} 1_{\{j>0\}}.$$

Putting both estimates together in (2.7) concludes the proof. □

Proposition 2.4 now follows by interpolation, and the proof of the model case is complete. □

2.3.2 Extension to the full supremum

From now on, we redefine m to be Fourier transform of an integrable function smoother than $1_{B(0,1)}$. Momentarily assume m satisfies

$$\|\sup_{1\le t\le 2} |(m_t \widehat{g}_j)^\vee|\|_{L^p} \lesssim (2^{j\alpha} 1_{\{j\le 0\}} + 2^{-j\varepsilon} 1_{\{j>0\}}) \|g_j\|_{L^p} \tag{2.9}$$

with $p = n/(n-\alpha)$, which we next show to be enough to conclude a bound as in Theorem 2.1. The proof of (2.9) is postponed to §§2.3.3.

Inequality (2.9) rescales as

$$\|\sup_{2^{-k}\le t\le 2^{-k+1}} |(m_t \widehat{g}_{j+k})^\vee|\|_{L^p} \lesssim (2^{j\alpha} 1_{\{j\le 0\}} + 2^{-j\varepsilon} 1_{\{j>0\}}) \|g_{j+k}\|_{L^p}. \tag{2.10}$$

In order to use this bound, break the full supremum over all possible scales and use the embedding $\ell^p \subseteq \ell^\infty$,

$$\sup_{t>0} |(m_t\widehat{g})^\vee| = \sup_{k\in\mathbb{Z}} \sup_{2^{-k}\le t\le 2^{-k+1}} |(m_t\widehat{g})^\vee| \le \Big(\sum_{k\in\mathbb{Z}} \sup_{2^{-k}\le t\le 2^{-k+1}} |(m_t\widehat{g})^\vee|^p\Big)^{1/p}.$$

Taking L^p norm and using (2.10), we see

$$\|\sup_{t>0} |(m_t\widehat{g})^\vee|\|_{L^p} \lesssim \sum_{j\in\mathbb{Z}} (2^{j\alpha}1_{\{j\le 0\}} + 2^{-j\varepsilon}1_{\{j>0\}})\Big(\sum_{k\in\mathbb{Z}} \|g_{j+k}\|_{L^p}^p\Big)^{1/p}.$$

Using the geometric decay to sum in $j \in \mathbb{Z}$ and recalling

$$\|g_{j+k}\|_{L^p} = \|(-\Delta)^{(1-\alpha)/2} f_{j+k}\|_{L^p} \lesssim 2^{(j+k)(1-\alpha)}\|f_{j+k}\|_{L^p},$$

we obtain

$$\Big(\sum_{k\in\mathbb{Z}} \|g_{j+k}\|_{L^p}^p\Big)^{1/p} \lesssim \|f\|_{\dot{B}^{1-\alpha}_{p,p}}.$$

We then claim

$$\|f\|_{\dot{B}^{1-\alpha}_{p,p}} \lesssim |f|_{\mathrm{BV}} \tag{2.11}$$

for $n > 1$ and $0 < \alpha < n/2$. This will follow from a Gagliardo–Nirenberg–Sobolev type inequality.

Proposition 2.8 ([CDDD03]). *Assume $\gamma > 1$ or $\gamma < 1 - 1/n$, and let (s, q) satisfy $(s-1)q'/n = \gamma - 1$ for some $1 < q \le \infty$, where $1/q + 1/q' = 1$. Then, for any $0 < \theta < 1$,*

$$\|f\|_{\dot{B}^t_{p,p}} \lesssim \|f\|^{1-\theta}_{\dot{B}^s_{q,q}} |f|^\theta_{\mathrm{BV}}$$

where $\frac{1}{p} = \frac{1-\theta}{q} + \theta$ and $t = (1-\theta)s + \theta$.

Indeed, taking $\gamma = 0$, $s = 1 - n/2$ and $\theta = 1 - 2\alpha/n$, which are admissible for $n > 1$ and $0 < \alpha < n/2$, one has

$$\|f\|_{\dot{B}^{1-\alpha}_{p,p}} \lesssim \|f\|^{1-\theta}_{\dot{B}^{1-n/2}_{2,2}} |f|^\theta_{\mathrm{BV}}.$$

Applying Bernstein's and Minkowski's inequalities as well as Littlewood–Paley theory, we see

$$\begin{aligned}\|f\|_{\dot{B}^{1-\frac{n}{2}}_{2,2}} &\sim \Big(\sum_{j\in\mathbb{Z}} 2^{2j(1-\frac{n}{2})}\|f_j\|^2_{L^2}\Big)^{1/2} \lesssim \Big(\sum_{j\in\mathbb{Z}} 2^{2j(1-\frac{n}{2})}2^{2jn(\frac{n-1}{n}-\frac{1}{2})}\|f_j\|^2_{L^{\frac{n}{n-1}}}\Big)^{1/2}\\ &= \Big(\sum_{j\in\mathbb{Z}} \|f_j\|^2_{L^{\frac{n}{n-1}}}\Big)^{1/2} \le \Big\|\Big(\sum_{j\in\mathbb{Z}} |f_j|^2\Big)^{1/2}\Big\|_{L^{\frac{n}{n-1}}} \sim \|f\|_{L^{\frac{n}{n-1}}}.\end{aligned}$$

Inequality (2.11) then follows from the Gagliardo–Nirenberg–Sobolev inequality [EG92, Theorem 5.6.1. (i)], and we conclude

$$\|\sup_{1\le t\le 2} |\mathcal{F}^{-1}(m_t\widehat{g})|\|_{L^p} \lesssim \|f\|^{1-\theta}_{L^{n/(n-1)}} |f|^\theta_{\mathrm{BV}} \lesssim |f|_{\mathrm{BV}}.$$

Thus it suffices to verify (2.9). This is done separately in the cases when m comes from a smooth kernel and when the maximal function is lacunary.

2.3.3 Smooth kernel

Define the smooth fractional maximal function as follows. Let $\epsilon > 0$. Let φ be a positive function with radial L^1 majorant such that $\widehat{\varphi}(\xi) \lesssim_\varphi (1+|\xi|)^{-n/2-\epsilon}$. For instance, any positive Schwartz function or even

$$\varphi(x) = (1-|x|^2)_+^\epsilon$$

with $\epsilon > 0$ will do (see Appendix B.5 in [Gra14a]). The subscript denotes the positive part as $f_+ = f \cdot 1_{\{f>0\}}$. Now we want to analyse M_α^φ, as defined in the introduction. A repetition of the proof of Proposition 2.7 gives the L^2 bound

$$\| \sup_{1\le t\le 2} |\mathcal{F}^{-1}((t|\xi|)^\alpha \widehat{\varphi}(t\xi)\widehat{g}_j)|\|_{L^2} \lesssim (1_{\{j\le 0\}}2^{j\alpha} + 1_{\{j>0\}}2^{j(-\frac{n}{2}+\alpha-\epsilon)})\|g_j\|_{L^2}.$$

The ϵ-decay gain in the above estimate continues to hold on $L^{n/(n-\alpha)}$, so the extra decay assumption (2.9) is satisfied for smooth convolution kernels. By §§2.3.2, Theorem 2.1 holds in this case.

2.3.4 Lacunary set of radii

Similarly, there is a gain in the L^2 estimate when we study the lacunary fractional maximal function. Now $m(\xi) = |\xi|^\alpha \widehat{1_{B(0,1)}}(\xi)$ and

$$c_n M_\alpha^{lac} f(x) = \sup_{k\in\mathbb{Z}} |2^{k\alpha - nk} \int_{B(x,2^k)} f(y)\,dy| \le \Big(\sum_{k\in\mathbb{Z}} |2^{k\alpha-nk} \int_{B(x,2^k)} f(y)\,dy|^p \Big)^{1/p}$$

so that it suffices to use a bound for a single dilate instead of a supremum bound. Thus, it is enough to use (2.8) to replace Proposition (2.7) by

$$\||\mathcal{F}^{-1}(m\widehat{g}_j)|\|_{L^2} \lesssim (2^{j\alpha}1_{\{j\le 0\}} + 2^{j(-\frac{n+1}{2}+\alpha)}1_{\{j>0\}})\|g_j\|_{L^2},$$

which has an extra 1/2-decay compared to Proposition 2.7. After interpolation, this leads to an ε-decay gain in the $L^{n/(n-\alpha)}$ estimate so that (2.9) (without supremum) and Theorem 2.1 for lacunary set of radii follow.

2.4 Proof of Theorem 2.2

Recall the definition (2.4). By the characterisation through finite differences described in §2, the sublinearity of S_α and by density, it suffices to prove

$$\|S_\alpha D^h f\|_{L^q} \lesssim \|f\|_{L^p}$$

for all Schwartz functions f uniformly in $h \in \mathbb{R}^n$.

Observe that by means of Fourier transform,

$$S_\alpha D^h f(x) = \sup_{t>0} \left| \mathcal{F}^{-1}\left(t^\alpha |\xi| \widehat{\sigma}(t\xi) \mathcal{F}(T_h f)(x)\right)\right|, \tag{2.12}$$

where T_h is the Fourier multiplier operator (2.5). As described in §§2.3.1, T_h is bounded on L^p for all $1 < p < \infty$ uniformly in $h \in \mathbb{R}^n$ by the Mikhlin–Hörmander multiplier theorem, so it plays no role in determining the boundedness range for $S_\alpha D^h$; for this reason, $T_h f$ is identified with f in the rest of this section.

2.4.1 The case $q \geq 2$

It is enough to consider the single scale version of the maximal function in (2.12): suppose we can prove

$$\|\sup_{1\leq t\leq 2}|\mathcal{F}^{-1}(t^{\alpha}|\xi|\widehat{\sigma}(t\xi)\widehat{f}_j)|\|_{L^q} \lesssim (2^{js_1}1_{\{j\leq 0\}} + 2^{-js_2}1_{\{j>0\}})\|f_j\|_{L^p} \tag{2.13}$$

for $s_1, s_2 > 0$. Then rescaling gives

$$\|\sup_{2^{-k}\leq t\leq 2^{-k+1}}|\mathcal{F}^{-1}(t^{\alpha}|\xi|\widehat{\sigma}(t\xi)\widehat{f_{j+k}})|\|_{L^q} \lesssim (2^{js_1}1_{\{j\leq 0\}} + 2^{-js_2}1_{\{j>0\}})\|f_{j+k}\|_{L^p}$$

under the relation $\frac{1}{q} = \frac{1}{p} - \frac{\alpha-1}{n}$, and arguing as in §§2.3.2

$$\begin{aligned}\|\sup_{t>0}|\mathcal{F}^{-1}(t^{\alpha}|\xi|\widehat{\sigma}(t\xi)\widehat{f})|\|_{L^q} &\lesssim \sum_{j\in\mathbb{Z}}(2^{js_1}1_{\{j\leq 0\}} + 2^{-js_2}1_{\{j>0\}})\Big(\sum_{k\in\mathbb{Z}}\|f_{j+k}\|_{L^p}^q\Big)^{1/q} \\ &\lesssim \|f\|_{L^p}\end{aligned}$$

where the last inequality follows from Minkowski's inequality ($q \geq p$); controlling ℓ^q norm by ℓ^2 norm, and applying Littlewood–Paley theory to see the inner sum as L^p norm of f. The sum in j converges as $s_1, s_2 > 0$. Hence it suffices to prove (2.13).

For low frequencies $j \leq 0$, we can use domination by the Hardy–Littlewood maximal function, Young's convolution inequality and Bernstein's inequality to see

$$\|\sup_{1\leq t\leq 2}|\mathcal{F}^{-1}(t^{\alpha}|\xi|\widehat{\sigma}(t\xi)\widehat{f}_j)|\|_{L^q} \lesssim \|M(-\Delta)^{1/2}f_j\|_{L^q} \lesssim 2^{j(1+\alpha)}\|f_j\|_{L^p}.$$

Hence it suffices to prove (2.13) for $j > 0$.

2.4.2 A local smoothing estimate

The Fourier transform of the spherical measure is

$$\widehat{\sigma}(\xi) = 2\pi|\xi|^{-\frac{n-2}{2}}J_{\frac{n-2}{2}}(2\pi|\xi|) = \sum_{\pm} a_{\pm}(\xi)e^{\pm 2\pi i|\xi|},$$

where the symbols $a_{\pm}$ are in the class $S^{-(n-1)/2}$, that is

$$|\partial_{\xi}^{\gamma}a_{\pm}(\xi)| \lesssim (1+|\xi|)^{-\frac{n-1}{2}-|\gamma|}$$

for all multi-indices $\gamma \in \mathbb{N}_0^n$ (c.f. [Ste93, Chapter VIII]). Hence

$$\mathcal{F}^{-1}(\widehat{\sigma}(t\xi)\widehat{f}) = \sum_{\pm}\int_{\mathbb{R}^n} e^{2\pi i(x\cdot\xi \pm t|\xi|)}a_{\pm}(t\xi)\widehat{f}(\xi)\,d\xi,$$

so that the connection to half-wave propagator $e^{it\sqrt{-\Delta}}f(x) := \int_{\mathbb{R}^n} e^{ix\cdot\xi}e^{it|\xi|}\widehat{f}(\xi)d\xi$ is evident. We will quote the following result:

Proposition 2.9 (Consequence of [BD15])**.** *For $n \geq 2$, $s \in \mathbb{R}$,*

$$\Big(\int_1^2 \|e^{it\sqrt{-\Delta}}f\|_{L^p_{s-s_p+\theta}(\mathbb{R}^n)}^p dt\Big)^{1/p} \lesssim \|f\|_{L^p_s(\mathbb{R}^n)}$$

holds for $0 \leq \theta < \frac{1}{p}$ and $s_p = (n-1)\big(\frac{1}{2} - \frac{1}{p}\big)$ whenever $p \geq \frac{2(n+1)}{n-1}$.

This can be found as Corollary 1.3 (i) in [GS09] knowing that the conjectured value of p_d in Table 1 of that paper has later been verified by [BD15].

Proposition 2.10. *Let g be a Schwartz function and $j > 0$. For any $\epsilon > 0$*

$$\| \sup_{1\le t\le 2} |\sigma_t * g_j| \|_{L^{n-1}} \lesssim_\epsilon 2^{j(\epsilon-1)} \|g_j\|_{L^{n-1}}.$$

Proof. For $j > 0$ and a smooth bump χ around $[1,2]$, we have

$$\begin{aligned} \| \sup_{1\le t\le 2} |\sigma_t * g_j| \|_{L^{n-1}(\mathbb{R}^n)} &\lesssim \|(1+\sqrt{-\partial_t^2})^r \chi \cdot \sigma_t * g_j\|_{L^{n-1}(\mathbb{R}^{n+1})} \\ &\lesssim 2^{j\left(r+s_p-\theta-\frac{n-1}{2}+\epsilon\right)} \|g_j\|_{L^{n-1}(\mathbb{R}^n)} \end{aligned}$$

where we used Sobolev embedding with $r > 1/(n-1)$, Proposition 2.9 with $p = n-1$ as well as Young's convolution inequality. Simplifying the exponent in accordance with Proposition 2.9 [1], we obtain the claim. $\square$

2.4.3 $L^p \to L^q$ estimates

To finish the proof of (2.13), we prove $L^p \to L^q$ estimates following the interpolation scheme of Lee [Lee03] enhanced with the sharp local smoothing estimate. Denote

$$S_j^* f(x) := \sup_{1\le t\le 2} |\mathcal{F}^{-1}(\widehat{\sigma}(t\xi)|\xi|\widehat{f}_j(\xi))(x)|,$$

where $\widehat{f}_j = \widehat{f}\psi_j$ still stands for Fourier localization at the level of a Littlewood–Paley piece of frequency 2^j.

Proposition 2.11. *Let P be the open convex polygon with vertices*

$$\begin{aligned} A &= \Big(\frac{n-2}{n}, \frac{2}{n}\Big), \quad B = \Big(\frac{n^2-2n-1}{n^2+1}, \frac{2(n-1)}{n^2+1}\Big) \\ C &= \Big(\frac{1}{n-1}, \frac{1}{n-1}\Big), \quad D = \Big(\frac{n-2}{n}, \frac{n-2}{n}\Big). \end{aligned}$$

Then

$$\|S_j^* f\|_{L^q} \lesssim 2^{-\varepsilon j} \|f_j\|_{L^p}$$

for some $\varepsilon > 0$ and all $j > 0$ provided that $(1/p, 1/q) \in P$.

Proof. Since $\mathrm{supp}(\widehat{\sigma}\cdot\psi_j(t\cdot)) \subset \{|\xi| \sim 2^j\}$, we can assume that $\widehat{f}$ is supported in an annulus around $|\xi| = 2^j$. We use the following bounds:

$$\begin{aligned} \|S_j^* f\|_{L^1} &\lesssim 2^{2j}\|f\|_{L^1} \\ \|S_j^* f\|_{L^\infty} &\lesssim 2^{2j}\|f\|_{L^1} \\ \|S_j^* f\|_{L^{n-1}} &\lesssim_\delta 2^{j\delta}\|f\|_{L^{n-1}}, \quad \text{for all } \delta > 0 \qquad (2.14) \\ \|S_j^* f\|_{L^2} &\lesssim 2^{-\frac{n-4}{2}j}\|f\|_{L^2} \\ \|S_j^* f\|_{L^{\frac{2(n+1)}{n-1}}} &\lesssim 2^{-j\frac{n^2-4n-3}{2n+2}}\|f\|_{L^2}. \end{aligned}$$

[1] The full strength of [BD15] is only needed when $n = 5$. When $n \geq 6$, the earlier results from [ŁW02] will already do.

To verify (2.14), use Proposition 2.10 as well as Young's convolution inequality to obtain

$$\|S_j^* f\|_{L^{n-1}} \lesssim_\delta 2^{-j(1-\delta)} \|(-\Delta)^{1/2} f\|_{L^{n-1}} \lesssim 2^{j\delta} \|f\|_{L^{n-1}}.$$

The other inequalities follow similarly, that is, by borrowing the corresponding bounds for the spherical maximal function (inequalities (1.7)–(1.10) in [Lee03]), and applying Young's convolution inequality. Interpolating the bounds above, we obtain the claimed proposition. □

For each $p > 1$, we want to find the values of α such that $(1/p, 1/q) \in P$ when $(\alpha - 1)/n = 1/p - 1/q$ and $q \geq 2$. When $q \geq 2$ is assumed, this happens when

$$\frac{n}{n-2} < p \leq \frac{n^2+1}{n^2-2n-1}, \quad \alpha < \frac{n^2-2n-1}{n-1} - \frac{2n}{p(n-1)}$$

or

$$\frac{n^2+1}{n^2-2n-1} < p \leq n-1, \quad \alpha < \frac{n-1}{p}.$$

This concludes the proof for the case $q \geq 2$. Notice that the restriction $q \geq 2$ is not dictated by validity of $L^p \to L^q$ estimates but it was required in order to upgrade the single scale bounds to bounds for the full maximal operator in §§2.4.1.

2.4.4 The case $q \leq 2$

Next we remove the assumption $q \geq 2$. Let

$$T^* f(x) = \sup_{t>0} |\mathcal{F}^{-1}((t|\xi|)^\alpha \widehat{\sigma}(t\xi) \widehat{f}(\xi))(x)|.$$

The operator S_α in (2.12) can be written

$$S_\alpha = T^* I_{\alpha-1} T_h f$$

where $\widehat{I_{\alpha-1} f} = |\xi|^{1-\alpha} \widehat{f}$ is the Riesz potential of order $\alpha - 1$ and T_h are as in (2.5). As discussed in §§2.3.1, T_h are bounded in L^p for all $p > 1$. Also, by the Hardy–Littlewood–Sobolev inequality $I_{\alpha-1}$ is bounded $L^p \to L^q$, for p, q obeying $\frac{\alpha-1}{n} = \frac{1}{p} - \frac{1}{q}$. Therefore, it is enough to analyse the operator T^*.

Let $m(\xi) = |\xi|^\alpha \widehat{\sigma}(\xi)$ and take a Littlewood–Paley function ψ (as in §2.2). We define $m_1 = \sum_{j>0} \psi_j m$ and $m_0 = \sum_{j \leq 0} \psi_j m$. Take T_j^* to be as T^* but m replaced by m_j. Then

$$T^* f \leq T_0^* f + T_1^* f.$$

We first bound T_0^*. A straightforward computation shows that m_0 is bounded and for any multi-index $\beta \in \mathbb{N}^n$ with $|\beta| = k$, $k \leq n+1$

$$|\partial_\xi^\beta m_0(\xi)| \lesssim |\xi|^{\alpha - k}$$

so that

$$\|(1 + |\cdot|)^{n+1} \mathcal{F}^{-1}(m_0)\|_{L^\infty} \lesssim 1.$$

Consequently

$$T_0^* f \lesssim Mf$$

and boundedness in any L^p with $p > 1$ follows from that of the Hardy–Littlewood maximal function.

To bound T_1^*, we use a part of Theorem B from [RdF86]:

Theorem 2.12 (Rubio de Francia [RdF86]). *Let m be a function in $C^{s+1}(\mathbb{R}^n)$ for some integer $s > n/2$ such that $|D^\gamma m(\xi)| \lesssim |\xi|^{-a}$, for all $|\gamma| \le s+1$. Suppose also that $a > \frac{1}{2}$. Then the maximal multiplier operator $T^* f := \sup_{t>0} |\mathcal{F}^{-1}(m(t\cdot)\widehat{f})|$ is bounded in L^r, for*

$$\frac{2n}{n+2a-1} < r \le 2.$$

Since $\sum_{j>0} \psi_j m$ is smooth and satisfies $|D^\gamma m(\xi)| \lesssim |\xi|^{-a}$, for all $\gamma \in \mathbb{N}^n$ with $a = \frac{n-1}{2} - \alpha$, we can apply the theorem to conclude the proof whenever

$$\frac{2n}{2n-2-2\alpha} < q \le 2, \quad a > \frac{1}{2}$$

which is equivalent to $p > \frac{n}{n-2}$ and $\alpha < \frac{n-2}{2} < \alpha(p)$. However, given $p > \frac{n}{n-2}$, the condition $\alpha < \frac{n-2}{2}$ is automatically satisfied whenever $q \le 2$. Hence $\alpha < \alpha(p)$ is an active constraint only when $q > 2$.

□

Chapter 3

Weak differentiability for fractional maximal function of general L^p functions on domains

Wir müssen wissen, wir werden wissen.

–*D.H.*

This chapter contains the paper [RSW], a collaboration between the author of this thesis, Olli Saari and Julian Weigt. Let $\Omega \subset \mathbb{R}^n$ be bounded a domain. We prove under certain structural assumptions that the fractional maximal operator relative to Ω maps $L^p(\Omega) \to W^{1,p}(\Omega)$ for all $p > 1$, when the smoothness index $\alpha \geq 1$. In particular, the results are valid in the range $p \in (1, n/(n-1)]$ that was previously unknown. As an application, we prove an endpoint regularity result in the domain setting.

3.1 Introduction

Regularity of the Hardy–Littlewood maximal function of a Sobolev function was first studied in [Kin97]. It was shown that the maximal operator preserves $W^{1,p}(\mathbb{R}^n)$ regularity for $p > 1$. This continues to hold true at the derivative level when $p = 1$ and $n = 1$ [Tan02, Kur15] and for radial functions [Lui18]. Extending such a statement to more general Sobolev functions of several variables is a difficult open problem, which has inspired many results in related topics. For instance, slightly stronger bounds have been proved for maximal operators with more special convolution kernels (see [CS13], [CFS15], [CGR19] and [PPSS18]), the continuity of the mapping has been studied in [Lui07] and [CMP17], and a part of the techniques used for continuity, also relevant for the current paper, have been extended to $p = 1$ in [HM10].

Another aspect of the problem is the fractional endpoint question proposed by Carneiro and Madrid [CM17]. The fractional maximal function is given by

$$M_\alpha f(x) = \sup_{r>0} \frac{r^\alpha}{|B(x,r)|} \int_{B(x,r)} |f(y)| \, dy,$$

and it defines a bounded operator $L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$ when $q = np/(n-p)$ and $p > 1$. This boundedness fails at the endpoint $p = 1$, but the question about boundedness of ∇M_α from $W^{1,1}(\mathbb{R}^n)$ to $L^{n/(n-\alpha)}(\mathbb{R}^n)$ has not been answered so far for $\alpha < 1$ (see [LM17], [BM19] and [BRS19] for related research and partial results). The case $\alpha \geq 1$ turned out to be very simple, and the reason can be traced back to the inequality

$$|\nabla M_\alpha f(x)| \leq c_{\alpha,n} M_{\alpha-1} f(x) \tag{3.1}$$

of Kinnunen and Saksman [KS03]. Carneiro and Madrid [CM17] noted that (3.1) together with the Gagliardo–Sobolev–Nirenberg inequality and the $L^p \to L^q$ bounds for the fractional maximal function imply the expected endpoint bound when $\alpha \geq 1$.

In the present paper, we study these problems in general open subsets of $\mathbb{R}^n$, which is a natural context for analysis from the point of view of potential theory and partial differential equations. Regularity of the local Hardy–Littlewood maximal function of a Sobolev function on an open $\Omega \subset \mathbb{R}^n$ was first studied by Kinnunen and Lindqvist [KL98], and a local variant of the inequality for the derivative of the fractional maximal function (3.1) was proved in [HKKT15]. This is our starting point, and for more thorough discussion of what was proved and what is unknown, we introduce some more notation.

If $\Omega \subset \mathbb{R}^n$ is an open set, the local fractional maximal function is defined as

$$M_\alpha^\Omega f(x) = \sup_{0<r<\operatorname{dist}(x,\Omega^c)} \frac{r^\alpha}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy.$$

As the boundary of Ω restricts the choice of r in the definition, one cannot expect (3.1) to trivially carry over to the local setting. Indeed, such a pointwise inequality is false in general (Example 4.1 in [HKKT15]). On the other hand, if one adds a correction term involving the surface measure of the sphere to the right hand side of (3.1), one obtains

$$|\nabla M_\alpha^\Omega f(x)| \leq c_{\alpha,n} \left(M_{\alpha-1} f(x) + \sup_{r>0} |r^{\alpha-1} \sigma_r * f(x)| \right), \tag{3.2}$$

which is valid in all domains. This was used in [HKKT15] to prove that L^p functions with $p > n/(n-1)$ large enough have $M_\alpha^\Omega f$ in a first order Sobolev class. The lower bound on p rules out functions too singular for an application of a spherical maximal function argument.

Our main theorem shows that under suitable assumptions on the domain Ω, the maximal function M_α^Ω maps $L^p(\Omega)$ into a first order Sobolev space for all $p > 1$.

Theorem 3.1. *Let $\Omega \subset \mathbb{R}^n$ be open, $n \geq 2$, $p > 1$ and $f \in L^p(\Omega)$. Then $M_\alpha^\Omega f$ is weakly differentiable and*

$$\|\nabla M_\alpha^\Omega f\|_{L^p(\Omega)} \leq C \|f\|_{L^p(\Omega)}$$

if any one of the following holds:

i. $\alpha > 1$ *and* Ω *is bounded.*

ii. $\alpha = 1$ *and* Ω^c *is convex.*

iii. $\alpha = 1$ *and* Ω *is bounded and satisfies a uniform interior ball condition (see Section 3.2.2 for the definition).*

iv. $\alpha = 1$ *and* $p > 1 + \frac{1}{n}$.

The constant C depends on the dimension, and in A and C it also depends on α and the domain.

Unlike [HKKT15], we are not able to prove an $L^p \to L^q$ smoothing effect on top of winning one derivative. However, our method does apply to singular functions in L^p spaces with $1 \leq p \leq n/(n-1)$ where the argument in [HKKT15] fails to give any result. In particular, we have the following endpoint regularity result, which was previously out of reach.

Corollary 3.2. *Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain. Then for all $f \in W^{1,1}(\Omega)$*

$$\|\nabla M_1^\Omega f\|_{L^{n/(n-1)}(\Omega)} \leq C\|f\|_{W^{1,1}(\Omega)}$$

where the constant C only depends on Ω and the dimension.

We briefly outline the proof of the main theorem. The maximal function on a domain behaves differently depending on whether the ball attaining the maximum touches the boundary or not. In case it does not, the local maximal function behaves like the global one, and the analysis is very similar. Otherwise it coincides with a linear averaging operator (3.5), which depends on the domain. These two parts are analyzed separately, and the main part of the proof is to establish L^p bounds for the derivative of (3.5). This leads to studying a domain dependent weighted spherical averaging operator (3.6).

Instead of resorting to maximal averages and the Bourgain–Stein theorem, an angular decomposition of the operator is carried out. The additional geometric information allows instead to establish good L^1 bounds that can be interpolated with trivial L^∞ bounds in order to obtain a domination of (3.6) by a converging sum of L^p bounded operators. Improving the L^1 bound over what follows from the behaviour of generic spherical means is crucial when aiming at L^p bounds for all $p > 1$. Such a conclusion cannot be drawn from mere polynomial decay of the Fourier transform of the weighted spherical measure in question, if no additional L^1 information is taken into account. Turning the focus from the Littlewood–Paley decomposition and L^2 methods to an angular decomposition and geometric estimates in L^1 is the leading insight of the proof.

The key idea in the L^1 estimates can be described as follows. Each domain Ω comes endowed with a family of sets (Figure 3.1)

$$\{P(y) : y \in \Omega\}, \quad P(y) = \{x \in \Omega : y \in \partial B(x, \operatorname{dist}(x, \Omega^c))\},$$

which can morally be used to dualize the spherical averaging operators (3.6) through Fubini's theorem. The L^1 bounds for the constituents in the angular decomposition of the spherical averaging operator correspond to weighted integrals over the pieces of $P(y)$. If Ω is a ball, then the sets $P(y)$ are ellipsoids with foci at the center of the ball and at y. In the cases of the complement of a ball and a half-space, the $P(y)$ take the simple forms of hyperboloids and paraboloids. One cannot hope for as explicit descriptions as that in more general domains, but all $P(y)$ are boundaries of convex sets. This observation is used extensively in the proof.

The structure of the paper is as follows. In the first section, we introduce notation and some tools that will be helpful throughout the proof. The first sections are about differentiating the maximal function on so-called unconstrained points and proving the weak differentiability of the maximal function conditionally to the L^p boundedness of the averaging operator (3.5). The rest of the paper is devoted to proving those L^p bounds by first computing a formula for the derivative and then carrying out the strategy sketched above. Finally, there is a concluding section with remarks on open problems and certain observations about the proof which might be of independent interest for future research.

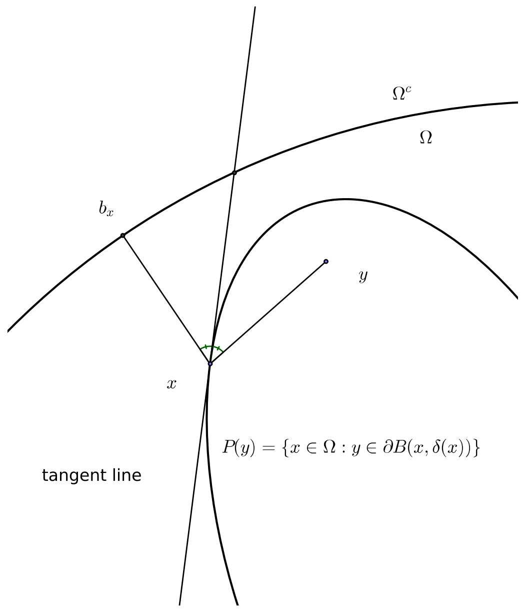

Figure 3.1: A set $P(y)$ and a tangent line.

Acknowledgement. The authors would like to thank Juha Kinnunen for turning our attention to the error term in [HKKT15]. Part of the research was done when the second author was visiting the Nonlinear PDE group at Aalto University, which he wishes to thank for its hospitality. JPGR was supported by the Deutscher Akademischer Austauschdienst (DAAD). OS was supported by Hausdorff Center for Mathematics DFG EXC 2047 and the project DFG SFB 1060. JW was supported by the Väisälä Foundation and the Academy of Finland.

3.2 Preliminaries

3.2.1 Notation

We let $n \geq 1$ denote the dimension. For a measurable set E, we let $|E|$ denote the n-dimensional Lebesgue measure. The k-dimensional Hausdorff measures are denoted by $\mathcal{H}^k$. An Euclidean ball with center $x \in \mathbb{R}^n$ and radius $r > 0$ is denoted by $B(x, r)$. A finite constant only depending on quantities that are not being kept track of is denoted by C. If $A \leq CB$ for such constant, we denote $A \lesssim B$ or write A is $\lesssim B$. We write $A \sim B$ if both $A \lesssim B$ and $B \lesssim A$ hold.

3.2.2 Domains

We always assume $\Omega \subset \mathbb{R}^n$ to be an open set, which we interchangeably call domain as the distinction obviously plays no role in this paper. We assume it to have non-empty complement. The distance function is denoted by $\delta(x) = \mathrm{dist}(x, \Omega^c)$. As Ω^c is closed, there exists at least one $b_x \in \Omega^c$ so that $|x - b_x| = \delta(x)$. We reserve the notation b_x for such a point, which need not be unique unless Ω^c is convex. The distance function $\delta : \Omega \to [0, \infty)$ is always 1-Lipschitz. The gradient exists almost everywhere by Rademacher's theorem, and it holds that

$$\nabla\delta(x) = \frac{x - b_x}{\delta(x)}. \tag{3.3}$$

This is because clearly the one sided directional derivative of $\delta(x)$ in the direction of $b_x - x$ always exists and is -1. Where the gradient exists, we can use $|\nabla\delta(x)| \leq 1$ to conclude that the directional derivative in all directions orthogonal to $x - b_x$ must be zero.

A domain is said to satisfy a uniform interior ball condition if there is an $R > 0$ so that for every point $b \in \partial\Omega$ there exists a ball $B(z, R) \subset \Omega$ so that $\partial B(x, R) \cap \partial\Omega = \{b\}$. All bounded C^2 domains satisfy this condition, but a domain satisfying a uniform interior ball condition might be non-smooth and have inwards-pointing cusps.

3.2.3 Function spaces on domains

Functions $f \in L^p(\Omega)$ are a priori only defined in the domain Ω, but we always extend them by zero to $\mathbb{R}^n$ without additional comments. The Sobolev class $W^{1,p}(\Omega)$ consists of functions $f \in L^p(\Omega)$ such that $|\nabla f| \in L^p(\Omega)$. The weak derivatives are defined using test functions in $C_c^\infty(\Omega)$.

For the application of the main theorem to the endpoint regularity problem, we need a Sobolev embedding theorem for domains. One concrete case we can deal with is that of a Lipschitz domain.

Proposition 3.3 (Section 4.4 in [EG92])**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set so that $\partial\Omega$ is Lipschitz. Then for every $1 \leq p < \infty$ there exists a bounded extension operator*

$$E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$$

such that $\operatorname{supp}(Ef) \subset B(x_0, 2\operatorname{diam}(\Omega))$ *for some $x_0 \in \Omega$ and all $f \in W^{1,p}(\Omega)$.*

By the boundary being Lipschitz, we mean that it can be covered by a finite number of open balls B_i so that for each i the domain $B_i \cap \Omega$ is the epigraph of a Lipschitz function.

The proposition together with the Gagliardo–Nirenberg–Sobolev inequality (see e.g. Section 4.5.1 in [EG92]) implies a rudimentary local Sobolev embedding

$$\|f\|_{L^{pn/(n-p)}(\Omega)} \leq C_{\Omega,p,n}\|f\|_{W^{1,p}(\Omega)} \tag{3.4}$$

valid for all $f \in W^{1,p}(\Omega)$ whenever Ω is a bounded open set with Lipschitz boundary. This is sufficient for our purposes.

3.2.4 Maximal function

For $\alpha \in [1, n)$, define the local fractional maximal function relative to Ω as

$$M_\alpha^\Omega f(x) = \sup_{0<r<\delta(x)} r^\alpha \fint_{B(x,r)} f(y)\, dy$$

whenever $f \in L^1_{\mathrm{loc}}(\Omega)$. We omit the superscript when Ω is the whole $\mathbb{R}^n$. In addition, we define for $\alpha \in \mathbb{R}$ the auxiliary linear operator

$$A_\alpha f(x) = \delta(x)^\alpha \fint_{B(z,\delta(x))} f(y)\, dy. \tag{3.5}$$

3.2.5 Constrained points

Let f be continuous. Fix $x \in \Omega$. Because the complement of Ω is non-empty, $\delta(x) \leq \operatorname{diam}(\Omega) < \infty$ and there exists a convergent sequence $r_j \in (0, \delta(x))$ with limit $r = \lim_{j\to\infty} r_j \in [0, \delta(x)]$ such that

$$M_\alpha^\Omega f(x) = \lim_{j\to\infty} r_j^\alpha \fint_{B(x,r_j)} f(y)\, dy = r^\alpha \fint_{B(x,r)} f(y)\, dy$$

if $r > 0$. If

$$M_\alpha^\Omega f(x) > \delta(x)^\alpha \fint_{B(x,\delta(x))} f(y)\, dy,$$

the sequence r_j must be chosen so that $r < \delta(x)$, and the point x is said to be *unconstrained*. All other points are called *constrained*.

3.3 The unconstrained part

The local maximal function behaves similarly to the global one in the unconstrained set, and we reduce the differentiability question of the unconstrained part accordingly to that of the global maximal function. This is the content of the following proposition.

Proposition 3.4. *Let $p > 1$, $\alpha \geq 1$ and $f \in L^p_{loc}(\Omega)$ be continuous. The set U of the unconstrained points is open, the maximal function $M_\alpha^\Omega f$ is weakly differentiable in U, and the pointwise bound*

$$|\nabla M_\alpha^\Omega f(x)| \leq c M_{\alpha-1} f(x)$$

holds for a constant c only depending on the dimension and α whenever $x \in U$.

Proof. Consider the fractional average function

$$A(z,r) := r^\alpha \fint_{B(z,r)} f(y)\, dy.$$

It is continuous in $(z,r) \in \Omega \times \mathbb{R}_+$. Fix now an unconstrained point x. By definition, there exists $\varepsilon > 0$ so that $M_\alpha^\Omega f(x) - A(x,\delta(x)) > \epsilon$. Moreover, there exists $\gamma > 0$ so that if $|(z,r) - (x,\delta(x))| < \gamma$, then $M_\alpha^\Omega f(x) - A(z,r) > \varepsilon/2$. Since $M_\alpha^\Omega f$ is lower semicontinuous, one can find for every z close enough to x a sequence $r_{z,j} \to r_z < \delta(x) - \gamma/2$ so that

$$M_\alpha^\Omega f(z) = \lim_{j\to\infty} r_{z,j}^\alpha \fint_{B(z,r_{z,j})} f(y)\, dy.$$

In particular, there is an open neighborhood U_x of x so that for all $z \in U_x$

$$M_\alpha^\Omega f(z) = M_\alpha(1_{B(x,\delta(x))} f)(z).$$

By Theorem 3.1 in [KS03],

$$|\nabla M_\alpha^\Omega f(x)| \leq C M_{\alpha-1} f(x)$$

follows. □

3.4 The full maximal function

Next we prove the differentiability of the local maximal function conditional to L^p bounds for the derivative of the averaging operator (3.5). This step morally follows from the lattice property of Sobolev functions, but as we only know the weak differentiability of $M_\alpha^\Omega f$ in the unconstrained set, some extra work is needed.

Lemma 3.5. *Let $p > 1$, $\alpha \geq 1$ and $\Omega \subset \mathbb{R}^n$ be such that ∇A_α and $M_{\alpha-1}$ are bounded $L^p(\Omega) \to L^p(\Omega)$. If $f \in L^p(\Omega)$, then the local fractional maximal function is weakly differentiable and*

$$\|\nabla M_\alpha^\Omega f\|_{L^p(\Omega)} \lesssim \|f\|_{L^p(\Omega)}.$$

Proof. Assume first that f is continuous and compactly supported. Following the arguments in [KS03], we infer that $M_\alpha^\Omega f$ can be seen as supremum over radii between a fixed upper and lower bound. The fractional averages are Lipschitz continuous with constants only depending on the radii, and hence their supremum is also Lipschitz. In particular, we know that $M_\alpha^\Omega f$ is continuous.

Denote by $g^+ = \max(g, 0)$ the positive part of a function g and write

$$M_\alpha^\Omega f = (M_\alpha^\Omega f - A_\alpha f)^+ + A_\alpha f.$$

By assumption, the second term admits the desired Sobolev bounds. To deal with the other term, let $\epsilon > 0$ and define

$$F_\epsilon(t) = \begin{cases} ((t-\epsilon)^2 + \epsilon^2)^{1/2} - \epsilon, & t > \epsilon \\ 0, & t \leq \epsilon. \end{cases}$$

These functions are of class $C^1(\mathbb{R})$ and converge pointwise to $t \mapsto (t)^+$ as $\epsilon \to 0$. Moreover, as $M_\alpha^\Omega f$ and $A_\alpha f$ are continuous, $E = \{x \in \Omega : F_\epsilon(M_\alpha^\Omega f(x) - A_\alpha f(x)) > 0\}$ has its closure contained in the open set of unconstrained points U. By Proposition 3.4, the assumption on A_α and the chain rule for Sobolev derivatives (4.2.2 in [EG92]), we obtain for all partial derivatives ∂_i

$$\partial_i F_\epsilon(M_\alpha^\Omega f - A_\alpha f) = (\partial_i M_\alpha^\Omega f - \partial_i A_\alpha) F'_\epsilon(M_\alpha^\Omega f - A_\alpha f).$$

Taking a test function φ and computing

$$\int_\Omega F_\epsilon(M_\alpha^\Omega f - A_\alpha f)\partial_i\varphi \, dx = \int_\Omega (\partial_i M_\alpha^\Omega f - \partial_i A_\alpha f) F'_\epsilon(M_\alpha^\Omega f - A_\alpha f)\varphi \, dx,$$

we see that taking the limit $\epsilon \to 0$ proves the claim for continuous and compactly supported f.

To deal with the general case, let $f \in L^p(\Omega)$ and let f_j be continuous and compactly supported functions converging to f in L^p norm. By L^p continuity of the fractional maximal operator, $M_\alpha^\Omega f_j \to M_\alpha^\Omega f$ in L^p. As we have proved the following inequality

$$\|\nabla M_\alpha^\Omega f_j\|_{L^p(\Omega)} \lesssim \|f_j\|_{L^p(\Omega)},$$

for continuous functions f_j, the sequence $M_\alpha^\Omega f_j$ is bounded in $W^{1,p}(\Omega)$. We can extract a weakly convergent subsequence. By taking limits along this sequence and using the uniqueness of distributional limit, we conclude the proof for general $f \in L^p(\Omega)$. □

As the main theorem is a direct consequence of the previous lemma, it remains to investigate the boundedness of the operator ∇A_α on $L^p(\Omega)$. The following sections are devoted to establishing the required L^p bounds when Ω is sufficiently well-behaved.

3.5 Constrained part

By a change of variables, we can write the averaging operator (3.5) as

$$A_\alpha f(x) = \delta(x)^\alpha \fint_{B(0,1)} f(x + y\delta(x))\, dy.$$

This operator is linear, and as we are aiming for L^p bounds, there is no loss of generality in restricting the attention to smooth functions. If x is a constrained point, then $M_\alpha^\Omega f(x) = A_\alpha f(x)$, which justifies our reference to A_α as the constrained part. Also, Lemma 3.5 showed that L^p bounds for the derivative of $A_\alpha f$ are enough to imply weak differentiability of the full maximal operator, so the maximal function does not play any role in what follows. A version of the following proposition was already proved in [HKKT15], but as we need a formula more precise than what they stated, we include the short proof for clarity.

Proposition 3.6. *Let $f \in C^\infty(\Omega)$. Then for almost every $x \in \Omega$*

$$|\nabla A_\alpha f(x)| \leq c_{n,\alpha}|A_{\alpha-1} f(x)| + c_n \delta(x)^{\alpha-1} \fint_{B(x,\delta(x))} \frac{|y - b_x|}{\delta(x)} f(y)\, d\mathcal{H}^{n-1}(y)$$

where $b_x \in \partial\Omega$ is a point such that $|b_x - x| = \delta(x)$.

Proof. Fix a point x. As $A_\alpha f(x) = \delta(x)^\alpha A_0 f(x)$, it holds that

$$\nabla A_\alpha f(x) = \alpha\delta(x)^{\alpha-1} A_0 f(x) \nabla\delta(x) + \delta(x)^\alpha (\nabla A_0 f)(x).$$

Since $|\nabla\delta(x)| \leq 1$ (cf. (3.3)), the first summand above is bounded by $A_{\alpha-1} f(x)$. Thus it suffices to analyze the gradient of $A_0 f$. Take the unit vector

$$e = \nabla A_0 f(x)/|\nabla A_0 f(x)|.$$

Then

$$\begin{aligned}
|\nabla(A_0 f)(x)| &= (e \cdot \nabla) A_0 f(x) \\
&= \fint_{B(0,1)} \big(e + y(e \cdot \nabla\delta(x))\big) \cdot \nabla f(x + \delta(x)y)\, dy \\
&= \frac{1}{\delta(x)} \fint_{B(0,1)} \operatorname{div}_y \big((e + y(e \cdot \nabla\delta(x))) f(x + \delta(x)y)\big)\, dy \\
&\quad - \frac{n e \cdot \nabla\delta(x)}{\delta(x)} \fint_{B(0,1)} f(x + \delta(x)y)\, dy =: \mathrm{I} + \mathrm{II}\,.
\end{aligned}$$

Since $|\nabla\delta(x)| \leq 1$, the contribution $\delta(x)^\alpha \cdot \mathrm{II}$ is pointwise bounded by $nA_{\alpha-1} f$. To estimate the other term, we apply Gauss's theorem to obtain

$$\begin{aligned}
\mathrm{I} &= \frac{c_n}{\delta(x)} \int_{\partial B(0,1)} y \cdot (e + y(e \cdot \nabla\delta(x))) f(x + \delta(x)y)\, d\mathcal{H}^{n-1}(y) \\
&= \frac{c_n}{\delta(x)} \fint_{\partial B(x,\delta(x))} \frac{(y - b_x) \cdot e}{\delta(x)} f(y) dy.
\end{aligned}$$

So we reach the inequality

$$|\nabla A_\alpha f(x)| \leq |\alpha - n||A_{\alpha-1}(x)| + c_n \delta(x)^{\alpha-1} \fint_{B(x,\delta(x))} \frac{|y - b_x|}{\delta(x)} f(y)\, d\mathcal{H}^{n-1}(y),$$

which proves the claim. $\square$

Because $A_{\alpha-1}f(x) \leq M^{\Omega}_{\alpha-1}f(x)$, and $M^{\Omega}_{\alpha-1}$ satisfies the right $L^p \to L^q$ bounds, we have reduced the matter to understanding the weighted spherical average

$$B_\alpha f(x) := \delta(x)^{\alpha-1} \fint_{B(x,\delta(x))} \frac{|y-b_x|}{\delta(x)} f(y)\, d\mathcal{H}^{n-1}(y) \tag{3.6}$$

on the right hand side of the conclusion of the previous proposition. The weight $|y - b_x|/\delta(x)$ measures the angle between $b_x - x$ and $y - x$ when $|y - b_x|/\delta(x)$ is small. We decompose the weighted spherical averaging operator according to the angle and location in the domain as follows. For $k \in \mathbb{Z}$, let

$$\Omega_k = \{x \in \Omega : 2^k \leq \delta(x) < 2^{k+1}\}$$

and for every point $x \in \Omega$ and integer $j \geq 0$

$$\omega_j(x) = \left\{ y \in \partial B(x, \delta(x)) : 2^{-j} < \frac{|y-b_x|}{\delta(x)} \leq 2^{-j+1} \right\}.$$

Define

$$S^k_j f(x) = 1_{\Omega_k}(x) \int_{\omega_j(x)} f(y)\, d\mathcal{H}^{n-1}(y).$$

Then

$$B_\alpha f(x) \lesssim \sum_{k\in\mathbb{Z}} \sum_{j=1}^{\infty} 2^{k(\alpha-n)-j} S^k_j f(x) \tag{3.7}$$

and it remains to prove bounds for S^k_j so that the right hand side sums up in L^p. This is done by interpolating bounds on L^∞ and L^1.

Proposition 3.7. *Let Ω be any domain. It holds that $\|S^k_j\|_{L^\infty\to L^\infty} \lesssim 2^{(n-1)(k-j)}$, and consequently $\|\sum_k 2^{k(1-n)} S^k_j\|_{L^\infty\to L^\infty} \lesssim 2^{-(n-1)j}$.*

Proof. This follows from $\mathcal{H}^{n-1}(\omega_j(x)) \lesssim 2^{(n-1)(k-j)}$. □

3.6 L^1 bounds

To prove L^1 bounds, we introduce some more notation. For each integer $j \geq 0$ and each point $y \in \Omega$, define

$$P_j(y) = \{x \in \Omega : y \in \omega_j(x)\}, \quad P(y) = \bigcup_{j=0}^{\infty} P_j(y). \tag{3.8}$$

In addition, let

$$A^k_j = \bigcup_{x\in\Omega_k} \omega_j(x). \tag{3.9}$$

Formally, certain weighted integrals over $P(y)$ give the adjoint operator of B_α. A naive change of order of integration is not justified in this case, but using the decomposition of B_α, we can make the idea precise. The following two propositions give effective description of $P(y)$ and provide a substitute for Fubini's theorem.

Proposition 3.8. *Let Ω be an open set and let $y \in \Omega$. Then*

$$E(y) = \{x \in \Omega : |x - y| \leq \delta(x)\}$$

is closed and convex set such that

$$P(y) = \partial E(y).$$

For each $x \in P(y)$, the supporting hyperplane at x bisects the angle between $y - x$ and $b_x - x$ and is normal to $b_x - y$.

Proof. Recall that $P(y)$ consists of the points with $\{x \in \Omega : |y - x| = \delta(x)\}$. For $x \in P(y)$, it holds that

$$x + \epsilon \frac{b_x - x}{|b_x - x|} \in E(y)^c,$$

and it is easy to see $\partial E(y) = P(y)$. Consider the hyperplane

$$\{z \in \mathbb{R}^n : |z - b_x| = |z - y|\}.$$

It divides the space into two half spaces $H_1 = \{z : |z - b_x| < |z - y|\}$ and $H_2 = \{z : |z - b_x| \geq |z - y|\}$. If $x \in P(y)$, then $E(y) \subset H_2$ and $x \in H_2$. Thus ∂H_2 is a supporting hyperplane for $E(y)$ at x. As every boundary point of $E(y)$ has a supporting hyperplane, $E(y)$ is convex. The remaining assertions readily follow from the definition of ∂H_2. □

Proposition 3.9. *Let Ω be a domain, $j \geq 0$ and k integers and $f \geq 0$ a bounded continuous function on Ω. Then*

$$\int_\Omega S_j^k f(x)\, dx \lesssim 2^j \int_{A_j^k} f(y) \mathcal{H}^{n-1}(P_j^k(y))\, dy,$$

where we let $P_j^k(y) = P_j(y) \cap \Omega_k$.

Note that $y \in A_j^k$ if and only if $P_j^k(y) \neq \emptyset$.

Proof. The parameter k plays no role in the following computation, but is included in the statement for future reference. Let $\varphi \geq 0$ be a smooth function of one variable with compact support in $(0, 1)$ and $\|\varphi\|_{L^1(\mathbb{R})} = 1$. Denote the ϵ-dilation by $\varphi_\epsilon(t) = \epsilon^{-1}\varphi(t\epsilon^{-1})$. For any fixed x, we define the set of relevant directions

$$\omega_j^{\text{dir}} = \delta(x)^{-1}(\omega_j(x) - x) \subset \partial B(0, 1).$$

As f is positive, the weak convergence

$$S_j^k f(x) = \int_{\omega_j(x)} f(y)\, d\mathcal{H}^{n-1}(y) \leq \lim_{\epsilon \to 0} \int_{x + \mathbb{R}\omega_j^{\text{dir}}(x)} f(y) \varphi_\epsilon\left(\delta(x) - |x - y|\right)\, dy$$

holds. Integrating over x, applying the dominated convergence theorem (this is justified, see the remark at the end of the argument), and using Fubini's theorem, we obtain

$$\int_{\Omega_k} S_j f(x) dx \lesssim \int_{A_j^k} f(y) \left(\lim_{\epsilon \to 0} \frac{|\{x \in \Omega_k : y \in \omega_j^{\epsilon-}(x)\}|}{\epsilon} \right) dy \tag{3.10}$$

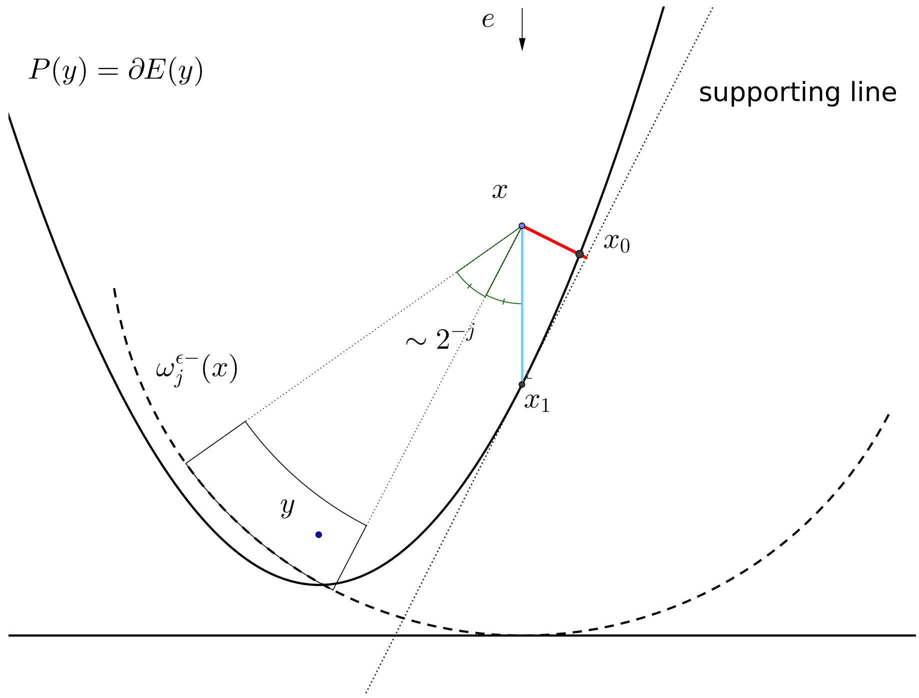

Figure 3.2: The construction to find x_0.

where the one-sided neighborhood is defined as

$$\omega_j^{\epsilon-}(x) = x + \omega_j^{\mathrm{dir}}(x)(\delta(x) - \epsilon, \delta(x)).$$

Next we estimate the limit expression in (3.10). As j and k are fixed, we can assume ϵ to be very small relative to them. Fix y. Let $x \in \Omega_k$. Assume that $y \in \omega_j^{\epsilon-}(x)$. Then

$$-\epsilon < |y - x| - \delta(x) < 0 \tag{3.11}$$

and by definition $x \in E(y)$.

Set

$$e = \frac{b_x - x}{|b_x - x|}$$

and let $r \in (0, \delta(x))$ be such that $x + re \in P(y)$. Next we give an upper bound for r. Because $y \in \omega_j^{\epsilon-}(x)$, it also holds that

$$\frac{y - x}{|y - x|} \in \omega_j^{\mathrm{dir}}(x).$$

The mapping

$$g(\rho) := |y - (x + \rho e)| - \delta(x + \rho e) = |y - x - \rho e| - \delta(x) + \rho$$

is Lipschitz and hence absolutely continuous.

For all $\rho \geq 0$ we have the lower bound

$$\begin{aligned} g'(\rho) = \partial_\rho[|y - (x + \rho e)| - \delta(x + \rho e)] &= -e \cdot \frac{y - x - \rho e}{|y - x - \rho e|} + 1 \\ &= 1 - \cos \measuredangle(b_x - x, y - x - \rho e) \\ &\geq 1 - \cos \measuredangle(b_x - x, y - x) \\ &\gtrsim 2^{-2j} \end{aligned}$$

The last inequality is due to $y \in \omega_j^{\epsilon-}(x)$. Recall that $g(0) \geq -\epsilon$ and $g(r) = 0$. Since g is absolutely continuous, we conclude

$$2^{-2j} r \lesssim \int_0^r g'(s)\, ds = g(r) - g(0) \leq \epsilon,$$

and

$$r \lesssim 2^{2j}\epsilon.$$

Denote $x_0 = x + re \in P(y)$. Consider the 2-plane containing x, y, b_x (and x_0). Its intersection with the convex body $E(y)$ provided by Proposition 3.8 is again a convex set E' in the plane. Let ℓ be its supporting line at x_0. Then

$$\measuredangle(b_x - x_0, y - x_0) \geq \measuredangle(b_x - x, y - x) \geq 2^{-j}$$
$$\sin \measuredangle(b_x - x_0, y - x_0) \leq \frac{|b_x - y|}{\delta(x) - C2^{2j}\epsilon} = \frac{|b_x - y|\delta(x)^{-1}}{1 - C\delta(x)^{-1}2^{2j}\epsilon} \leq \sin 2^{-j+2}$$

for ϵ small enough. By Proposition 3.8 this means that $y - x_0$ makes an angle $\sim 2^{-j}$ with ℓ, and hence so does $x - x_0$. Let e' be the unit vector perpendicular to ℓ and $e' \cdot (y - x) < 0$. Then there is

$$s \lesssim |x - x_0| \sin 2^{-j} \lesssim 2^j \epsilon$$

so that $x + se' \in \ell$. Since $x \in E'(y)$ and ℓ intersects $E'(y)$ only in $\partial E'(y)$, there is $s' < s$ with $x + se' \in \partial E'(y)$, which means

$$\operatorname{dist}(x, P(y)) \lesssim 2^j \epsilon. \tag{3.12}$$

Since $x_0 \in P_j^k(y)$ we also have

$$\operatorname{dist}(x, P_j^k(y)) \lesssim 2^{2j}\epsilon.$$

Finally, let $N(\epsilon') = \{x \in P(y) : \operatorname{dist}(x, P_j^k(y)) \leq \epsilon'\}$. Then

$$\begin{aligned}
\lim_{\epsilon \to 0} & \frac{|\{x \in \Omega_k : y \in \omega_j^{\epsilon-}(x) \leq \epsilon\}|}{\epsilon} \\
&\leq \lim_{\epsilon' \to 0} \lim_{\epsilon \to 0} \frac{|\{x \in \Omega_k : \operatorname{dist}(x, P_j^k(y) \cap N(\epsilon')) \leq c_n 2^j \epsilon\}|}{\epsilon} \\
&\lesssim \lim_{\epsilon' \to 0} 2^j \mathcal{H}^{n-1}(P(y) \cap N(\epsilon')) = 2^j \mathcal{H}^{n-1}(P_j^k(y)),
\end{aligned}$$

where the second inequality follows, for instance, by Theorem 3.2.39 in [Fed69]. The integrable majorant of the sequence above that was needed for the application of the dominated convergence theorem before can be obtained by an application of coarea formula. This completes the proof. □

These two propositions are enough to conclude a general L^1 bound for the pieces S_j^k. This bound can be refined further, when additional regularity on the domain Ω is assumed.

Proposition 3.10. *Let Ω be an open set. Then $\|S_j^k\|_{L^1 \to L^1} \lesssim 2^{k(n-1)+j}$.*

Proof. If $x \in P(y) \cap \Omega_k$, then $|x-y| = \mathrm{dist}(x, \Omega^c) \le 2^{k+1}$. Hence $P(y) \cap \Omega_k \subset B(y, 2^{k+1})$. Recall that $P(y) = \partial E(y)$ and that $E(y)$ is convex. Thus $P(y) \cap \Omega_k \subset \partial(B(y, 2^{k+1}) \cap E(y))$ where $B(y, 2^{k+1}) \cap E(y)$ is convex. Since the perimeter of $B(y, 2^{k+1})$ dominates the perimeters of all convex sets with non-empty interior contained in it, we can conclude

$$\begin{aligned}\mathcal{H}^{n-1}(P_j^k(y)) \le \mathcal{H}^{n-1}(P(y) \cap B(y, 2^{k+1})) &\le \mathcal{H}^{n-1}(\partial(B(y, 2^{k+1}) \cap E(y))) \\ &\le \mathcal{H}^{n-1}(\partial B(y, 2^{k+1})) \lesssim 2^{k(n-1)}.\end{aligned}$$

Now the claim follows from Proposition 3.9. □

Remark 3.11. In case Ω is bounded and $\partial\Omega$ is C^2 smooth, the estimate for $\mathcal{H}^{n-1}(P_j^k(y))$ can be refined as follows. If $x \in P_j^k(y)$, then $|y - b_x| \le \delta(x) 2^{-j+1}$. This implies $\mathrm{dist}(y, \partial\Omega) \le \delta(x) \cdot 2^{-j+1}$ and further

$$|b_y - b_x| \le |b_y - y| + |y - b_x| \le 4\delta(x) \cdot 2^{-j}.$$

As the inward-pointing unit normal N_Ω at the boundary is well-defined and Lipschitz,

$$|N_\Omega(b_y) - N_\Omega(b_x)| \lesssim \mathrm{diam}(\Omega) 2^{-j}.$$

Because $N(b_z) = (z - b_z)/|z - b_z|$, this implies

$$\left| N_\Omega(b_y) - \frac{(x-y)}{\delta(x)} \right| \le |N_\Omega(b_y) - N_\Omega(b_x)| + \frac{|y - b_x|}{\delta(x)} \lesssim \mathrm{diam}(\Omega) \cdot 2^{-j}.$$

Therefore, all vectors $x - y$ with $y \in \omega_j(x)$ are within an angle $\sim \tilde{c}(\Omega) \cdot 2^{-j}$ of $N_\Omega(b_y)$. Hence the set $P_j^k(y)$ is contained in a cylinder of height $\sim 2^k$ and basis $\sim \tilde{c}(\Omega) \cdot 2^{k-j}$. By the inequality for perimeters of convex sets as in the proof of Proposition 3.10

$$\mathcal{H}^{n-1}(P_j^k(y)) \lesssim c(\Omega) 2^k \cdot 2^{(k-j)(n-2)}.$$

This dependency on j is sharp even for very flat domains as can be seen by letting Ω be a smoothed out $B(0, 10) \cap \{x_1 \ge 0\}$ and $y = 2^{-j} e_1$ and $k \le 0$.

However, as the estimate on $\mathcal{H}^{n-1}(P_j^k(y))$ is not the narrow gap of the proof of our main theorem, we do not pursue this aspect further.

The estimate $\mathcal{H}^{n-1}(P_j^k(y)) \lesssim 2^{k(n-1)}$ cannot be improved in general. If the boundary of the domain is a single point, the equality is achieved up to a constant. However, focusing on the whole $P_j(y)$ instead of single pieces $P_j^k(y)$, one can obtain a different estimate at cost of worsening the dependency on j. The following proposition is useful for small values of j, and it holds in very general domains.

Proposition 3.12. *Let Ω be an open set and $y \in \Omega$. Then*

$$\int_{P_j(y)} \frac{1}{\mathrm{dist}(x, y)^{n-1}} \, d\mathcal{H}^{n-1}(x) \lesssim 2^j$$

with the constant independent of y. In particular,

$$\| \sum_k 2^{k(1-n)} S_j^k \|_{L^1 \to L^1} \lesssim 2^{2j}$$

Proof. We have

$$\int_{P_j(y)} \frac{1}{\operatorname{dist}(x,y)^{n-1}}\, d\mathcal{H}^{n-1}(x) \lesssim \liminf_{\epsilon\to 0} \frac{1}{\epsilon} \int_{\{x\in E(y)^c:\operatorname{dist}(x,P_j(y))\le\epsilon\}} \frac{1}{\operatorname{dist}(x,y)^{n-1}}\, dx.$$

Given any point $x \in P_j(y)$ and a line $l_x = \{y+t(x-y) : t \in \mathbb{R}\}$, we see that by Proposition 3.8 the line makes an angle $\sim 2^{-j}$ with $P_j(y)$, and hence

$$\mathcal{H}^1(l_x \cap \{z \in E(y)^c : \operatorname{dist}(z, P_j(y)) \le \epsilon\}) \lesssim 2^j\epsilon.$$

The first claimed bound for the integral follows immediately from passing to polar coordinates with origin at y.

To prove the second claim, note that by Proposition 3.9

$$\begin{aligned}\int_\Omega \sum_k 2^{k(1-n)} S_j^k f(x)\, dx &\lesssim 2^j \int_\Omega f(y) \left(\sum_k 2^{k(1-n)}\mathcal{H}^{n-1}(P_j^k(y))\right) dy \\ &\lesssim 2^j \int_\Omega f(y) \left(\int_{P_j(y)} \frac{1}{\operatorname{dist}(x,y)^{n-1}}\, d\mathcal{H}^{n-1}(x)\right) dy \\ &\lesssim 2^{2j}\|f\|_{L^1},\end{aligned}$$

where the last step was an application of the first claim. □

3.7 L^p bounds and geometry

To conclude bounds for the operator B_α, we have to sum up all the pieces in the decomposition. In order to make this work, one has to ensure that there is enough decay in j and k. Although the L^1 bounds do not sum up, interpolation with the better L^∞ bounds provides us with enough decay in the angle parameter j. If Ω is bounded, we can take advantage of the $L^p(\Omega)$ spaces being nested and use the decay in the scale parameter k near the boundary to complete the proof with no smoothness assumptions on the boundary of the domain. This is possible only when we do not attempt to prove scalable estimates that would capture $L^p \to L^q$ smoothing beyond one derivative gain.

Theorem 3.13. *Let Ω be a bounded open set, $p, \alpha > 1$. Then*

$$\|B_\alpha\|_{L^p(\Omega)\to L^p(\Omega)} \lesssim \operatorname{diam}(\Omega)^{\alpha-1}$$

where the implicit constant only depends p, α and the dimension.

Proof. Let $S_j = \sum_k 2^{k(\alpha-n)} S_j^k$ so that $B_\alpha = \sum_j 2^{-j} S_j$. Then by Proposition 3.10

$$\begin{aligned}\|S_j\|_{L^1(\Omega)\to L^1(\Omega)} \le \sum_{k=-\infty}^{\log\operatorname{diam}(\Omega)+1} 2^{k(\alpha-n)}\|S_j^k\|_{L^1(\Omega)\to L^1(\Omega)} \\ \lesssim \sum_{k=-\infty}^{\log\operatorname{diam}(\Omega)+1} 2^{k(\alpha-1)}2^j \lesssim 2^j \operatorname{diam}(\Omega)^{\alpha-1}.\end{aligned}$$

By Proposition 3.7

$$\|2^{-j}S_j\|_{L^\infty(\Omega)\to L^\infty(\Omega)} \lesssim 2^{-nj}$$

and by interpolation we obtain

$$\|2^{-j}S_j\|_{L^p(\Omega)\to L^p(\Omega)} \lesssim 2^{-\frac{(p-1)n}{p}j}\operatorname{diam}(\Omega)^{\alpha-1}.$$

As the exponent is negative, we can sum up in j to conclude the proof. □

To deal with the critical case $\alpha = 1$ where our estimates have the correct scaling, we have to take into account finer properties of the boundary, as the estimation as rough as above leads to a logarithmic blow-up of the k-sum at the boundary.

Proposition 3.14. *Let Ω be an open set.*

- *If Ω satisfies the interior ball condition with R, then for all $y \in \Omega$ and $x \in P(y)$ with $\delta(x) \leq R$, it holds that*

$$\delta(x)(1 - \frac{\delta(x)}{R})(1 - \cos\beta) \leq \operatorname{dist}(y, \partial\Omega) \tag{3.13}$$

  where $\beta = \measuredangle(b_x - x, y - x)$.

- *If Ω^c is convex, then*

$$\delta(x)(1 - \cos\beta) \leq \operatorname{dist}(y, \partial\Omega). \tag{3.14}$$

Proof. Take $x \in \Omega$ and $y \in \partial B(x, \delta(x))$ and let β be the angle between $b_x - x$ and $y - x$. Because Ω satisfies a uniform interior ball condition, there is an $R > 0$ independent of x and y so that we can find a ball $B(z, R) \subset \Omega$ with $z = x + (x - b_x)R/\delta(x)$ so that $\overline{B(z,R)} \cap \partial\Omega = \{b_x\}$. The Pythagorean identity reads

$$\begin{aligned}
|z - y|^2 &= (\delta(x)\sin\beta)^2 + (R - \delta(x)(1 - \cos\beta))^2 \\
&= R^2(1 - 2\frac{\delta(x)}{R}(1 - \frac{\delta(x)}{R})(1 - \cos\beta)) \\
&\leq R^2(1 - \frac{\delta(x)}{R}(1 - \frac{\delta(x)}{R})(1 - \cos\beta))^2.
\end{aligned}$$

Let w be the closest point to y in $\partial B(z, R)$. Since z, y and w are on the same line, we get

$$\begin{aligned}
\operatorname{dist}(y, \partial\Omega) &\geq |y - w| = |z - w| - |z - y| \\
&\geq R - R(1 - \frac{\delta(x)}{R}(1 - \frac{\delta(x)}{R})(1 - \cos\beta)) \\
&= \delta(x)(1 - \frac{\delta(x)}{R})(1 - \cos\beta)
\end{aligned}$$

as claimed. If Ω^c is convex, then the interior ball condition is satisfied with $R = \infty$, whence the second claim follows.

□

Theorem 3.15. *Let Ω be an open set. Let $\alpha = 1$ and $p > 1$. Then*

- *If Ω is bounded and satisfies the interior ball condition, then*

$$\|B_1\|_{L^p(\Omega)\to L^p(\Omega)} \lesssim \log\left(\frac{\operatorname{diam}(\Omega)}{R} + 1\right)$$

  where R is the radius from the interior ball condition.

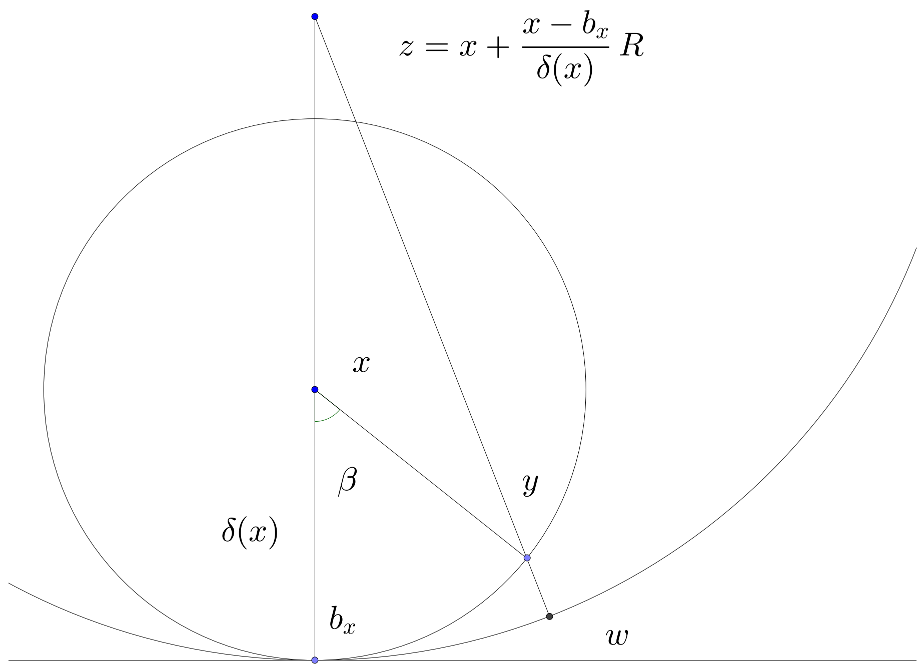

Figure 3.3: The balls and points appearing in the proof of Proposition 3.14.

- *If Ω^c is convex, then*
$$\|B_1\|_{L^p(\Omega)\to L^p(\Omega)} \lesssim 1$$
and the operator norm only depends on the dimension and p.
- *If Ω is merely open, then*
$$\|B_1\|_{L^p(\Omega)\to L^p(\Omega)} \lesssim 1$$
under the restriction $p > 1 + \frac{1}{n}$.

Proof. Proposition 3.10 implies
$$\int_\Omega 2^{k(1-n)-j} S_j^k f(x)\, dx \lesssim \int f(y) 1_{A_j^k}(y)\, dy. \tag{3.15}$$
Recall the definition (3.9). There are only $\sim \log(\operatorname{diam}(\Omega)/R+1)$ values of k so that $R/8 \leq 2^k \leq 2\operatorname{diam}(\Omega)$. For k such that $2^{k+3} \leq R$, we can use the first item in Proposition 3.14 to see that for fixed y, the set $P_j^k(y)$ is non-empty only for k such that $2^{-2j+k} \lesssim \operatorname{dist}(y, \partial\Omega)$. On the other hand, the upper bound
$$\operatorname{dist}(y, \partial\Omega) \leq |y - b_x| \lesssim 2^{k-j}$$
is always valid, so $P_j^k(y)$ is non-empty only for for $2^{-2j+k} \lesssim \operatorname{dist}(y, \partial\Omega) \lesssim 2^{-j+k}$. Consequently,
$$A_j^k \subset \{y \in \Omega : 2^{-2j+k} \lesssim \operatorname{dist}(y, \partial\Omega) \lesssim 2^{-j+k}\}.$$
For any y, there are only $\lesssim j$ values k such that the set above is non-empty, and hence by (3.15)
$$\|\sum_k 2^{k(1-n)-j} S_j^k\|_{L^1\to L^1} \lesssim \log\left(\frac{\operatorname{diam}(\Omega)}{R} + 1\right) + j.$$

Interpolation as in the proof of Theorem 3.13 implies the claim.

To prove the second item, just note that the convexity assumption on the complement means sending $R \to \infty$ so that $2^{k+3} \leq R$ always holds. To prove the third item, we study S_j as in the proof of Theorem 3.13 and replace the L^1 bound from Proposition 3.10 by that from Proposition 3.12. □

Corollary 3.16. *Let Ω be a domain, $p > 1$ and $f \in L^p$. Then $A_\alpha f(x)$ from* (3.5) *is weakly differentiable and*

$$\|\nabla A_\alpha f\|_{L^p(\Omega)} \lesssim \|f\|_{L^p(\Omega)}$$

if any one of the following holds:

- *$\alpha > 1$ and Ω is bounded.*
- *$\alpha = 1$ and Ω is bounded and satisfies a uniform interior ball condition.*
- *$\alpha = 1$ and Ω^c is convex.*

The constant depends on the domain, α and the dimension.

Proof. By linearity, it suffices to prove the norm inequality for smooth functions. By Proposition 3.6, it suffices to bound B_α from (3.6). This follows from Theorem 3.13 and Theorem 3.15 □

Theorem 3.1 follows from Corollary 3.16 and Lemma 3.5.

3.8 Remarks

3.8.1 Role of the domain

It is not clear if the conditions on the domain in the hypothesis of Theorem 3.1 are necessary. One may ask if

$$\|\nabla M_1^\Omega\|_{L^p(\Omega)\to L^p(\Omega)} \lesssim 1$$

holds for all domains Ω and all $p > 1$. We are not aware of any counterexamples so far. Since M_0^Ω does satisfy an $L^p(\Omega)$ bound independent of the domain, the question is about the behaviour of B_1 (see Theorem 3.15) in general domains. We point out that one avenue for improving the L^p bounds for B_1 could be to replace the strong L^1 bounds for S_j^k by weak type bounds in order to improve the operator norm bound with respect to j.

3.8.2 Endpoint regularity in domains

Corollary 3.2 follows from Theorem 3.1, since

$$\|\nabla M_1^\Omega f\|_{L^{n/(n-1)}(\Omega)} \lesssim \|f\|_{L^{n/(n-1)}(\Omega)} \lesssim \|f\|_{W^{1,1}(\Omega)}.$$

Here we used the main theorem and (3.4). The same observation was done by [CM17] to notice that the fractional endpoint regularity problem follows from inequality (3.1) as $\alpha \geq 1$ in the full space $\mathbb{R}^n$. The domain case was not known before as the inequality (3.1) should have been replaced by (3.2). This amounts to changing the Hardy–Littlewood maximal function to the spherical maximal function in the display above. That one is not bounded in $L^{n/(n-1)}$, so the argument breaks down. However, using Theorem 3.1, we can complete the argument in certain domains Ω.

To the best of our knowledge, the fractional endpoint regularity problem has not been studied in domains before. It is hence natural to ask

Question 3.17. What must be assumed about an open set $\Omega \subset \mathbb{R}^n$ so that

$$\|\nabla M_\alpha^\Omega f\|_{L^{\alpha/(n-\alpha)}(\Omega)} \lesssim \|f\|_{W^{1,1}(\Omega)}$$

for $\alpha \in (0,1]$?

Our main theorem gives some information on the case $\alpha = 1$, but the remaining values of α remain open. The values $\alpha > 1$ can be dealt with using a spherical maximal function argument with no additional assumptions. The remaining values of α are probably way harder to handle as the endpoint regularity question is completely open even in the full space.

Finally, we remark that the techniques used to get results for smooth kernels as in [BRS19] are insensitive to the ambient domain, because one does not use precise information about the maximizing radius. The arguments there only rely on sublinearity of maximal functions. Hence a $W^{1,1}$ variant of Theorem 1.1 in [BRS19] easily extends to the domain setting. Indeed, fixing $\alpha \in (0,1)$, letting Ω be any Sobolev extension domain, $\Omega_\epsilon = \{x \in \Omega : \operatorname{dist}(x, \Omega^c) \le \epsilon\}$ and m a local maximal function with kernel compactly supported and smooth enough as in [BRS19], one can invoke Theorem 3 in Section 5.8.2 in [Eva10] to reduce the problem to proving

$$\lim_{\epsilon \to 0} \sup_{h \in B(0,\epsilon/2)} \int_{\Omega_\epsilon} \left| \frac{mf(x+h) - mf(x)}{|h|} \right|^{\frac{n}{n-\alpha}} dx \lesssim \|f\|_{W^{1,1}(\Omega)}^{\frac{n}{n-\alpha}}.$$

As $f \in W^{1,1}(\Omega)$ coincides with its extension $Ef \in W^{1,1}(\mathbb{R}^n)$ for all $x \in \Omega$, the integral on the left hand side can be controlled by a maximal multiplier as in [BRS19] acting on $Ef(\cdot + h) - Ef(\cdot)$. Then the claim follows from Theorem 3.1 in [BRS19] and the assumed boundedness of $E : W^{1,1}(\Omega) \to W^{1,1}(\mathbb{R}^n)$.

3.8.3 Smoothing for cube maximal functions

An equally interesting variant of the local fractional maximal function is the one defined by taking averages over cubes instead of balls

$$M_\alpha^{\Omega,\text{cube}} f(x) = \sup_{r>0, Q(x,r) \subset \Omega} r^\alpha \fint_{Q(x,r)} f(y)\, dy.$$

As the faces of the cubes are completely flat, there are no L^p bounds for the maximal function

$$\sup_{r>0} \fint_{\partial Q(x,r)} f(y)\, d\mathcal{H}^{n-1}(y), \tag{3.16}$$

and this was singled out as the principal reason why the methods in [HKKT15] do not extend to the case of cubical fractional maximal function.

Although we avoid the use quantites of the type (3.16), our proof is also inapplicable to the cubical case. There are two obvious obstructions:

- Let Ω be the upper half-plane. Take $\delta > 0$ and define f as the characteristic function of $[-\delta, \delta] \times [0, \delta^s]$ for some $s \ge 1$. Varying s and sending $\delta \to 0$, we see that

  $$\|B_1 f\|_{L^p} \lesssim \|f\|_{L^p}$$

  cannot hold for any $p < \infty$.

- As a detail in the proof, one can note that the analogues of the sets $P(y)$ from (3.8) defined relative to cubes might have full measure. The role of curvature, or lack of it, manifests in the 2^j factor in the statement of Proposition (3.9).

On the other hand, it seems that the problems with the cubical maximal function are not only a matter of lack of curvature. As the remarks above show, there are domains where averages over flat surfaces cause problems. However, if the geometry of the domain is very special, this kind of phenomena can be ruled out. The following observation gives an example.

Proposition 3.18. *Let $\Omega = \{(x, y) \in \mathbb{R}^2 : x < y\}$. Then*

$$\|\nabla M_\alpha^{\Omega,cube} f\|_{L^p} \lesssim \|f\|_{L^p}$$

for all $f \in L^p$.

Sketch of proof. The reduction to the cubical analogue of (3.7) follows by the lines of the spherical proof. Then it suffices to note that the decomposition in j and k is unnecessary, and an L^p bound for $p > 1$ follows by Minkowski's inequality and a change of variables. □

The exact behaviour of the cubical local fractional maximal function in more general domains remains an interesting open problem.

3.8.4 Scalable estimates

The method of the proof of Theorem 3.1 forced us to prove $L^p \to W^{1,p}$ estimates for the derivative of the fractional maximal function. Such estimates can only hold true in bounded domains or for $\alpha = 1$, and in bounded domains they are weaker than the expected $L^p \to \dot{W}^{1,\frac{np}{p-(\alpha-1)}}$ estimates, only known for $p > n/(n-1)$ by [HKKT15]. We do not pursue this possible improvement direction here, although we believe it to be an interesting open problem.

Part II

Pointwise restriction of the Fourier transform

Chapter 4

Maximal restriction estimates and the maximal function of the Fourier transform

> Se todo mundo diz que é assim,
> melhor eu inventar um mundo
> novo.
>
> *–M.C.A.*

This chapter contains the paper [Ram18]. We prove inequalities concerning the restriction of the strong maximal function of the Fourier transform to the circle, providing an answer to a question left open by Müller, Ricci and Wright. We employ methods similar in spirit to the classical proofs of the two-dimensional restriction theorem, with the addition of a suitable trick to help us linearise our maximal function. In the end, we comment on how to use the same linearisation trick in combination with Vitturi's duality argument to obtain sharper high-dimensional results for the Hardy–Littlewood maximal function.

4.1 Introduction

Restriction estimates for the Fourier transform have been a very active topic within harmonic analysis for over the past 40 years. Basically, one inquires whether an inequality of the form

$$\|\widehat{f}|_S\|_{L^q(S,\mathrm{d}\sigma)} \leq C_{p,d}\|f\|_{L^p(\mathbb{R}^d)} \tag{4.1}$$

can hold on a hypersurface S, where σ stands for the standard surface measure on S, which is the same as the arclength measure for the case of plane curves. Here we shall focus on compact hypersurfaces S with nonvanishing curvature, the typical example being the sphere $\mathbb{S}^{d-1}$. By taking examples of functions (either the so called *Knapp* examples or constant functions; see, e.g., [Tao04, Section 4]), one finds out that a *necessary* condition for such inequalities to hold is that

$$1 \leq p < \frac{2d}{d+1} \text{ and } p' \geq \frac{d+1}{d-1}q, \tag{4.2}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. The *restriction conjecture* then asserts that the above conditions are also *sufficient*.

The first manifestation of such a restriction principle, in a range smaller than (4.2), was perhaps the result of Fefferman and Stein (see [Fef70, page 28]), where an estimate in all dimensions for $q = 2$ was proven, this estimate being sharpened to the optimal range of p for such q by Tomas [Tom75], who credits Stein for the endpoint result. For the sphere (and, in general, for compact hypersurfaces with nonvanishing curvature), it reads that

$$\|\widehat{f}|_S\|_{L^2(S,d\sigma)} \leq C_{p,d}\|f\|_{L^p(\mathbb{R}^d)},$$

whenever $1 \leq p \leq \frac{2(d+1)}{d+3}$.

Regarding ranges of exponents, for dimension $d \geq 3$, Problem (4.1) is still open, with new technology being developped continously to improve ranges of exponents; see, for instance, [Tao04, Gut14, Gut18, HR18, Wan18] for further information and more recent developments in this subject.

For dimension 2, however, Problem (4.1) has been completely solved, as we observe that the conditions can be rewritten as follows:

$$1 \leq p < \frac{4}{3},\, p' \geq 3q. \tag{4.3}$$

In the non-endpoint case $p' > 3q$, the result is due to Fefferman [Fef70, page 33], and the endpoint to Zygmund [Zyg74] and Carleson and Sjölin [CS72]. Later, Sjölin [Sjö74] also extended these results to other classes of curves.

In [MRW19], D. Müller, F. Ricci and J. Wright consider a slight strenghtening of the restriction properties of the Fourier transform in two dimensions: namely, they prove a maximal version of restriction estimates and conclude a differentiation result. Here, we shall state the result only in the case of $\mathbb{S}^1$, for simplicity:

Theorem 4.1. *[Müller, Ricci, Wright [MRW19]; 2016] Let $\mathbb{S}^1$ be the unit circle in $\mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{C}$ be a L^p function. Assume that $1 \leq p < \frac{8}{7}$. Then, with respect to arclength measure, almost every point in $\mathbb{S}^1$ is a Lebesgue point for $\widehat{f}$ and the regularised value of $\widehat{f}$ at x coincides with the restriction operator $\mathcal{R}f(x)$ for almost every $x \in \mathbb{S}^1$.*

The purpose of this note is to improve ranges of exponents of such maximal restriction results. Explicitly, our main result is:

Theorem 4.2. *Theorem 4.1 extends to $1 \leq p < \frac{4}{3}$.*

We remark that Theorems 4.1 and 4.2 hold also for any C^2 compact convex curve with nonvanishing curvature in place of $\mathbb{S}^1$. In order to keep the presentation clean, however, we have opted for presenting it only in the circle case.

The strategy in [MRW19] passes through a maximal function with absolute values *outside* the integral, and then uses Hölder inequality. Namely, they focus on maximal functions of the form

$$\mathcal{M}f(x) = \sup_{\substack{R \text{ axis parallel,} \\ \text{centered at} x}} \left| \int \chi_R(y)\widehat{f}(y)\,\mathrm{d}y \right|,$$

where $\chi_R \in \mathcal{S}(\mathbb{R})$ is a smooth bump function adapted to the rectangle R. They then prove that, for the *whole* restriction range $1 \le p < \frac{4}{3}$ and $p' \ge 3q$,

$$\|\mathcal{M}f\|_{L^q(\mathrm{d}\sigma)} \le C_{p,\Gamma}\|f\|_{L^p(\mathbb{R}^2)},$$

where σ stands again for the arclength measure on the curve Γ. Finally, in order to prove Theorem 4.1, the authors bound the maximal function

$$M_{\mathcal{R}}f(t) = \sup_{\substack{R \text{ axis parallel,} \\ \text{centered at} x}} \int \chi_R(y)|\widehat{f}(y)|\,\mathrm{d}y \tag{4.4}$$

by $(\mathcal{M}h(x))^{1/2}$, where $h = f * \tilde{f}$, with $\tilde{f}(x,y) = f(-x,-y)$.

It is crucial to notice, however, that this approach *cannot* imply the full range of restriction bounds: one needs $p < \frac{8}{7}$ in order to be able to use that $\|f * \tilde{f}\|_{\tilde{p}} \le \|f\|_p^2$, for $\tilde{p} < \frac{4}{3}$. We bypass this problem by introducing an additional linearisation of $M_{\mathcal{R}}f$.

Namely, for fixed g with $\|g\|_\infty = 1$ and measurable choice R of axis-parallel rectangles, define the linearised maximal operator

$$M_{g,R}f(x) = \int_{\mathbb{R}^2} |R(x)|^{-1}1_{R(x)}(y-x)\widehat{f}(y)g(y)\,\mathrm{d}y \tag{4.5}$$

acting initially, say, on functions in $L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$. The key point here is to view $M_{g,R}$ as an operator acting *only* on f, with $g \in L^\infty$ being fixed. Setting $g(y) = \frac{\overline{\widehat{f}}}{|\widehat{f}|}$ where $\widehat{f} \ne 0$, and zero otherwise implies, together with a suitable choice of a measurable R, that $M_{g,R}f(t) \ge \frac{1}{2}M_{\mathcal{R}}f$. It is therefore sufficient to estimate (4.5) from $L^r(\mathbb{R}^2)$ to some $L^q(\mathbb{S}^1)$, where $1 \le r < \frac{4}{3}$, with bounds independent of $g \in L^\infty, \|g\|_\infty \le 1$, as setting the aforementioned g and R allows us to conclude Theorem 4.2. Bounding $M_{g,R}$ is the basic goal of Lemmata 4.4 and 4.5.

Following [MRW19], M. Vitturi [Vit17] and V. Kovač and D. Oliveira e Silva [KOeS18] have proved, as a consequence of $p' = 4$ being admissible for the restriction estimate, results in dimensions ≥ 3 : they have obtained that, in the same range of exponents as in Theorem 4.6, one gets *pointwise convergence* $\chi_\varepsilon * \widehat{f} \to \widehat{f}$ for $\sigma-$almost every point on the sphere $\mathbb{S}^{d-1}$, where $\chi_\varepsilon(y) = \frac{1}{\varepsilon^n}\chi(y/\varepsilon)$, and $\chi \in \mathcal{S}(\mathbb{R}^d)$. Although this is already present in [Vit17] and in both cases the techniques also imply the same result for $\chi = 1_{B(0,1)}$, the ideas in [KOeS18] represent a stronger, quantitative form of such a theorem, as they consider *variation norms* instead of suprema.

Our second result is also an improvement on Vitturi's techniques, yet in another direction:

Theorem 4.3. *Let $d \ge 3$ and $S \subset \mathbb{R}^d$ be a compact hypersurface with nonvanishing curvature. If $f \in L^p(\mathbb{R}^d)$, $1 \le p \le \frac{4}{3}$, then σ_S-almost every point of S is a Lebesgue point of $\widehat{f}$, and the regularised value of $\widehat{f}$ at x coincides with the restriction operator $\mathcal{R}f(x)$ for σ_S-almost every $x \in S$, where σ_S stands for the surface measure on S.*

The argument to prove Theorem 4.3 is similar to the one employed to treat Theorem 4.2, and we postpone it to the end of this manuscript.

Finally, a few weeks after this article's first version, Vjekoslav Kovač [Kov19] proved a general abstract maximal restriction principle. Among other results, his main Theorem implies that, whenever inequality (4.1) holds with $p < q$, then

$$\| \sup_{t>0} |\widehat{f} * \chi_{B_t}(x)| \|_{L^q(S,\mathrm{d}\sigma)} \leq C_{p,q,d} \|f\|_{L^p(\mathbb{R}^d)}, \tag{4.6}$$

where B_t denotes the ball of center at the origin and radius t. Although his results hold in a larger range of exponents in higher dimensions than the one in Theorem 4.3, there seems to be no direct way to prove our main theorems as corollaries of his techniques. This is due to the fact that his proofs depend heavily on properties of the Fourier transform of $\widehat{f} * \chi_{B_t}$. In addition, Theorem 4.2 deals with *strong maximal functions* in contrast to the usual Hardy–Littlewood maximal function, which imposes further technical complications.

On the other hand, Kovač's methods explore directly oscillatory properties of Fourier transforms of measures. As a consequence, he obtains results concerning not only maximal functions, but also *maximal variations* of averages of the Fourier transform. Our techniques do not seem to achieve any results in the variational case.

It is of interest, however, whether one can establish (4.6) with $\sup_{t>0} |\widehat{f}| * \chi_{B_t}$ on the left hand side. This seems to be a more challenging problem, as a direct combination of the techniques in [Kov19] with the ones in the present paper does not yield any result.

4.2 Main Argument

Call a measurable function a in $\mathbb{R}^d$ *bump function adapted to an axis parallel rectangle* R centered at the origin if

$$|a| \leq |R|^{-1} 1_R.$$

Convolution with such a bump function satisfies a pointwise bound by the strong Hardy Littlewood maximal function, uniformly in the rectangle. The following lemma concerns an adjoint of a linearised maximal operator, combined with a Fourier transform.

Lemma 4.4. *For each $x \in \mathbb{R}^d$ let a_x be a convolution product of k bump functions, $k \geq 1$, adapted to (possibly different) axis parallel rectangles. Assume further that $a_x(y)$, as function in (x, y), is in $L^\infty(x, L^1(y))$. Let T be defined on functions $f \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ by*

$$Tf(\xi) = \int_{\mathbb{R}^d} \widehat{a}_x(\xi) e^{2\pi i x\cdot\xi} f(x)\,\mathrm{d}x.$$

Then, for some universal constant C depending on k and d only,

$$\|Tf\|_2 \leq C\|f\|_2.$$

Proof. We set up a duality argument, testing Tf against an arbitrary function $\widehat{\varphi} \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. We have, by Fubini and Plancherel,

$$\int_{\mathbb{R}^d} \overline{\widehat{\varphi}(\xi)} \int_{\mathbb{R}^d} \widehat{a}_x(\xi) e^{2\pi i x\cdot\xi} f(x)\,\mathrm{d}x\mathrm{d}\xi = \int_{\mathbb{R}^d} f(x) \int_{\mathbb{R}^d} \overline{\varphi(y)} a_x(y-x)\,\mathrm{d}y\mathrm{d}x$$

Identifying on the right-hand-side a $k-$fold convolution of bump functions acting on φ, we estimate the last display by

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x)|M^k(\varphi)(x)\,\mathrm{d}x \le \|f\|_2\|M^k\varphi\|_2 \le C\|f\|_2\|\widehat{\varphi}\|_2 \ ,$$

where we have used the strong maximal theorem and Plancherel again. Since $\widehat{\varphi}$ was arbitrary, this proves Lemma 4.4.

□

The hypotheses in the next Lemma are motivated by the parameterised circle

$$z(t) = (\cos(2\pi t), \sin(2\pi t)).$$

By the addition theorem for the sine function, we have

$$|\det(z'(t), z'(s))| = 4\pi^2|\sin(2\pi(s-t))|.$$

Note the vanishing of the determinant when the two tangent vectors become parallel or anti-parallel. Note further that one can recover $t \neq s \in I := [0,1)$ from

$$x := z(t) + z(s).$$

Namely, $x/2$ is the midpoint between $z(t)$ and $z(s)$, and these two points on the circle are mirror symmetric relative to the line through this midpoint and the origin. This determines the two points $t \neq s$, up to permutation. Define, therefore, the *upper triangle*

$$\Delta = \{(t,s) \in I \times I : t > s\}.$$

Lemma 4.5. *Let $z : \mathbb{R} \to \mathbb{R}^2$ be a smooth one-periodic curve such that for all $(t,s) \in \Delta$*

$$|\det(z'(t), z'(s))| \ge c|\sin(2\pi(t-s))| \tag{4.7}$$

and such that the map

$$(t,s) \to z(t) + z(s) \tag{4.8}$$

is a bijection from Δ onto a bounded set $\Omega \subset \mathbb{R}^2$. With $a_{z(t)}$ a bump function for every $t \in I$ such that $a_{z(t)}(x)$ is in $L^\infty(t, L^1(x))$, consider an operator acting on functions in $L^4(I)$ as follows:

$$Tf(\xi) = \int_I \widehat{a}_{z(t)}(\xi)e^{2\pi\xi\cdot z(t)}f(t)\,\mathrm{d}t.$$

Then we have for all $1 \le p < 2$ with some constant depending only on p:

$$\|Tf\|_{2p'} \le C_p\|f\|_{\frac{2p}{3-p}},$$

with the obvious interpretation when $p' = \infty$. Notice, moreover, that the reciprocals $\left(\frac{1}{2p'}, \frac{3-p}{2p}\right)$ of the aforementioned exponents lie on the line segment joining $(1/4, 1/4)$ and $(0,1)$.

Proof. To reduce to Lemma 4.4, we need to pass to a two dimensional integral. We follow the idea of Carleson-Sjölin and consider the square

$$Tf(\xi)^2 = \int_{I\times I} \widehat{a}_{z(t)}(\xi)\widehat{a}_{z(s)}(\xi)e^{2\pi\xi\cdot(z(t)+z(s))}f(t)f(s)\,\mathrm{d}t\mathrm{d}s.$$

The integral is twice the analoguous integral over the triangle Δ, where we change coordinates by the bijective map (4.8) to obtain

$$Tf(\xi)^2 = 2\int_{\Omega} \widehat{b}_x(\xi)e^{2\pi i\xi\cdot x}h(x)\,\mathrm{d}x.$$

Here we have unambiguously defined, for (t,s) in the triangle,

$$\widehat{b}_{z(t)+z(s)} := \widehat{a}_{z(t)}\widehat{a}_{z(s)},$$

$$h(z(t)+z(s)) := f(t)f(s)|\det(z'(t),z'(s))|^{-1}.$$

Note that the determinant here is the Jacobian determinant of the map (4.8).

It is now easy to prove, by interpolation, that for $1 \le p \le 2$ we have

$$\|Tf\|_{2p'}^{2p} = \|(Tf)^2\|_{p'}^{p} \le C\|h\|_p^p\,.$$

Namely, $p = 2$ follows directly from Lemma 4.4 applied to a function supported on Ω, and $p = 1$ is trivial since $\|\widehat{b}_x\|_\infty \le C$. To conclude the proof of the lemma, we invert the change of variables to estimate the right-hand-side for $1 \le p < 2$:

$$\int_\Omega |h(x)|^p\,\mathrm{d}x = \int_\Delta |f(t)f(s)|^p|\det(z'(t),z'(s))|^{1-p}\,\mathrm{d}t\mathrm{d}s \le C_p\||f|^p\|^2_{\frac{2}{3-p}} = C_p\|f\|^{2p}_{\frac{2p}{3-p}}\,.$$

Here, the last inequality follows from the Hardy–Littlewood–Sobolev inequality for fractional integrals. Namely, we estimate with (4.7) on the triangle:

$$|\det(z'(t),z'(s))|^{1-p} \le C\sum_{k=-2}^{2}|t-s-k|^{1-p},$$

and we note that each summand leads to a translated fractional integral. □

Proof of Theorem 4.2. As mentioned in the introduction, fix $g \in L^\infty$. We introduce the bump function

$$a_x(y) := |R(x)|^{-1}1_{R(x)}(y)\overline{g(x-y)},$$

and abuse the notation for the operator in (4.5) as

$$M_{g,R}f(t) = \int_{\mathbb{R}^2} \overline{a_{z(t)}}(y-z(t))\widehat{f}(y)\,\mathrm{d}y.$$

This is just a composition of the operator in (4.5) with a parametrisation, so we identify them. By Plancherel, similarly to the proof of Lemma 4.4, we have

$$M_{g,R}f(t) = \int_{\mathbb{R}^2} \overline{\widehat{a}_{z(t)}}(\xi)e^{-2\pi i\xi\cdot z(t)}f(\xi)\,\mathrm{d}y.$$

The adjoint operator then becomes

$$M^*_{g,R}(\psi)(\xi) = \int_I \widehat{a}_{z(t)}(\xi) e^{2\pi i \xi \cdot z(t)} \psi(t)\, \mathrm{d}t.$$

By Lemma 4.5, this is bounded from $L^{\frac{2p}{3-p}}$ to $L^{2p'}$ for $p < 2$. We set now $r = (2p')'$. By a computation, $\frac{2p}{3-p} = (r'/3)'$. With this notation, we have that $M_{g,R}$ is bounded from $L^r(\mathbb{R}^2)$ to $L^{r'/3}(\mathbb{S}^1)$ for all $r < \frac{4}{3}$. By now taking $g(z) = \frac{\overline{\widehat{f}(z)}}{|\widehat{f}(z)|}$ and a suitable measurable choice of rectangles R, we retrieve Theorem 4.2. Notice, moreover, that this implies $L^r(\mathbb{R}^2) \to L^q(\mathbb{S}^1)$ estimates in the optimal two-dimensional restriction range $1 \leq r < \frac{4}{3}, r' \geq 3q$. $\square$

4.3 The high-dimensional result

Just like we employed our techniques to deal with the two-dimensional case, we adapt the arguments by M. Vitturi [Vit17] to achieve high-dimensional estimates. We briefly sketch on how to do it.

Theorem 4.6. *Let*

$$\mathfrak{M}f(x) = \sup_{0<\varepsilon\leq 1} \fint_{B(0,\varepsilon)} |\widehat{f}(x+y)| \mathrm{d}y.$$

Under the hypotheses of Theorem 4.3, it holds that

$$\|\mathfrak{M}f\|_{L^q(S, \mathrm{d}\sigma_S)} \leq C_{p,q,d} \|f\|_{L^p(\mathbb{R}^d)},$$

where $1 \leq p \leq \frac{4}{3}$ *and* $p' \geq \frac{d+1}{d-1} q$.

Proof. Let $g \in L^\infty(\mathbb{R}^d)$ such that $\|g\|_\infty = 1$. We will take, in the end, $g(z) = \frac{\overline{\widehat{f}(z)}}{|\widehat{f}(z)|}$, just like in the proof of Theorem 4.2. We write the auxiliary bilinear operator

$$\mathfrak{M}(f;g)(x) = \sup_{0\leq\varepsilon\leq 1} \left| \fint_{B(0,\varepsilon)} \widehat{f}(x+y) g(x+y)\, \mathrm{d}y \right|.$$

We subsequently define $\mathfrak{A}_{\varepsilon(\cdot),g} f(x) = \fint_{B(0,\varepsilon(x))} \widehat{f}(x+y) g(x+y)\, \mathrm{d}y$ to be the linearised operator for suitable measurable ε. Its adjoint has the form

$$\mathfrak{A}^*_{\varepsilon(\cdot),g} h(\xi) = \int_S G(x,\xi) e^{2\pi i \xi \cdot x} h(x)\, \mathrm{d}\sigma_S(x),$$

where $G(x,\xi) = \mathcal{F}(g(x+\cdot)\chi_{B(0,\varepsilon(x))})(\xi)$. Following Vitturi's arguments and the ones in the proof of Theorem 4.2, it is enough to prove the following estimate:

$$\|\mathfrak{A}^*_{\varepsilon(\cdot),g} h\|_{L^4(\mathbb{R}^d)} \leq C_{q,d} \|h\|_{L^{q_d'}(S, \mathrm{d}\sigma_S)},$$

where $q_d = 4\frac{d-1}{d+1}$. Now we write the L^4 norm as a (square root of a) L^2 norm of the convolution of the Fourier transform $(\mathfrak{A}^*_{\varepsilon(\cdot),g} h)\widehat{\ }$ with itself. With this in mind, one gets from a calculation that

$$(\mathfrak{A}^*_{\varepsilon(\cdot),g} h)\widehat{\ }(\eta) = g(\eta) \int_S h(x) \chi_{B(0,\varepsilon(x))}(\eta - x)\, \mathrm{d}\sigma_S(x) =: g(\eta) T_{\varepsilon(\cdot)} h(\eta).$$

We are then able to bound

$$|(\mathfrak{A}^*_{\varepsilon(\cdot),g}h)\widehat{\ } * (\mathfrak{A}^*_{\varepsilon(\cdot),g}h)\widehat{\ }(\rho)| \le |(T_{\varepsilon(\cdot)}|h|) * (T_{\varepsilon(\cdot)}|h|)(\rho)|.$$

The operator on the right hand is estimated directly by Vitturi's proof, and therefore we conclude the desired bounds from the ones in [Vit17].

□

4.4 Acknowledgements

The author is deeply indebted to his doctoral advisor, Prof. Dr. Christoph Thiele, for the time devoted to thorough inspection of the arguments, careful understanding of ideas and several suggestions that led to the final version of this manuscript. He would also like to acknowledge the interest in this project manifested by some of his colleagues, among which he would like to thank especially Marco Fraccaroli and Gennady Uraltsev for discussions about related work, as well as Diogo Oliveira e Silva and Vjekoslav Kovač, which, besides sharing their recent project [KOeS18], also took the time carefully read and make valuable suggestions on an early version of this work. Finally, the author thanks the referee for numerous remarks and corrections.

Chapter 5

Low-dimensional maximal restriction principles for the Fourier transform

No creo en una tercera
alternativa: creo en muchas.

–G.G.M.

This chapter contains the paper [Ram19b]. Following the ideas from the previous chapter, we prove abstract maximal results for the Fourier transform. Our results deal mainly with maximal operators of convolution-type and $r-$average maximal functions. As a by-product of our techniques we obtain spherical maximal restriction estimates, as well as restriction estimates for $2-$average maximal functions, answering thus points left open by V. Kovač and Müller, Ricci and Wright.

5.1 Introduction

The classical *restriction problem* for the Fourier transforms asks for the largest possible range of exponents $1 \le p, q \le +\infty$ so that an inequality of the form

$$\|\widehat{f}|_S\|_{L^q(S)} \le C_{p,q,d,S}\|f\|_{L^p(\mathbb{R}^d)} \tag{5.1}$$

holds for any function $f \in \mathcal{S}(\mathbb{R}^d)$. Here, S is taken to be a subset of $\mathbb{R}^d$, endowed with a suitable measure.

The existence of such *a priori* inequalities allows one to define restrictions of Fourier transforms of L^p functions to smaller sets in the L^q-sense. Recently, effort has been put into extending this definition to a *pointwise* sense: one has to look instead at

$$\|\mathcal{M}(\widehat{f})\|_{L^q(S)} \le C_{p,q,d,S}\|f\|_{L^p(\mathbb{R}^d)},$$

where $\mathcal{M}$ is a suitable maximal operator. In [MRW19], the authors prove, for the first time, such a statement about restriction to curves. Their techniques adapt the ones in [CS72] to the maximal context. The works of Vitturi [Vit17], Kovač and Oliveira e Silva [KOeS18] and Ramos [Ram18] have subsequently dealt with this problem, extending the maximal restriction property to higher dimensions, considering variational versions of it and sharpening the results in [MRW19].

More recently, Kovač [Kov19] proved a general, abstract principle for such pointwise statements to hold. One of his results is that, whenever restriction estimates like (5.1) hold with $p < q$, and whenever $\mu : \mathcal{B}(\mathbb{R}^d) \to \mathbb{C}$ is a complex measure such that $|\nabla\widehat{\mu}(\xi)| \le D(1+|\xi|)^{-1-\eta}$, for some $\eta > 0$, then

$$\|\sup_{t>0} |\widehat{f} * \mu_t(x)|\|_{L^q(S)} \le C_{p,q,S,\eta} \cdot D\|f\|_{L^p(\mathbb{R}^d)}. \tag{5.2}$$

Here, $\mu_t(E) := \mu(t^{-1}E)$. Note that $\mathrm{d}\mu = \chi_{B(0,1)}(x)\mathrm{d}x$ satisfies the Fourier decay condition above in *any* dimension, which generalizes the results of Vitturi [Vit17], Müller, Ricci, Wright [MRW19] and Kovač and Oliveira e Silva [KOeS18].

The purpose of this note is to employ the techniques in [Ram18] to extend inequality (5.2) in low-dimensional cases not covered by Kovač's techniques. Additionally, we simplify the techniques in [Ram18] in order to extend a result from [MRW19].

5.1.1 Two-dimensional results

In (5.2), the main requirement on the measure μ that $|\nabla\widehat{\mu}(\xi)| \lesssim_{\eta,\mu} (1+|\xi|)^{-1-\eta}$, for some $\eta > 0$, is only satisfied by the *spherical measure* $\mathrm{d}\mu = \mathrm{d}\sigma_{\mathbb{S}^{d-1}}$ if $d \ge 4$. Therefore, Kovač's result does not yield bounds for lower-dimensional restrictions of spherical maximal functions of the Fourier transform. This was our motivation for the first result of this paper.

Theorem 5.1. *Let μ be a positive, finite Borel measure defined in $\mathbb{R}^2$, and suppose that the maximal function*

$$M_\mu g(x) := \sup_{t>0} |g| * \mu_t(x).$$

is bounded from $L^r(\mathbb{R}^2) \to L^r(\mathbb{R}^2)$, whenever $r > 2$. Then the following bound holds:

$$\|M_\mu(\widehat{f})\|_{L^q(\mathbb{S}^1)} \le C_{p,\mu}\|f\|_{L^p(\mathbb{R}^2)},$$

where $1 \le p < \frac{4}{3}, p' \ge 3q$.

In Proposition 5.9 at the end of this note, we prove that Kovač's [Kov19] assumptions on the measure *imply* ours. The spherical maximal function in dimensions 2, 3 is an example that shows, as elaborated in Section 5.4.1, that Theorems 5.1 and 5.4 are strictly stronger.

On the other hand, in [MRW19], the authors, in the end of their manuscript, make use of the maximal function

$$M_2(h) := M(|h|^2)^{1/2},$$

where $Mf(x) = \sup_{r>0} \fint_{B(x,r)} |f|$ denotes the usual Hardy–Littlewood maximal function, to prove results about Lebesgue points of the Fourier transform on curves in the range $1 \le p < \frac{8}{7}$. In [Ram18], this author circumvents this problem by considering a suitable linearization instead of working with M_2. Our next result combines the two approaches:

Theorem 5.2. *Let $1 \le r \le 2$. Define the maximal functions $M_r h(x) := (M(|h|^r)(x))^{1/r}$. The following bound holds:*

$$\|M_r(\widehat{f})\|_{L^q(\mathbb{S}^1)} \le C_{p,r}\|f\|_{L^p(\mathbb{R}^2)},$$

where $1 \le p < \frac{4}{3}, p' \ge 3q$.

The main feature in the proofs of these Theorems is the linearization method employed in [Ram18] together with Lemmata 5.6 and 5.7. These, on the other hand, provide a way to bypass the interpolation scheme employed in [Ram18, Lemmata 1 and 2]. Also, in the case where one takes $\mathrm{d}\mu = \mathrm{d}\sigma_{\mathbb{S}^1}$ to be the arc-length measure in the circle the interpolation idea fails due to the lack of $L^2(\mathbb{R}^2)$ bounds for maximal functions, whereas working directly with the aid of the Hausdorff–Young inequality gives us the result, as long as the measure we consider satisfies the above conditions. By the celebrated result of Bourgain [Bou86], this is *exactly* the case for the circular maximal function in dimension 2.

In Section 5.4.3, we present two different kinds of counterexamples, in order to impose restrictions on r so that Theorem 5.2 can hold. Both the examples yield the same $r \leq 4$ bound, whereas Theorem 5.2 only works in the $r \leq 2$ case. One is led to pose the following question:

Question 5.3. *Can the two-dimensional full range of maximal restriction inequalities hold for $M_s, 2 < s \leq 4$?*

5.1.2 Three-dimensional results

In dimension 3, our main theorems deal with the Tomas-Stein exponent case, in both the context of measures as well as in the context of M_r-maximal functions:

Theorem 5.4. *Let Let μ be a positive, finite Borel measure defined in $\mathbb{R}^3$, and suppose that the maximal function*

$$M_\mu g(x) := \sup_{t>0} |g| * \mu_t(x).$$

is bounded from $L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$. Then the following bound holds:

$$\|M_\mu(\widehat{f})\|_{L^2(\mathbb{S}^2)} \leq C_{p,\mu} \|f\|_{L^p(\mathbb{R}^3)},$$

where $1 \leq p \leq \frac{4}{3}$.

Theorem 5.5. *Let $1 \leq r < 2$. Then the following bound holds:*

$$\|M_r(\widehat{f})\|_{L^2(\mathbb{S}^2)} \leq C_{p,r} \|f\|_{L^p(\mathbb{R}^3)},$$

where $1 \leq p \leq \frac{4}{3}$. Aditionally, the quadratic maximal function M_2 satisfies that

$$\|M_2(\widehat{f})\|_{L^2(\mathbb{S}^2)} \leq C_{p,r} \|f\|_{L^p(\mathbb{R}^3)},$$

whenever $1 \leq p < \frac{4}{3}$.

We prove these results in Section 5.3 by merging the ideas in Theorems 5.1 and 5.2 with Vitturi's method. As a by-product, the counterexamples built in Section 5.1.1 provide us with the restriction that $s \leq 2$ in order for Theorem 5.5 to hold. In particular, a further use of one of these counterexamples in higher dimensions gives us as a direct corollary that the only dimensions in which a full-range restriction result for the *strong* maximal function

$$M_{\mathcal{S}} f(x) := \sup_{\substack{R \text{ axis parallel,} \\ \text{centered at } x}} \fint_R |f|$$

of the Fourier transform could hold are $d = 2, 3$. We talk about this property in more detail in Proposition 5.12.

5.1.3 Notation

In what follows, we denote $A \lesssim B$ to mean that $A \le C \cdot B$, for some universal constant $C > 0$. If we let C depend on a parameter α, we write $A \lesssim_\alpha B$. We suppress this notation in case the specific dependence on α is not important. We also normalize the Fourier transform as $\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi} f(x)\,\mathrm{d}x$. Finally, we often write $⨍_B g := \frac{1}{|B|}\int_B g$ for the average of g over a set B.

Acknowledgements. The author would like to thank Felipe Gonçalves for helpful discussions which led to the proof of Theorem 5.1 and to a great deal of inspiration for the other results in this manuscript. He is also indebted to his supervisor Prof. Dr. Christoph Thiele for reading this manuscript and giving advice on how to improve the presentation. The author also acknowledges finantial support by the Deutscher Akademischer Austauschdient (DAAD).

5.2 Proof of Theorems 5.1 and 5.2

5.2.1 Proof of Theorem 5.1

The basic outline of the proof is essentially the same as in the proof of [Ram18, Theorem 1]. After using the Kolmogorov–Saliverstov linearization method and letting $g(z) = \frac{\overline{\widehat{f}(z)}}{|\widehat{f}(z)|}$, it suffices to prove bounds for

$$M_{\mu,g,t(\cdot)}f(x) = \int_{\mathbb{R}^2} \widehat{f}(x-y)g(x-y)\mathrm{d}\mu_{t(x)}(y).$$

Here, we actually regard $M_{\mu,g,t(\cdot)}$ as an operator with a *fixed* g, prove bounds for it and then substitute the chosen g above. An application of Plancherel's Theorem implies that

$$M_{\mu,g,t(\cdot)}f(x) = \int_{\mathbb{R}^2} f(\xi)e^{2\pi i x\cdot\xi}\widehat{A_x}(\xi)\mathrm{d}\xi,$$

where $\mathrm{d}A_x(y) := g(x-y)\,\mathrm{d}\mu_{t(x)}(y)$. A dualization argument then implies that Theorem 5.1 is equivalent to proving

$$M^*_{\mu,g,t(\cdot)}h(\xi) = \int_{\mathbb{S}^1} h(\omega)e^{-2\pi i\xi\cdot\omega}\widehat{A_\omega}(\xi)\mathrm{d}\sigma_{\mathbb{S}^1}(\omega)$$

to be bounded from $L^{q'}(\mathbb{S}^1) \to L^{p'}(\mathbb{R}^2)$. Just like in the proof of [Ram18, Lemma 2], we write $\|M^*_{\mu,g,t(\cdot)}h\|_{p'} = \|(M^*_{\mu,g,t(\cdot)}h)^2\|_{p'/2}^{1/2}$. Expanding the square gives

$$(M^*_{\mu,g,t(\cdot)}h)^2(\xi) = \int_{\mathbb{S}^1\times\mathbb{S}^1} h(\omega)\,h(\omega')\,e^{-2\pi i(\omega+\omega')\cdot\xi}\widehat{A_\omega}(\xi)\widehat{A_{\omega'}}(\xi)\,\mathrm{d}\sigma_{\mathbb{S}^1}(\omega)\,\mathrm{d}\sigma_{\mathbb{S}^1}(\omega').$$

We perform two changes of variable: first, we parametrise the circle by $z(r) = (\cos(2\pi r), \sin(2\pi r))$. After that, we take a pair of points $(t,s), t > s$, into the point $x := z(t) + z(s)$. This map is easily seen to be a bijection from

$$\Delta := \{(t,s) \in [0,1)^2, t > s\} \text{ to } B_2(0) \subset \mathbb{R}^2.$$

After a calculation, we rewrite our operator as

$$(M^*_{\mu,g,t(\cdot)}h)^2(\xi) = 2\int_{B_2(0)} H(x)e^{-2\pi i x\cdot\xi}\widehat{B_x}(\xi)\,\mathrm{d}x,$$

where

$$\widehat{B_x}(\xi) := \widehat{A_{z(t)}}(\xi)\widehat{A_{z(s)}}(\xi), \tag{5.3}$$
$$H(x) := \frac{h(z(s))h(z(t))}{|\det(z'(s), z'(t))|} = \frac{h(z(s))h(z(t))}{4\pi^2|\sin(2\pi(s-t))|}.$$

Notice that the factor 2 multiplying the integral comes from considering twice the contribution from the upper triangle. The representation for our squared operator leads us to our main Lemma, which is a generalization of [Ram18, Lemma 2]:

Lemma 5.6. *Let, for every $x \in \mathbb{R}^2$, $B_x = \mu_{t_1(x)} * \cdots \mu_{t_k(x)}$ be the convolution product of dilates $\mu_{t_1(x)}, ..., \mu_{t_k(x)}$ of a finite Borel measure such that*

$$\|M_\mu\|_{r\to r} < +\infty,\ \forall r > 2. \tag{5.4}$$

Assume, in addition, that the map $x \mapsto B_x$ is in $L^\infty(M(\mathbb{R}^2))$, where $M(\mathbb{R}^2)$ denotes the space of finite Borel measures on $\mathbb{R}^2$. If

$$Tf(\xi) = \int_{\mathbb{R}^2} \widehat{B_x}(\xi)e^{-2\pi i x\cdot\xi}f(x)\,\mathrm{d}x,$$

then T maps $L^p(\mathbb{R}^2)$ to $L^{p'}(\mathbb{R}^2)$ boundedly for $1 \le p < 2$.

Proof. We write, for an arbitrary function $g \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$,

$$\langle Tf, g\rangle = \int_{\mathbb{R}^2} \overline{g}(\xi) \int_{\mathbb{R}^2} \widehat{B_x}(\xi)e^{-2\pi i x\cdot\xi}f(x)\,\mathrm{d}x\,\mathrm{d}\xi.$$

By Fubini and Plancherel, this equals, in turn,

$$\int_{\mathbb{R}^2} f(x)\widehat{\overline{g}} * B_x(x)\mathrm{d}x.$$

By the definition of B_x, property (5.4) and the Hausdorff-Young inequality, we bound the absolute value of the integral above by

$$\int_{\mathbb{R}^2} |f(x)||M^k_\mu(\widehat{\overline{g}})|(x)\,\mathrm{d}x \le \|f\|_p\||M^k_\mu(\widehat{\overline{g}})|\|_{p'} \le (C_\mu)^k\|f\|_p\|\widehat{\overline{g}}\|_{p'} \le (C_\mu)^k\|f\|_p\|g\|_p.$$

This proves the asserted bound for T. □

Notice that the function B_x in (5.3) satisfies the hypotheses of Lemma 5.6. Notice also that $p'/2 > 2$. After applying the Lemma above we are left with

$$\|(M^*_{\mu,g,t(\cdot)}h)^2\|^{1/2}_{p'/2} \lesssim \|H\|^{1/2}_{L^{(p'/2)'}(B_2(0))}.$$

To conclude the proof, we revert from H back to a product to estimate the right-hand-side for $1 \le (p'/2)' =: \eta < 2$:

$$\begin{aligned}\int_{B_2(0)} |H(x)|^\eta \,\mathrm{d}x &= \int_\Delta |h(z(t))h(z(s))|^\eta \cdot (4\pi^2|\sin(2\pi(s-t))|)^{1-\eta}\,\mathrm{d}t\mathrm{d}s \\ &\le C_p \||f|^\eta\|^2_{\frac{2}{3-\eta}} = C_p\|f\|^{2\eta}_{\frac{2\eta}{3-\eta}} = C_p\|f\|_{(p'/3)'}.\end{aligned} \tag{5.5}$$

Here, the last inequality follows from the Hardy–Littlewood–Sobolev inequality for fractional integrals. Indeed, we can bound

$$4\pi^2|\sin(2\pi(s-t))|^{1-\eta} \lesssim \sum_{j=-2}^{2} |t-s-j|^{1-\eta},$$

and then notice that each summand on the right hand side leads to a translated fractional integral. The result follows for the range $1 \le p < \frac{4}{3}, p' \ge 3q$ by interpolating this bound with the $L^1(\mathbb{R}^2) \to L^\infty(\mathbb{S}^1)$ bound, which follows in turn from the Riemann-Lebesgue Lemma and finiteness of the measure μ. $\square$

5.2.2 Proof of Theorem 5.2

In the same spirit as above, proving Theorem 5.2 is equivalent to proving bounds for

$$M_{r,g,t(\cdot)}f(x) := \int_{\mathbb{R}^2} \widehat{f}(x-y)g_x(x-y)\chi_{t(x)}(y)\mathrm{d}y,$$

where we will take, in the aftermath,

$$g_x(z) = \frac{\overline{\widehat{f}(z)}|\widehat{f}(z)|^{r-2}}{|B_{t(x)}(0)|^{1/r-1}\cdot \|\widehat{f}\|^{r-1}_{L^r(B_{t(x)}(x))}}.$$

With the above choice, the integral defining $M_{r,g,t(\cdot)}$ equals $\left(\int_{B_{t(x)}(0)} |\widehat{f}(x-y)|^r \mathrm{d}y\right)^{1/r}$. We denote a L^1–normalized dilation of characteristic function of the unit ball as $\chi_a(x) := (1/a^2)\cdot\chi(x/a)$. We then write the adjoint as

$$M^*_{r,g,t(\cdot)}h(\xi) = \int_{\mathbb{S}^1} h(\omega)e^{-2\pi i\omega\cdot\xi}\widehat{\mathcal{A}_\omega}(\xi)\mathrm{d}\sigma_{\mathbb{S}^1}(\omega),$$

with $\mathcal{A}_x(y) = g_x(x-y)\chi_{t(x)}(y)$. As before, we calculate $(M^*_{r,g,t(\cdot)})^2$ and change variables. It suffices to bound

$$(M^*_{r,g,t(\cdot)h})^2(\xi) = 2\int_{B_2(0)} H(x)e^{-2\pi i x\cdot\xi}\widehat{\mathcal{B}_x}(\xi)\,\mathrm{d}x =: T_r H(\xi), \tag{5.6}$$

where, again,

$$H(x) := \frac{h(z(s))h(z(t))}{|\det(z'(s),z'(t))|} = \frac{h(z(s))h(z(t))}{4\pi^2|\sin(2\pi(s-t))|},$$

and

$$\widehat{\mathcal{B}_x}(\xi) := \widehat{\mathcal{A}_{z(t)}}(\xi)\widehat{\mathcal{A}_{z(s)}}(\xi). \tag{5.7}$$

Of course, $z(s)+z(t) = x$. The next Lemma is the main tool for bounding (5.6), in order to employ the previous techniques:

Lemma 5.7. *Let $u \in \mathcal{S}(\mathbb{R}^2)$. Suppose that we are given a measurable function $\mathcal{A} : \mathbb{R}^2 \to L^{r'}(\mathbb{R}^2)$ so that $\sup_{x\in\mathbb{R}^2} |B_x|^{1/r} \|\mathcal{A}_x\|_{L^{r'}} < +\infty$ and support$(\mathcal{A}_x) \subset B_x$ for some ball B_x centered at the origin. If we define $\mathcal{B}_x$ as in equation (5.7), then it holds that*

$$u * \mathcal{B}_x(\theta) \le C \cdot M_r(M_r u)(\theta), \forall \theta \in \mathbb{R}^2,$$

where C is independent *of $x \in B_2(0)$.*

Proof. We denote first $\pi_1(x), \pi_2(x) \in \mathbb{S}^1$ the points such that $\pi_1(x)+\pi_2(x) = x$. The above convolution is

$$u * \mathcal{A}_{\pi_1(x)} * \mathcal{A}_{\pi_2(x)}(\theta).$$

It suffices to prove that $u * \mathcal{A}_{\pi_1(x)}(\eta) \le C \cdot M_r u(\eta)$, as the same argument holds for the convolution with $\mathcal{A}_{\pi_2(x)}$. We write

$$\begin{aligned} u * \mathcal{A}_{\pi_1(x)}(\eta) &= \int_{B_x} u(\eta - s)\mathcal{A}_{\pi_1(x)}(s)\mathrm{d}s \le (\sup_{z\in\mathbb{R}^2} \|\mathcal{A}_z\|_{L^{r'}})\|u(\eta - \cdot)\|_{L^r(B_x)} \\ &\lesssim |B_x|^{-1/r}\|u(\eta - \cdot)\|_{L^r(B_x)} \le M_r u(\eta), \end{aligned}$$

where we have used Hölder's inequality and the properties of $\mathcal{A}$. □

With Lemma 5.7, we are set to employ the techniques of the proof of Lemma 5.6. In fact, we let $G \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$, and take $\mathcal{B}_x$ as defined in equation (5.7) with $\mathcal{A}_x(y) = g_x(x - y)\chi_{t(x)}(y)$. By a direct computation – due to the dualization nature of our choice – to check that this $\mathcal{A}$ satisfies the hypotheses of Lemma 5.7. Therefore, we estimate the pairing:

$$\begin{aligned} \langle T_r H, G\rangle &= \int_{\mathbb{R}^2} \overline{G}(\xi) \left(\int_{B_2(0)} H(x)e^{-2\pi i x\cdot\xi}\widehat{\mathcal{B}_x}(\xi)\,\mathrm{d}x \right) \mathrm{d}\xi \\ &= \int_{B_2(0)} H(x) \cdot \widehat{\overline{G}} * \mathcal{B}_x(x)\,\mathrm{d}x \le \int_{B_2(0)} |H(x)| \cdot M_r(M_r\widehat{\overline{G}})(x)\mathrm{d}x \\ &\le \|H\|_{L^p(B_2(0))}\|M_r(M_r\widehat{\overline{G}})\|_{p'} \le (C_r)^2\|\widehat{\overline{G}}\|_{p'}\|H\|_{L^p(B_2(0))} \le \tilde{C}_{r,p}\|G\|_p\|H\|_{L^p(B_2(0))}. \end{aligned}$$

We have, similarly as before, used Fubini and Plancherel Theorems together with Lemma 5.7 in the second line, and Hölder's inequality in combination with boundedness of M_r in $L^{p'}$ (as $p' > 2 \ge r$) and the Hausdorff–Young inequality.

We conclude, by density, that $\|T_r H\|_{p'} \le \tilde{C}_{r,p}\|H\|_p$, $1 \le p < 2$. Now one resumes from the calculation in (5.5), and our previous considerations allow us to finish, once one notices that the $L^1(\mathbb{R}^2) \to L^\infty(\mathbb{S}^1)$ boundedness in this case is also a direct consequence of the Riemann-Lebesgue lemma. □

5.3 Proof of Theorems 5.4 and 5.5

5.3.1 Proof of Theorem 5.4

The strategy here is a modification of the scheme of proof in [Vit17]. There, one uses an integral representation for the convolution of Fourier transforms. Here, as we are working with measures and not functions, such a representation only becomes available to some measures through delta calculus. We bypass this difficulty by an argument similar to the one in the proofs of Theorems 5.1 and 5.2.

Explicitly, we start by linearizing our operator through

$$\mathcal{M}_{\mu,g,t(\cdot)}f(x) = \int_{\mathbb{R}^3} f(\xi)e^{2\pi i x\cdot\xi}\widehat{S_x}(\xi)\,\mathrm{d}\sigma(x),$$

where $\mathrm{d}S_x(y) = g(x-y)\,\mathrm{d}\mu_{t(x)}(y)$, $\|g\|_\infty \le 1$. Again, we will take $g(z) = \frac{\overline{\widehat{f}(z)}}{|\widehat{f}(z)|}$ afterwards. The desired inequality translates into proving that

$$\|\mathcal{M}_{\mu,g,t(\cdot)}f\|_{L^4(\mathbb{R}^3)} \lesssim \|f\|_{L^2(\mathbb{S}^2)}.$$

We write the L^4-norm above as $\|(\mathcal{M}_{\mu,g,t(\cdot)}f)^2\|_2^{1/2}$, and evaluate the L^2-norm by duality: for any $h \in L^2(\mathbb{R}^3)$, $\|h\|_2 \le 1$, we have

$$\begin{aligned}
&\langle(\mathcal{M}_{\mu,g,t(\cdot)}f)^2, h\rangle = \\
&= \int_{\mathbb{R}^3}\left(\int_{(\mathbb{S}^2)^2} f(x_1)f(x_2)e^{-2\pi i(x_1+x_2)\cdot\xi}\widehat{S_{x_1}}(\xi)\widehat{S_{x_2}}(\xi)\,\mathrm{d}\sigma(x_1)\,\mathrm{d}\sigma(x_2)\right)h(\xi)\,\mathrm{d}\xi \\
&= \int_{\mathbb{S}^2\times\mathbb{S}^2} f(x_1)f(x_2)\left(\int_{\mathbb{R}^3} h(\xi)e^{-2\pi i(x_1+x_2)\cdot\xi}\widehat{S_{x_1}}(\xi)\widehat{S_{x_2}}(\xi)\,\mathrm{d}\xi\right)\mathrm{d}\sigma(x_1)\mathrm{d}\sigma(x_2),
\end{aligned} \tag{5.8}$$

where we used Fubini's theorem to exchange integrals. Another application of Fubini's theorem in the innermost integral gives us that

$$\begin{aligned}
&\int_{\mathbb{R}^3\times\mathbb{R}^3} g(x_1-y_1)g(x_2-y_2)\widehat{h}((x_1+x_2)-(y_1+y_2))\,\mathrm{d}\mu_{t(x_1)}(y_1)\,\mathrm{d}\mu_{t(x_2)}(y_2) = \\
&= \int_{\mathbb{R}^3} h(\xi)e^{-2\pi i(x_1+x_2)\cdot\xi}\widehat{S_{x_1}}(\xi)\widehat{S_{x_2}}(\xi)\,\mathrm{d}\xi.
\end{aligned}$$

It is relatively simple to bound this integral: the integrand is pointwise bounded by

$$\int_{\mathbb{R}^3\times\mathbb{R}^3} |\widehat{h}((x_1+x_2)-(y_1+y_2))|\,\mathrm{d}\mu_{t(x_1)}(y_1)\,\mathrm{d}\mu_{t(x_2)}(y_2) \le M_\mu(M_\mu)(\widehat{h})(x_2+x_1),$$

where we used the definition of our maximal function associated to μ. Thus, the integral we wish to estimate is bounded by

$$\int_{\mathbb{S}^2\times\mathbb{S}^2} |f(x_1)||f(x_2)|M_\mu(M_\mu)(\widehat{h})(x_2+x_1)\,\mathrm{d}\sigma(x_1)\,\mathrm{d}\sigma(x_2).$$

By the Tomas-Stein theorem in dimension 3, as stated in [Vit17, Equation 2.3], the quantity above is at most a constant times

$$\|f\|_{L^2(\mathbb{S}^2)}^2\|(M_\mu)^2(\widehat{h})\|_{L^2(\mathbb{R}^3)} \le (C_\mu)^2\|f\|_{L^2(\mathbb{S}^2)}^2.$$

Along with the previous considerations, it is exactly what we wanted to prove. □

5.3.2 Proof of Theorem 5.5

The general idea here is similar to the proofs above, so we move somewhat faster through it. In fact, we consider the maximal operator M_2 first. Like before, we define the linearization of this operator as

$$M_{2,g,t(\cdot)}f(x) := \int_{\mathbb{R}^3} \widehat{f}(x-y)g_x(x-y)\chi_{t(x)}(y)\mathrm{d}y,$$

where, in the end, $\tilde{g}_x$ is to be taken as

$$\tilde{g}_x(z) = \frac{\overline{\widehat{f}(z)}}{|B_{t(x)}(0)|^{-1/2}\cdot \|\widehat{f}\|_{L^2(B_{t(x)}(x))}}.$$

Like in the cases before, we fix $\tilde{g}_x$ with certain properties and then substitute the above to get our results. The formal adjoint of this operator is given by

$$M^*_{2,g,t(\cdot)}h(\xi) = \int_{\mathbb{S}^2} h(\omega)e^{-2\pi i\omega\cdot\xi}\widehat{\mathcal{S}_\omega}(\xi)\mathrm{d}\sigma_{\mathbb{S}^2}(\omega),$$

with $\mathcal{S}_x(y) = \tilde{g}_x(x-y)\chi_{t(x)}(y)$. This leads us to estimate, as before, the inner product $\langle (M^*_{2,g,t(\cdot)}h)^2, F\rangle$. The calculation is entirely analogous to the one in (5.8), and we are led to estimate the function

$$\int_{\mathbb{R}^3} F(\xi)e^{-2\pi i(\omega_1+\omega_2)\cdot\xi}\widehat{\mathcal{S}_{\omega_1}}(\xi)\widehat{\mathcal{S}_{\omega_2}}(\xi)\,\mathrm{d}\xi.$$

An application of Fubini's theorem, along with the calculations from the proofs of Theorems 5.2 and 5.4 yield *pointwise* bounds for this integral by the iterated maximal function $M_2(M_2(\widehat{F}))(\omega_1+\omega_2)$. This summarizes as

$$|\langle (M^*_{2,g,t(\cdot)}h)^2, F\rangle| \le \int_{\mathbb{S}^2\times\mathbb{S}^2} |h(\omega_1)h(\omega_2)|M_2(M_2(\widehat{F}))(\omega_1+\omega_2)\,\mathrm{d}\sigma(\omega_1)\,\mathrm{d}\sigma(\omega_2). \tag{5.9}$$

In order to finish, we need to apply the following Lemma:

Lemma 5.8. *Let $2 \le p \le \infty$. There is a constant $C = C(p)$ such that, for all $v \in L^2(\mathbb{S}^2)$ and $W \in L^p(\mathbb{R}^3)$, it holds that*

$$\left|\int_{\mathbb{S}^2\times\mathbb{S}^2} v(\omega_1)v(\omega_2)W(\omega_1+\omega_2)\,\mathrm{d}\sigma(\omega_1)\,\mathrm{d}\sigma(\omega_2)\right| \le C\|v\|^2_{L^2(\mathbb{S}^2)}\|W\|_{L^p(\mathbb{R}^3)}.$$

Proof. We define the operator

$$T_{v_1}W(\omega_1) = \int_{\mathbb{S}^2} v_1(\omega_2)W(\omega_1+\omega_2)\,\mathrm{d}\sigma(\omega_2)$$

and note it satisfies the two following estimates:

- For $p=\infty$, the estimate $\|T_{v_1}W\|_{L^2(\mathbb{S}^2)} \lesssim \|v_1\|_{L^2(\mathbb{S}^2)}\|W\|_\infty$ follows by duality and triangle and Hölder's inequality.

- For $p = 2$, the estimate $\|T_{v_1} W\|_{L^2(\mathbb{S}^2)} \lesssim \|v_1\|_{L^2(\mathbb{S}^2)} \|W\|_2$ follows from the Tomas-Stein restriction theorem (see, e.g., [Fos15]), as stated in [Vit17]. In fact, for any two v_1, v_2, we have

  $$\|(v_1 \,\mathrm{d}\sigma) * (v_2 \,\mathrm{d}\sigma)\|_2 = \|\widehat{(v_1 \,\mathrm{d}\sigma)}\widehat{(v_1 \,\mathrm{d}\sigma)}\|_2 \leq \|\widehat{(v_1 \,\mathrm{d}\sigma)}\|_4 \|\widehat{(v_2 \,\mathrm{d}\sigma)}\|_4 \lesssim \|v_1\|_{L^2(\mathbb{S}^2)} \|v_2\|_{L^2(\mathbb{S}^2)}.$$

  The asserted inequality follows then by duality.

The considerations above show that T_{v_1} satisfies $L^\infty(\mathbb{R}^3) \to L^2(\mathbb{S}^2)$ and $L^2(\mathbb{R}^3) \to L^2(\mathbb{S}^2)$ estimates. By interpolation, it must also satisfy $L^p(\mathbb{R}^3) \to L^2(\mathbb{S}^2)$ estimates, with norm at most $\lesssim \|v_1\|_{L^2(\mathbb{S}^2)}$. By duality, this assertion is equivalent to

$$\left| \int_{\mathbb{S}^2 \times \mathbb{S}^2} v_1(\omega_1) v_2(\omega_2) W(\omega_1 + \omega_2) \,\mathrm{d}\sigma(\omega_1) \,\mathrm{d}\sigma(\omega_2) \right| \leq C \|v_1\|_{L^2(\mathbb{S}^2)} \|v_2\|_{L^2(\mathbb{S}^2)} \|W\|_{L^p(\mathbb{R}^3)}.$$

By setting $v_1 = v_2$ one obtains the Lemma. $\square$

To finish the proof, we apply Lemma 5.8 in (5.9) with $\eta > 2$. Using that M_2 is bounded in L^η and the Hausdorff–Young inequality gives

$$|\langle (M^*_{2,g,t(\cdot)} h)^2, F\rangle| \lesssim \|h\|^2_{L^2(\mathbb{S}^2)} \|\widehat{F}\|_\eta \leq \|h\|^2_{L^2(\mathbb{S}^2)} \|F\|_{\eta'}.$$

It is straightforward to check that this last inequality is equivalent to $M^*_{2,g,t(\cdot)} h$ being bounded from L^2 to $L^{2\eta}$. As $\eta > 2$ was arbitrary, we finish this part of the proof.

In order to deal with $1 \leq r < 2$, we use the pointwise domination $M_r f \leq M_2 f, 1 \leq r \leq 2$. Thus the only missing point in the proof above is the endpoint $(\frac{4}{3}, 2)$. A combination of the proofs of Theorems 5.2 and 5.4 gives us estimates in the endpoint case, in the same spirit as above. This time, the application of Lemma 5.8 might be circumvented, as M_r is bounded in L^2. We skip the details. $\square$

5.4 Comments, generalizations and remarks

5.4.1 Maximal operators of convolution-type and multiplier theorems

Theorems 5.1 and 5.2 deal with maximal functions related to a measure $\mathrm{d}\mu$. There, the key assumption is that these maximal functions must be bounded "near" L^2. As mentioned before, V. Kovač's result [Kov19] has a seemingly different assumption on the measure. For his purposes, it is important that the measure is finite – implied by the fact that the measure is complex – and that the gradient of its Fourier transform satisfies a decay of the type

$$|\nabla \widehat{\mu}(\xi)| \leq C(1 + |\xi|)^{-1-\eta}, \text{ for some } \eta > 0.$$

The next proposition shows that Kovač's hypotheses actually *imply* ours. We mention that this result is far from new, with a similar version appearing in [SMS85]. For the convenience of the reader, we quickly review the results from [RdF86]:

Proposition 5.9. *Let $T^* f(x) = \sup_{t>0} |\mathcal{F}^{-1}(m(t\cdot)\widehat{f})|$. Suppose that*

$$|m(\xi)| \lesssim (1 + |\xi|)^{-a}, \; |\nabla m(\xi)| \lesssim (1 + |\xi|)^{-b},$$

with $a + b > 1$. Then $T^* : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ boundedly.

Proof. Letting $\psi_0 : \mathbb{R}^n \to \mathbb{R}$ be a (radial) smooth function supported in the annulus $\{y\colon 1/2 \leq |y| \leq 2\}$ so that

$$\sum_{j\in\mathbb{Z}} \psi_0(2^j\xi) = 1,\ \forall \xi \neq 0,$$

we define $m_j(\xi) := m(\xi)\psi_0(2^j\xi)$. By letting T_j^* denote the maximal multiplier operator associated to each of these multipliers, we have

$$T^* f \leq T_0^* f + \sum_{j\leq 0} T_j^* f.$$

Here, we let $\sum_{j>0} m_j(\xi) = \phi_0(\xi)$ and define the operator T_0^* to be the maximal multiplier operator associated to ϕ_0. As ϕ_0 is a smooth function with compact support, this operator is bounded pointwise by a maximal function. We then move on to estimate each factor $T_j^* f$ individually: we bound the supremum by

$$\sup_{t>0} |\mathcal{F}^{-1}(m_j(t\cdot)\widehat{f})(x)|^2 \leq \left(\int_0^\infty |T_{j,t}f(x)\cdot \tilde{T}_{j,t}f(x)|\,\frac{\mathrm{d}t}{t}\right),$$

where $\widehat{T_{j,t}f}(\xi) = m_j(t\xi)\widehat{f}(\xi)$, $\widehat{\tilde{T}_{j,t}f}(\xi) = \tilde{m}_j(t\xi)\widehat{f}(\xi)$, with $\tilde{m}_j(\xi) = \xi\cdot\nabla m_j(\xi)$. We estimate then

$$\|T_j^* f\|_2^2 = \left(\int_0^\infty \int_{\mathbb{R}^n} |m_j(t\xi)\widehat{f}(\xi)|^2 \,\mathrm{d}\xi\,\frac{\mathrm{d}t}{t}\right)^{1/2} \times \left(\int_0^\infty \int_{\mathbb{R}^n} |\tilde{m}_j(t\xi)\widehat{f}(\xi)|^2 \,\mathrm{d}\xi\,\frac{\mathrm{d}t}{t}\right).$$

The integrals above exist only for $2^j t|\xi| \in [1/2, 2]$. Therefore, using the decay properties of $m, \tilde{m}$, we obtain

$$\|T_j^* f\|_2^2 \lesssim 2^{ja} 2^{j(b-1)} \|f\|_2^2 = 2^{j(a+b-1)}\|f\|_2^2.$$

As we supposed that $a + b > 1$, the series above is summable in $j < 0$, which completes the proof. □

Theorem 5.1 not only recovers a version of the two-dimensional results from Kovač, but also allows us to extend them, as mentioned before, to a larger class of maximal functions. For instance, Bourgain's circular maximal function fulfills the conditions to Theorem 5.1, whereas the gradient

$$|\nabla \widehat{\sigma_{\mathbb{S}^1}}(\xi)| \sim |\xi|^{-1/2}$$

for non-trivial sets of $|\xi| \to \infty$ in two dimensions, so that Kovač's result does not apply. Also, the spherical maximal function in dimension three satisfies that

$$|\nabla \widehat{\sigma_{\mathbb{S}^2}}(\eta)| \sim |\eta|^{-1}$$

on a non-trivial set of $|\eta| \to \infty$, but, as $|\widehat{\sigma_{\mathbb{S}^2}}(\eta)| = O(|\eta|^{-1})$, it is still possible to use Proposition 5.9 to conclude the L^2-boundedness of this operator, which is all we need to conclude.

5.4.2 The spherical maximal functions and previous maximal restriction results

In [Ram18], this author proves a full range 2-dimensional maximal restriction estimate for the *strong maximal function.* Namely, the main theorem there is that

$$\|M_{\mathcal{S}}(\widehat{f})\|_{L^q(\mathbb{S}^1)} \lesssim_p \|f\|_{L^p(\mathbb{R}^2)},$$

with $M_{\mathcal{S}}g(x) = \sup_{\substack{R \text{ axis parallel,} \\ \text{centered at} x}} \fint_R |g|$. One might ask is whether Theorem 5.1 implies the result above through a pointwise domination, as the spherical maximal function dominates the usual Hardy–Littlewood maximal function. Our next result shows that the answer is *no* in all dimensions larger than 1.

Proposition 5.10. *Let $d \geq 2$. Then there exists $f \in \cap_{p\geq 1} L^p(\mathbb{R}^d)$ such that*

$$\text{ess sup}_{x\in\mathbb{R}^d} \frac{M_{\mathcal{S}}f(x)}{M_{\mathbb{S}^{d-1}}f(x)} = +\infty.$$

Proof. Let first $d \geq 3$. In these cases, the counterexample is much simpler. In fact, we take $f = \chi_{Q(0,1)}$, the characteristic of the unit cube. It is a simple calculation to verify that $M_{\mathcal{S}}f(x) \gtrsim \frac{1}{|x|}$ whenever $|x| \gg 1$. Also, one obtains in a fairly straightforward manner that $M_{\mathbb{S}^{d-1}}f(x) \lesssim \frac{1}{|x|^{d-1}}$, $|x| \gg 1$. As $d-1 > 1$, f is a sought-after counterexample.

In dimension $d = 2$ matters are subtler. Let $g_n(x_1, x_2) = \chi_{[0,1]}(x_1)\chi_{[0,1/n^{20}]}(x_2)$. We take a sequence $(y_n, r_n) \in \mathbb{R}^2 \times \mathbb{R}_+$ such that

- $r_{n+1} = 10^n r_n$, $r_1 = 1$;
- $y_{n+1} = (r_1 + 2(r_2 + \cdots + r_n) + r_{n+1}, 0)$.

We then set up the function $f(x) = \sum_{n=1}^{\infty} g_n(x - y_n)$. This function is clearly in any L^p space. We estimate the strong and spherical maximal functions for x in a strip near y_n.

Effectively, let $x \in S_n := y_n + [-10^n, 10^n] \times [0, 1/n^{20}]$. Similarly as in the high dimensional case, $M_{\mathcal{S}}f(x) \gtrsim \frac{1}{|x-y_n|}$. Now we split the spherical maximal function into two parts as

$$M_{\mathbb{S}^1}f = \max\{M_{\mathbb{S}^1,\geq r_n}f, M_{\mathbb{S}^1,<r_n}f\}. \tag{5.10}$$

Here, $M_{\mathbb{S}^1,\geq t}g$ stands for the maximal function obtained by only taking radii larger than t, and define analogously $M_{\mathbb{S}^1,<t}f$. By the properties of the radii r_n and the way we defined y_n,

$$M_{\mathbb{S}^1,\geq r_n}f(x) \lesssim \frac{1}{r_n}.$$

Also, for the local part we obtain

$$M_{\mathbb{S}^1,<r_n}f(x) \lesssim \frac{1}{n^{10}|x-y_n|}.$$

Substituting these inequalities in the quotient, using (5.10), we get

$$\frac{M_{\mathcal{S}}f(x)}{M_{\mathbb{S}^1}f(x)} \gtrsim \min\left\{n^{10}, \frac{r_n}{|x-y_n|}\right\}.$$

Notice that $r_n = 10^{\frac{n(n-1)}{2}}$ and that $|x - y_n| \lesssim 10^n$ in S_n. We have found a set of measure $\gtrsim 10^{n/2}$ where the desired quotient is at least n^{10}. But these sets are mutually disjoint, which readily implies that the L^∞ norm of the quotient is not finite. $\square$

5.4.3 Theorems 5.2 and 5.5 and a Knapp-like counterexample

In this Section, we adapt the classical Knapp counterexample to obtain constraints on s, in order for versions of Theorems 5.2 and 5.5 to hold for a family of strong maximal functions:

Proposition 5.11. *Let*

$$M_{\mathcal{S},s}\, g = \left(\sup_{\substack{R \text{ axis parallel,} \\ \text{centered at } x}} \fint_R |g|^s \right)^{1/s}$$

denote the $s-$strong maximal function, in either two or three dimensions. Suppose that

$$\|M_{\mathcal{S},s}\, \widehat{g}\|_{L^q(\mathbb{S}^1)} \lesssim \|g\|_{L^p(\mathbb{R}^2)},$$

whenever $1 \le p < \frac{4}{3}$ and $3q \le p'$. Then $s \le 4$.

Analogously, suppose that

$$\|M_{\mathcal{S},s}\, \widehat{g}\|_{L^2(\mathbb{S}^2)} \lesssim \|g\|_{L^p(\mathbb{R}^3)},$$

for all $1 \le p \le 4/3$. Then $s \le 2$.

Before we move on to the proof, we remark that a combination of the proofs of Theorems 5.2, 5.5 and the ideas in [Ram18] attains that

$$\|M_{\mathcal{S},s}\widehat{f}\|_{L^q(\mathbb{S}^1)} \lesssim \|f\|_{L^p(\mathbb{R}^2)}, \text{ whenever } 1 \le p < \frac{4}{3}, p' \ge 3q \text{ and } s \le 2,$$

and

$$\|M_{\mathcal{S},s}\widehat{g}\|_{L^2(\mathbb{S}^2)} \lesssim \|g\|_{L^p(\mathbb{R}^3)}, \text{ whenever } 1 \le p \le \frac{4}{3} \text{ and } s < 2.$$

We spare the details, for their proofs are essentially the same as the ones presented.

Proof. We begin with the two-dimensional part. Let $\widehat{f_t}(\xi_1, \xi_2) = \chi_{(-t,t)}(\xi_1)\chi_{(1-t^2,1)}(\xi_2)$. We call this the *box-Knapp example.* It is easy to compute that

$$\|f_t\|_{L^p(\mathbb{R}^2)} = C \cdot t^{3-3/p}, \; \forall t > 0.$$

On the other hand, we estimate the maximal function $M_{\mathcal{S},s}(\widehat{f_t})$ from bellow as follows. Fix a small angle $\theta_0 > 0$. Then, for $\theta \in (\pi/4, \pi/2 - \theta_0)$, there is a constant $c(\theta_0)$ so that $\cos(\theta), 1 - \sin(\theta) \ge c(\theta_0)$. We estimate:

$$M_{\mathcal{S},s}(\widehat{f_t})(e^{i\theta}) \ge \left(\fint_{(-t,\cos(\theta))\times(\sin(\theta),1)} \chi_{(-t,t)\times(1-t^2,t)} \right)^{1/s} \gtrsim \frac{t^{3/s}}{\cos(\theta)^{1/s}(1-\sin(\theta))^{1/s}} \gtrsim t^{3/s}.$$

This is the estimate we need, for then

$$\|M_{\mathcal{S},s}(\widehat{f_t})\|_{L^q(\mathbb{S}^1)} \gtrsim \left(\int_{\pi/4}^{\pi/2-\theta_0} t^{3q/s} \mathrm{d}\theta \right)^{1/q} \gtrsim_{\theta_0} t^{3/s}.$$

Putting together yields that

$$\forall\, 0 < t \ll 1,\ t^{3/s} \lesssim t^{3-3/p} \iff \frac{1}{s} + \frac{1}{p} - 1 \geq 0.$$

If $s = 4 + \varepsilon$, then $1/p \geq \frac{3+\varepsilon}{4+\varepsilon} \iff p \leq \frac{4+\varepsilon}{3+\varepsilon} < \frac{4}{3}$, and the restriction estimates cannot hold in the full two-dimensional range.

For the three-dimensional part, we let $\widehat{F_t}(\eta_1, \eta_2, \eta_3) = \chi_{B^2(0,t)}(\eta_1, \eta_2)\chi_{(-1,1)}(\eta_3)$, and call this a *long-Knapp example.* Again, a computation shows that

$$\|F_t\|_{L^p(\mathbb{R}^3)} = \tilde{C} t^{2-2/p},\ \forall t > 0.$$

In this case, we bound $M_{\mathcal{S},s}(\widehat{F_t})$ from below by the $s-$average over a rectangle of dimensions $t \times t \times 4$ centered at each point $x \in \mathbb{S}^2$. In a spherical region of positive $\mathcal{H}^2-$measure, we have

$$M_{\mathcal{S},s}(\widehat{F_t}) \gtrsim t^{1/s} \Rightarrow t^{1/s} \lesssim t^{2-2/p},\ \forall t \text{ small} \iff \frac{1}{s} - 2 + \frac{2}{p} > 0.$$

Again, if $s > 2$, then p is forced to be strictly less than $4/3$. □

With the long-Knapp example, we prove the following:

Proposition 5.12. *The only dimensions in which maximal restriction estimates for $M_{\mathcal{S}} := M_{\mathcal{S},1}$ can hold in the full range are $d = 2, 3$.*

Proof. By an argument using long-Knapp example from above, in order for the full range of maximal restriction estimates of the kind

$$\|M_{\mathcal{S},s}(\widehat{f})\|_{L^q(\mathbb{S}^{d-1})} \lesssim \|f\|_{L^p(\mathbb{R}^d)} \tag{5.11}$$

to hold in the same regime as the already known restriction estimates, we must have $s \leq \frac{2(n+1)}{(n-1)^2}$. This number is less than 1 if $n \geq 5$. Also, using the results from [Gut18] (see also [HR18] for further developments), we know that the restriction estimates from 5.1 in dimension 4 for the sphere hold as long as $p' > 2.8$. Thus, in order for 5.11 to hold in the full range for $d = 4$, we need $s \leq \frac{2.8}{3} < 1$. In particular, this implies that $M_{\mathcal{S}}$ cannot be bounded in the full range, except for when $d = 2$ or $d = 3$. □

As proved in [Ram18], these estimates do hold in the case of the two-dimensional problem. An interesting question is the validity of the same bounds in dimension 3. Nevertheless, an affirmative answer would trivially imply the three-dimensional restriction conjecture, which is still not completely settled.

Note that the long-Knapp example, if translated to 2 dimensions, provides us with the *exact same* bounds as we have achieved. In fact, one achieves that, for $\widehat{\tilde{f}}_t(\xi_1, \xi_2) = \chi_{(-t,t)}(\xi_1)\chi_{[-1,1]}(\xi_2)$,

$$t^{1/s} \lesssim \|M_{\mathcal{S},s}(\widehat{\tilde{f}_t})\|_{L^q(\mathbb{S}^1)} \lesssim \|\tilde{f}_t\|_{L^p(\mathbb{R}^2)} \lesssim t^{1-1/p} \iff p' \geq s \Rightarrow s \leq 4.$$

Thus, we get no improvement from changing the counterexample's nature. Furthermore, if we replace the strong maximal function by the Hardy–Littlewood maximal function in *any* dimension, the long-Knapp and the box-Knapp examples deliver the same bounds for s:

$$\frac{1}{s} - \frac{1}{p'} \geq 0 \iff s \leq p'.$$

For the three-dimensional Tomas-Stein exponent case, we get the same $s \leq 4$ bound as in the two dimensions. One inquires whether there is any fundamental difference between the strong and the Hardy–Littlewood maximal functions in this context. Our counterexamples seem to hint at an intrinsic geometric distinction.

The three-dimensional Theorem 5.5 is sharp, in the sense that we have attained an almost exact characterization of when the maximal restriction estimates work. The only remaining case is the $s = 2, p = 4/3$ case. We suspect that the inequality should fail in that endpoint.

Part III

Carleson theorem and Hilbert transform along curves

Chapter 6

The Hilbert transform along the parabola, the polynomial Carleson theorem and oscillatory singular integrals

Aha!

–C.T.

This chapter contains the paper [Ram19a]. We make progress on an interesting problem on the boundedness of maximal modulations of the Hilbert transform along the parabola. Namely, if we consider the multiplier arising from it and restrict it to lines, we prove uniform L^p bounds for maximal modulations of the associated operators. Our methods consist of identifying where to use effectively the polynomial Carleson theorem, and where we can take advantage of the presence of oscillation to obtain decay through the TT^* method.

6.1 Introduction

6.1.1 Historical background

We define the Hilbert transform along the parabola as

$$\mathcal{H}_2 f(x,y) = \text{p.v.} \int_{\mathbb{R}} f(x-t, y-t^2) \frac{\mathrm{d}t}{t}, \tag{6.1}$$

where we let $f \in \mathcal{S}(\mathbb{R}^2)$. This operator has an anisotropic symmetry and has been considered in the wider framework of anisotropically homogeneous operators, dating back to the work of Fabes and Rivière [Fab66] in the 1970's. In this context, L^p estimates for such operators imply additional regularity of solutions of certain associated parabolic partial differential equations.

For the particular case of the Hilbert transform along the parabola, the works of Nagel, Rivière and Wainger [NRW74, NRW76] prove that it is indeed bounded in $L^p(\mathbb{R}^2)$. Their results provide, in fact, L^p-bounds for higher dimensional generalizations of this operator. Possible generalizations have been further explored in the nilpotent groups case by Christ [Chr85], as well as the question of weak-type endpoint estimates in the work of Christ and

Stein [CS87] and, more recently, in the work of Seeger, Tao and Wright [STW04].

Parallelly to that, the theory of *maximally modulated* Calderón-Zygmund operators also developped in the last 50 years. Indeed, in 1966, in order to prove almost everywhere convergence of Fourier series in L^2, Carleson [Car66] considers the operator

$$Cf(x) := \sup_{N\in\mathbb{R}} \left| \int_{\mathbb{R}} f(x-t)e^{iNt}\,\frac{\mathrm{d}t}{t} \right| = \sup_{N\in\mathbb{R}} |H(e^{iN(\cdot)}f)|(x).$$

This is now called the *Carleson operator*. After Carleson's paper, many works have been dedicated to sharpening and perfecting his proof. Hunt [Hun68] extended, in 1967, Carleson's result to all L^p spaces, $p \in (1,+\infty)$, and Fefferman [Fef70] and Lacey and Thiele [LT00] provided different proofs of the same result. All of the proofs above share, however, the property of employing a time-frequency decomposition to encompass translation, dilation and modulation symmetries of the Carleson operator.

Inspired by that result, E. M. Stein [Ste95] posed the following problem: if instead of linear phases, we take suprema over *polynomial* phases, do we still have L^p bounds? Namely, if one considers the operator

$$f \mapsto \sup_{\deg P\leq n} \left| \int_{\mathbb{R}} f(x-t)e^{iP(t)}\,\frac{\mathrm{d}t}{t} \right|,$$

is it bounded in $L^p(\mathbb{R}), p \in (1,+\infty)$? A first step in this direction is the work of Stein and Wainger [MSW01], where they consider a restricted supremum over polynomials *without* the linear term. Unlike the proofs of bounds for the Carleson operator, this does not rely on a time-frequency decomposition directly, but on a dyadic decomposition and TT^* method to exploit oscillatory integral estimates.

In subsequent works, Lie [Lie09] treated the case of weak-type $(2,2)$ bounds for the operator above if $n = 2$, and considered in [Lie11b] the general $n \geq 1$ case, in the one dimensional setting. More recently, Zorin-Kranich [ZK17] extended the analysis of the operator above for *higher* dimensions and Calderón-Zygmund operators with fairly general conditions. Their techniques, however, resort more to time-frequency methods in the style of Fefferman [Fef70] rather than the TT^* strategy of Stein and Wainger.

Pierce and Yung [BPY15] considered a hybrid version of the two parallel kinds of results we discussed. In particular, they consider operators of the form

$$f(x,y) \mapsto \sup_{P\in\mathcal{P}} \left| \int_{\mathbb{R}^d} f(x-t, y-|t|^2)e^{iP(t)}K(t)\,\mathrm{d}t \right|,$$

where K is a suitable Calderón-Zygmund kernel and $\mathcal{P}$ some finite-dimensional subspace of polynomials. They obtain L^p estimates for certain subspaces that avoid linear and some quadratic terms, as long as $d \geq 2$. Subsequently to it, Guo, Pierce, Roos and Yung [GPRY17] considered the $d = 2$ case by taking a partial supremum for curves like (t, t^d) and $P(t) = N \cdot t^m$. For these results, as well as the ones in [BPY15], the strategy resembles that of Stein and Wainger, in the sense that the main tools are still dyadic decompositions, TT^* estimates and suitable oscillatory integral estimates to obtain decay.

Continuing this line of thought, the following question arises naturally in [GPRY17]:

Question 6.1. *For $f \in \mathcal{S}(\mathbb{R}^2)$, is the* parabolic *Carleson operator*

$$\mathcal{C}_2 f(x,y) := \sup_{N,M\in\mathbb{R}} \left| \int_{\mathbb{R}} f(x-t, y-t^2) e^{iNt+iMt^2} \frac{\mathrm{d}t}{t} \right| \tag{6.2}$$

bounded in $L^2(\mathbb{R}^2)$?

This is nothing but a supremum of the Hilbert transform along the parabola of all possible modulations of f. In other terms, this operator admits a representation as

$$\sup_{N,M} |\mathcal{H}_2(e^{iN(\cdot)_1+iM(\cdot)_2} f)(x,y)| = \sup_{N,M} |\mathcal{F}^{-1}(m_2(\cdot + (N,M))\widehat{f})(x,y)|,$$

where $m_2\left(\frac{\xi}{2\pi}, \frac{\eta}{\sqrt{2\pi}}\right) = \int_{\mathbb{R}} e^{i(\xi\cdot t + \eta\cdot t^2)} \frac{\mathrm{d}t}{t}$, and $\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^2} f(y) e^{-2\pi i \xi\cdot y}\,\mathrm{d}y$ denotes the Fourier transform. Here and henceforth we abuse notation and denote by m_2 the dilation given above by $m_2(\cdot/2\pi, \cdot/\sqrt{2\pi})$.

Partial progress in Question 6.1 has been made by Roos [Roo17], where he extends the techniques from Lacey and Thiele [LT00] to the anisotropic case. The main obstacle to apply his result to Question 6.1 is the fact that m_2 is only Hölder continuous of exponent < 1 along $\mathbb{R}$, while Roos needs regularity of his multipliers greater than three times the anisotropic degree of homogeneity. Interestingly enough, Zorin-Kranich mentions in [ZK17] that it should be possible to extend his results on the polynomial Carleson operator to the anisotropic context. This would, however, still not imply the validity of Question 6.1, as the techniques from Zorin-Kranich only yield bounds for symbols with regularity at least equal to the homogeneity degree.

6.1.2 Main results

We are interested in the restriction of $m_2(\xi, \eta)$ to lines. That is, we consider the family of one-dimensional functions given by $m_{a,b}(\eta) = m_2(a\eta + b, \eta)$, $a, b \in \mathbb{R}$. In order to consider also horizontal lines, we define $m_{+\infty,b}(\eta) := m_2(\eta, b)$. They define, via Fourier inversion, a family of linear operators $T_{a,b}f(x) := \mathcal{F}^{-1}(m_{a,b}\widehat{f})(x)$ in dimension 1. Our main result deals with maximal modulations of these operators – or, analogously, maximal translations in the multiplier side:

Theorem 6.2. *Let $\mathcal{C}_{a,b}f(x) = \sup_{N\in\mathbb{R}} |T_{a,b}(e^{iN\cdot} f)(x)|$. Then it holds that*

$$\sup_{a\in\mathbb{R}\cup\{+\infty\}} \|\sup_{b\in\mathbb{R}} \mathcal{C}_{a,b}\|_{p\to p} < +\infty,\ \forall p \in (1, +\infty).$$

Intuitively, Theorem 6.2 represents taking a very thin strip around the line $(a\eta + b, \eta)$ and a function with Fourier transform supported there and calculating $\mathcal{C}_2$ with this additional restriction.

It is not hard to see that Theorem 6.2 follows if the answer to Question 6.1 is affirmative. Indeed, for any line $\ell \subset \mathbb{R}^2$ as above consider a strip S_δ of width $\delta > 0$ with direction ℓ. Consider also the set of lines $\ell' \sim \ell$ parallel to ℓ. If we consider functions F_δ such that their Fourier transform is supported on S_δ, is essentially constant along the perpendicular direction to ℓ and equals $\widehat{g}$ on ℓ, we have, formally,

$$\|\sup_{\ell'\sim\ell} \mathcal{C}_{\ell'} g\|_{L^2(\mathbb{R})} \le \lim_{\delta\to 0} \|\mathcal{C}_2(F_\delta)\|_2 \le C \limsup_{\delta\to 0} \|F_\delta\|_2 = C\|g\|_{L^2(\mathbb{R})}. \tag{6.3}$$

We elaborate more on (6.3) in Section 6.4.

Our first task is to pass from the rather complicated formulation in Theorem 6.2 to a formulation with which we can work more directly. This is the main content of the next proposition, which we prove in Section 6.2.

Proposition 6.3. *Let $[u]^{1/2}$ denote either $|u|^{1/2}$ or sign$(u)|u|^{1/2}$. Suppose that the operators*

$$\mathfrak{C}^R f(x) = \sup_{N,b\in\mathbb{R}} \left| \int_{-R}^{R} f(x-t) e^{iNt} e^{ib[t+1]^{1/2}} \frac{\mathrm{d}t}{t} \right| \tag{6.4}$$

are bounded in L^p, for $1 < p < +\infty$, independently of the truncating parameter $R > 0$. Then Theorem 6.2 holds.

We still have to bound the operator arising from the proposition above. The following result asserts the boundedness of the second maximal operator.

Theorem 6.4. *Let $\mathfrak{C}^R$ be defined as (6.4) above. It holds that*

$$\|\mathfrak{C}^R f\|_p \leq C_p \|f\|_p,$$

for all $f \in L^p(\mathbb{R})$ and all $p \in (1, +\infty)$, and C_p independent *of $R > 0$.*

The proof of Theorem 6.4 is the main novelty of this article. In order to prove it, we employ two different ideas. More specifically, we first prove that we can regard the parameter b as belonging to a fixed dyadic scale $\sim 2^k$, as long as we prove summable decay in $|k|$. We then break the interval of integration defining $\mathfrak{C}^A f$ into distinct regimes of intervals, namely mainly the ones where oscillatory behaviour is strong enough to enable the use of the TT^* method, and the ones where the phase is mimics a polynomial Carleson operator, as considered in [Lie11b, ZK17].

The crucial point of using the polynomial Carleson theorem together with an application of TT^* to prove Theorem 6.4 is that this technique is optimal. That is, for smaller scales, the oscillatory integral estimates used in the TT^* method only give us a bound not decaying with $b \sim 2^k$. Truncating at a high power of b does not appear randomly. On the other hand, one asks whether it is possible to use a comparison to a polynomial Carleson operator *directly* at least in the interval $[-1/2, 1/2]$. We finish our discussion of the proof of Theorem 6.2 by proving that it is impossible unless letting the degree of the polynomial tend to infinity.

6.1.3 Notation

Some remarks are in order to facilitate the reader's understanding:

i. We denote by $C > 0$ a constant that may change from line to line;

ii. We write throughout the paper $A \lesssim B$ to mean that $A \leq C \cdot B$, for some constant C. If C depends on some parameter δ in a relevant way, we write $A \lesssim_\delta B$;

iii. ψ generally denotes positive bump functions with some partition of unity property, whereas ϕ, φ denote usually phase functions in oscillatory integrals;

iv. Finally, unless otherwise stated, we consider the functions in the proofs below to belong to $\mathcal{S}(\mathbb{R})$ and extend the respective bounds by density.

6.2 Proof of Proposition 6.3

In order to prove Theorem 6.2, we first reduce the analysis to simpler operators. We mention that, if we let go of the uniformity of the estimates on the line, the comparisons in this section can be made much looser. Therefore, the emphasis is on getting bounds *independent* on the paramers when comparing.

6.2.1 From two to one parameter

We reduce the task of proving Theorem 6.2 to the analysis of a one-parameter family of operators. We must consider *all* lines in the analysis, and therefore also $m_{+\infty,b}(\xi) = m_2(\xi, b)$ must be considered as a multiplier. Nevertheless, an analysis identical to the one undertaken below shows that the operator

$$\sup_{b,N} |\mathcal{F}^{-1}(m_{+\infty,b}(\widehat{\mathcal{M}_N f}))|$$

is just a quadratic Carleson operator, so, by the results in [Lie09, Lie11b, ZK17], we need not include it in our discussion. The main tool to our reduction will be the following Lemma:

Lemma 6.5. *Let $\{g_{a,b}\} \subset L^1(\mathbb{R})$ be a family of positive functions with $|g_{a,b}| \leq h_a$ pointwise for another family $\{h_a\}$, which is uniformly bounded in L^1. I.e., $\sup_a \|h_a\|_1 < +\infty$. It holds that*

$$\sup_a \| \sup_{b,N} |g_{a,b} * (\mathcal{M}_N f)|\|_p \leq C\|f\|_p,$$

with $\mathcal{M}_N f(x) = e^{2\pi i N x} f(x)$.

Proof. By Young's convolution inequality,

$$\| \sup_{b,N} |g_{a,b} * (M_N f)|\|_{L^p} \leq \|h_a * |f|\|_{L^p} \leq \|h_a\|_1 \|f\|_{L^p} \leq \left(\sup_a \|h_a\|_1 \right) \|f\|_{L^p} =: C\|f\|_{L^p}.$$

□

We will especially use it in the following form: if two families of maximally modulated operators $\mathcal{O}_i^{a,b} f(x) := \sup_N |O_i^{a,b}(\mathcal{M}_N f)(x)|$, $i = 1, 2$ satisfy

$$|\mathcal{O}_1^{a,b} f(x) - \mathcal{O}_2^{a,b} f(x)| \leq |f| * h_a(x),$$

with $h_a \in L^1(\mathbb{R})$ as in Lemma 6.5 above, then bounding $\mathcal{O}_1^{a,b}$ in L^p uniformily in a is equivalent to bounding $\mathcal{O}_2^{a,b}$ uniformily in a. With this in mind, we rewrite our multipliers as

$$m_2(2a\eta + b, \eta) = \int_{\mathbb{R}} e^{i(2a\eta t + bt + \eta t^2)} \frac{\mathrm{d}t}{t} = e^{i \cdot a^2 \cdot \eta} \int_{\mathbb{R}} e^{i(t-a)^2 \eta + ibt} \frac{\mathrm{d}t}{t},$$

We further rewrite the operators $T_{2a,b}$ using Fourier inversion:

$$T_{2a,b} f(x) = \int_{\mathbb{R}} f(x + a^2 - (t-a)^2) e^{ibt} \frac{\mathrm{d}t}{t}.$$

Notice that the a^2 term only contributes as a translation in the $x-$variable, so we consider the simpler operators

$$\tilde{T}_{2a,b}f(x) := \int_{\mathbb{R}} f(x-(t-a)^2)e^{ibt}\frac{\mathrm{d}t}{t}.$$

By translating t by a and changing variables $s = t^2$ – after breaking the integral into $\mathbb{R}_+$ and $\mathbb{R}_-$ parts –, we get from the observation above that bounding $\sup_{N,b}|\tilde{T}_{2a,b}(\mathcal{M}_N f)|$ is equivalent to bounding

$$A_a f(x) = \sup_{N,b}\left|\int_0^{+\infty}(\mathcal{M}_N f)(x-s)\left(\frac{e^{ibs^{1/2}}}{s^{1/2}(s^{1/2}-a)} - \frac{e^{-ibs^{1/2}}}{s^{1/2}(s^{1/2}+a)}\right)\mathrm{d}s\right|.$$

6.2.2 Reduction to model operators

We now look more closely into this family of operators. First, we rewrite the kernel defining A_a as

$$\begin{aligned}\left(\frac{e^{ibs^{1/2}}}{s^{1/2}(s^{1/2}-a)} - \frac{e^{-ibs^{1/2}}}{s^{1/2}(s^{1/2}+a)}\right) &= \frac{1}{2s^{1/2}}\cdot\frac{e^{ibs^{1/2}}-e^{-ibs^{1/2}}}{s^{1/2}+a} + \frac{a}{s^{1/2}}\cdot\frac{e^{ibs^{1/2}}}{s-a^2}\\ &=: K^b_{1,a}(t) + K^b_{2,a}(t).\end{aligned}$$

If $a = 0$, we have $K^b_{2,0}(t) = 0$, whereas $K^b_{1,0}(t)$ becomes

$$\frac{e^{ibt^{1/2}}-e^{-ibt^{1/2}}}{2t}.$$

We write, for the time being, the maximal operator we are left with as

$$\sup_{N,b}\left|\int_0^{+\infty}(\mathcal{M}_N f)(x-t)\frac{e^{ibt^{1/2}}-e^{-ibt^{1/2}}}{2t}\,\mathrm{d}t\right|. \tag{6.5}$$

For the $a \neq 0$ cases, we notice that $A_{-a}(f) = A_a f$, so we suppose without loss of generality that $a > 0$. We bound the kernel pointwise in a suitable neighborhood of the origin, and compare it to another operator away. Especifically, for $0 \le t \le a^2$, we have

$$|K^b_{1,a}(t)| \le \min(1, t^{-1/2})\cdot\frac{1}{t^{1/2}+a}.$$

It is then easy to see that $\int_0^{a^2}\min(1,t^{-1/2})\cdot\frac{1}{t^{1/2}+a}\mathrm{d}t$ is bounded independently of $a > 0$. By Lemma 6.5, we are left with the $t \ge a^2$ portion, where we estimate

$$\left|K^b_{1,a}(t) - \frac{e^{ibt^{1/2}}-e^{-ibt^{1/2}}}{2t}\right| \le \frac{a}{t(t^{1/2}+a)}.$$

It is again straightforward to see that $\int_{a^2}^{\infty}\frac{a}{t(t^{1/2}+a)}\mathrm{d}t < 10$, $\forall a > 0$. We have again reduced the analysis to the operator in (6.5).

We now address the $K^b_{2,a}$ part. We split the integral defining $K^b_{2,a} * (\mathcal{M}_N f)$ into three regimes: $[0,+\infty) = [0, a^2/2] \cup (a^2/2, 3a^2/2) \cup [3a^2/2, +\infty)$. For each of them, we have:

- $|K^b_{2,a}(t)| \le \frac{4}{a^2}$ for $t \in [0, a^2/2]$. Therefore, convolution with this part can be controlled by Lemma 6.5;

- For $t \ge 3a^2/2$, we estimate $|K^b_{2,a}(t)| \le \frac{a}{t^{1/2}(t-a^2)}$, where a change of variables leads us to conclude that $\int_{3a^2/2}^{+\infty} \frac{a}{t^{1/2}(t-a^2)} \mathrm{d}t < 10$;

- For the singular middle interval, we compare:

$$\left| K^b_{2,a}(t) - \frac{e^{ibt^{1/2}}}{t-a^2} \right| \le \frac{1}{t^{1/2}(t^{1/2}+a)}.$$

  By a change of variables $t \mapsto s^2$, one sees that $\int_{a^2/2}^{3a^2/2} \frac{\mathrm{d}t}{2t^{1/2}(t^{1/2}+a)} = \log\left(\frac{\sqrt{3/2}+1}{\sqrt{1/2}+1}\right)$. By observing the operator to which we compared, we conclude that it is enough to control the L^p norm of $\sup_{N,b}\left|\int_{-a^2/2}^{a^2/2} f(x-t)e^{iNt}e^{ib\sqrt{t+a^2}}\frac{\mathrm{d}t}{t}\right|$ independently of the parameter a. changing variables, it is easy to see that this expression equals $\mathfrak{C}^{1/2}(f_a)(x/a^2)$, where $f_{a^2}(y) = f(a^2y)$. If Theorem 6.4 holds, then the L^p norm of this expression is bounded independently of a.

In order to finish the proof of Proposition 6.3, we must conclude boundedness of the operator

$$\sup_{N,b}\left|\int_0^{+\infty} (\mathcal{M}_N f)(x-t)\frac{e^{ibt^{1/2}} - e^{-ibt^{1/2}}}{2t}\,\mathrm{d}t\right|$$

given Theorem 6.4. In fact, we first need an auxiliary result:

Proposition 6.6. *Suppose Theorem 6.4 holds. Then the maximal functions*

$$f \mapsto \sup_{N,b}\left|\int_{\mathbb{R}} f(x-t)e^{iNt}e^{ib[t]^{1/2}}\frac{\mathrm{d}t}{t}\right|$$

are both bounded in L^p.

Proof. We consider $\mathfrak{C}^{+\infty} := \mathfrak{C}$ as in Theorem 6.4. If we define

$$\mathfrak{C}_a f(x) = \sup_{N,b}\left|\int_{\mathbb{R}} f(x-t)e^{iNt}e^{ib[t+a]^{1/2}}\frac{\mathrm{d}t}{t}\right|,$$

the dilation symmetries of $\mathfrak{C}$ imply that $\mathfrak{C}f(x) = \mathfrak{C}_a(f_{1/a})(ax)$, where $f_{1/a}(y) = f(y/a)$. This plainly implies

$$\|\mathfrak{C}_a f\|_p = a^{1/p}\|\mathfrak{C}(f_a)\|_p \lesssim a^{1/p}\|f_a\|_p = \|f\|_p. \tag{6.6}$$

Now, a direct computation together with dominated convergence shows that

$$\liminf_{a\to 0} \mathfrak{C}_a f(x) \ge \sup_{N,b}\left|\int_{\mathbb{R}} f(x-t)e^{iNt}e^{ib[t]^{1/2}}\frac{\mathrm{d}t}{t}\right|$$

for all smooth f with compact support. The proposition is then proved by Fatou's lemma and (6.6). ☐

We rewrite the operators from Proposition 6.6 as

$$\sup_{N,b}\left|\int_0^{+\infty}((\mathcal{M}_N f)(x-t)e^{ib[t]^{1/2}}-(\mathcal{M}_N f)(x+t)e^{ib[-t]^{1/2}})\,\frac{\mathrm{d}t}{t}\right|$$

If $[u]^{1/2}=\operatorname{sign}(u)|u|^{1/2}$, the integrand equals

$$(\mathcal{M}_N f)(x-t)e^{ibt^{1/2}}-(\mathcal{M}_N f)(x+t)e^{-ibt^{1/2}},$$

and it becomes $((\mathcal{M}_N f)(x-t)-(\mathcal{M}_N f)(x+t))e^{ibt^{1/2}}$ in case $[u]^{1/2}=|u|^{1/2}$. Notice now that

$$\begin{aligned}(\mathcal{M}_N f)(x-t)\cdot(e^{ibt^{1/2}}-e^{-ibt^{1/2}}) &= (\mathcal{M}_N f)(x-t)e^{ibt^{1/2}}-(\mathcal{M}_N f)(x+t)e^{-ibt^{1/2}}\\ &\quad -((\mathcal{M}_N f)(x-t)-(\mathcal{M}_N f)(x+t))e^{-ibt^{1/2}},\end{aligned}$$

so that the operator from (6.5) is bounded by the sum of the two in Proposition 6.6. This finishes the reduction to Theorem 6.4.

6.3 Proof of Theorem 6.4

In order to deal with the operators $\mathfrak{C}^R$, we use the Kolmogorov-Seliverstov linearization method. In fact, by suitably choosing, we find two functions $b,N:\mathbb{R}\to\mathbb{R}_+$, taking on only finitely many values, so that

$$|\mathfrak{C}^R_{b,N}f(x)|=\left|\int_{-R}^{R}f(x-t)e^{ib(x)[t+1]^{1/2}}e^{iN(x)t}\frac{\mathrm{d}t}{t}\right|\geq\frac{1}{2}\mathfrak{C}^R f(x).$$

Our goal is to bound this operator *independently* of both b and N, as well as $R>0$. We omit therefore b,N,R in order to clean up notation. The first step is to split the analysis of this operator into two parts:

$$\begin{aligned}|\mathfrak{C}f(x)| &\leq 1_{\{b(x)\leq 10\}}\mathfrak{C}f(x)+1_{\{b(x)>10\}}\mathfrak{C}f(x)\\ &=: \mathfrak{C}^1 f(x)+\mathfrak{C}^2 f(x).\end{aligned}$$

Part 1: Analysis of $\mathfrak{C}^1$. We split the interval $[-R,R]$ of the integral defining $\mathfrak{C}^1$ as

$$[-R,-\min\{R,b(x)^{-2}\}]\cup(-\min\{R,b(x)^{-2}\},\min\{R,b(x)^{-2}\})\cup[\min\{R,b(x)^{-2}\},R].$$

In the middle interval, the aim is to simply approximate the phase $b(x)[t+1]^{1/2}$ by $b(x)$. In more effective terms, the difference

$$\left|\int_{-\min\{R,b(x)^{-2}\}}^{\min\{R,b(x)^{-2}\}}f(x-t)e^{iN(x)t}e^{ib(x)[t+1]^{1/2}}\,\frac{\mathrm{d}t}{t}-e^{ib(x)}\int_{-\min\{R,b(x)^{-2}\}}^{\min\{R,b(x)^{-2}\}}f(x-t)e^{iN(x)t}\,\frac{\mathrm{d}t}{t}\right|$$

is bounded pointwise by

$$b(x)\int_{-\min\{R,b(x)^{-2}\}}^{\min\{R,b(x)^{-2}\}}|f(x-t)||[t+1]^{1/2}-1|\,\frac{\mathrm{d}t}{|t|}.$$

Notice that the difference $h(t) = |[t+1]^{1/2} - 1|$ satisfies that

$$h(t) \leq \begin{cases} 4|t|, & \text{if } t \in [-1/2, 2]; \\ 4, & \text{if } t \in [-2, -1/2]; \\ 4|t|^{1/2}, & \text{if } |t| \geq 2. \end{cases}$$

The function $h(t)/|t|$ admits then a radial majorizer $H(t)$ whose integral is at most a multiple of $\min\{R, b(x)^{-2}\}^{1/2}$. Because of the multiplying $b(x)$ factor in front, this integral is pointwise bounded by an absolute constant times the Hardy-Littlewood maximal function of f at the point x. It is well-known (cf. Grafakos [Gra14b, Section 6.3]) that the maximally truncated version of the Carleson operator is bounded in L^p. The L^p boundedness for $\mathfrak{C}^1$ restricted to this middle interval then follows from L^p boundedness of the maximal function.

For the two outer intervals, the main idea is to use the TT^* method to get summable decay in the scales. This is a crucial idea in this argument, and this will be emphasized by its incidence in this section.

Namely, we suppose that $R > b(x)^{-2}$, as this part of the analysis gets trivialized in case $R \leq b(x)^{-2}$. Let $\psi_0 : \mathbb{R} \to \mathbb{R}$ be a positive, smooth bump function supported in $[1/2, 2]$ such that

$$\sum_{j \in \mathbb{Z}} \psi_0\left(\frac{\xi}{2^j}\right) \equiv 1, \ \forall \xi \in \mathbb{R} \setminus \{0\}.$$

We analyze the integral only over the interval $[b(x)^{-2}, R]$, as the other part the analysis is entirely analogous. By a computation analogous to the one performed above to control the middle interval, we obtain that the integral defining this operator over the interval $[b(x)^{-2}, R]$ is, modulo error terms amounting to maximal function, equal to

$$\sum_{j \geq 0} \int_{\mathbb{R}} f(x-t) e^{iN(x)t} e^{ib(x)[t+1]^{1/2}} \psi_0(2^{-j} b(x)^2 t) \frac{\mathrm{d}t}{t} =: \sum_{j \geq 0} \mathfrak{S}^j f(x). \tag{6.7}$$

Some remarks are in order about the operators $\mathfrak{S}^j$. First of all, these operators are pointwise bounded by an absolute constant times the Hardy–Littlewood maximal function, due to the space localization $1/2 \leq 2^{-j} b(x)^2 t \leq 2$ imposed in the integral. Therefore, we immediately get

$$\|\mathfrak{S}^j f\|_{1,+\infty} \lesssim \|f\|_1, \ \|\mathfrak{S}^j f\|_\infty \lesssim \|f\|_\infty.$$

In order to conclude bounds on the sum in (6.7), it suffices to prove that $\|\mathfrak{S}^j f\|_2 \lesssim 2^{-\tau j}\|f\|_2$, for some $\tau > 0$. Indeed, by interpolating with the endpoint estimates above, we obtain that there is $\tau_p > 0$ such that

$$\|\mathfrak{S}^j f\|_p \lesssim_p 2^{-\tau_p j}\|f\|_p, \ \forall p \in (1, +\infty).$$

Finally, the L^p-norm of the expression in (6.7) is controlled, by triangle inequality, by

$$\sum_{j \geq 0} \|\mathfrak{S}^j f\|_p \lesssim_p \sum_{j \geq 0} 2^{-\tau_p j}\|f\|_p \lesssim_p \|f\|_p.$$

We focus hence on the extra decay for the L^2 bounds. That is the content of the following proposition.

Proposition 6.7. *Let $\mathfrak{S}^j$ be defined as above. It holds that*

$$\|\mathfrak{S}^j f\|_2 \lesssim 2^{-j/200}\|f\|_2,$$

for all $f \in L^2(\mathbb{R})$.

Proof. We first write the operator $\mathfrak{S}^j f(x) = S^j_{N(x),b(x)} * f(x)$, where we define

$$S^j_{N(x),b(x)}(s) = e^{iN(x)s}e^{ib(x)[s+1]^{1/2}}\psi_0(2^{-j}b(x)^2 s)\,\frac{1}{s}.$$

Now, in order to compute the L^2 norm of $\mathfrak{S}^j$, we compute instead its composition with its adjoint. It admits an expansion as

$$\mathfrak{S}^j(\mathfrak{S}^j)^* f(x) = \int_{\mathbb{R}} (S^j_{N(x),b(x)} * \tilde{S}^j_{N(y),b(y)})(x-y)f(y)\,\mathrm{d}y.$$

Here, we let $\tilde{S}^j_{N(x),b(x)}(z) = \overline{S^j}_{N(x),b(x)}(-z)$. A computation shows, on the other hand, that $|(S^j_{N(x),b(x)} * \tilde{S}^j_{N(y),b(y)})(\xi)|$ equals

$$\left|\int_{\mathbb{R}} e^{i(N(x)-N(y))s}e^{i(b(x)[s+1]^{1/2}-b(y)[s-\xi+1]^{1/2})}\,\frac{\psi_0(2^{-j}b(x)^2 s)}{s}\,\frac{\psi_0(2^{-j}b(y)^2(s-\xi))}{s-\xi}\,\mathrm{d}s\right|.$$

We change variables in this last integral to simplify the analysis. Effectively, assume, without loss of generality, that $b(y) < b(x)$. We let $s' = 2^{-j}b(x)^2 s$, and denote $\xi' = 2^{-j}b(y)^2\xi$. The integral whose absolute value we would like to estimate rewrites then as

$$\frac{b(y)^2}{2^j}\int_{\mathbb{R}} e^{i2^j b(x)^{-2}(N(x)-N(y))s'}e^{iR_{\xi',j,b}(s')}\,\frac{\psi_0(s')}{s'}\,\frac{\psi_0(hs'-\xi')}{hs'-\xi'}\,\mathrm{d}s',$$

where we let $h := \frac{b(y)^2}{b(x)^2} < 1$, and consider the phase function given by

$$R_{\xi',j,b}(s') = 2^{j/2}([s'+2^{-j}b(x)^2]^{1/2} - [hs'-\xi'+2^{-j}b(y)^2]^{1/2}).$$

This is the oscillatory integral we would like to estimate. The following lemma is the tool to directly do it.

Lemma 6.8. *Let $\Psi : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be a smooth function supported in $\{(s',\xi') \in \mathbb{R}^2 : s', hs'-\xi' \in [1/2,2]\}$, for some fixed positive parameter $h \le 1$. It holds that, for all $v, \xi' \in \mathbb{R}$ and $j \ge 0$,*

$$\begin{aligned}&\left|\int_{\mathbb{R}} e^{ivs'}\cdot e^{iR_{\xi',j,b}(s')}\Psi(\xi',s')\mathrm{d}s'\right| \\ &\lesssim \sup_{\xi'\in[-2,2]}\|\Psi(\xi',\cdot)\|_{C^2}\left(1_{[-2^{-j/100},2^{-j/100}]}(\xi') + 2^{-\frac{j}{9}}1_{[-4,4]}(\xi')\right),\end{aligned} \tag{6.8}$$

where the implicit constant does not depend on $h \in (0,1]$.

Proof. The proof follows the essential principle that, for ξ' small enough, we cannot expect much more from the integral than the trivial triangle inequality bound, and if ξ' is non-small, the oscillation of the phase $R_{\xi',j,b}$ starts providing cancellation, and therefore decay. In fact, if $|\xi'| \le 2^{-j/100}$, we use triangle inequality as pointed out, and one readily obtains the first term on the right hand side of the statement.

If, nonetheless, $|\xi'| \ge 2^{-j/100}$, we have to prove some sort of lower bound on the derivatives of the (completed) phase

$$\varphi(s') = vs' + R_{\xi',j,b}(s').$$

For that purpose, we consider the vector

$$\mathcal{Q}(s') = \begin{pmatrix} \varphi''(s') \\ -\frac{2}{3}\varphi'''(s') \end{pmatrix}$$

of second and third derivatives of the phase. The aim is to prove this is bounded from below by some positive power of 2^j, so that stationary phase considerations give us the desired decay. In order to prove this bound, we adopt a strategy already present in [GPRY17] and [GHLR17]. Namely, the idea used multiple times there is to rewrite the vector $\mathcal{Q}$ as a certain (invertible) matrix applied at a vector. If we prove sufficiently good bounds on the determinant and norms of the objects involved, we should get good enough bounds on the original $\mathcal{Q}$. A more precise version of this principle is the following Lemma.

Lemma 6.9. *Let A be a $n \times n$ invertible matrix. It holds that*

$$|A \cdot x| \ge |\det(A)| \|A\|^{1-n} \cdot |x|.$$

Proof. Assume, by homogeneity, that $\|A\| = 1$. Let us show first that $\|A^{-1}\| \le \frac{1}{\det(A)}$. It is simple to see that the eigenvalues of AA^* are all contained in $[0,1]$. If $\lambda(A)$ is the smallest eigenvalue of AA^*, it holds that $1 \ge \lambda(A) \ge \det(AA^*)$. On the other hand,

$$\|A^{-1}\| = \sup_{\|v\|=1} \langle A^{-1}v, A^{-1}v\rangle^{1/2} \le \sup_{\|v\|=1} \langle v, (A^*)^{-1}A^{-1}v\rangle^{1/2} \le \lambda(A)^{-1/2}.$$

Both imply that $\|A^{-1}\| \le \frac{1}{|\det(AA^*)|^{1/2}} \le \frac{1}{|\det(A)|}$, which implies our first claim. Now, we simply write

$$|x| = |A^{-1} \cdot A \cdot x| \le \|A^{-1}\| \cdot |A \cdot x| \le \|A\|^{n-1} \cdot |\det(A)|^{-1} \cdot |A \cdot x|,$$

in order to conclude the proof. □

With this Lemma in hands, we simply need to notice that $\mathcal{Q}(s')$ equals

$$2^{j/2} \begin{pmatrix} 1 & 1 \\ (s' + 2^{-j}b(x)^2)^{-1} & h(hs' - \xi' + 2^{-j}b(y)^2)^{-1} \end{pmatrix} \cdot \begin{pmatrix} (s' + 2^{-j}b(x)^2)^{-1/4} \\ h^2(hs' - \xi' + 2^{-j}b(y)^2)^{-1/4} \end{pmatrix}.$$

$$=: 2^{j/2}\mathcal{M}(s') \cdot \mathcal{V}(s').$$

By the fact that $j \ge 0, b(x) \le 10$, we see that $\|\mathcal{M}(s')\| \lesssim 1$, as well as

$$|\det(\mathcal{M}(s'))| \gtrsim \frac{|\xi'|}{|s' + 2^{-j}b(x)^2||hs' - \xi' + 2^{-j}b(x)^2|} \gtrsim 2^{-j/100}$$

and $|\mathcal{V}(s')| \gtrsim 1$. Lemma 6.9 gives that

$$|\mathcal{Q}(s')| \gtrsim 2^{j/3}.$$

Therefore, there is $i \in \{2,3\}$ so that $|\varphi^{(i)}(s')| \gtrsim 2^{j/3}$. By Proposition 2 in Chapter VIII of [Ste93], we get that, for $|\xi'| \gtrsim 2^{-j/100}$, the integral from the Lemma is bounded by a multiple of

$$\|\Psi(\xi',\cdot)\|_{C^2} \cdot 2^{-j/9}.$$

This gives the second summand in the statement of Lemma 6.8, and therefore finishes the proof. □

By Lemma 6.8, we obtain that $|(S^j_{N(x),b(x)} * \tilde{S}^j_{N(y),b(y)})(\xi)|$ is bounded by an absolute constant times

$$\frac{\mu(x,y)^2}{2^j}\left(1_{(-2^{-j/100},2^{-j/100})}\left(\frac{\mu(x,y)^2\xi}{2^j}\right) + 2^{-j/9}1_{(-4,4)}\left(\frac{\mu(x,y)^2\xi}{2^j}\right)\right),$$

where $\mu(x,y) = \min\{b(x),b(y)\}$. Substituting into the formula of $\mathfrak{S}^j(\mathfrak{S}^j)^*f$, we obtain that for any $g \in L^2(\mathbb{R})$,

$$|\langle \mathfrak{S}^j(\mathfrak{S}^j)^*f, g\rangle| \lesssim (2^{-j/100} + 2^{-j/9}) \cdot \left(\int_{\mathbb{R}} Mf(x)\cdot|g(x)|\mathrm{d}x + \int_{\mathbb{R}} Mg(x)|f(x)|\,\mathrm{d}x\right).$$

By L^2 boundedness of the maximal function, this is less than an absolute constant times $2^{-j/100}\|f\|_2\|g\|_2$. As a consequence, it follows that

$$\|\mathfrak{S}^j(\mathfrak{S}^j)^*f\|_2 \lesssim 2^{-j/100}\|f\|_2.$$

Therefore,

$$\|\mathfrak{S}^j f\|_2 \lesssim 2^{-j/200}\|f\|_2.$$

This finishes the proof of the proposition. □

Part 2: Analysis of $\mathfrak{C}^2$. For this part, we need a slightly more sophisticated approximation to the phase function. We split the interval of integration as

$$[-R,-2] \cup (-2,-1/2] \cup (-1/2,-b(x)^{-\frac{1}{6}}) \cup [-b(x)^{-\frac{1}{6}}, b(x)^{-\frac{1}{6}}] \cup (b(x)^{-\frac{1}{6}}, 1/2) \cup [1/2, R].$$

We define the approximation polynomial

$$P_{b(x)}(t) = b(x)\cdot\left(1 + \frac{t}{2} - \frac{t^2}{4} + \frac{3t^3}{8} - \frac{15t^4}{16} + \frac{105t^5}{32}\right),$$

and note that

$$|b(x)[t+1]^{1/2} - P_{b(x)}(t)| \lesssim b(x)\cdot t^6$$

for $t \in [-1/2, 1/2]$. This follows directly by noting that $P_{b(x)}$ is nothing but the Taylor polynomial of order 5 for the function $\sqrt{t+1}$ around the origin. Using this fact, we compare the integral defining our operator restricted to the middle interval:

$$\left|\int_{-b(x)^{-1/6}}^{b(x)^{-1/6}} f(x-t)e^{iN(x)t}e^{ib(x)[t+1]^{1/2}}\,\frac{\mathrm{d}t}{t} - \int_{-b(x)^{-1/6}}^{b(x)^{-1/6}} f(x-t)e^{iN(x)t}e^{iP_{b(x)}(t)}\,\frac{\mathrm{d}t}{t}\right| \tag{6.9}$$

is bounded by

$$Cb(x)\cdot\int_{-b(x)^{-1/6}}^{b(x)^{-1/6}}|f(x-t)|t^5\,\mathrm{d}t\le Cb(x)^{1/6}\int_{-b(x)^{-1/6}}^{b(x)^{-1/6}}|f(x-t)|\,\mathrm{d}t\lesssim Mf(x).$$

We have to resort to the full version of the polynomial Carleson theorem in [Lie11b] and [ZK17] to bound the second integral in (6.9) in L^p: the operator to which we compared is bounded by a maximally truncated polynomial Carleson operator of degree ≤ 5, which is L^p-bounded by the aforementioned references. As the difference is bounded in L^p, the analysis for $[-b(x)^{-1/6}, b(x)^{-1/6}]$ is complete.

It remains to bound the operators relative to the integral over the "outer" layers. We first begin by analyzing the outermost intervals

$$(-1/2,-b(x)^{-1/6})\cup(b(x)^{-1/6},1/2).$$

As the proof for both of them is essentially the same, we focus on the positive interval $(b(x)^{-1/6}, 1/2)$.

Let ψ_0 be a smooth bump with the properties as in Part 1. Up to a maximal function error, bounding the integral defining $\mathfrak{C}^2$ over the interval $(b(x)^{-1/6}, 1/2)$ is equivalent to bounding

$$\tilde{\mathfrak{C}}f(x):=\int f(x-t)e^{iN(x)t}e^{ib(x)[t+1]^{1/2}}\cdot\phi_{b(x)}(t)\,\frac{\mathrm{d}t}{t},$$

where $\phi_{b(x)}(t)=\phi_{\lfloor\log_2 b(x)\rfloor}(t):=\sum_{j=(2-\frac{1}{6})\lfloor\log_2 b(x)\rfloor}^{2\lfloor\log_2 b(x)\rfloor-3}\psi_0(2^{-j}2^{2\lfloor\log_2 b(x)\rfloor}t)$. This holds by the fact that the smooth cutoff function $\phi_{b(x)}$ approximates well the characteristic function of $(b(x)^{-1/6}, 1/2)$ due to the properties of ψ_0.

Our main goal now is to achieve exponential decay in $\lfloor\log_2 b(x)\rfloor$. This will be enough for our purposes, as the operator $\tilde{C}$ behaves well in the sets where $b\sim 2^k$. Explicitly, we compute:

$$\begin{aligned}|\tilde{\mathfrak{C}}f(x)|&\le\left(\sum_{k\ge3}\left|1_{b_k(x)\in(1,2]}\int f(x-t)e^{iN(x)t}e^{i2^kb_k(x)[t+1]^{1/2}}\phi_k(t)\,\frac{\mathrm{d}t}{t}\right|^p\right)^{1/p}\\&=:\left(\sum_{k\ge3}|\mathfrak{C}_kf(x)|^p\right)^{1/p},\end{aligned}$$

where $b_k(x)=\frac{b(x)}{2^k}$. It suffices then to bound $\|\mathfrak{C}_kf\|_p\lesssim k\cdot 2^{-\alpha_p\cdot k}\|f\|_p$, with $\alpha_p>0$. In order to obtain that bound, we decompose each of the $\mathfrak{C}_k$ further as

$$\mathfrak{C}_kf(x)=\sum_{j=\frac{11k}{6}}^{2k}\mathfrak{C}_k^jf(x),$$

where

$$\mathfrak{C}_k^jf(x):=1_{b_k(x)\in(1,2]}\int f(x-t)e^{iN(x)t}e^{i2^kb_k(x)[t+1]^{1/2}}\psi_0(2^{j-2k}t)\,\frac{\mathrm{d}t}{t}.$$

Our main Proposition to get decay in j for these operators reads as follows.

Proposition 6.10. *There exists $\beta > 0$ such that for all $k \geq 1$ and all $j \in ((2-\frac{1}{6})k, 2k)$,*

$$\|\mathfrak{C}_k^j f\|_2 \lesssim \left(2^{-\beta j} + 2^{\beta\cdot(\frac{7}{3}k - \frac{399}{300}j)}\right) \|f\|_2.$$

It is direct to notice that the pointwise bound

$$|\mathfrak{C}_k^j f(x)| \lesssim Mf(x)$$

holds independently of $j \in [11k/6, 2k]$. This estimate implies automatically the endpoint

$$\|\mathfrak{C}_k^j f\|_\infty \lesssim \|f\|_\infty,\ \|\mathfrak{C}_k^j f\|_{1,\infty} \lesssim \|f\|_1 \tag{6.10}$$

bounds, so that interpolating between Proposition 6.10 and estimate (6.10) gives us the existence of $\beta_p > 0$ such that for all $k \geq 1$ and all $j \in [11k/6, 2k]$,

$$\|\mathfrak{C}_k^j f\|_p \lesssim \left(2^{-\beta_p j} + 2^{\beta_p\cdot(\frac{7}{3}k - \frac{399}{300}j)}\right) \|f\|_p.$$

Proof of Proposition 6.10. In order to get decay on $\mathfrak{C}_k^j f := \Phi^{k,j}_{N(x),b_k(x)} * f(x)$ in L^2, it suffices to bound $\mathfrak{C}_k^j(\mathfrak{C}_k^j)^* f$ instead. A computation gives us that

$$\mathfrak{C}_k^j(\mathfrak{C}_k^j)^* f(x) = 1_{b_k(x)\in(1,2]} \int_{\mathbb{R}} (\Phi^{k,j}_{N(y),b_k(y)} * \tilde{\Phi}^{k,j}_{N(x),b_k(x)})(x-y)(1_{b_k(y)\in(1,2]} f)(y)\,\mathrm{d}y,$$

with $\tilde{\Phi}^{k,j}_{N(\cdot),b_k(\cdot)}(\xi) := \overline{\Phi^{k,j}_{N(\cdot),b_k(\cdot)}}(-\xi)$. The convolution inside this integral is explicitly given by

$$\int_{\mathbb{R}} e^{i(N(x)-N(y))s} \cdot e^{2^k\cdot i(b_k(x)\sqrt{s+1} - b_k(y)\sqrt{s-\xi+1})} \cdot \frac{\psi_0(2^{2k-j}s)}{s} \frac{\psi_0(2^{2k-j}(s-\xi))}{s-\xi}\,\mathrm{d}s,$$

times a modulating factor depending on ξ but not on s. As our goal is to bound the absolute value of this expression, we can safely ignore it. In order to bound this expression, we assume, without loss of generality, that $b_k(y) \leq b_k(x)$. By changing variables $s = 2^{j-2k} b_k(x)^{-2} s'$ and letting $\xi = 2^{j-2k} b_k(y)^{-2}\xi'$, we rewrite it as

$$\frac{b_k(y)^2}{2^{j-2k}} \cdot \int_{\mathbb{R}} e^{i(\tilde{N}(x)-\tilde{N}(y))s'} \cdot e^{i\tilde{R}_{\xi',j,k}(s')} \frac{\psi_0(b_k(x)^{-2}s')}{s'} \frac{\psi_0(b_k(y)^{-2}(hs'-\xi'))}{hs'-\xi'}\,\mathrm{d}s',$$

where

$$\tilde{R}_{\xi',j,k}(s') = 2^{j/2}\cdot\left(\sqrt{s' + 2^{2k-j}b_k(x)^2} - \sqrt{hs' - \xi' + 2^{2k-j}b_k(y)^2}\right),$$

$\tilde{N}$ is a measurable function and $h := \left(\frac{b_k(y)}{b_k(x)}\right)^2 \leq 1$. Notice that the function

$$\frac{\psi_0(b_k(x)^{-2}s')}{s'} \frac{\psi_0(b_k(y)^{-2}(hs'-\xi'))}{hs'-\xi'}$$

is smooth, bounded and supported in $s' \in [1/4, 4]$ with bounded C^3 norm, as $b_k(x), b_k(y) \in (1,2)$. This allows us to focus on the oscillatory nature of the phase.

The next lemma is the tool we need to bound this kernel pointwise.

Lemma 6.11. *Let $\Psi : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be a smooth function supported in $\{(s', \xi') \in \mathbb{R}^2 : s', hs'-\xi' \in [1/4, 4]\}$, for some fixed positive parameter $h \le 1$. It holds that, for all $v, \xi' \in \mathbb{R}$ and $j \in [11k/6, 2k]$,*

$$\left| \int_{\mathbb{R}} e^{ivs'} \cdot e^{i\tilde{R}_{\xi',j,k}(s')} \Psi(\xi', s') \mathrm{d}s' \right| \\ \lesssim \sup_{\xi' \in [-4,4]} \|\Psi(\xi', \cdot)\|_{C^2} \left(1_{[-2^{-j/100}, 2^{-j/100}]}(\xi') + 2^{\frac{7}{3}k - \frac{399}{300}j} 1_{[-4,4]}(\xi') \right), \tag{6.11}$$

where the implicit constant does not depend on $h \in (0, 1]$.

Proof of Lemma 6.11. This is an application of the stationary phase principle:

If $|\xi'| \le 2^{-\frac{1}{100}j}$, we bound the integral by taking the modulus inside, and we get the first summand $1_{[-2^{-\frac{j}{100}}, 2^{-\frac{j}{100}}]}(\xi')$ on the right hand side.

If, on the other hand, $|\xi'| \ge 2^{-\frac{j}{100}}$, then we denote for shortness $\phi(s') = vs' + \tilde{R}_{\xi',j,k}(s')$. We consider the vector

$$Q(s') = \begin{pmatrix} \phi''(s') \\ -\frac{2}{3}\phi'''(s') \end{pmatrix}.$$

Our aim is to prove that the norm $|Q(s')| \gtrsim 2^{\frac{399}{300}j - \frac{7}{3}k}$, as this implies that either the second or the third derivative of the phase $\phi(s')$ have this same, what enables us then to use stationary phase to conclude the proof. For that purpose, we write the vector $Q(s')$ alternatively as

$$\frac{2^{j/2}}{4} \cdot \begin{pmatrix} 1 & 1 \\ (s' + 2^{2k-j} b_k(x)^2)^{-1} & h(hs' - \xi' + 2^{2k-j} b_k(y)^2)^{-1} \end{pmatrix} \cdot V(s') = \frac{2^{j/2}}{4} \cdot M(s') \cdot V(s'),$$

where $V(s') = \begin{pmatrix} (s' + 2^{2k-j} b_k(x)^2)^{-\frac{3}{2}} \\ (hs' - \xi' + 2^{2k-j} b_k(y)^2)^{-\frac{3}{2}} \end{pmatrix}$. By the fact that $s', hs' - \xi' \in [1/4, 4]$, we see that

$$|\det(M(s'))| \gtrsim \frac{|\xi'|}{2^{4k-2j}} \ge 2^{(2-\frac{1}{100})j - 4k}.$$

It is also straightforward to see that the supremum norm $\|M(s')\| \lesssim 1$. By Lemma 6.9, the representation formula for $Q(s')$ and the fact that $|V(s')| \gtrsim 2^{\frac{3j}{2} - 3k}$, it holds that

$$|Q(s')| \gtrsim 2^{j/2} \cdot 2^{(2-\frac{1}{100})j - 4k} \cdot 2^{(\frac{3j}{2} - 3k)} = 2^{(4-\frac{1}{100})j - 7k}.$$

As this implies that either the second or third derivatives of the function $\phi(s')$ above are bounded from below by $2^{\frac{399}{100}j - 7k}$. By the stationary phase principle as stated in [Ste93, Proposition VIII.2], the oscillatory integral

$$\left| \int_{\mathbb{R}} e^{i\phi(s')} \Psi(\xi', s') \, \mathrm{d}s' \right| \lesssim \|\Psi(\xi', \cdot)\|_{C^2} 2^{\frac{7}{3}k - \frac{399}{300}j},$$

whenever $|\xi'| \ge 2^{-j/100}$. This gives us the second summand in the statement of Lemma 6.11, and therefore the result. ☐

In order to finish the proof of the proposition, we notice that the oscillatory kernel given from the convolution defining the kernel of $\mathfrak{C}_k^j(\mathfrak{C}_k^j)^*$ fits the framework of Lemma 6.11. Therefore, by using that $b_k(y), b_k(x) \in (1,2]$, we obtain

$$|\mathfrak{C}_k^j(\mathfrak{C}_k^j)^* f(x)| \lesssim \frac{1}{2^{j-2k}} \int_{-4\cdot 2^{\frac{99j}{100}-2k}}^{4\cdot 2^{\frac{99j}{100}-2k}} |f(x-y)|\,\mathrm{d}y$$
$$+ \frac{2^{\left(\frac{7k}{3}-\frac{399}{300}j\right)}}{2^{j-2k}} \int_{-10\cdot 2^{j-2k}}^{10\cdot 2^{j-2k}} |f(x-y)|\,\mathrm{d}y \lesssim 2^{-j/100} Mf(x) + 2^{\left(\frac{7k}{3}-\frac{399}{300}j\right)} Mf(x).$$

In particular, by boundedness of the maximal function,

$$\|\mathfrak{C}_k^j(\mathfrak{C}_k^j)^* f\|_2 \lesssim \left(2^{-j/100} + 2^{\left(\frac{7k}{3}-\frac{399}{300}j\right)}\right) \|f\|_2.$$

This implies directly that

$$\|\mathfrak{C}_k^j f\|_2 \lesssim \left(2^{-j/200} + 2^{\frac{1}{2}\left(\frac{7k}{3}-\frac{399}{300}j\right)}\right) \|f\|_2.$$

This finishes the proof of the proposition by taking $\beta = \frac{1}{200}$. □

With the proposition in hands, our previous considerations yield that there is $\beta_p > 0$ so that

$$\|\mathfrak{C}_k^j f\|_p \lesssim_p \left(2^{-\beta_p\cdot j} + 2^{\beta_p\cdot\left(\frac{7k}{3}-\frac{399}{300}j\right)}\right) \|f\|_p.$$

Summing for $j \in [11k/6, 2k]$ yields

$$\|\mathfrak{C}_k f\|_p \le \sum_{j=11k/6}^{2k} \|\mathfrak{C}_k^j f\|_p \lesssim_p \sum_{j=11k/6}^{2k} \left(2^{-\beta_p\cdot j} + 2^{\beta_p\cdot\left(\frac{7k}{3}-\frac{399}{300}j\right)}\right) \|f\|_p$$
$$\lesssim_p k\cdot\left(2^{-\frac{11\beta_p}{6}k} + 2^{\beta_p\cdot\left(\frac{7k}{3}-\frac{4389}{1800}k\right)}\right)\|f\|_p \lesssim_p k\cdot 2^{-\alpha_p k}\|f\|_p,$$

for $\alpha_p = \frac{189}{1800}\cdot\beta_p$. This implies, on the other hand, that

$$\|\tilde{\mathfrak{C}} f\|_p \le \left(\sum_{k\ge 3} \|\mathfrak{C}_k f\|_p^p\right)^{1/p} \lesssim \left(\sum_{k\ge 3} k^p\cdot 2^{-p\alpha_p k}\right)^{1/p} \|f\|_p \lesssim \|f\|_p,$$

which concludes the proof of boundedness for the intervals $(-1/2, -b(x)^{-1/6}) \cup (b(x)^{-1/6}, 1/2)$.

We briefly remark on the necessity of a large degree approximation of the Taylor polynomial in the phase. Indeed, redoing the argument above shows that choosing a Taylor polynomial of degree d splits naturally the integration interval as $(-1/2, -b(x)^{-1/(d+1)}) \cup [-b(x)^{-1/(d+1)}, b(x)^{-1/(d+1)}] \cup (b(x)^{-1/(d+1)}, 1/2)$. In the middle interval, the same comparison holds as before, whereas the bounds for each scale in the outer intervals are *not* altered.

That is, we still obtain a $2^{\beta\left(\frac{7k}{3}-\frac{399j}{300}\right)}$ factor, which we wish to decay exponentially with $k > 0$. For that, we must have, necessarily, $j > \frac{7}{4}k$. But from our definitions, in the intervals $(b(x)^{-1/(d+1)}, 1/2)$, we obtain $j - 2k > \frac{-k}{d+1} \iff j > \frac{(2d+1)k}{d+1}$. In order for this last factor to be $> \frac{7}{4}$, we need $d \geq 4$. The choice $d = 5$ comes about in order to relax the tightness in the various steps of the proof.

For the interval $[-2, -1/2]$ where the singularity of the phase $[t+1]^{1/2}$ lies, one notices that the kernel $\frac{1}{t}$ has upper and lower bounds, so that the integral

$$\left| \int_{-2}^{-1/2} f(x-t) e^{iN(x)t} e^{ib(x)[t+1]^{1/2}} \frac{\mathrm{d}t}{t} \right| \lesssim \int_{-2}^{-1/2} |f(x-t)| \, \mathrm{d}t \lesssim Mf(x).$$

We are then left with the outermost intervals $[-R, -2] \cup [1/2, R]$. As the analysis is virtually the same in both cases – and as the phase $b(x) \cdot [t+1]^{1/2}$ does not change sign in either of the intervals –, we focus on the positive interval $[1/2, R]$.

We decompose the operator $\mathfrak{C}^2$ directly this time, without the need to identify the scale of b. Explicitly, we write the integral

$$\int_{1/2}^{R} f(x-t) e^{iN(x)t} e^{ib(x)[t+1]^{1/2}} \frac{\mathrm{d}t}{t},$$

modulo error terms that amount to a constant times the Hardy–Littlewood maximal function of f, as

$$\sum_{j \geq 2\log_2 b(x) - 2}^{\log_2 R} \int_{\mathbb{R}} f(x-t) e^{iN(x)t} e^{ib(x)[t+1]^{1/2}} \psi_0(2^{-j} b(x)^2 t) \frac{\mathrm{d}t}{t} = \sum_{j \geq 1} \mathfrak{T}^j f(x). \tag{6.12}$$

Notice that we encompass the fact that $j \geq 2\log_2 b(x) - 2$ already in the definition of $\mathfrak{T}^j$. We must now only prove exponential decay in j in the L^p norms of $\mathfrak{T}^j f$. The proof of this fact follows essentially the same line as before: after all the reductions, we need to look at an oscillatory integral representing the kernel of $\mathfrak{T}^j (\mathfrak{T}^j)^*$. This is given by a similar oscillatory integral to the one before, i.e.,

$$\frac{b(y)^2}{2^j} \int_{\mathbb{R}} e^{i(\tilde{N}(x) - \tilde{N}(y))s'} \cdot e^{i\tilde{R}_{\xi',j,k}(s')} \frac{\psi_0(s')}{s'} \frac{\psi_0(hs' - \xi')}{hs' - \xi'} \, \mathrm{d}s',$$

where

$$\tilde{\tilde{R}}_{\xi',j,k}(s') = 2^{j/2} \cdot \left(\sqrt{s' + 2^{-j} b(x)^2} - \sqrt{hs' - \xi' + 2^{-j} b(y)^2}\right),$$

$\tilde{N}$ is a measurable function and $h := \left(\frac{b(y)}{b(x)}\right)^2 \leq 1$. What changes now are the estimates we can achieve with the stationary phase method. Now, the vector

$$\tilde{Q}(s') = \begin{pmatrix} \tilde{\phi}''(s') \\ -\frac{2}{3}\tilde{\phi}'''(s') \end{pmatrix} = 2^{j/2} \tilde{M}(s') \cdot \tilde{V}(s')$$

has slightly different properties: it is easy to see that

$$\tilde{M}(s') = \begin{pmatrix} 1 & 1 \\ (s' + 2^{-j} b(x)^2)^{-1} & h(hs' - \xi' + 2^{-j} b(y)^2)^{-1} \end{pmatrix},$$

and therefore $\|\tilde{M}(s')\| \lesssim 1$ still, but now, as $b(x)^2 2^{-j} \lesssim 1$, the determinant bounds change to

$$|\det(\tilde{M}(s'))| \gtrsim |\xi'| \geq 2^{-j/100}.$$

Also, we can only ensure that $|\tilde{V}(s')| \gtrsim 1$. This implies, by Lemma 6.9, that $|\tilde{Q}(s')| \gtrsim 2^{j/3}$. Stationary phase and the considerations as in Lemma 6.11 give the bound

$$\|\mathfrak{T}^j f\|_2 \lesssim 2^{-j/200}\|f\|_2.$$

It is direct to conclude from the definition that $|\mathfrak{T}^j f| \lesssim Mf(x)$, so that interpolation gives the existence of $\theta_p > 0$ so that $\|\mathfrak{T}^j f\|_p \lesssim 2^{-\theta_p j}\|f\|_p$. as $b(x) \geq 10$, we see that $j \geq 1$, so that the L^p norm of the sum in (6.12) is pointwise bounded by

$$\sum_{j\geq 1} 2^{-\theta_p j}\|f\|_p \lesssim_p \|f\|_p.$$

This concludes the analysis of L^p bounds of $\mathfrak{C}^2$ and therefore the proof of Theorem 6.4. □

As mentioned in the introduction, one wonders whether the analysis for the intervals $(-1/2, -b(x)^{-1/6}) \cup (b(x)^{-1/6}, 1/2)$ in the proof of Theorem 6.4 can be suppressed by using a better polynomial approximation. The next proposition proves that employing our approximation technique is impossible without being forced to allow the degree of the polyonomial to depend on b:

Proposition 6.12. *Suppose that, for each $b \geq 1$, we are given a polynomial $P_b(t)$ such that*

$$\int_{-1/2}^{1/2} |b\sqrt{t+1} - P_b(t)| \frac{\mathrm{d}t}{|t|} \leq 1.$$

Then $\lim_{b\to\infty}$ *deg* $(P_b) = +\infty$.

Proof of Proposition 6.12. In order for the integral

$$\int_{-1/2}^{1/2} |b\sqrt{t+1} - P_b(t)| \frac{\mathrm{d}t}{|t|}$$

to be finite, we must have that $P_b(0) = b$. The condition on the polynomials given by the proposition then becomes

$$\int_{-1/2}^{1/2} \left|(\sqrt{t+1} - 1) - \left(\frac{P_b(t)}{b} - 1\right)\right| \frac{\mathrm{d}t}{|t|} \leq \frac{1}{b}.$$

The last inequality reveals that (a) $\frac{P_b(t)}{b} - 1 \to \sqrt{t+1} - 1$ in $L^1([-1/2, 1/2], \frac{\mathrm{d}t}{|t|})$; (b) The sequence $\frac{\left(\frac{P_b(t)}{b} - 1\right)}{t}$ is bounded in $L^1(-1/2, 1/2)$.

Now we suppose that there is an upper bound on the degrees of the polynomials P_b. As the sequence $\frac{\left(\frac{P_b(t)}{b} - 1\right)}{t}$ lies in a finite-dimensional polynomial space, all norms are equivalent. In particular, the sum of coefficients norm is bounded by the $L^1(-1/2, 1/2)$ norm. This and (b) give us that, denoting this norm by $\|\cdot\|_{\text{coeff}}$,

$$\left\|\frac{\left(\frac{P_b(t)}{b} - 1\right)}{t}\right\|_{\text{coeff}} \leq C, \ \forall b \geq 1.$$

But, from (a), we know that $\frac{\left(\frac{P_b(t)}{b}-1\right)}{t} \to (\sqrt{t+1}-1)/t$ in $L^1(-1/2,1/2)$. We then extract a subsequence of $\{b_k\}_k$ so that (a) $\frac{\left(\frac{P_{b_k}(t)}{b_k}-1\right)}{t} \to \frac{\sqrt{t+1}-1}{t}$ *almost everywhere* in $(-1/2,1/2)$; and (b) the coefficients of $\frac{\left(\frac{P_{b_k}(t)}{b_k}-1\right)}{t}$ converge. But this is already a contradiction, for then, as the degree is bounded and coefficients converge, $\frac{\left(\frac{P_{b_k}(t)}{b_k}-1\right)}{t}$ should converge to a polynomial pointwise, which is clearly not the case for $\frac{\sqrt{t+1}-1}{t}$. $\square$

6.4 Comments and remarks

6.4.1 Question 6.1 and Theorem 6.2

As briefly sketched in the introduction, if the answer to Question 6.1 is *affirmative*, then Theorem 6.2 holds. We explain this relationship in greater detail here. As discussed before, we let ℓ be a line in $\mathbb{R}^2$. Without loss of generality, we suppose it is given by an equation of the form $(a\eta+b,\eta)$, as the remaining case of having equation (η, C), C constant, is completely analogous. Let $\mathcal{C}_{[\ell]} = \sup_{b\in\mathbb{R}} \mathcal{C}_{a,b}$ be the operator associated to the equivalence class of lines $\ell' \sim \ell$, where two lines are equivalent if they have the same slope.

We fix $g \in \mathcal{S}(\mathbb{R})$ and write $\xi = \theta \cdot v_\ell + t \cdot v_\ell^\perp$, $(\theta,t) \in \mathbb{R}^2$ and $v_\ell = (a,1)$. If we fix $N_{(\cdot,\cdot)} : \mathbb{R}^2 \to \mathbb{R}$ a measurable function and F_δ such that $\widehat{F_\delta}(\theta\cdot v_\ell + t\cdot v_\ell^\perp) = \widehat{g}(\theta)\psi_\delta(t)$, we have

$$\mathcal{C}_2(F_\delta)(z) \geq \left| \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \widehat{g}(\theta) e^{2\pi i\theta\langle z, v_\ell\rangle} m_2(\theta\cdot v_\ell + t\cdot v_\ell^\perp + N_z)\,\mathrm{d}\theta \right) e^{2\pi i t\langle z, v_\ell^\perp\rangle} \psi_\delta(t)\,\mathrm{d}t \right|.$$

Choose $\psi_\delta = \frac{1}{\delta^{1/2}}\varphi\left(\frac{x}{\delta}\right)$, with $1_{[-1,1]} \leq \varphi \leq 1_{[-2,2]}$ smooth. Let

$$G_{[\ell]}(z,t) = \int_{\mathbb{R}} \widehat{g}(\theta) e^{2\pi i\theta\langle z,v_\ell\rangle} m_2(\theta\cdot v_\ell + t\cdot v_\ell^\perp + N_z)\,\mathrm{d}\theta.$$

Since $g \in \mathcal{S}(\mathbb{R})$, it is easy to see by the dominated convergence theorem that $G_{[\ell]}(z,t) \to G_{[\ell]}(z) := \int_{\mathbb{R}} \widehat{g}(\theta) e^{2\pi i\theta\langle z,v_\ell\rangle} m(\theta\cdot v_\ell + N_z)\,\mathrm{d}\theta$ pointwise, as m_2 is bounded and continuous in $\mathbb{R}^2$. By choosing N suitably, we can make $|G_{[\ell]}(z)| \geq \frac{1}{2} C_{[\ell]} g(\langle z, v_\ell\rangle)$, $\forall z \in \mathbb{R}^2$. Reasoning again with dominated convergence gives

$$\left| \int_{\mathbb{R}} G_{[\ell]}(z,t) e^{2\pi i t\langle z,v_\ell^\perp\rangle} \frac{\psi_\delta(t)}{\delta^{1/2}}\,\mathrm{d}t - \int_{\mathbb{R}} G_{[\ell]}(z) e^{2\pi i t\langle z,v_\ell^\perp\rangle} \frac{\psi_\delta(t)}{\delta^{1/2}}\,\mathrm{d}t \right| \to 0$$

as $\delta \to 0$. Moreover, each of the integrals above is bounded as a function of z. These considerations imply that, for $R > 0$ fixed,

$$\|\mathcal{C}_2(F_\delta)\|_{L^2(B_R)} + o_\delta^R(1) \geq \frac{1}{2} \|\delta^{1/2}\widehat{\varphi}(\delta\langle\cdot, v_\ell^\perp\rangle) C_{[\ell]} g(\langle\cdot, v_\ell\rangle)\|_{L^2(B_R)} \geq \frac{1}{5}\|C_{[\ell]}g\|_{L^2(-\frac{R}{2},\frac{R}{2})}. \tag{6.13}$$

We use here $o_\delta^R(1)$ to denote a quantity that goes to 0 as $\delta \to 0$, with R fixed. But, assuming the answer of Question 6.1 to be affirmative,

$$\|\mathcal{C}_2(F_\delta)\|_{L^2(B_R)} \leq C\|F_\delta\|_{L^2(\mathbb{R}^2)} = C\|\widehat{F_\delta}\|_{L^2(\mathbb{R}^2)}.$$

By using the explicit representation of F_δ and the choice of ψ, we get that the right hand side converges to $C\|g\|_{L^2(\mathbb{R})}$ as $\delta \to 0$. Putting together, we have

$$\|C_{[\ell]}g\|_{L^2\left(-\frac{R}{2},\frac{R}{2}\right)} \leq 5C \cdot \|g\|_{L^2(\mathbb{R})}.$$

Notice that C is independent of both $R, [\ell]$. By taking $R \to \infty$ in this last inequality one obtains Theorem 6.2.

6.4.2 Hilbert transform along more general curves

Throughout this article, we have investigated the case of the Hilbert transform along the parabola (t, t^2). There is, however, no reason not to consider more general *monomial* curves of the form (t, t^m). For those, it is natural to expect that the reductions performed in Section 6.2 carry through, and that effectively one needs to bound the operator

$$f \mapsto \sup_{N,b} \left| \int_{-1/2}^{1/2} f(x-t) e^{iNt} \cdot e^{ib[t+1]^{1/m}} \frac{\mathrm{d}t}{t} \right|,$$

where $[u]^r$ represents either $|u|^r$ or $\operatorname{sign}(u)|u|^r$. The proof in Section 6.3 is not particular to the $m = 2$ case, and therefore can be adapted to prove that these operators are bounded in L^p. The reduction to these operators is not as direct as the quadratic case, though. For the case of higher degrees, one would have to use a form of decomposition as in the recent article by Guo [Guo17]. Following this idea, it should be possible to exploit the *polynomial* case $(t, Q(t))$, $Q \in \operatorname{Poly}(\mathbb{R}\colon \mathbb{R})$. In order to keep the exposition short, we do not investigate these questions further.

6.4.3 Oscillatory integrals and the proof of Theorem 6.2

The proof of boundedness of the operators $\mathfrak{C}^R$ highlights what seems to be a general principle: if we are given a function $\eta : \mathbb{R} \to \mathbb{R}$ whose derivative is singular in a neighbourhood of the origin, but sufficiently regular (together with its higher degree derivatives) everywhere else, then the maximal function

$$f \mapsto \sup_{\substack{P \in \mathcal{P}_d \\ N \in \mathbb{R}}} \left| \int_{\mathbb{R}} f(x-t) e^{iP(t) + iN\cdot\eta(t+1)} \frac{\mathrm{d}t}{t} \right|, \tag{6.14}$$

where $\mathcal{P}_d$ is the space of polynomials of degrees $\leq d$, should be bounded in L^2. In this article, we have explored the case $d = 1$, where $\eta(t) = |t|^{1/2}$ or $\eta(t) = \operatorname{sign}(t)|t|^{1/2}$. We notice, however, that by taking further derivatives of the phase, defining approximation polynomials with higher degrees and running the basic strategy we set here, there should not stand any barrier to prove the case of general $d > 1$. This suggests the existence of an underlying principle for a more general class of functions η whose decay are sufficiently controllable. We currently believe this principle is intimately related to Question 6.1.

6.5 Acknowledgements

The author is indebted to his doctoral advisor Prof. Dr. Christoph Thiele for having suggested investigating Theorem 6.2. He also thanks Joris Roos, for helpful discussions, in both early and late stages of this project, and Pavel Zorin-Kranich for discussions on the main results of this manuscript and insights on his article [ZK17]. Finally, the author ackowledges finantial support by the Deutscher Akademischer Austauschdienst (DAAD).

Bibliography

[AB09] Artur Avila and Jairo Bochi. On the subadditive ergodic theorem. 01 2009.

[AdR] Juan Arias-de Reyna. *Pointwise convergence of Fourier series*, volume 1785 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York.

[AdR02] Juan Arias-de Reyna. Pointwise convergence of fourier series. *Journal of the London Mathematical Society*, 65(1):139–153, 2002.

[AK10] Daniel Aalto and Juha Kinnunen. The discrete maximal operator in metric spaces. *J. Anal. Math.*, 111(1):369–390, 2010.

[APL06] J Aldaz and Javier Pérez Lázaro. Functions of bounded variation, the derivative of the one dimensional maximal function, and applications to inequalities. *Transactions of the American Mathematical Society*, 359, 01 2006.

[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.

[BD15] Jean Bourgain and Ciprian Demeter. The proof of the l^2 decoupling conjecture. *Ann. of Math. (2)*, 182(1):351–389, 2015.

[BHS] David Beltran, Jonathan Hickman, and Christopher D. Sogge. Wolff-type inequalities and sharp local smoothing estimates for wave equations on manifolds. `arxiv.org/abs/1801.06910` (2018).

[Bir31] George D Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931.

[BL76] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction.* Springer-Verlag, Berlin-New York, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[BM19] David Beltran and José Madrid. Endpoint sobolev continuity of the fractional maximal function in higher dimensions. *arXiv preprint arXiv:1906.00496*, 2019.

[Bou86] J. Bourgain. Averages in the plane over convex curves and maximal operators. *J. Analyse Math.*, 47:69–85, 1986.

[Bou13] Jean Bourgain. On the schrödinger maximal function in higher dimension. *Proceedings of the Steklov Institute of Mathematics*, 280(1):46–60, 2013.

[BPY15] L B. Pierce and Po-Lam Yung. A polynomial carleson operator along the paraboloid. *Revista Matemática Iberoamericana*, 05 2015.

[BRS19] David Beltran, João Pedro Ramos, and Olli Saari. Regularity of fractional maximal functions through fourier multipliers. *Journal of Functional Analysis*, 276(6):1875 – 1892, 2019.

[Car66] Lennart Carleson. On convergence and growth of partial sums of fourier series. *Acta Math.*, 116:135–157, 1966.

[Car80] Lennart Carleson. Some analytic problems related to statistical mechanics. In John J. Benedetto, editor, *Euclidean Harmonic Analysis*, pages 5–45, Berlin, Heidelberg, 1980. Springer Berlin Heidelberg.

[CDDD03] Albert Cohen, Wolfgang Dahmen, Ingrid Daubechies, and Ronald DeVore. Harmonic analysis of the space BV. *Rev. Mat. Iberoamericana*, 19(1):235–263, 2003.

[CFS15] Emanuel Carneiro, Renan Finder, and Mateus Sousa. On the variation of maximal operators of convolution type ii. *Revista Matematica Iberoamericana*, 34:741–768, 12 2015.

[CGR19] Emanuel Carneiro and Cristian González-Riquelme. Gradient bounds for radial maximal functions. *arXiv preprint arXiv:1906.01487*, 2019.

[Chr85] Michael Christ. Hilbert transforms along curves: I. nilpotent groups. *Annals of Mathematics*, 122(3):575–596, 1985.

[CLS19] E Compaan, R Lucà, and G Staffilani. Pointwise convergence of the schr\" odinger flow. *arXiv preprint arXiv:1907.11192*, 2019.

[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. *Journal of Functional Analysis*, 273(10):3262–3294, 2017.

[CS72] Lennart Carleson and Per Sjölin. Oscillatory integrals and multiplier problem for the disc. *Studia Mathematica*, 44(3):287–299, 1972.

[CS87] Michael Christ and Elias M. Stein. A remark on singular calderón-zygmund theory. *Proceedings of the American Mathematical Society*, 99(1):71–75, 1987.

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

[DGL17] Xiumin Du, Larry Guth, and Xiaochun Li. A sharp schrödinger maximal estimate in $\mathbb{R}^2$. *Annals of Mathematics*, pages 607–640, 2017.

[Duo01] Javier Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe.

[DZ19] Xiumin Du and Ruixiang Zhang. Sharp L^2 estimates of the schrödinger maximal function in higher dimensions. *Annals of Mathematics*, 189(3):837–861, 2019.

[EG92] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[Fab66] Rivière N. Fabes, E. Singular integrals with mixed homogeneity. *Studia Mathematica*, 27(1):19–38, 1966.

[Fed69] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[Fef70] Charles Fefferman. Inequalities for strongly singular convolution operators. *Acta Math.*, 124:9–36, 1970.

[Fef73] Charles Fefferman. Pointwise convergence of fourier series. *Annals of Mathematics*, 98(3):551–571, 1973.

[Fos15] Damiano Foschi. Global maximizers for the sphere adjoint fourier restriction inequality. *J. Funct. Anal.*, 268(3):690–702, 2015.

[FZ72] Herbert Federer and William Ziemer. The lebesgue set of a function whose distribution derivatives are p-th power summable. *Indiana University Mathematics Journal*, 22, 01 1972.

[GHLR17] Shaoming Guo, Jonathan Hickman, Victor Lie, and Joris Roos. Maximal operators and hilbert transforms along variable non-flat homogeneous curves. *Proceedings of the London Mathematical Society*, 115(1):177–219, 2017.

[GPRY17] Shaoming Guo, Lillian B Pierce, Joris Roos, and Po-Lam Yung. Polynomial carleson operators along monomial curves in the plane. *The Journal of Geometric Analysis*, 27(4):2977–3012, 2017.

[Gra14a] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[Gra14b] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[GS09] Gustavo Garrigós and Andreas Seeger. On plate decompositions of cone multipliers. *Proc. Edinb. Math. Soc. (2)*, 52(3):631–651, 2009.

[GS10] Gustavo Garrigós and Andreas Seeger. A mixed norm variant of Wolff's inequality for paraboloids. In *Harmonic analysis and partial differential equations*, volume 505 of *Contemp. Math.*, pages 179–197. Amer. Math. Soc., Providence, RI, 2010.

[Guo17] Shaoming Guo. A remark on oscillatory integrals associated with fewnomials. *New York J. Math*, 23:1733–1738, 2017.

[Gut14] Larry Guth. A restriction estimate using polynomial partitioning. *Journal of the American Mathematical Society*, 29, 07 2014.

[Gut18] Larry Guth. Restriction estimates using polynomial partitioning ii. *Acta Math.*, 221(1):81–142, 09 2018.

[Hed72] Lars Inge Hedberg. On certain convolution inequalities. *Proceedings of the American Mathematical Society*, 36(2):505–510, 1972.

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

[HKNT13] Toni Heikkinen, Juha Kinnunen, Juho Nuutinen, and Heli Tuominen. Mapping properties of the discrete fractional maximal function in metric measure spaces. *Kyoto J. Math.*, 53(3):693–712, 2013.

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

[HNS11] Yaryong Heo, Fëdor Nazarov, and Andreas Seeger. Radial Fourier multipliers in high dimensions. *Acta Math.*, 206(1):55–92, 2011.

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

[HR18] Jonathan Hickman and Keith M Rogers. Improved fourier restriction estimates in higher dimensions. *arXiv preprint arXiv:1807.10940*, 2018.

[Hun68] R. Hunt. On the convergence of fourier series. *Orthogonal Expansions and their Contin-uous Analogues (Proceedings of Conference, Edwardsville,Illinois, 1967)*, pages 235–255, 1968.

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

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

[KOeS18] Vjekoslav Kovač and Diogo Oliveira e Silva. A variational restriction theorem. *arXiv preprint arXiv:1809.09611*, 2018.

[Kol23] Andrey Kolmogorov. Une série de fourier-lebesgue divergente presque partout. *Fundamenta Mathematicae*, 4(1):324–328, 1923.

[Kov19] Vjekoslav Kovač. Fourier restriction implies maximal and variational fourier restriction. *Journal of Functional Analysis*, 2019.

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

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

[KW82] Yitzhak Katznelson and Benjamin Weiss. A simple proof of some ergodic theorems. *Israel Journal of Mathematics*, 42(4):291–296, Dec 1982.

[Lac04] Michael T Lacey. Carleson's theorem: Proof, complements, variations. *Publicacions matematiques*, pages 251–307, 2004.

[Lee03] Sanghyuk Lee. Endpoint estimates for the circular maximal function. *Proc. Amer. Math. Soc.*, 131(5):1433–1442, 2003.

[Lie09] Victor Lie. The (weak-L^2) boundedness of the quadratic carleson operator. *Geometric and Functional Analysis*, 19(2):457–497, Sep 2009.

[Lie11a] Victor Lie. On the boundedness of the carleson operator near l^1. *Revista Matemática Iberoamericana*, 29, 12 2011.

[Lie11b] Victor Lie. The polynomial carleson operator. *arXiv preprint arXiv:1105.4504*, 2011.

[LM17] Hannes Luiro and José Madrid. The Variation of the Fractional Maximal Function of a Radial Function. *International Mathematics Research Notices*, 11 2017.

[LS13] Sanghyuk Lee and Andreas Seeger. Lebesgue space estimates for a class of Fourier integral operators associated with wave propagation. *Math. Nachr.*, 286(7):743–755, 2013.

[LT00] Michael Lacey and Christoph Thiele. A proof of boundedness of the carleson operator. *Mathematical Research Letters*, 7(4):361–370, 2000.

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

[Lui18] Hannes Luiro. The variation of the maximal function of a radial function. *Ark. Mat.*, 56(1):147–161, 04 2018.

[Luz18] Nuno Luzia. A simple proof of the strong law of large numbers with rates. *Bulletin of the Australian Mathematical Society*, 97(3):513–517, 2018.

[ŁW02] Izabella Łaba and Thomas Wolff. A local smoothing estimate in higher dimensions. *J. Anal. Math.*, 88:149–171, 2002. Dedicated to the memory of Tom Wolff.

[Mel03] Antonios D. Melas. The best constant for the centered hardy-littlewood maximal inequality. *Annals of Mathematics*, 157(2):647–688, 2003.

[MP15] José Madrid Padilla. Sharp inequalities for the variation of the discrete maximal function. *Bulletin of the Australian Mathematical Society*, 95, 12 2015.

[MRW19] Detlef Müller, Fulvio Ricci, and James Wright. A maximal restriction theorem and lebesgue points of functions in $\mathcal{F}(L^p)$. *Revista Matemática Iberoamericana*, 4 2019.

[MSS92] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge. Wave front sets, local smoothing and Bourgain's circular maximal theorem. *Ann. of Math. (2)*, 136(1):207–218, 1992.

[MSW01] Elias M. Stein and Stephen Wainger. Oscillatory integrals related to carleson's theorem. *Mathematical Research Letters*, 8, 09 2001.

[MW74] Benjamin Muckenhoupt and Richard L. Wheeden. Weighted norm inequalities for fractional integrals. *Transactions of the American Mathematical Society*, 192:261–274, 1974.

[NRW74] Alexander Nagel, Nestor Rivière, and Stephen Wainger. On hilbert transforms along curves. *Bull. Amer. Math. Soc.*, 80(1):106–108, 01 1974.

[NRW76] Alexander Nagel, Nestor M. Riviere, and Stephen Wainger. On hilbert transforms along curves. ii. *American Journal of Mathematics*, 98(2):395–403, 1976.

[PPSS18] Carlos Pérez, Tiago Picon, Olli Saari, and Mateus Sousa. Regularity of maximal functions on hardy–sobolev spaces. *Bulletin of the London Mathematical Society*, 50(6):1007–1015, 2018.

[Ram17] João Pedro Ramos. Sharp total variation results for maximal functions. *Annales Academiae Scientiarum Fennicae Mathematica*, 44, 03 2017.

[Ram18] João PG Ramos. Maximal restriction estimates and the maximal function of the fourier transform. *arXiv preprint arXiv:1810.06971, to appear in the Procedings of the American Mathematical Society*, 2018.

[Ram19a] João PG Ramos. The hilbert transform along the parabola, the polynomial carleson theorem and oscillatory singular integrals. *arXiv preprint arXiv:1908.01833*, 2019.

[Ram19b] João PG Ramos. Low-dimensional maximal restriction principles for the fourier transform. *arXiv preprint arXiv:1904.10858*, 2019.

[RdF86] José L. Rubio de Francia. Maximal functions and Fourier transforms. *Duke Math. J.*, 53(2):395–404, 1986.

[Rie32] Frédéric Riesz. Sur un Théorème de Maximum de Mm. Hardy et Littlewood. *Journal of the London Mathematical Society*, s1-7(1):10–13, 01 1932.

[Roo17] Joris Roos. Bounds for anisotropic carleson operators. *Journal of Fourier Analysis and Applications*, pages 1–32, 2017.

[RSW] João Pedro Ramos, Olli Saari, and Julian Weigt. Weak differentiability for fractional maximal function of general l^p functions on domains.

[Sch97] W. Schlag. A generalization of Bourgain's circular maximal theorem. *J. Amer. Math. Soc.*, 10(1):103–122, 1997.

[Sjö74] Per Sjölin. Fourier multipliers and estimates of the fourier transform of measures carried by smooth curves in r. *Studia Mathematica*, 51:169–182, 1974.

[SMS85] Christopher Sogge and Elias M. Stein. Averages of functions over hypersurfaces in n. *Inventiones mathematicae*, 82:543–556, 10 1985.

[Sog91] Christopher D. Sogge. Propagation of singularities and maximal functions in the plane. *Invent. Math.*, 104(2):349–376, 1991.

[SS97] Wilhelm Schlag and Christopher D. Sogge. Local smoothing estimates related to the circular maximal theorem. *Math. Res. Lett.*, 4(1):1–15, 1997.

[Ste61] E. M. Stein. On limits of sequences of operators. *Annals of Mathematics*, 74(1):140–170, 1961.

[Ste76] Elias M. Stein. Maximal functions. I. Spherical means. *Proc. Nat. Acad. Sci. U.S.A.*, 73(7):2174–2175, 1976.

[Ste93] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[Ste95] Elias M. Stein. Oscillatory integrals related to radon - like transforms. 1995.

[Str77] Robert S. Strichartz. Restrictions of fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Math. J.*, 44(3):705–714, 09 1977.

[STW04] Andreas Seeger, Terence Tao, and James Wright. Singular maximal functions and radon transforms near l. *American Journal of Mathematics*, 126(3):607–647, 2004.

[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.

[Tao04] Terence Tao. *Some Recent Progress on the Restriction Conjecture*, pages 217–243. Birkhäuser Boston, Boston, MA, 2004.

[Tem13] Faruk Temur. On regularity of the discrete hardy-littlewood maximal function. 03 2013.

[Tom75] Peter A. Tomas. A restriction theorem for the fourier transform. *Bull. Amer. Math. Soc.*, 81(2):477–478, 03 1975.

[Veg88] Luis Vega. Schrödinger equations: pointwise convergence to the initial data. *Proceedings of the American Mathematical Society*, 102(4):874–878, 1988.

[Vit17] Marco Vitturi. A note on maximal fourier restriction for spheres in all dimensions. *arXiv preprint arXiv:1703.09495*, 2017.

[Wan18] Hong Wang. A restriction estimate in $\mathbb{R}^3$ using brooms. *arXiv preprint arXiv:1802.04312*, 2018.

[Wol00] T. Wolff. Local smoothing type estimates on L^p for large p. *Geom. Funct. Anal.*, 10(5):1237–1288, 2000.

[ZK17] Pavel Zorin-Kranich. Maximal polynomial modulations of singular integrals. *arXiv preprint arXiv:1711.03524*, 2017.

[Zyg74] A. Zygmund. On fourier coefficients and transforms of functions of two variables. *Studia Mathematica*, 50(2):189–201, 1974.