Geometric control of oscillatory integrals

by

David Beltran Portalés

A thesis submitted to
The University of Birmingham
for the degree of
Doctor of Philosophy

School of Mathematics
The University of Birmingham
MAY 2017

Abstract

The aim of this thesis is to provide a geometric control of certain oscillatory integral operators. In particular, if T is an oscillatory Fourier multiplier, a pseudodifferential operator associated to a symbol $a \in S^m_{\rho,\delta}$ or a Carleson-like operator, we obtain a weighted L^2 inequality of the type

$$\int |Tf|^2 w \leqslant C \int |f|^2 \mathcal{M}_T w.$$

Here C is a constant independent of the weight function w, and the operator $\mathcal{M}_T$, which depends on the corresponding T, has an explicit geometric character. In the case of oscillatory Fourier multipliers and of Carleson-like operators we also determine auxiliary geometric operators g_1 and g_2 and establish a *pointwise* estimate of the type

$$g_1(Tf)(x) \leqslant C g_2(f)(x).$$

Finally, we include a careful study of a method developed by Bourgain and Guth in Fourier restriction theory, that allows making progress on the Fourier restriction conjecture from their conjectured multilinear counterparts. Our conjectured progress via multilinear estimates has been recently obtained by Guth.

Acknowledgements

I am indebted to my supervisor Jon Bennett for his continuous guidance and support over the last few years. Besides suggesting me very interesting problems, Jon has always encouraged me to follow my own questions, helping me to become a more mature and independent mathematician. Without his wisdom, enthusiasm and valuable conversations, this thesis would have never been possible. I am also extremely grateful to him for giving me the opportunity to attend several international conferences and, of course, for the financial support associated with his project, funded by the European Research Council.

I also would like to thank all the current and former members of the analysis group in Birmingham. It has been a great pleasure and a very enlightening experience to be part of this young, energetic and live group.

Of course, I would like to express my gratitude to all my friends, the old and the more recent ones, for being around and bringing moments of great fun.

Finally, I couldn't finish without mentioning my parents, brother and family for their unconditional support.

Contents

Notation

We typically use the letter C to denote a constant, that may change from line to line, and whose dependence on the relevant parameters will be specified when necessary. We shall write $A \lesssim B$ if there exists a constant C such that $A \leqslant CB$. The relations $A \gtrsim B$ and $A \sim B$ are defined similarly.

Given a cube $Q \subset \mathbb{R}^d$ and $k \in \mathbb{N}$, we denote by kQ the concentric cube whose sidelength is k times that of Q. In the case $k = 2$, we write $\bar{Q}$ instead of $2Q$.

Let (X, μ) be a measure space. Given a set $E \subset X$, we denote by $\mu(E)$ the measure of E, which in the case of the Lebesgue measure we denote by $|E|$. We say that a property holds *almost everywhere* in a set X, and we use the notation $a.e.\, x \in X$, if it holds except for subsets of X of measure zero.

For $1 \leqslant p < \infty$, we define $L^p(X, \mu)$ as the space of measurable functions $f : X \to \mathbb{C}$ such that

$$\|f\|_p := \Big(\int_X |f|^p d\mu \Big)^{1/p} < \infty.$$

The space $L^\infty(X, \mu)$ corresponds to those functions satisfying

$$\|f\|_\infty := \sup\{C \geqslant 0 : \mu(\{x \in X : |f(x)| > C\}) > 0\} < \infty.$$

Given a fixed p, its conjugate exponent p' is defined by the relation $\frac{1}{p} + \frac{1}{p'} = 1$.

We define the weak-L^p spaces $L^{p,\infty}(X, \mu)$ as the space of measurable functions f :

$X \to \mathbb{C}$ such that

$$\|f\|_{p,\infty} := \sup_{\lambda>0} \mu(\{x \in X : |f(x)| > \lambda\})\lambda^p < \infty.$$

Observe that $L^p(X,\mu) \subset L^{p,\infty}(X,\mu)$. In this thesis, X will typically be $\mathbb{R}^d$, and when $d\mu = dx$ is the Lebesgue measure, we use the notation $L^p(\mathbb{R}^d)$ or simply L^p. For a weight function w, that is, a nonnegative locally integrable function, and $d\mu = wdx$, we use the notation $L^p(w)$.

We denote by M the Hardy–Littlewood maximal function, defined by

$$Mf(x) = \sup_{B \ni x} \frac{1}{|B|} \int_B |f(y)|dy,$$

where B is a ball in $\mathbb{R}^d$ containing the point x.

Given a multi-index $\gamma = (\gamma_1, \ldots, \gamma_d) \in \mathbb{N}^d$ and a function $f : \mathbb{R}^d \to \mathbb{C}$, we write $x^\gamma = x_1^{\gamma_1} \cdots x_d^{\gamma_d}$ and

$$D^\gamma f(x) = \partial_x f(x) = \frac{\partial^{|\gamma|} f}{\partial x_1^{\gamma_1} \cdots \partial x_d^{\gamma_d}},$$

where $|\gamma| = \gamma_1 + \cdots + \gamma_d$. Given $x \in \mathbb{R}$, we denote by $\lfloor x \rfloor$ its integer part.

We say function f belongs to the Schwartz class $\mathcal{S}(\mathbb{R}^d)$ if $f \in C^\infty(\mathbb{R}^d)$ and

$$\sup_{x \in \mathbb{R}^d} |x^\alpha D^\beta f(x)| < \infty$$

for all $\alpha, \beta \in \mathbb{N}^d$.

The Fourier transform of a function $f \in \mathcal{S}(\mathbb{R}^d)$ is defined as

$$\widehat{f}(\xi) := \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)dx.$$

Introduction

This thesis has its origins in a long-standing conjecture of Stein for the disc multiplier. In 1978, at the Williamstown conference in Harmonic Analysis, Stein [125] suggested the possibility that a two-weight inequality of the type

$$\int_{\mathbb{R}^d} |Tf|^2 w \leqslant C \int_{\mathbb{R}^d} |f|^2 \mathcal{M}w \tag{†}$$

might hold for any weight function w with constant C independent of w, where T denotes the disc multiplier, that is $\widehat{Tf} = \chi_{B(0,1)}\widehat{f}$, and $\mathcal{M}$ is a variant of the universal maximal function

$$\mathcal{N}w(x) := \sup_{R \ni x} \frac{1}{|R|} \int_R w;$$

here the supremum is taken over all arbitrary rectangles in $\mathbb{R}^d$ containing the point x. This conjecture had as supporting evidence the connection made by Fefferman [50] between the disc multiplier and Besicovitch sets, which allowed him to prove that the disc multiplier is unbounded on $L^p(\mathbb{R}^d)$ for $p \neq 2$ when $d \geqslant 2$; we note that for $d = 1$ the study of the disc multiplier reduces to that of the Hilbert transform. This question (†), which was also raised by Córdoba [34] in the more general context of Bochner–Riesz multipliers, is still very much open. Positive results were obtained in the case of radial weights by Carbery, Romera and Soria [21], and numerous authors have contributed with partial progress [19, 29, 23, 8, 82]. If true, such a conjecture would be striking, as it involves control via a weighted inequality of a highly oscillatory and cancellative operator by a positive maximal function.

Motivated by the above conjecture of Stein, Bennett, Carbery, Soria and Vargas [8] established a version of this conjecture on the circle. That work was followed by that of Bennett and Harrison [10], who studied weighted L^2 inequalities for certain oscillatory kernels on the real line. Later, Bennett [7] took a Fourier multiplier perspective on such questions on the real line. In all cases, the authors managed to control those oscillatory operators by positive, geometrically-defined maximal functions.

One of the main results of this thesis is a higher dimensional version of the result in [7], for broader classes of oscillatory Fourier multipliers. This is the content of Chapter 2, which is based on the joint work with Bennett in [5]. The classes of multipliers under study are modelled by $m_{\alpha,\beta} := \frac{e^{i|\xi|^\alpha}}{(1+|\xi|^2)^{\beta/2}}$ for any $\alpha, \beta \in \mathbb{R}$. As in the one-dimensional case, the controlling maximal functions are positive operators and involve fractional averages over certain approach regions. Also, the maximal functions are closely related to certain Kakeya-type maximal operators, very much in the spirit of Stein's conjecture.

Our weighted L^2 inequalities follow from a stronger *pointwise* result. In particular, we are able to identify two auxiliary operators, g_1 and g_2, such that the estimate

$$g_1(T_m f)(x) \leqslant C g_2(f)(x) \tag{$\star$}$$

holds, where T_m denotes the operator associated to the multiplier m. A weighted estimate of the type (†) for T_m may be then obtained from those for the auxiliary operators g_1 and g_2. In our case, g_1 and g_2 are novel square functions of Littlewood–Paley type that reflect on the geometric properties of the multipliers under study. We remark that our results on the multipliers $m_{\alpha,\beta}$ have obvious interpretations in the setting of oscillatory convolution kernels and dispersive partial differential equations.

A classical non-translation-invariant generalisation of the Fourier multipliers is given by the pseudodifferential operators. Given a smooth function $a \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$, referred to as the *symbol*, define the associated pseudo-differential operator T_a by

$$T_a f(x) := \int_{\mathbb{R}^d} e^{ix\cdot\xi} a(x,\xi) \widehat{f}(\xi) d\xi,$$

where $f \in \mathcal{S}$. We focus ourselves in the symbol classes $S^m_{\rho,\delta}$, introduced by Hörmander in [69]. We say that $a \in S^m_{\rho,\delta}$ if it satisfies the differential inequalities

$$|\partial_x^\nu \partial_\xi^\sigma a(x,\xi)| \lesssim (1+|\xi|)^{m-\rho|\sigma|+\delta|\nu|}$$

for all multi-indices $\nu, \sigma \in \mathbb{N}^d$, where $m \in \mathbb{R}$ and $0 \leqslant \delta, \rho \leqslant 1$. Observe that the model oscillatory multipliers $m_{\alpha,\beta} \in S^{-\beta}_{1-\alpha,0}$ for $0 \leqslant \alpha \leqslant 1$. Thus, the symbol classes $S^m_{\rho,0}$ constitute a generalisation of the classes of multipliers studied in Chapter 2 for $0 \leqslant \rho \leqslant 1$.

In Chapter 3 we study how to extend the techniques presented for the multiplier case to this pseudodifferential operator context. With additional appropriate applications of the symbolic calculus and the Cotlar–Stein almost orthogonality principle, we are able to control the operators T_a, where $a \in S^m_{\rho,\delta}$, by maximal operators via weighted L^2 inequalities of the type (†). This constitutes the second main result of this thesis, which may be found in [3]. In contrast to the multiplier case, our proof does not follow from a *pointwise* estimate of the type ($\star$). The question of obtaining pointwise control remains open, except for the case $m \leqslant d(\rho-1)/2$, where techniques closer to Calderón–Zygmund theory may be applied.

We remark that the weighted estimates of the type (†) obtained for the oscillatory Fourier multipliers and the Hörmander symbol classes allow one to recover the optimal Lebesgue space bounds for such objects via the appropriate bounds on the controlling maximal function $\mathcal{M}$ in each case.

In Chapter 4, we address the question of obtaining pointwise and weighted control for the Carleson operator, a crucial operator in harmonic analysis related to the almost everywhere convergence of Fourier series. This is motivated by a future line of investigation, which consists in obtaining control for maximal multiplier operators. Given a multiplier m and writing $m_t(\xi) := m(t\xi)$ for any $t > 0$, we define its maximal multiplier operator as

$$T_m^* f(x) := \sup_{t>0} |(m_t \widehat{f})^{\vee}(x)|.$$

Obtaining control for the operator T_m^* in the context of the multipliers $m_{\alpha,\beta}$ would provide control for a central operator in partial differential equations such as the maximal Schrödinger operator. A first attempt towards answering this general question is to study what happens when considering easier families of multipliers. If, for instance, one considers a multiplier m of global bounded variation on $\mathbb{R}$, it is easy to observe that $T_m^* f(x) \lesssim |f(x)| + |\mathcal{C}f(x)|$, where $\mathcal{C}$ denotes the Carleson operator. Obtaining control for $\mathcal{C}$ provides control for T_m^* in this case.

The weighted theory for the Carleson operator is much closer in spirit to that of Calderón–Zygmund operators, as it implicitly uses the fact that the Carleson operator is bounded in $L^p(\mathbb{R})$ for $1 < p < \infty$. Certain *pointwise* control for the Carleson operator along the lines of ($\star$) was obtained by Rubio de Francia, Ruiz and Torrea [118], who established that

$$M^{\#}(\mathcal{C}f)(x) \leqslant C(M(f^s)(x))^{1/s}$$

for any $s > 1$; here $M^{\#}$ denotes the Fefferman–Stein sharp maximal function.

In the context of Calderón–Zygmund operators, more sophisticated variants of the above pointwise estimate have been highly effective very recently. For instance, they play a central role in Lerner's alternative proof of the A_2-conjecture [85], previously resolved by Hytönen [75]. Further developments in that direction have led to *pure pointwise* estimates for Calderón–Zygmund operators, in which the auxiliary operator g_1 on the left hand side of $(\star)$ is entirely absent; see [87, 32, 79, 86].

Following the ideas of Lerner [86], and inspired by the work of Di Plinio and Lerner [42], we have obtained a pure pointwise estimate for the Carleson operator. Namely, we show that

$$|\mathcal{C}f(x)| \leqslant C\mathcal{A}_{r,\mathcal{S}}f(x) := \sum_{Q\in\mathcal{S}} \left(\frac{1}{|Q|}\int_Q |f|^r\right)^{1/r} \chi_Q(x)$$

for any $1 < r < \infty$, where $\mathcal{S}$ is a family of cubes Q satisfying certain almost-disjointness properties. Such a pointwise estimate allows one to deduce weighted estimates for the Carleson operator from those for the operator $\mathcal{A}_{r,\mathcal{S}}$. In particular we obtain weighted L^p inequalities of the type (†) with controlling maximal function $\mathcal{M} = M^{\lfloor p\rfloor+1}$, which denotes the $(\lfloor p\rfloor + 1)-$fold composition of M. This improves on the previously known maximal operator $\mathcal{M}w = (M(w^s))^{1/s}$, where $1 < s < \infty$. Our weighted estimate is along the lines of that of Pérez [108] for Calderón–Zygmund operators. This constitutes the third main result of this thesis, and most of the content in Chapter 4 may be found in [4].

Finally, in Chapter 5 we include a minor contribution in the context of the Fourier restriction conjecture, a problem of central importance in harmonic analysis due to its strong interdisciplinary flavour and numerous applications. The aim of this conjecture is to study whether the Fourier transform of a function may be meaningfully restricted to a m-dimensional manifold S in $\mathbb{R}^d$. In the late 1960's, Stein made the remarkable observation that under certain appropriate curvature hypotheses on S, there exists a $p_0(S) > 1$ for which this restriction is possible for any $f \in L^p$, $1 \leqslant p \leqslant p_0(S)$. These results may be deduced from estimates of the type

$$\|\widehat{gd\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|g\|_{L^p(d\sigma)},$$

where $d\sigma$ denotes the induced Lebesgue measure on S and g is a function defined on S.

The latest progress towards establishing the sharp range of exponents of p, q for which the above estimate holds has been achieved by considering multilinear analogues of the problem. If S is a hypersurface of nonvanishing Gaussian curvature, the progress in the multilinear problem achieved by Bennett, Carbery and Tao [9], combined with a recent method developed by Bourgain and Guth [17], provided some of the best recent results on the restriction conjecture. We study the method of Bourgain and Guth, and we establish a conjectural theorem that quantifies what impact the optimal conjectured multilinear estimates would have on the linear problem. This anticipated progress has been recently achieved by Guth [61] via the algebraic technique of polynomial partitioning.

Structure of the thesis

This thesis is organised as follows, with the main results being contained in Chapters 2, 3 and 4. Some appendices are included at the end for completeness.

Chapter 1

We give a quick overview of classical and modern weighted harmonic analysis related to Calderón–Zygmund theory. This encompasses classical tools such as the sharp maximal function and the more novel sparse operator approach. We also revisit some standard Littlewood–Paley theory and how it may be used to deal with the classical Hörmander–Mikhlin multiplier operators.

Chapter 2

We provide *pointwise* and weighted L^2 control for oscillatory Fourier multipliers. Given α, β in $\mathbb{R}$, the multipliers under study satisfy the differential inequalities

$$|D^\gamma m(\xi)| \lesssim |\xi|^{-\beta+|\gamma|(\alpha-1)}$$

in $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$ for every multi-index $\gamma \in \mathbb{N}^d$ with $|\gamma| \leqslant \lfloor \frac{d}{2} \rfloor + 1$. They are controlled by positive, geometrically-defined maximal functions, which involve fractional averages over certain approach regions. This is joint work with J. Bennett and it is mostly based on the published work [5].

Chapter 3

We study pseudodifferential operators associated to the Hörmander symbol classes $S^m_{\rho,\delta}$. These symbol classes are non-translation invariant generalisations of the above classes of multipliers for $0 \leqslant \alpha \leqslant 1$. We control them by the same maximal functions that in the multiplier case via weighted L^2 inequalities. This chapter is mostly based on the submitted work [3].

Chapter 4

We provide sharp pointwise and weighted L^p estimates for a family of maximally modulated Calderón–Zygmund operators. This class of operators encompasses a wide variety of operators, such as Calderón–Zygmund operators or the Carleson operator. We use the machinery of dyadic sparse operators, which has proved to be highly effective in recent years. Most of the content of this chapter may be found on the accepted work [4].

Chapter 5

Following a method of Bourgain and Guth [17], we establish a conjectural theorem for the Fourier restriction conjecture. This theorem establishes progress on the Fourier restriction conjecture provided optimal estimates are obtained for their multilinear counterparts. The anticipated progress of this theorem has recently been confirmed by Guth in [61].

Chapter 1

Background

In this chapter we collect several classical and modern results to which we shall appeal throughout this thesis. We claim no originality here, and it must be seen as a preliminary chapter encompassing an overview of different results.

1.1 Classical weighted theory

The theory of weighted inequalities has been classically attached to that of the Hardy–Littlewood maximal function and Calderón–Zygmund operators. The development of this research area originates in the 1970s, with fundamental work of Muckenhoupt and others. There is a vast literature in weighted inequalities; here we only intend to give a brief overview. We refer to the standard references [45, 57, 55, 38] for a more detailed introduction to this topic.

One of the first questions studied in weighted theory was to characterise the nonnegative, locally integrable functions w so that the Hardy–Littlewood maximal function M extends to a bounded operator on $L^p(w)$ for $1 < p < \infty$, that is whether

$$\int_{\mathbb{R}^d} (Mf)^p w \leqslant C_{p,d}(w) \int_{\mathbb{R}^d} |f|^p w$$

holds for some finite constant $C_{p,d}(w)$. The answer to this question was given by Muckenhoupt [102], who showed that $M : L^p(w) \to L^p(w)$ is a bounded operator for $1 < p < \infty$ if and only if w is an A_p weight.

Definition 1.1.1. For $1 < p < \infty$, we say that $w \in A_p$ if

$$[w]_{A_p} := \sup_{Q \subset \mathbb{R}^d} \left(\frac{1}{|Q|} \int_Q w \right) \left(\frac{1}{|Q|} \int_Q w^{-1/(p-1)} \right)^{p-1} < \infty,$$

where the supremum is taken over all cubes in $\mathbb{R}^d$. The quantity $[w]_{A_p}$ is known as the A_p constant (or characteristic) of w.

Similarly, Muckenhoupt [102] characterised those weights w for which $M : L^1(w) \to L^{1,\infty}(w)$ is a bounded operator. In this case, the answer is given by the weights satisfying the A_1 condition.

Definition 1.1.2. We say that $w \in A_1$ if there exists a constant $C > 0$ such that

$$Mw(x) \leqslant Cw(x) \qquad \text{a.e.} \quad x \in \mathbb{R}^d.$$

The infimum of such constants C is denoted by $[w]_{A_1}$, and it is known as the A_1 constant (or characteristic) of w.

The A_p condition for $1 < p < \infty$ first appeared in the work of Rosenblum [117], whilst the A_1 condition has a precedent in the work of Fefferman and Stein [47]. Classical examples of A_p weights are the power weights $w(x) = |x|^a$ for $-d < a < d(p-1)$ if $1 < p < \infty$ and for $-d < a \leqslant 0$ if $p = 1$. The A_p classes of weights are increasing in p, that is $A_p \subset A_q$ for $1 \leqslant p < q$.

Similar questions were asked for other classical operators in harmonic analysis, such as Calderón–Zygmund operators.

Definition 1.1.3. A Calderón-Zygmund operator T on $\mathbb{R}^d$ is a L^2 bounded operator that may be represented as

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy, \quad x \notin \operatorname{supp} f,$$

where the kernel K satisfies

(i) $|K(x,y)| \leqslant \frac{C}{|x-y|^d}$ for all $x \neq y$;

(ii) $|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \leqslant C\frac{|x-x'|^\delta}{|x-y|^{d+\delta}}$ for some $0 < \delta \leqslant 1$ when $|x-x'| < |x-y|/2$.

Of course prototypical examples of Calderón–Zygmund operators are the Hilbert and the Riesz transforms. Hunt, Muckenhoupt and Wheeden [71] showed that the A_p condition also characterises the weights for which the Hilbert transform H is a bounded operator on $L^p(w)$ for $1 < p < \infty$ from $L^1(w)$ to $L^{1,\infty}(w)$. This reconciles with a result of Helson–Szegö [67] in the case $p = 2$. The A_p condition also suffices to ensure boundedness of Calderón–Zygmund operators on $L^p(w)$ for $1 < p < \infty$ and from $L^1(w)$ to $L^{1,\infty}(w)$ and it is necessary in certain cases, such as for the Riesz transforms; see the classical work of Coifman and Fefferman [30] or the standard references [45, 129].

The rapid development of the one-weight theory quickly led to the study of two-weight inequalities. The question in this case is to characterise the pair of weights (u,v) for which the two-weight inequality

$$\int_{\mathbb{R}^d} (Mf)^p u \lesssim \int_{\mathbb{R}^d} |f|^p v$$

holds. The natural analogue to the A_p condition for a pair of weights (u,v), given by

$$[u,v]_{A_p} = \sup_{Q \subset \mathbb{R}^n} \left(\frac{1}{|Q|}\int_Q u\right)\left(\frac{1}{|Q|}\int_Q v^{-1/(p-1)}\right)^{p-1} < \infty, \tag{1.1.1}$$

is necessary but not sufficient to guarantee that M is bounded from $L^p(v)$ to $L^p(u)$ for $1 < p < \infty$. However, it is a necessary and sufficient condition in the case of weak-type estimates. That is, there exists a constant C such that

$$u(\{x \in \mathbb{R}^d : Mf(x) > \lambda\}) \leqslant \frac{C}{\lambda^p} \int_{\mathbb{R}^d} |f|^p v$$

holds if and only if $(u, v) \in A_p$ for $1 \leqslant p < \infty$, where we naturally say that $(u, v) \in A_1$ if there exists a constant C such that $Mu(x) \leqslant Cv(x)$ a.e. $x \in \mathbb{R}^d$. These weak-type results may be found in the work of Fefferman and Stein [47], and Muckenhoupt [102]. The question of finding a necessary and sufficient condition on a pair of weights (u, v) in the case of strong-type inequalities is a lot harder. We shall discuss more on this at the end of Chapter 4.

In what follows, we focus on looking for sufficient conditions on a pair of weights (u, v) for two-weight inequalities to hold. In particular, we look for an operator $w \to \mathcal{M}w$ such that the pair of weights $(w, \mathcal{M}w)$ is admissible for any weight w. That is, given an operator U, one would like to identify an operator $\mathcal{M}$ such that a two-weight inequality of the type

$$\int_{\mathbb{R}^d} |Uf|^p w \leqslant C_{p,d} \int_{\mathbb{R}^d} |f|^p \mathcal{M}w \tag{1.1.2}$$

holds for any weight w, where the constant $C_{p,d}$ is independent of w. The first instance of such an inequality goes back to the work of Fefferman and Stein [47], which ensures that in the case of the Hardy–Littlewood maximal function $U = M$, it suffices to take $\mathcal{M} = M$ in (1.1.2).

Of course this question may be addressed for any operator. In the context of Calderón–Zygmund operators, Córdoba and Fefferman [33] showed that for every $s > 1$ and $1 < p < \infty$, there is a constant $C < \infty$ such that

$$\int_{\mathbb{R}^d} |Tf|^p w \leqslant C_T \int_{\mathbb{R}^d} |f|^p M_s w \tag{1.1.3}$$

holds for any weight w, where $M_s w := (Mw^s)^{1/s}$. [1] Observe that given a general two-weight inequality of the type (1.1.2), one may use a duality argument and Hölder's inequality to deduce

$$\begin{aligned}
\|Uf\|_{\tilde{q}} &= \sup_{\|w\|_{(\tilde{q}/p)'}=1} \left(\int_{\mathbb{R}^d} |Uf|^p w \right)^{1/p} \\
&\lesssim \sup_{\|w\|_{(\tilde{q}/p)'}=1} \left(\int_{\mathbb{R}^d} |f|^p \mathcal{M} w \right)^{1/p} \\
&\leqslant \sup_{\|w\|_{(\tilde{q}/p)'}=1} \left(\int_{\mathbb{R}^d} |f|^q \right)^{1/q} \left(\int_{\mathbb{R}^d} (\mathcal{M} w)^{(q/p)'} \right)^{\frac{1}{(q/p)'} \frac{1}{p}} \\
&\leqslant \sup_{\|w\|_{(\tilde{q}/p)'}=1} \|\mathcal{M}\|^{1/p}_{(\tilde{q}/p)' \to (q/p)'} \|f\|_q \left(\int_{\mathbb{R}^d} |w|^{(\tilde{q}/p)'} \right)^{\frac{1}{(\tilde{q}/p)'} \frac{1}{p}} \\
&\leqslant \|\mathcal{M}\|^{1/p}_{(\tilde{q}/p)' \to (q/p)'} \|f\|_q,
\end{aligned} \tag{1.1.4}$$

for $q, \widetilde{q} \geqslant p$. This mechanism serves in many cases to obtain Lebesgue space bounds for the operator U from those for the controlling maximal function $\mathcal{M}$, provided we have an inequality of the type (1.1.2). This will be the case in Chapters 2 and 3 in order to deduce Lebesgue space bounds for Fourier multipliers and pseudodifferential operators.

In the case of Calderón–Zygmund operators, one may not obtain Lebesgue space bounds for T via the inequality (1.1.3), as the implicit constant C_T depends already on the *unweighted* Lebesgue space bounds for T. However, the above mechanism provides a concept of optimality on the maximal function $\mathcal{M}$. Observe that the estimate (1.1.3) leads, via (1.1.4), to

$$\|T\|_{q \to q} \leqslant C \|M_s\|^{1/p}_{(q/p)' \to (q/p)'} \tag{1.1.5}$$

for $q \geqslant p$. As M_s fails to be bounded on L^q for $1 < q \leqslant s$, one would not *recover* the full range of Lebesgue space bounds for T. This suggests that there is scope to improve

[1]This may also be seen as a consequence of the A_p theory, since $M_s w \in A_1 \subset A_p$ for $p > 1$, with constant independent of w, and $w \leqslant M_s w$.

the inequality (1.1.3). Such an improvement was achieved by Wilson [142] in the range $1 < p \leqslant 2$ and by Pérez [107] in the whole range $1 < p < \infty$, who showed that for $1 < p < \infty$, there is a constant $C_T < \infty$, depending on the *unweighted* bounds of T, such that

$$\int_{\mathbb{R}^d} |Tf|^p w \leqslant C_T \int_{\mathbb{R}^d} |f|^p M^{\lfloor p \rfloor + 1} w \tag{1.1.6}$$

holds for any weight w. The operator $M^{\lfloor p \rfloor + 1}$ is bounded on L^q, $1 < q < \infty$, for any p. Thus, this is optimal in the sense of L^q bounds in views of (1.1.4), as one would recover the L^q boundedness of T for the whole range $p \leqslant q < \infty$ if the constant C_T were independent of the *unweighted* bounds. Furthermore, their result is best possible in the sense that the inequality (1.1.6) fails if $M^{\lfloor p \rfloor + 1}$ is replaced by $M^{\lfloor p \rfloor}$. It should be noted that for each $s > 1$ and $k \geqslant 1$, the pointwise estimate $M^k w(x) \leqslant C M_s w(x)$ holds for some constant C independent of w.

Such *sharp* weighted inequalities have also been obtained for operators close to the Calderón-Zygmund theory, like fractional integrals [109], commutators [110] and vector-valued singular integrals [111].

We remark that for the case $p = 1$ these types of two-weight inequalities may be asked in the context of weak-type estimates. As outlined above, the sufficiency of the A_1 condition in this context, together with the trivial observation that $(w, Mw) \in A_1$, yields

$$w(\{x \in \mathbb{R}^d : Mf(x) > \lambda\}) \leqslant \frac{C}{\lambda^p} \int_{\mathbb{R}^d} |f|^p Mw.$$

Muckenhoupt and Wheeden raised the question of whether this inequality also holds for the Hilbert transform and more general Calderón–Zygmund operators. This question was open for a long time, and it was eventually disproved by Reguera and Thiele [115]; see also the previous work of Reguera [114].

1.2 Orlicz maximal functions

In this section we present some concepts related to the theory of Orlicz spaces. This played a fundamental role in the proof of Pérez of the weighted inequality (1.1.6), and also in developing the theory of more general two-weight inequalities for Calderón–Zygmund operators. We will make use of this in Chapter 4. For the standard definitions below we refer the reader to [38] and the references therein.

Let A be a Young function, that is, $A : [0,\infty) \to [0,\infty)$ is a continuous, convex, increasing function with $A(0) = 0$ and such that $A(t) \to \infty$ as $t \to \infty$. We say that a Young function A is doubling if there exists a positive constant C such that $A(2t) \leqslant CA(t)$ for $t > 0$. For each cube $Q \subset \mathbb{R}^d$, we define the Luxemburg norm of f over Q by

$$\|f\|_{A,Q} = \inf\left\{\lambda > 0 : \frac{1}{|Q|}\int_Q A\left(\frac{|f(y)|}{\lambda}\right) dy \leqslant 1\right\}.$$

The Orlicz maximal function associated to the Young function A is defined by

$$M_A f(x) = \sup_{Q \ni x} \|f\|_{A,Q} \tag{1.2.1}$$

for all locally integrable functions f, where the supremum is taken over all cubes Q in $\mathbb{R}^d$ containing x.

We define the complementary Young function $\bar{A}$ to be the Legendre transform of A, that is

$$\bar{A}(t) = \sup_{s>0}\{st - A(s)\}, \quad t > 0.$$

We have that $\bar{A}$ is also a Young function, and it satisfies

$$t \leqslant A^{-1}(t)\bar{A}^{-1}(t) \leqslant 2t$$

for $t > 0$. There is a version of Hölder's inequality in terms of these function space norms,

$$\frac{1}{|Q|}\int_Q f(x)g(x)dx \leqslant \|f\|_A\|f\|_{\bar{A}}.$$

Pérez [108] characterised the Young functions A such that M_A is bounded on L^p for $1 < p < \infty$ and established that the L^p boundedness is equivalent to certain weighted inequalities for M_A and related maximal operators. Such a characterisation is given by the B_p condition.

Definition 1.2.1. Let $1 < p < \infty$. We say that a doubling Young function A satisfies the B_p condition, and we denote it by $A \in B_p$, if there is a constant $c > 0$ such that

$$\int_c^\infty \frac{A(t)}{t^p}\frac{dt}{t} \approx \int_c^\infty \left(\frac{t^{p'}}{\bar{A}(t)}\right)^{p-1}\frac{dt}{t} < \infty. \tag{1.2.2}$$

Observe that for $p < q$ we have $B_p \subset B_q$. The characterisation is given by the following theorem.

Theorem 1.2.2 ([108])**.** *Let $1 < p < \infty$. Let A and B be doubling Young functions satisfying $\bar{B}(t) = A(t^{p'})$. Then the following are equivalent:*

(i) $B \in B_p$.

(ii) There is a constant $c > 0$ such that

$$\int_c^\infty \left(\frac{t}{A(t)}\right)^{p-1}\frac{dt}{t} < \infty.$$

(iii) There is a constant $C < \infty$ such that

$$\int_{\mathbb{R}^d} (M_B f)^p \leqslant C\int_{\mathbb{R}^d} f^p$$

for all non-negative functions f.

(iv) There is a constant $C < \infty$ such that

$$\int_{\mathbb{R}^d} (M_B f)^p u \leqslant C \int_{\mathbb{R}^d} f^p M u$$

for all non-negative functions f and any weight u.

(v) There is a constant $C < \infty$ such that

$$\int_{\mathbb{R}^d} (Mf)^p \frac{u}{(M_A w)^{p-1}} \leqslant C \int_{\mathbb{R}^d} f^p \frac{Mu}{w^{p-1}} \tag{1.2.3}$$

for all non-negative functions f and any weights u, w.

A classical result from Coifman and Rochberg [31] asserts that for any locally integrable function w such that $Mw(x) < \infty$ a.e. and $0 < \delta < 1$, the function $(Mw)^\delta(x)$ is an A_1 weight with A_1-constant independent of w. As Pérez [107] remarks, this result still holds when one replaces the Hardy-Littlewood maximal function by the maximal operator M_A.

Proposition 1.2.3. *Let A be a Young function. If $0 < \delta < 1$, then $(M_A w)^\delta \in A_1$ with A_1 constant independent of w. In particular,*

$$M\left((M_A w)^\delta\right)(x) \leqslant C_d \frac{1}{1-\delta} (M_A w)^\delta (x)$$

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

One may find a proof of this result in [38] (Proposition 5.32). We give an alternative proof following the classical approach from [31] in the Appendix C for completeness.

1.3 Sparse operators

One possible (and classical) approach to proving that a Calderón–Zygmund operator is a bounded operator on $L^p(w)$ for all $w \in A_p$ and $1 < p < \infty$ is via the sharp maximal function, introduced by Fefferman and Stein [48].

Definition 1.3.1. Given $f \in L^1_{loc}(\mathbb{R}^d)$, the sharp maximal function of f is defined by

$$M^{\#}(f)(x) := \sup_{Q \ni x} \inf_{c \in \mathbb{R}} \frac{1}{|Q|} \int_Q |f(z) - c| dz,$$

where the supremum is taken over cubes Q in $\mathbb{R}^d$ containing the point x. This definition is equivalent to the more classical one, in which $c = \frac{1}{|Q|} \int_Q f(y) dy$.

The idea behind this approach is to establish a *pointwise* estimate of the type

$$M^{\#}(Tf)(x) \leqslant C_T \widetilde{\mathcal{M}} f(x) \tag{1.3.1}$$

for a suitable operator $\widetilde{\mathcal{M}}$, which is typically a variant of the Hardy–Littlewood maximal function. Weighted estimates for T follow then from those for $M^{\#}$ and $\widetilde{\mathcal{M}}$; we develop this further in Section 1.5. This was the approach used by Córdoba and Fefferman in [33] to deduce (1.1.3), and it was successfully employed later by many authors in different contexts, allowing to deduce, for instance, that $\|Tf\|_{L^p(w)} \lesssim \|f\|_{L^p(w)}$ for $w \in A_p$ and $1 < p < \infty$.[2]

One of the big open problems in weighted harmonic analysis was to determine the sharp dependence of the operator norm $\|T\|_{L^p(w)}$ in terms of the A_p characteristic of the weight. This question, commonly referred to as the A_2-*conjecture* was recently solved by

[2] Again, and similarly to the case of the weighted inequality (1.1.3), the constant C_T depends on the *unweighted* bounds for T. Thus, it is not possible to use (1.3.1) to obtain boundedness of T in Lebesgue spaces.

Hytönen [75] in the general case of Calderón–Zygmund operators; the specific cases of the Hilbert and Riesz transforms were previously obtained by Petermichl [112, 113].

Theorem 1.3.2 ([75])**.** *Let T be a Calderón–Zygmund operator in $\mathbb{R}^d$. Then*

$$\|Tf\|_{L^2(w)} \leqslant C(T,d)[w]_{A_2}\|f\|_{L^2(w)}, \tag{1.3.2}$$

and the dependence on $[w]_{A_2}$ is sharp.

The proof of this theorem has been simplified over the last few years thanks to the fundamental work of Lerner [84, 85, 87, 86] and others, leading to a better understanding of Calderón–Zygmund operators and related objects. Lerner's approach consists in controlling Calderón–Zygmund operators by simple, geometric objects, for which an estimate of the type (1.3.2) follows by elementary means. To define such simple objects we need to recall some standard definitions.

Let $\mathcal{D}$ be a general dyadic grid, that is a collection of cubes such that

(i) any $Q \in \mathcal{D}$ has sidelength 2^k, $k \in \mathbb{Z}$;

(ii) for any $Q, R \in \mathcal{D}$, we have $Q \cap R \in \{Q, R, \varnothing\}$;

(iii) the cubes of a fixed sidelength 2^k form a partition of $\mathbb{R}^d$.

We say that $\mathcal{S}$ is a *sparse* family of cubes if for any cube $Q \in \mathcal{S}$ there is a measurable subset $E(Q) \subset Q$ such that $|Q| \leqslant 2|E(Q)|$ and the sets $\{E(Q)\}_{Q\in\mathcal{S}}$ are pairwise disjoint.

Given a sparse family $\mathcal{S}$ and $1 \leqslant r < \infty$, we define a sparse operator by

$$\mathcal{A}_{r,\mathcal{S}}f(x) := \sum_{Q\in\mathcal{S}} \left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} |f|^r\right)^{1/r} \chi_Q(x). \tag{1.3.3}$$

Lerner proved in [84] that Banach space norms for T follow from those for the sparse operators $\mathcal{A}_{1,\mathcal{S}}$.

Theorem 1.3.3 ([84])**.** *Let X be a Banach function space over $\mathbb{R}^d$ equipped with Lebesgue measure. Let T be a Calderón–Zygmund operator. Then*

$$\|Tf\|_X \leqslant C(T,d) \sup_{\mathcal{D},\mathcal{S}} \|\mathcal{A}_{1,\mathcal{S}} f\|_X,$$

where the supremum is taken over all dyadic grids $\mathcal{D}$ and all sparse families $\mathcal{S} \subset \mathcal{D}$. The constant $C(T,d)$ depends on $\|T\|_{L^1 \to L^{1,\infty}}$.

This leads to an alternative proof for Theorem 1.3.2. Bounds for the operators $\mathcal{A}_{1,\mathcal{S}}$, and more generally $\mathcal{A}_{r,\mathcal{S}}$, may be obtained with rather elementary techniques, see for instance [84, 42]. In particular, they are bounded on $L^p(w)$ for $w \in A_p$ and $1 < p < \infty$, and it is possible to obtain good quantitative control of the operator norm in terms of the A_p characteristic of the weight; for instance, linear dependence on the $[w]_{A_2}$ constant in the case of $\mathcal{A}_{1,\mathcal{S}}$.

Theorem 1.3.3 was subsequently refined, and it was simultaneously observed by Lerner and Nazarov [87] and Conde–Alonso and Rey [32] that for every $f \in C_c^\infty(\mathbb{R}^d)$ there exists a sparse family of cubes $\mathcal{S}$ such that

$$|Tf(x)| \leqslant C(T,d)\mathcal{A}_{1,\mathcal{S}} f(x). \tag{1.3.4}$$

This belongs to the framework (1.3.1), where the sharp maximal function is now entirely absent. The proof for such a pointwise control has been further simplified by Lacey [79] and Lerner [86]. The most recent proof of Lerner [86] is phrased in a more general context than that of Calderón–Zygmund operators. Given a sublinear operator T, he introduced the *grand maximal function* $\mathcal{N}_T$, defined by

$$\mathcal{N}_T f(x) := \sup_{Q \ni x} \operatorname*{ess\,sup}_{z \in Q} |T(f\chi_{\mathbb{R}^d \setminus 3Q})(z)|; \tag{1.3.5}$$

here the supremum is taken over all cubes $Q \subset \mathbb{R}^d$ containing x.[3]

Theorem 1.3.4 ([86]). *Assume that T is a sublinear operator of weak type (q,q) and $\mathcal{N}_T$ is of weak type (r,r), where $1 \leqslant q \leqslant r < \infty$. Then, for every compactly supported $f \in L^r(\mathbb{R}^d)$, there exists a sparse family $\mathcal{S}$ such that for a.e. $x \in \mathbb{R}^d$,*

$$|Tf(x)| \leqslant C\mathcal{A}_{r,\mathcal{S}}f(x),$$

where $C = C(d,q,r)(\|T\|_{L^q \to L^{q,\infty}} + \|\mathcal{N}_T\|_{L^r \to L^{r,\infty}})$.

In the case of Calderón–Zygmund operators, the grand maximal function $\mathcal{N}_T$ is shown to be of weak-type $(1,1)$ through the maximal truncated operator. In particular it is relatively easy to show [86] that for all $x \in \mathbb{R}^d$,

$$\mathcal{N}_T f(x) \leqslant C(T,d)Mf(x) + T^* f(x),$$

where M denotes the Hardy–Littlewood maximal function and

$$T^* f(x) = \sup_{\varepsilon > 0} \Big| \int_{|y-x|>\varepsilon} K(x,y) f(y) dy \Big|.$$

The $L^1 - L^{1,\infty}$ boundedness for $\mathcal{N}_T$ follows then from that of M and T^*, leading to the pointwise estimate (1.3.4).

Finally, we remark that the proof of the norm estimate in Theorem 1.3.3 relied on an improved version of the *pointwise* estimate (1.3.1). This requires the notion of local mean oscillation, which is a refinement of the concept of the sharp maximal function; see [131, 26, 56, 54] for other historical refinements. In particular, this allows one to exploit

[3] The use of $\bar{Q}$ in the definition of $\mathcal{A}_{r,\mathcal{S}}$ and of $3Q$ in the definition of $\mathcal{N}_T$ is quite conventional; the important underlying feature is that away from a fixed dilate of Q, one may apply the smoothing properties of the Calderón–Zygmund kernels. The choice of $\bar{Q}$ or $3Q$ in each case is taken to be consistent with the referenced literature.

that Calderón–Zygmund operators are of weak-type $(1,1)$.

Given a measurable function f and a cube Q, the local mean oscillation of f on Q is defined by

$$\omega_\lambda(f;Q) = \inf_{c\in\mathbb{R}}((f-c)\chi_Q)^*(\lambda|Q|)$$

for $0 < \lambda < 1$, where f^* denotes the non-increasing rearrangement of f.

The median value of f over a cube Q, denoted by $m_f(Q)$, is a nonunique real number such that

$$|\{x \in Q : f(x) > m_f(Q)\}| \leqslant |Q|/2 \quad \text{and} \quad |\{x \in Q : f(x) < m_f(Q)\}| \leqslant |Q|/2.$$

Lerner proved the following local mean oscillation decomposition in [83]; see [76] for the following refined version.

Theorem 1.3.5 ([76]). *Let f be a measurable function on $\mathbb{R}^d$ and Q_0 be a fixed cube. Then there exists a sparse family of cubes $\mathcal{S} \subset \mathcal{D}(Q_0)$ such that*

$$|f(x) - m_f(Q_0)| \leqslant 2\sum_{Q\in\mathcal{S}} \omega_{\frac{1}{2^{d+2}}}(f;Q)\chi_Q(x)$$

for a.e. $x \in Q_0$.

The local mean oscillation of a Calderón–Zygmund operator satisfies the estimate

$$\omega_\lambda(Tf;Q) \lesssim \left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} |f|\right) + \sum_{m=0}^{\infty} \frac{1}{2^{m\delta}} \left(\frac{1}{|2^m Q|}\int_{2^m Q} |f|\right).$$

This constitutes a refined version of the inequality (1.3.1).

1.4 Littlewood–Paley theory

Square functions have played a pivotal role in the study of many operators in harmonic analysis. One of their common roles is to capture manifestations of orthogonality in L^p spaces for $p \neq 2$. The study of those functions has its roots in the work of Littlewood and Paley [92] and the later development of such a theory is named after them. We refer to work of Stein [127], [128], [126] for a standard real-variable treatment of Littlewood–Paley theory.

An application of Plancherel's theorem quickly reveals that if a family of functions $\{f_j\}_j$ defined on $\mathbb{R}^d$ have Fourier transforms $\widehat{f_j}$ supported in disjoint sets, then the functions are orthogonal, that is

$$\|\sum_j f_j\|_2^2 = \sum_j \|f_j\|_2^2.$$

This orthogonality does not hold when 2 is replaced by another exponent $p \neq 2$. The role of classical Littlewood–Paley theory is to provide a substitute for this principle when $p \neq 2$. To this end, we consider certain *discrete* and *continuous* square functions.

Let $P : \mathbb{R}^d \to \mathbb{R}$ be a smooth function such that $\mathrm{supp}(\widehat{P}) \subseteq \{\xi \in \mathbb{R}^d : |\xi| \sim 1\}$. For any $k \in \mathbb{Z}$, let P_k be defined by $\widehat{P}_k(\xi) = \widehat{P}(2^{-k}\xi)$ and let Δ_k be the operator given by $\widehat{\Delta_k f}(\xi) = \widehat{P}_k(\xi)\widehat{f}(\xi)$. Here we assume that the functions $\{\widehat{P}_k\}_{k\in\mathbb{Z}}$ define a partition of unity, that is

$$\sum_{k\in\mathbb{Z}} \widehat{P}(2^{-k}\xi) = 1 \tag{1.4.1}$$

for $\xi \neq 0$. Consider the square function

$$S(f)(x) := (\sum_{k\in\mathbb{Z}} |\Delta_k f(x)|^2)^{1/2}.$$

The main result of Littlewood–Paley theory is that the square function S characterises

L^p spaces [4] for $1 < p < \infty$, that is

$$\|Sf\|_p \sim \|f\|_p. \tag{1.4.2}$$

Observe that the case $p = 2$ amounts to an application of Plancherel's theorem. The estimates for S are very closely related to Calderón–Zygmund theory; see the standard references cited above.

The square function S satisfies the following two-weight L^2 estimates, from which the characterisation (1.4.2) follows. The forward estimate is a consequence of a more general result of Wilson [143]; we refer to the PhD thesis of Harrison [66] for a careful explanation of how to deduce the above estimate from the work of Wilson. We remark that for this result, the condition (1.4.1) imposed on P is not required.

Proposition 1.4.1 ([143, 66])**.**

$$\int_{\mathbb{R}^d} (Sf)^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 Mw.$$

The reverse estimate is slightly less standard and corresponds to a d-dimensional version of a result in [10].

Proposition 1.4.2 ([10])**.**

$$\int_{\mathbb{R}^d} |f|^2 w \lesssim \int_{\mathbb{R}^d} (Sf)^2 M^3 w.$$

We also need to consider the *continuous* square function

$$s_\phi(f)(x) = \Big(\int_0^\infty |f * \phi_t(x)|^2 \frac{dt}{t} \Big)^{1/2},$$

[4] Littlewood–Paley theory also may be used to characterise other function spaces such as Besov spaces or Triebel–Lizorkin spaces; see for instance [57].

where ϕ is a smooth function such that $\operatorname{supp}(\widehat{\phi}) \subseteq \{\xi \in \mathbb{R}^d : |\xi| \sim 1\}$. In order to obtain a reverse estimate for s_ϕ, one also needs to impose

$$\int_0^\infty \widehat{\phi}(t\xi)\frac{dt}{t} = 1; \quad \xi \neq 0. \tag{1.4.3}$$

The square function s_ϕ satisfies the same estimates as S.

Proposition 1.4.3.

$$\int_{\mathbb{R}^d} s_\phi(f)(x)^2 w(x)dx \lesssim \int_{\mathbb{R}^d} |f(x)|^2 Mw(x)dx \tag{1.4.4}$$

and

$$\int_{\mathbb{R}^d} |f(x)|^2 w(x)dx \lesssim \int_{\mathbb{R}^d} s_\phi(f)(x)^2 M^3 w(x)dx. \tag{1.4.5}$$

There is an equivalence between the continuous and the discrete square functions given by

$$\int_0^\infty |f * \phi_t(y)|^2 \frac{dt}{t} = \sum_{k\in\mathbb{Z}} \int_{2^k}^{2^{k+1}} |f * \phi_t(y)|^2 \frac{dt}{t} \sim \int_1^2 \sum_{k\in\mathbb{Z}} |f * \phi_{\theta 2^k}(y)|^2 d\theta. \tag{1.4.6}$$

The discrete square function

$$S_\theta(f)^2(y) = \sum_{k\in\mathbb{Z}} |f * \phi_{\theta 2^k}(y)|^2$$

satisfies the same estimates as S uniformly in $\theta \in [1,2]$ via an elementary scaling argument. The above equivalence between discrete and continuous square functions allows us to deduce weighted L^2 inequalities for s_ϕ from those for S; see also [5] for a more direct proof of the estimate (1.4.5).

Proof. By (1.4.6), Fubini's theorem and Proposition 1.4.1 we have

$$\int_{\mathbb{R}} s_\phi(f)^2(x)w(x)dx \lesssim \int_1^2 \int_{\mathbb{R}} S_\theta(f)^2(x)w(x)dxd\theta \lesssim \int_{\mathbb{R}} |f(x)|^2 Mw(x)dx.$$

Similarly, by Proposition 1.4.2, averaging over $\theta \in [1,2]$ and (1.4.6),

$$\int_{\mathbb{R}} |f(x)|^2 w(x)dx \lesssim \int_{\mathbb{R}} \int_1^2 S_\theta(f)^2(x)d\theta M^3 w(x)dx \lesssim \int_{\mathbb{R}} \int_0^\infty |f * \phi_t(x)|^2 \frac{dt}{t} M^3 w(x)dx.$$

□

1.5 Hörmander–Mikhlin multipliers

Littlewood–Paley theory has shown to be highly effective in the context of Fourier multiplier theorems. A classical example is that of Hörmander–Mikhlin multipliers.

Theorem 1.5.1. *Let $m : \mathbb{R}^d \backslash \{0\} \to \mathbb{C}$ and denote by T_m the associated multiplier operator. Assume that either,*

- *Mikhlin formulation:*

$$|D^\gamma m(\xi)| \lesssim |\xi|^{-|\gamma|}$$

  for all $\gamma \in \mathbb{N}^d$ with $|\gamma| \leqslant \lfloor \frac{d}{2} \rfloor + 1$.

- *Hörmander formulation (classical derivatives):*

$$\sup_{r>0} r^{|\gamma|} \Big(\frac{1}{r^d} \int_{r \leqslant |\xi| \leqslant 2r} |D^\gamma m(\xi)|^2 d\xi \Big)^{1/2} < \infty \tag{1.5.1}$$

  for all $\gamma \in \mathbb{N}^d$ with $|\gamma| \leqslant \lfloor \frac{d}{2} \rfloor + 1$.

- *Hörmander formulation (Sobolev spaces):*

$$\sup_{r>0} \|m(r\cdot)\Psi\|_{H^\sigma} < \infty$$

  for some $\sigma > d/2$, where Ψ is a suitable smooth function with compact support away *from the origin and H^σ denotes the inhomogeneous Sobolev space. Equivalently*

$$\sup_{r>0} r^\theta r^{-d/2} \|m\Psi(r^{-1}\cdot)\|_{\dot{H}^\theta} < \infty \tag{1.5.2}$$

for all $0 \leqslant \theta \leqslant \sigma$ and some $\sigma > d/2$, where $\dot{H}^\theta$ denotes the homogeneous Sobolev space.

Then m is an $L^p(\mathbb{R}^d)$ multiplier for $1 < p < \infty$, that is $\|T_m f\|_p \lesssim \|f\|_p$.

This may be proved with the classical discrete square functions from the previous section and using that m satisfies good decay estimates adapted to dyadic annuli. Perhaps more enlightening for us is Stein's approach [129] to prove the above theorem. This approach appeared in Section 1.3 in the context of Calderón–Zygmund operators with the pointwise estimate (1.3.1). On a more abstract level, given an operator U, it consists in identifying auxiliary operators g_1 and g_2 for which we have the *pointwise* estimate

$$g_1(Uf)(x) \lesssim g_2(f)(x). \tag{1.5.3}$$

Given such an estimate one may then deduce bounds on U from bounds on the operators g_1 and g_2. More specifically, if one has

$$\|f\|_X \lesssim \|g_1(f)\|_Y \quad \text{and} \quad \|g_2(f)\|_Y \lesssim \|f\|_Z, \tag{1.5.4}$$

for suitable normed spaces X, Y, Z, then the pointwise estimate (1.5.3) quickly reveals that

$$\|Uf\|_X \lesssim \|g_1(Uf)\|_Y \lesssim \|g_2(f)\|_Y \lesssim \|f\|_Z; \tag{1.5.5}$$

that is, U is bounded from Z to X.[5]

[5] Of course this requires that the norm $\|\cdot\|_Y$ is increasing in the sense that $f_1 \lesssim f_2 \implies \|f_1\|_Y \lesssim \|f_2\|_Y$.

In the setting of Hörmander–Mikhlin multipliers, Stein established inequality (1.5.3) with g_1 and g_2 being square functions of Littlewood–Paley type. The relevant square functions here are

$$g_1(f)(x) \equiv g(f)(x) := \left(\int_0^\infty \left|\frac{\partial u}{\partial t}(x,t)\right|^2 t dt\right)^{1/2}$$

where $u : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ denotes the Poisson integral of the function f on $\mathbb{R}^d$, and

$$g_2(f)(x) \equiv g_\lambda^*(f)(x) := \left(\int_{\mathbb{R}^{d+1}} |\nabla u(y,t)|^2 \Big(1 + \frac{|x-y|}{t}\Big)^{-d\lambda} \frac{dydt}{t^{d-1}}\right)^{1/2}.$$

As these square functions satisfy the same bounds as s_ϕ, that is,

$$\|g(f)\|_p \sim \|f\|_p \sim \|g_\lambda^*(f)\|_p$$

for $2 \leqslant p < \infty$ and $\lambda > 1$, the Hörmander–Mikhlin multiplier theorem follows from the pointwise estimate (1.5.3) for $U = T_m$. This is possible because the implicit constant in such an estimate does not depend on any *a priori* bounds for T_m; this is in contrast to (1.3.1), where the implicit constant depends on the *unweighted* bounds for T.

We should remark that with suitable weighted estimates for square functions closely related to g and g_λ^*, one may show that

$$\int_{\mathbb{R}^d} |T_m f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^5 w \tag{1.5.6}$$

for any weight w. We note that this weighted estimate is stronger than the Hörmander–Mikhlin multiplier theorem, as we may deduce L^p bounds for T_m from those for M^5 via the general mechanism (1.1.4).

Chapter 2

Oscillatory Fourier Multipliers

In this chapter we obtain *pointwise* and weighted control for broad classes of highly oscillatory Fourier multipliers on $\mathbb{R}^d$, which satisfy regularity hypotheses adapted to fine (subdyadic) scales. We introduce novel variants of the classical Littlewood–Paley–Stein g-functions adapted to those fine scales, that allow us to obtain pointwise estimates of the type (1.5.3). This approach is very much in the spirit to that of Stein's for Hörmander–Mikhlin multipliers presented in Section 1.5.

As a consequence, we obtain weighted L^2 inequalities that allow us to control such multipliers by positive geometrically-defined maximal functions, which involve fractional averages over certain approach regions. Our framework applies to solution operators for dispersive PDE, such as the time-dependent free Schrödinger equation, and other highly oscillatory convolution operators that fall well beyond the scope of the Calderón–Zygmund theory.

The content of this chapter is joint work with J. Bennett, and may be found in [5]. It builds on previous results of Bennett, Carbery, Soria and Vargas [8], Bennett and Harrison [10] and Bennett [7].

2.1 Classes of multipliers

We begin by describing the weaker Mikhlin-type formulation of our classes of multipliers. Given $\alpha, \beta \in \mathbb{R}$, consider the class of multipliers m on $\mathbb{R}^d$, with support in the set $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$, satisfying the *Miyachi* condition

$$|D^\gamma m(\xi)| \lesssim |\xi|^{-\beta+|\gamma|(\alpha-1)} \tag{2.1.1}$$

for every multi-index $\gamma \in \mathbb{N}^d$ with $|\gamma| \leqslant \lfloor \frac{d}{2} \rfloor + 1$. This class is modelled by the examples

$$\widetilde{m}_{\alpha,\beta}(\xi) := \frac{e^{i|\xi|^\alpha}}{|\xi|^\beta},$$

first studied by Hirschman [68], and later by Wainger [141], Fefferman [49], Fefferman and Stein [48], Miyachi [98, 99] and others. We note that these multipliers often correspond to highly-oscillatory convolution kernels; see for example [121] or the forthcoming Corollary 2.6.1.

The support condition on m is desirable here, as to impose the *same* power-like behaviour (2.1.1) as $|\xi| \to 0$ *and* $|\xi| \to \infty$ would be artificial, at least for $\alpha \neq 0$; for example the specific multiplier $\widetilde{m}_{\alpha,\beta}$ naturally satisfies (2.1.1) for $|\xi|^\alpha \geqslant 1$, but $|D^\gamma m(\xi)| \lesssim |\xi|^{-\beta-|\gamma|}$ for $|\xi|^\alpha \leqslant 1$. We postpone the discussion on multipliers defined on the whole of $\mathbb{R}^d \backslash \{0\}$ satisfying such "two-sided" conditions to Section 2.6.2. The presence of a distinguished (unit) scale here is indeed quite conventional, as may be seen in the formulation of the symbol classes $S^m_{\rho,\delta}$ in Chapter 3. The advantage of imposing a support condition rather than a global estimate of the form $|D^\gamma m(\xi)| \lesssim (1+|\xi|)^{-\beta+|\gamma|(\alpha-1)}$ is that it also has content for $\alpha < 0$.

Our results will naturally apply to broader classes of multipliers than that given by the pointwise condition (2.1.1). We may formulate a Hörmander-type version along the lines of (1.5.1). In order to describe this, we must first introduce the notion of an *α-subdyadic ball.*

Definition 2.1.1. Let $\alpha \in \mathbb{R}$. A euclidean ball B in $\mathbb{R}^d$ is *α-subdyadic* if $\mathrm{dist}(B,0)^\alpha \geqslant 1$ and

$$r(B) \sim \mathrm{dist}(B,0)^{1-\alpha}, \tag{2.1.2}$$

where $r(B)$ denotes the radius of B.

Observe that for $\alpha \neq 0$, typically $r(B) \ll \mathrm{dist}(B,0)$, making it natural to refer to such balls as *subdyadic* (or *α-subdyadic*). In the case $\alpha = 0$ this corresponds effectively to a decomposition into balls of radius r laying in dyadic annuli of width r; this is morally equivalent to the classical decomposition in dyadic annuli.

The Hörmander-type formulation for our classes of multipliers is the following. We consider multipliers m with support in $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$ satisfying the weaker condition

$$\sup_B \ \mathrm{dist}(B,0)^{\beta+(1-\alpha)|\gamma|}\Big(\frac{1}{|B|}\int_B |D^\gamma m(\xi)|^2 d\xi\Big)^{1/2} < \infty \tag{2.1.3}$$

for all $\gamma \in \mathbb{N}^d$ with $|\gamma| \leqslant \lfloor \frac{d}{2} \rfloor + 1$. Here the supremum is taken over all α-subdyadic balls.

As may be expected, the condition (2.1.3) may be weakened still further to

$$\sup_B \mathrm{dist}(B,0)^{\beta+(1-\alpha)\theta}|B|^{-1/2}\|m\Psi_B\|_{\dot{H}^\theta} < \infty, \tag{2.1.4}$$

for all $0 \leqslant \theta \leqslant \sigma$ and some $\sigma > d/2$, uniformly over normalised bump functions Ψ_B adapted to an α-subdyadic ball B. By a *normalised bump function* we mean a smooth function Ψ in $\mathbb{R}^d$, supported in the unit ball, such that $\|D^\gamma\Psi\|_\infty \leqslant 1$ for all multi-indices γ with $|\gamma| \leqslant N$. Here N is a fixed large number, which for our purposes should be taken to exceed d. Given a euclidean ball B in $\mathbb{R}^d$, a *normalised bump function adapted to B* is a function of the form $\Psi_B := \Psi \circ A_B^{-1}$, where Ψ is a normalised bump function and A_B is the affine transformation mapping the unit ball onto $\widetilde{B}$, where $\widetilde{B}$ denotes the concentric double of B.

The above condition is easily seen to reduce to the classical Hörmander condition (1.5.2) when $\alpha = \beta = 0$. Observe that (2.1.1) implies (2.1.3), which in turn implies the more general condition (2.1.4) by the Leibniz formula.

The reason to introduce α-subdyadic balls - and therefore the Hörmander-type formulation - is that the multipliers m satisfying the differential inequalities (2.1.1) are effectively constant on such balls. A manifestation of that principle is that it is possible to prove, with rather elementary techniques, the following weighted estimate for functions whose Fourier transform is supported in an α-subdyadic ball B.

Proposition 2.1.2. *Let $\alpha, \beta \in \mathbb{R}$. Let m be a multiplier on $\mathbb{R}^d$, supported in $\{|\xi|^\alpha \geqslant 1\}$ and satisfying the condition* (2.1.1)*. Let f_B be a function such that $\widehat{f}_B$ is supported in a α-subdyadic ball B. Then*

$$\int_{\mathbb{R}^d} |T_m f_B|^2 w \lesssim \mathrm{dist}(B,0)^{-2\beta} \int_{\mathbb{R}^d} |f_B|^2 Mw,$$

where the implicit constant is independent on the ball B.

Proof. Let ψ_B be a smooth function such that $\widehat{\psi}_B$ equals 1 on B and vanishes outside $\widetilde{B}$. Assume, as we may, that $|D^j \widehat{\psi}_B(\xi)| \lesssim r(B)^{-|j|}$ for any multi-index $j \in \mathbb{N}^d$ uniformly in B. As $\widehat{f}_B$ is supported in B, then $f_B = f_B * \psi_B$ and $T_m f_B = f_B * T_m \psi_B$. By the Cauchy–Schwarz inequality and Fubini's theorem,

$$\begin{aligned}
\int_{\mathbb{R}^d} |T_m f_B(x)|^2 w(x) dx &= \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} f_B(y) T_m \psi_B(x-y) dy \Big|^2 w(x) dx \\
&\leqslant \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |f_B(y)|^2 |T_m \psi_B(x-y)| dy \Big) \|T_m \psi_B\|_1 w(x) dx \\
&= \|T_m \psi_B\|_1 \int_{\mathbb{R}^d} |f_B(y)|^2 |T_m \psi_B| * w(y) dy.
\end{aligned}$$

Integrating by parts,

$$\begin{aligned}|T_m\psi_B(z)| &= \Big|\int_{\widetilde{B}} e^{iz\cdot\xi} m(\xi)\widehat{\psi}_B(\xi)d\xi\Big| \\ &= \frac{1}{|z|^{2N}}\Big|\int_{\widetilde{B}} (\Delta_\xi)^N(e^{iz\cdot\xi})m(\xi)\widehat{\psi}_B(\xi)d\xi\Big| \\ &= \frac{1}{|z|^{2N}}\Big|\int_{\widetilde{B}} e^{iz\cdot\xi}(\Delta_\xi)^N(m\widehat{\psi}_B)(\xi)d\xi\Big| \\ &\leqslant \frac{1}{|z|^{2N}}\int_{\widetilde{B}} |(\Delta_\xi)^N(m\widehat{\psi}_B)(\xi)|d\xi,\end{aligned}$$

for all $z \neq 0$. By Leibniz's rule, the regularity condition (2.1.1) on the multiplier m, the regularity of $\widehat{\psi}_B$, and that $r(B) \sim \mathrm{dist}(B,0)^{1-\alpha} \sim |\xi|^{1-\alpha}$, a straightforward computation shows that

$$|(\Delta_\xi)^N(m\widehat{\psi}_B)(\xi)| \lesssim |\xi|^{-\beta+2N(\alpha-1)}.$$

Thus,

$$|T_m\psi_B(z)| \lesssim \frac{1}{|z|^{2N}}\int_{\widetilde{B}} |\xi|^{-\beta+2N(\alpha-1)} \lesssim \frac{\mathrm{dist}(B,0)^{-\beta+2N(\alpha-1)}r(B)^d}{|z|^{2N}} = \frac{\mathrm{dist}(B,0)^{-\beta}r(B)^d}{|r(B)z|^{2N}}$$

for all $z \neq 0$. As $T_m\psi_B$ is trivially bounded by

$$|T_m\psi_B(z)| \leqslant \int_{\widetilde{B}} |m(\xi)\widehat{\psi}_B(\xi)|d\xi \lesssim \mathrm{dist}(B,0)^{-\beta}r(B)^d,$$

we have the bound

$$|T_m\psi_B(z)| \lesssim \frac{\mathrm{dist}(B,0)^{-\beta}r(B)^d}{(1+|r(B)z|^2)^N}$$

for all $z \in \mathbb{R}^d$. Choosing $N > d/2$ so that the function on the right hand side is integrable, we have that $\|T_m\psi_B\|_1 \lesssim \mathrm{dist}(B,0)^{-\beta}$, and $|T_m\psi_B| * w \lesssim \mathrm{dist}(B,0)^{-\beta}Mw$, from which we may conclude that

$$\int_{\mathbb{R}^d} |T_m f_B(x)|^2 w(x) dx \lesssim \operatorname{dist}(B,0)^{-2\beta} \int_{\mathbb{R}^d} |f_B(y)|^2 Mw(y) dy.$$

$\square$

2.2 Pointwise and weighted control

In this section we present the two main results of this chapter, that is, the *pointwise* and weighted estimates for the classes of multipliers described in the previous section.

To this end, we introduce the square function

$$g_{\alpha,\beta}(f)(x) := \Big(\int_{\Gamma_\alpha(x)} |f * \phi_t(y)|^2 \frac{dy}{t^{(1-\alpha)d+2\beta}} \frac{dt}{t} \Big)^{1/2}, \tag{2.2.1}$$

where $\phi_t(x) := t^{-d}\phi(x/t)$ for $t > 0$, and

$$\Gamma_\alpha(x) := \{(y,t) \in \mathbb{R}^d \times \mathbb{R}_+ : t^\alpha \leqslant 1, |y - x| \leqslant t^{1-\alpha}\}.$$

The function ϕ in the definition of $g_{\alpha,\beta}$ is a smooth function satisfying the uniformity condition (1.4.3) and such that $\widehat{\phi}$ is supported in $\{1 \leqslant |\xi| \leqslant 2\}$ for $\alpha \geqslant 0$, and in $\{1/2 \leqslant |\xi| \leqslant 1\}$ for $\alpha < 0$. The main purpose of this is to ensure that $g_{\alpha,\beta}(f) \equiv 0$ whenever $\widehat{f}$ is supported in $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \leqslant 1\}$. This feature, which also relies on the restriction $t^\alpha \leqslant 1$ in the definition of $\Gamma_\alpha(x)$, makes $g_{\alpha,\beta}$ well-adapted to the support hypothesis imposed on the multipliers that we consider. In particular, we have that $g_{\alpha,\beta}(T_m f) \equiv g_{\alpha,\beta}(T_{m'} f)$ whenever m and m' agree on $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$.

The nature of the region $\Gamma_\alpha(x)$ varies depending on the value of α. For $\alpha \neq 0$ the region $\Gamma_\alpha(x)$ is very different from the classical cone $\Gamma_0(x)$. In particular, when $\alpha > 1$ it becomes an "inverted cone", allowing tangential approach to infinite order, and when

$\alpha < 0$, it is perhaps best interpreted as an "escape" region since $t \geqslant 1$. These regions appeared in [8, 10, 7] in the context of maximal operators and dimension $d = 1$.

By close analogy with the classical g_λ^* we also introduce the more robust square function

$$g^*_{\alpha,\beta,\lambda}(f)(x) = \Big(\int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |f * \phi_t(y)|^2 \Big(1 + \frac{|x-y|}{t^{1-\alpha}}\Big)^{-d\lambda} \frac{dy}{t^{(1-\alpha)d+2\beta}} \frac{dt}{t} \Big)^{1/2},$$

which is manifestly a pointwise majorant of $g_{\alpha,\beta}$. It should be observed that $g_{0,0}$ and $g^*_{0,0,\lambda}$ are minor variants of the classical g and g^*_λ defined in Section 1.5, and very close to the square function s_ϕ. The square functions $g_{\alpha,\beta}$ and $g^*_{\alpha,\beta,\lambda}$ are efficient auxiliary operators in order to obtain a pointwise estimate for the multipliers under study.

Theorem 2.2.1. *Let $\alpha, \beta \in \mathbb{R}$ and m be a multiplier satisfying* (2.1.4). *Then*

$$g_{\alpha,\beta}(T_m f)(x) \lesssim g^*_{\alpha,0,\lambda}(f)(x), \tag{2.2.2}$$

with $\lambda = 2\sigma/d > 1$.

We note that it is not necessary to impose a support condition on the multiplier m in Theorem 2.2.1 thanks to the Fourier support property of the function ϕ in the definition of $g_{\alpha,\beta}$.

Theorem 2.2.1, along with the general mechanism (1.5.5), allows one to find bounds for the multipliers (2.1.4) provided suitable forward and reverse bounds for $g^*_{\alpha,0,\lambda}$ and $g_{\alpha,\beta}$ (respectively) may be found. In particular, we may deduce the following weighted estimate.

Corollary 2.2.2. *Let $\alpha, \beta \in \mathbb{R}$ and m be a multiplier supported in $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$ satisfying* (2.1.4). *Then*

$$\int_{\mathbb{R}^d} |T_m f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\alpha,\beta} M^4 w \tag{2.2.3}$$

for any weight w, where

$$\mathcal{M}_{\alpha,\beta}w(x) = \sup_{(y,r)\in\Gamma_\alpha(x)} \frac{1}{|B(y,r)|^{1-2\beta/d}} \int_{B(y,r)} w.$$

A one-dimensional version of the above weighted estimates in the context of multipliers of bounded variation over certain *subdyadic* intervals was previously obtained by Bennett in [7].

The maximal operator $\mathcal{M}_{\alpha,\beta}$ may be interpreted as a fractional Hardy–Littlewood maximal operator associated with the region Γ_α. Naturally, the case $\alpha = 0$ corresponds to the classical (uncentered) fractional Hardy–Littlewood maximal function. For $0 < \alpha < 1$ the maximal functions $\mathcal{M}_{\alpha,\beta}$ are closely-related to those considered by Nagel and Stein [103] in a different context. In this case, the maximal functions still have a local behaviour. For $\alpha > 1$, the maximal functions become highly non local, as the nature of the approach region Γ_α allows the supremum to be attained in very small balls which are very far away from the point x. This is consistent with its interpretations in the setting of dispersive PDE that we shall discuss in Section 2.6.2.

In dimensions larger than one, the maximal operators $\mathcal{M}_{\alpha,\beta}$ are relatives of the *Nikodym* (or *Kakeya*) maximal operators. In particular, elementary considerations reveal the pointwise bound

$$\mathcal{M}_{\alpha,\beta}f \gtrsim \mathcal{N}_{\alpha,\beta}f, \tag{2.2.4}$$

where

$$\mathcal{N}_{\alpha,\beta}f(x) := \sup_{0<r^\alpha\leqslant 1} \sup_{\substack{T\in\mathcal{T}_\alpha(r)\\ T\ni x}} \frac{r^{2\beta}}{|T|} \int_T |f|$$

and $\mathcal{T}_\alpha(r)$ denotes the collection of cylindrical tubes T of length $r^{1-\alpha}$ and cross-sectional radius r in $\mathbb{R}^d$. The inequality (2.2.4) follows merely by covering each $T \in \mathcal{T}_\alpha(r)$ by $O(r^{-\alpha})$ balls of radius r, and noting their positions. We note that the weighted estimate (2.2.3) is very much in the spirit of Stein's conjecture for the disc multiplier.

Corollary 2.2.2 provides us with an opportunity to comment on the optimality of Theorem 2.2.1 and the maximal functions $\mathcal{M}_{\alpha,\beta}$. Combining it with the general mechanism (1.1.4), one has that

$$\|T_m\|_{p\to q} \lesssim \|\mathcal{M}_{\alpha,\beta}\|_{(q/2)'\to(p/2)'}^{1/2}$$

for any $2 \leqslant p \leqslant q < \infty$. This allows to deduce bounds for the multipliers T_m from those for $\mathcal{M}_{\alpha,\beta}$. The optimal bounds on $\mathcal{M}_{\alpha,\beta}$ (see Section 2.5) may be reconciled with the optimal bounds on the specific multipliers $m_{\alpha,\beta}$ (see Miyachi [98]) in this way.

2.3 Proof of the pointwise estimate

As it is described in Section 1.4, the classical square functions $g_{0,0}$ and $g^*_{0,0,\lambda}$ are able to detect "orthogonality across dyadic frequency scales", but effectively no finer; for this reason they are commonly referred to as dyadic. This is manifested in the "decouplings"

$$g_{0,0}\left(\sum \Delta_k f\right)^2(x) \lesssim \sum g^*_{0,0,\lambda}(\Delta_k f)^2(x) \lesssim g^*_{0,0,\lambda}\left(\sum \Delta_k f\right)^2(x), \tag{2.3.1}$$

where Δ_k is a frequency projection onto the dyadic annulus $A_k = \{\xi \in \mathbb{R}^d : |\xi| \sim 2^k\}$.

The square functions $g_{\alpha,\beta}$ and $g^*_{\alpha,\beta,\lambda}$, which we refer to as *subdyadic* when $\alpha \neq 0$, detect orthogonality across subdyadic scales, leading to a decoupling of the form (2.3.1) associated with suitable families $\mathcal{B}$ of subdyadic balls. This will play a crucial role in our proof of Theorem 2.2.1.

Let $\mathcal{B}$ be a family of α-subdyadic balls B, with bounded overlap, and supporting a regular partition of unity $\{\widehat{\psi}_B\}_{B\in\mathcal{B}}$ on $\{|\xi|^\alpha \geqslant 1\}$. By regular we mean that $\mathrm{supp}(\widehat{\psi}_B) \subseteq \widetilde{B}$ and

$$|D^\gamma \widehat{\psi}_B(\xi)| \lesssim r(B)^{-|\gamma|} \tag{2.3.2}$$

for all multi-indices γ with $|\gamma| \leqslant N$, uniformly in B; for technical reasons that will become apparent later, we shall actually assume that $\mathrm{supp}(\widehat{\psi}_B)$ is contained in a concentric dilate of B with some fixed dilation factor strictly less than 2. This partition of unity gives rise to the reproducing formula

$$f = \sum_{B \in \mathcal{B}} f * \psi_B, \tag{2.3.3}$$

whenever $\mathrm{supp}(\widehat{f}) \subseteq \{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$. For general α and $\mathcal{B}$, elementary geometric considerations reveal that each dyadic annulus A_k will be covered by $O(2^{\alpha dk})$ elements of $\mathcal{B}$ of radius $O(2^{(1-\alpha)k})$.

The following explicit "lattice-based" example of a cover $\mathcal{B}$ and partition $\{\widehat{\psi}_B\}$ will be of use to us later on.

Example 2.3.1. Let $\eta \in \mathcal{S}(\mathbb{R}^d)$ have Fourier support in the annulus $\{|\xi| \sim 1\}$ and be such that

$$\sum_{k \in \mathbb{Z}} \widehat{\eta}_k(\xi) = 1$$

for all $\xi \neq 0$, where $\widehat{\eta}_k(\xi) := \widehat{\eta}(2^{-k}\xi)$. Thus $\{\widehat{\eta}_k\}$ forms a partition of unity on $\mathbb{R}^d \backslash \{0\}$ with $\mathrm{supp}(\widehat{\eta}_k) \subseteq \{|\xi| \sim 2^k\}$ for each $k \in \mathbb{Z}$. Next let $\nu \in \mathcal{S}(\mathbb{R}^d)$ have Fourier support in $\{|\xi| \lesssim 1\}$ be such that

$$\sum_{\ell \in \mathbb{Z}^d} \widehat{\nu}(\xi - \ell) = 1$$

for all $\xi \in \mathbb{R}^d$. For each $k \in \mathbb{Z}$ and $\ell \in \mathbb{Z}^d$ let $\widehat{\nu}_k(\xi) := \widehat{\nu}(2^{-(1-\alpha)k}\xi)$ and $\widehat{\nu}_{k,\ell}(\xi) := \widehat{\nu}_k(\xi - 2^{(1-\alpha)k}\ell)$. Thus for $\widehat{\zeta}_{k,\ell}(\xi) := \widehat{\eta}_k(\xi)\widehat{\nu}_{k,\ell}(\xi)$ we have

$$\sum_{\ell \in \mathbb{Z}^d} \sum_{k \in \mathbb{Z}} \widehat{\zeta}_{k,\ell}(\xi) = 1$$

on $\{|\xi|^\alpha \geqslant 1\}$. Finally we choose, as we may, a family of balls $\mathcal{B}$ and functions $\{\psi_B\}$ so that for each $B \in \mathcal{B}$ there is $(k, \ell) \in \mathbb{Z} \times \mathbb{Z}^d$ for which $\psi_B = \zeta_{k,\ell}$ and $\mathrm{diam}(\mathrm{supp}(\widehat{\zeta}_{k,\ell})) \sim r(B)$. By construction $\{\widehat{\psi}_B\}$ forms a partition of unity on $\{|\xi|^\alpha \geqslant 1\}$ of the type required, provided the implicit constants are chosen suitably. [1]

The square functions $g_{\alpha,\beta}$ and $g^*_{\alpha,\beta,\lambda}$ decouple such subdyadic frequency decompositions.

Proposition 2.3.2. *Let $\alpha, \beta \in \mathbb{R}$. If $\lambda > 1$, then*

$$g_{\alpha,\beta}\Big(\sum_{B\in\mathcal{B}} f * \psi_B\Big)(x)^2 \lesssim \sum_{B\in\mathcal{B}} g^*_{\alpha,\beta,\lambda}(f * \psi_B)(x)^2. \tag{2.3.4}$$

Also, if we assume that $\{\psi_B\}_{B\in\mathcal{B}}$ is as in Example 2.3.1,

$$\sum_{B\in\mathcal{B}} g^*_{\alpha,\beta,\lambda}(f * \psi_B)(x)^2 \lesssim g^*_{\alpha,\beta,\lambda}\Big(\sum_{B\in\mathcal{B}} f * \psi_B\Big)(x)^2. \tag{2.3.5}$$

These decoupling and re-coupling inequalities, together with the reproducing formula (2.3.3), immediately reduce the proof of Theorem 2.2.1 to functions localised at a subdyadic frequency scale, that is to prove that

$$g^*_{\alpha,\beta,\lambda}(T_m(f * \psi_B))(x) \lesssim g^*_{\alpha,0,\lambda}(f * \psi_B)(x) \tag{2.3.6}$$

holds uniformly in $B \in \mathcal{B}$ for $\lambda = 2\sigma/d > 1$. Note that putting this altogether we may quickly deduce Theorem 2.2.1.

Proof of Theorem 2.2.1. Let $\mathcal{B}$ and $\{\psi_B\}_{B\in\mathcal{B}}$ be as in Example 2.3.1. As m is supported in $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$, we may write $T_m f = \sum_{B\in\mathcal{B}} T_m f * \psi_B$. By the decoupling estimate (2.3.4),

$$g_{\alpha,\beta}(T_m f)(x)^2 = g_{\alpha,\beta}\Big(\sum_{B\in\mathcal{B}} T_m f * \psi_B\Big)(x)^2 \lesssim \sum_{B\in\mathcal{B}} g^*_{\alpha,\beta,\lambda}(T_m f * \psi_B)(x)^2.$$

[1] This two-stage decomposition example is implicitly used in the theory of pseudodifferential operators, as it may be extracted from Stein [129]. This will play an important role in Chapter 3.

As $T_m f * \psi_B = T_m(f * \psi_B)$, by the localised pointwise estimate (2.3.6),

$$g_{\alpha,\beta}(T_m f)(x)^2 \lesssim \sum_{B \in \mathcal{B}} g^*_{\alpha,0,\lambda}(f * \psi_B)(x)^2.$$

An application of the re-coupling inequality (2.3.5) allows to conclude

$$g_{\alpha,\beta}(T_m f)(x)^2 \lesssim g^*_{\alpha,0,\lambda}\Big(\sum_{B \in \mathcal{B}} f * \psi_B\Big)(x)^2 = g^*_{\alpha,0,\lambda}(f)(x)^2,$$

as required. □

We devote the rest of this section to proving the inequalities (2.3.4), (2.3.5) and (2.3.6).

2.3.1 Decoupling subdyadic frequency decompositions

Before proceeding with the proof of the decoupling estimate (2.3.4), we need to introduce the auxiliary square function

$$g_{\alpha,\beta,\Phi}(f)(x) = \Big(\int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |f * \phi_t(y)|^2 \Phi\Big(\frac{x-y}{t^{1-\alpha}}\Big) \frac{dy}{t^{(1-\alpha)d+2\beta}} \frac{dt}{t}\Big)^{1/2},$$

where Φ is a Schwartz function such that $\Phi(x) \geqslant c$ for $|x| \leqslant 1$ and $\operatorname{supp}(\widehat{\Phi}) \subseteq \{\xi \in \mathbb{R}^d : |\xi| \leqslant 1\}$.[2] Note that, up to constant factors, $g_{\alpha,\beta,\Phi}$ is a pointwise majorant of $g_{\alpha,\beta}$, and is pointwise majorised by $g^*_{\alpha,\beta,\lambda}$ for any $\lambda > 0$.

By (2.3.3) we have

$$g_{\alpha,\beta,\Phi}(f)(x)^2 = \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} \Big|\sum_{B \in \mathcal{B}} \psi_B * \phi_t * f(y)\Big|^2 \Phi\Big(\frac{x-y}{t^{1-\alpha}}\Big) \frac{dy}{t^{(1-\alpha)d+2\beta}} \frac{dt}{t}.$$

On multiplying out the square and using the Fourier transform, the inner (spatial) integral

[2]Observe that such a function Φ can be constructed by $\Phi = |\Theta|^2 \geqslant 0$, with $\Theta \in \mathcal{S}(\mathbb{R}^d)$ satisfying $\Theta(0) \neq 0$ and $\operatorname{supp}(\widehat{\Theta})$ compact.

in this expression becomes

$$\int_{\mathbb{R}^d}\sum_{B,B'\in\mathcal{B}}(\psi_B*\phi_t*f)(y)\overline{(\psi_{B'}*\phi_t*f)(y)}\Phi\Big(\frac{x-y}{t^{1-\alpha}}\Big)\frac{dy}{t^{(1-\alpha)d+2\beta}}$$
$$=\int_{\mathbb{R}^d}\sum_{B,B'\in\mathcal{B}}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\widehat{\psi}_B(\xi)\overline{\widehat{\psi}_{B'}(\eta)}\widehat{\phi}(t\xi)\overline{\widehat{\phi}(t\eta)}\widehat{f}(\xi)\overline{\widehat{f}(\eta)}e^{iy\cdot(\xi-\eta)}\Phi\Big(\frac{x-y}{t^{1-\alpha}}\Big)d\xi d\eta\frac{dy}{t^{(1-\alpha)d+2\beta}}$$
$$=\sum_{B,B'\in\mathcal{B}}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\widehat{\psi}_B(\xi)\overline{\widehat{\psi}_{B'}(\eta))}\widehat{\phi}(t\xi)\overline{\widehat{\phi}(t\eta)}\widehat{f}(\xi)\overline{\widehat{f}(\eta)}e^{ix\cdot(\xi-\eta)}\widehat{\Phi}(t^{1-\alpha}(\xi-\eta))d\xi d\eta\frac{1}{t^{2\beta}}.$$

The support conditions on $\widehat{\phi}$, $\widehat{\psi}_B$ and $\widehat{\Phi}$ ensure that the summand above vanishes unless $r(B)\sim r(B')\sim t^{\alpha-1}$ and $\operatorname{dist}(B,B')\lesssim t^{\alpha-1}$. In particular, since $\mathcal{B}$ consists of balls of bounded overlap, for each such B there are boundedly many B' satisfying these constraints. Consequently,

$$g_{\alpha,\beta,\Phi}(f)(x)^2$$
$$=\int_{t^\alpha\leqslant 1}\int_{\mathbb{R}^d}\sum_{\substack{B,B'\in\mathcal{B}\\ r(B)\sim r(B')\sim t^{\alpha-1}\\ \operatorname{dist}(B,B')\lesssim t^{\alpha-1}}}(\psi_B*\phi_t*f)(y)\overline{(\psi_{B'}*\phi_t*f)(y)}\Phi\Big(\frac{x-y}{t^{1-\alpha}}\Big)\frac{dy}{t^{(1-\alpha)d+2\beta}}\frac{dt}{t},$$

which by the Cauchy–Schwarz inequality yields

$$g_{\alpha,\beta,\Phi}(f)(x)^2\lesssim\int_{t^\alpha\leqslant 1}\int_{\mathbb{R}^d}\sum_{B\in\mathcal{B}}|\psi_B*\phi_t*f(y)|^2\Phi\Big(\frac{x-y}{t^{1-\alpha}}\Big)\frac{dy}{t^{(1-\alpha)d+2\beta}}\frac{dt}{t}$$
$$=\sum_{B\in\mathcal{B}}g_{\alpha,\beta,\Phi}(f*\psi_B)(x)^2,$$

and thus the decoupling estimate (2.3.4) is proved.

2.3.2 Re-coupling subdyadic frequency decompositions

Here we prove the re-coupling estimate (2.3.5) for the specific family of balls $\mathcal{B}$ and partition $\{\widehat{\psi}_B\}$ described in Example 2.3.1. While it may hold more generally, this lattice-based choice allows us to appeal to the following elementary lemma, which may be viewed as a certain local version of Bessel's inequality.

Lemma 2.3.3. *Suppose that the functions $\nu_{k,\ell} \in \mathcal{S}(\mathbb{R}^d)$, with $k \in \mathbb{Z}$ and $\ell \in \mathbb{Z}^d$, are as in Example 2.3.1. Then*

$$\sum_{\ell \in \mathbb{Z}^d} |f * \nu_{k,\ell}(x)|^2 \lesssim |f|^2 * |\nu_k|(x) \tag{2.3.7}$$

uniformly in k.

Proof. By scaling it suffices to establish (2.3.7) with $k = 0$. Noting that $\nu_0 = \nu$, observe that $f * \nu_{0,\ell}(x) = e^{2\pi i \ell \cdot x} \widehat{h}_x(\ell)$ where $h_x(y) = f(y)\nu(x - y)$. Hence by Parseval's identity, the Poisson summation formula and the Cauchy–Schwarz inequality,

$$\begin{aligned}\sum_{\ell \in \mathbb{Z}^d} |f * \nu_{0,\ell}(x)|^2 &= \int_{[0,1]^d} \Big| \sum_{\ell \in \mathbb{Z}^d} \widehat{h}_x(\ell) e^{2\pi i \ell \cdot y} \Big|^2 dy = \int_{[0,1]^d} \Big| \sum_{m \in \mathbb{Z}^d} h_x(y + m) \Big|^2 dy \\ &\leqslant \int_{[0,1]^d} \sum_{m \in \mathbb{Z}^d} |f(y+m)|^2 |\nu(x - y - m)| \sum_{m' \in \mathbb{Z}^d} |\nu(x - y - m')| dy.\end{aligned}$$

Since

$$\sum_{m' \in \mathbb{Z}^d} |\nu(x - m')| \lesssim 1$$

uniformly in x, we have

$$\sum_{\ell \in \mathbb{Z}^d} |f * \nu_\ell(x)|^2 \lesssim \sum_{m \in \mathbb{Z}^d} \int_{[0,1]^d} |f(y+m)|^2 |\nu(x - y - m)| dy = |f|^2 * |\nu|(x),$$

as required. □

We may now establish the re-coupling estimate (2.3.5) for the partition defined in Example 2.3.1. For ease of notation we let $R_t^\lambda(x) := t^{(\alpha-1)d}(1 + t^{\alpha-1}|x|)^{-d\lambda}$. Observe first that since $\psi_B = \zeta_{k,\ell} = \eta_k * \nu_{k,\ell}$,

$$\sum_{B \in \mathcal{B}} g^*_{\alpha,\beta,\lambda}(f * \psi_B)(x)^2 = \sum_{k \in \mathbb{Z}} \sum_{\ell \in \mathbb{Z}^d} \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |f * \phi_t * \eta_k * \nu_{k,\ell}(y)|^2 R_t^\lambda(x - y) \frac{dy}{t^{2\beta}} \frac{dt}{t}$$

$$= \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} \sum_{2^k \sim t^{-1}} \sum_{\ell \in \mathbb{Z}^d} |f * \phi_t * \eta_k * \nu_{k,\ell}(y)|^2 R_t^\lambda(x-y) \frac{dy}{t^{2\beta}} \frac{dt}{t},$$

where we have also used the Fourier support properties of $\widehat{\phi}_t$ to note that $\phi_t * \eta_k * \nu_{k,\ell} \neq 0$ only if $2^k \sim t^{-1}$. Applying Lemma 2.3.3, followed by the Cauchy–Schwarz inequality, we have

$$\sum_{\ell \in \mathbb{Z}^d} |f * \phi_t * \eta_k * \nu_{k,\ell}(y)|^2 \lesssim |f * \phi_t * \eta_k|^2 * |\nu_k|(y) \lesssim |f * \phi_t|^2 * |\eta_k| * |\nu_k|(y)$$

uniformly in k, t and y, and hence by Fubini's theorem,

$$\sum_{B \in \mathcal{B}} g^*_{\alpha,\beta,\lambda}(f * \psi_B)(x)^2 \lesssim \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |f * \phi_t(y)|^2 \sum_{2^k \sim t^{-1}} |\eta_k| * |\nu_k| * R_t^\lambda(x-y) \frac{dy}{t^{2\beta}} \frac{dt}{t}.$$

Lemma A.1 in Appendix A yields the elementary inequality

$$\sum_{2^k \sim t^{-1}} |\eta_k| * |\nu_k| * R_t^\lambda(x) \lesssim R_t^\lambda(x),$$

which holds uniformly in x and t satisfying $t^\alpha \leqslant 1$, completing the proof of (2.3.5).

2.3.3 The pointwise estimate at a subdyadic frequency scale

Now that Proposition 2.3.2 has been established, to conclude the proof of Theorem 2.2.1 is enough to show that

$$g^*_{\alpha,\beta,2\sigma/d}(T_m(f * \psi_B))(x) \lesssim g^*_{\alpha,0,2\sigma/d}(f * \psi_B)(x)$$

uniformly in $B \in \mathcal{B}$. The argument we present is similar to that given in [127] in the classical setting. We begin by introducing an auxiliary function φ_B, chosen so that its Fourier transform is supported in $\widetilde{B}$ and is equal to 1 on $\operatorname{supp} \widehat{\psi}_B$. For uniformity purposes we also assume, as we may, that

$$|D^j \widehat{\varphi}_B(\xi)| \lesssim r(B)^{-|j|} \tag{2.3.8}$$

for all multi-indices j, uniformly in $B \in \mathcal{B}$. Observe that, up to a constant factor depending only on the uniform implicit constants in (2.3.8), $\widehat{\varphi}_B$ is a normalised bump function adapted to B. We begin by writing

$$\begin{aligned}
&g^*_{\alpha,\beta,2\sigma/d}(T_m(f * \psi_B))(x)^2 \\
&\qquad = \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |T_m(f * \varphi_B * \psi_B) * \phi_t(y)|^2 R_t^{2\sigma/d}(x-y) \frac{dy}{t^{2\beta}} \frac{dt}{t} \\
&\qquad = \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |(T_m \varphi_B) * f * \psi_B * \phi_t(y)|^2 R_t^{2\sigma/d}(x-y) \frac{dy}{t^{2\beta}} \frac{dt}{t} \\
&\qquad \leqslant \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |T_m \varphi_B(z)||f * \psi_B * \phi_t(y-z)| dz \Big)^2 R_t^{2\sigma/d}(x-y) \frac{dy}{t^{2\beta}} \frac{dt}{t}.
\end{aligned}$$

For each t, we split the range of integration of the innermost integral in two parts, $|z| \leqslant t^{1-\alpha}$ and $|z| \geqslant t^{1-\alpha}$. For the term corresponding to $|z| \leqslant t^{1-\alpha}$, we use the Cauchy–Schwarz inequality, Plancherel's theorem, and the hypothesis (2.1.4) with $\sigma = 0$ to obtain

$$\begin{aligned}
\Big(\int_{|z| \leqslant t^{1-\alpha}} |T_m \varphi_B(z)||f * \psi_B * \phi_t(y-z)| dz \Big)^2 &\lesssim t^{2\beta} t^{(\alpha-1)d} \int_{|z| \leqslant t^{1-\alpha}} |f * \psi_B * \phi_t(y-z)|^2 dz \\
&\lesssim t^{2\beta} \int_{\mathbb{R}^d} R_t^{2\sigma/d}(z) |f * \psi_B * \phi_t(y-z)|^2 dz;
\end{aligned}$$

observe that the support hypothesis on $\widehat{\phi}$ and $\widehat{\psi}_B$ ensure $r(B) \sim t^{\alpha-1}$. Similarly, in $|z| \geqslant t^{1-\alpha}$,

$$\begin{aligned}
&\Big(\int_{|z| \geqslant t^{1-\alpha}} |T_m \varphi_B(z)||f * \psi_B * \phi_t(y-z)| dz \Big)^2 \\
&\qquad \leqslant \Big(\int_{\mathbb{R}^d} |T_m \varphi_B(z)|^2 |z|^{2\sigma} dz \Big) \Big(\int_{|z| \geqslant t^{1-\alpha}} \frac{1}{|z|^{2\sigma}} |f * \psi_B * \phi_t(y-z)|^2 dz \Big)
\end{aligned}$$

$$\lesssim t^{2\beta}t^{(\alpha-1)d}t^{2(1-\alpha)\sigma}\int_{|z|\geqslant t^{1-\alpha}}\frac{1}{(t^{1-\alpha}+|z|)^{2\sigma}}|f*\psi_B*\phi_t(y-z)|^2dz$$
$$\lesssim t^{2\beta}t^{(\alpha-1)d}\int_{\mathbb{R}^d}(1+t^{\alpha-1}|z|)^{-2\sigma}|f*\psi_B*\phi_t(y-z)|^2dz.$$

Putting together the above estimates we obtain

$$\begin{aligned}
&g^*_{\alpha,\beta,2\sigma/d}(T_m(f*\psi_B))(x)^2\\
&\qquad\lesssim\int_{t^\alpha\leqslant 1}\int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}R_t^{2\sigma/d}(z)|f*\psi_B*\phi_t(y-z)|^2dz\Big)R_t^{2\sigma/d}(x-y)dy\frac{dt}{t}\\
&\qquad=\int_{t^\alpha\leqslant 1}\int_{\mathbb{R}^d}|f*\psi_B*\phi_t(y)|^2R_t^{2\sigma/d}*R_t^{2\sigma/d}(x-y)dy\frac{dt}{t}\\
&\qquad\lesssim g^*_{\alpha,0,2\sigma/d}(f*\psi_B)(x)^2,
\end{aligned}$$

where the last inequality follows since $\sigma > d/2$ and $R_t^\lambda * R_t^\lambda(x) \lesssim R_t^\lambda(x)$ for $\lambda > 1$; see Appendix A. This concludes the proof of Theorem 2.2.1.

2.4 Proof of the weighted estimate

The proof of the weighted estimate (2.2.3) follows from the pointwise estimate via the mechanism described in (1.5.5), provided we establish weighted estimates for the square functions $g_{\alpha,\beta}$ and $g^*_{\alpha,0,\lambda}$.

The reverse weighted bound for $g_{\alpha,\beta}$ is the most interesting one since it involves the maximal function $\mathcal{M}_{\alpha,\beta}$.

Theorem 2.4.1. *Let $\alpha,\beta\in\mathbb{R}$, and f be a function such that* $\mathrm{supp}(\widehat{f})\subseteq\{\xi\in\mathbb{R}^d : |\xi|^\alpha\geqslant 1\}$. *Then*

$$\int_{\mathbb{R}^d}|f|^2w\lesssim\int_{\mathbb{R}^d}g_{\alpha,\beta}(f)^2\mathcal{M}_{\alpha,\beta}M^4w.$$

In order to prove Theorem 2.4.1, we make use of the following elementary lemma.

Lemma 2.4.2.

$$\int_{\mathbb{R}^d} f(x)h(x)dx \lesssim R^d \int_{\mathbb{R}^d} \int_{|y-x|\leqslant \frac{1}{R}} f(y)dy \sup_{z:|z-x|\leqslant \frac{1}{R}} h(z)\, dx \tag{2.4.1}$$

uniformly in $R > 0$ and nonnegative functions f, h on $\mathbb{R}^d$.

Proof. To simplify notation, we prove the one-dimensional case of (2.4.1); the d-dimensional case follows by applying the one-dimensional in each variable. Observe that we may decompose the integral as

$$\begin{aligned}\int_{\mathbb{R}} f(x)h(x)dx &= \sum_{k\in\mathbb{Z}} \int_{-1/R}^{1/R} f\Big(x+u+\frac{2k}{R}\Big)h\Big(x+u+\frac{2k}{R}\Big)dx \\ &= \sum_{k\in\mathbb{Z}} \int_{|y-u-\frac{2k}{R}|\leqslant\frac{1}{R}} f(y)h(y)dy \\ &\leqslant \sum_{k\in\mathbb{Z}} \int_{|y-u-\frac{2k}{R}|\leqslant\frac{1}{R}} f(y)dy \sup_{z:|z-u-\frac{2k}{R}|\leqslant\frac{1}{R}} h(z)\end{aligned}$$

for any $|u| \leqslant \frac{1}{R}$. Averaging over u,

$$\begin{aligned}\int_{\mathbb{R}} f(x)h(x)dx &\leqslant \sum_{k\in\mathbb{Z}} 2R \int_{-1/R}^{1/R} \int_{|y-u-\frac{2k}{R}|\leqslant\frac{1}{R}} f(y)dy \sup_{z:|z-u-\frac{2k}{R}|\leqslant\frac{1}{R}} h(z)\, du \\ &= 2R \sum_{k\in\mathbb{Z}} \int_{-1/R+2k/R}^{1/R+2k/R} \int_{|y-x|\leqslant\frac{1}{R}} f(y)dy \sup_{z:|z-x|\leqslant\frac{1}{R}} h(z)\, dx \\ &= 2R \int_{\mathbb{R}} \int_{|y-x|\leqslant\frac{1}{R}} f(y)dy \sup_{z:|z-x|\leqslant\frac{1}{R}} h(z)\, dx,\end{aligned}$$

as required. □

Proof of Theorem 2.4.1. We begin by using classical Littlewood–Paley theory in the form of (1.4.5) to write

$$\int_{\mathbb{R}^d} |f(x)|^2 w(x)dx \lesssim \int_0^\infty \int_{\mathbb{R}^d} |f * \phi_t(y)|^2 M^3 w(y)dy \frac{dt}{t}.$$

The support conditions on $\widehat{\phi}$ and $\widehat{f}$ reduce the range for the t-integration to those t such that $0 < t^\alpha \leqslant 1$. Choosing $\varphi \in \mathcal{S}$ such that $\widehat{\varphi} = 1$ on the support of $\widehat{\phi}$ and $\operatorname{supp}(\widehat{\varphi}) \subseteq \{\xi \in \mathbb{R}^d : \frac{1}{4} \leqslant |\xi| \leqslant 4\}$ allows us to write $f * \phi_t(y) = f * \phi_t * \varphi_t(y)$. Combining this with applications of the Cauchy–Schwarz inequality and Fubini's theorem gives

$$\begin{aligned}\int_{\mathbb{R}^d} |f(x)|^2 w(x)dx &\lesssim \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |f * \phi_t(y)|^2 (|\varphi_t| * M^3 w)(y) dy \frac{dt}{t} \\ &\lesssim \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} |f * \phi_t(y)|^2 A_t^* M^3 w(y) dy \frac{dt}{t},\end{aligned}$$

where $A_t^* w(x) := \sup_{r \geqslant t} A_r w(x)$ and

$$A_t w(x) := \frac{1}{|B(x,t)|} \int_{B(x,t)} w.$$

Observe that $A_t^* w \lesssim A_t A_t^* w \leqslant A_t M w$, so applying Lemma 2.4.2 at scale $R = t^{\alpha-1}$ yields

$$\begin{aligned}&\int_{\mathbb{R}^d} |f(x)|^2 w(x)dx \\ &\quad\lesssim \int_{t^\alpha \leqslant 1} \int_{\mathbb{R}^d} \int_{|y-x| \leqslant t^{1-\alpha}} |f * \phi_t(y)|^2 \frac{dy}{t^{(1-\alpha)d+2\beta}} \sup_{z:|z-x| \leqslant t^{1-\alpha}} t^{2\beta} A_t M^4 w(z)\, dx \frac{dt}{t} \\ &\quad\leqslant \int_{\mathbb{R}^d} \int_{t^\alpha \leqslant 1} \int_{|y-x| \leqslant t^{1-\alpha}} |f * \phi_t(y)|^2 \frac{dy}{t^{(1-\alpha)d+2\beta}} \frac{dt}{t} \mathcal{M}_{\alpha,\beta} M^4 w(x) dx,\end{aligned}$$

where the last inequality follows by taking the supremum in t, since

$$\sup_{t^\alpha \leqslant 1} \sup_{z:|z-x| \leqslant t^{1-\alpha}} t^{2\beta} A_t M^4 w(z) = \mathcal{M}_{\alpha,\beta} M^4 w(x),$$

by the definition of $\mathcal{M}_{\alpha,\beta}$. □

The forward estimate for $g^*_{\alpha,0,\lambda}$ is more classical in nature than its reverse counterpart above, and it is a simple consequence of Section 1.4. We also refer to [127] for an analogous result for g^*_λ.

Theorem 2.4.3. *Let $\alpha \in \mathbb{R}$ and $\lambda > 1$. Then*

$$\int_{\mathbb{R}^d} g^*_{\alpha,0,\lambda}(f)^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 w.$$

Proof. By Fubini's theorem,

$$\int_{\mathbb{R}^d} g^*_{\alpha,0,\lambda}(f)(x)^2 w(x)dx = \int_{\mathbb{R}^d} \int_{t^\alpha \leqslant 1} |f * \phi_t(y)|^2 R_t^\lambda * w(y) dy \frac{dt}{t}.$$

Since

$$\sup_t R_t^\lambda * w \lesssim Mw$$

for $\lambda > 1$, we have

$$\int_{\mathbb{R}^d} g^*_{\alpha,0,\lambda}(f)(x)^2 w(x)dx \lesssim \int_{\mathbb{R}^d} \int_0^\infty |f * \phi_t(y)|^2 \frac{dt}{t} Mw(y)dy,$$

which by an application of classical Littlewood–Paley theory in the form of (1.4.4) results in

$$\int_{\mathbb{R}^d} g^*_{\alpha,0,\lambda}(f)(x)^2 w(x)dx \lesssim \int_{\mathbb{R}^d} |f(y)|^2 M^2 w(y)dy,$$

as required. □

As may be expected, it is possible to obtain similar weighted L^2 estimates for $g^*_{\alpha,\beta,\lambda}$ for other values of β by minor modifications of the above argument.

Corollary 2.2.2 trivially follows now from applying Theorems 2.2.1, 2.4.1 and 2.4.3.

Proof of Corollary 2.2.2. Applying Theorem 2.4.1 to $T_m f$, which trivially satisfies that $\widehat{T_m f}$ is supported in $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$, the pointwise estimate (2.2.2) and Theorem 2.4.3

we have that

$$\begin{aligned}\int_{\mathbb{R}^d} |T_m f|^2 w &\lesssim \int_{\mathbb{R}^d} g_{\alpha,\beta}(T_m f)^2 \mathcal{M}_{\alpha,\beta} M^4 w \\ &\lesssim \int_{\mathbb{R}^d} g^*_{\alpha,0,\lambda}(f)^2 \mathcal{M}_{\alpha,\beta} M^4 w \\ &\lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\alpha,\beta} M^4 w,\end{aligned}$$

as required. $\square$

2.5 $L^p(\mathbb{R}^d) - L^q(\mathbb{R}^d)$ boundedness results

In this section we establish the relevant Lebesgue space bounds satisfied by the operators $\mathcal{M}_{\alpha,\beta}$ and how to use the weighted inequalities (2.2.3) and Theorem 2.4.1 to obtain bounds for T_m and $g_{\alpha,\beta}$ respectively. We have the following bounds for the maximal operator $\mathcal{M}_{\alpha,\beta}$.

Theorem 2.5.1. *Let $1 < p \leqslant q \leqslant \infty$ and $\alpha, \beta \in \mathbb{R}$. If $\alpha > 0$ and*

$$\beta \geqslant \frac{\alpha d}{2q} + \frac{d}{2}\Big(\frac{1}{p} - \frac{1}{q}\Big),$$

or $\alpha = 0$ and

$$\beta = \frac{d}{2}\Big(\frac{1}{p} - \frac{1}{q}\Big),$$

or $\alpha < 0$ and

$$\beta \leqslant \frac{\alpha d}{2q} + \frac{d}{2}\Big(\frac{1}{p} - \frac{1}{q}\Big),$$

then $\mathcal{M}_{\alpha,\beta}$ is bounded from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$.

This theorem is a straightforward adaptation of the one-dimensional case in [7]; the proof is included at the end of this section.

As described at the end of Section 2.2, the mechanism (1.1.4) allows us to deduce bounds for T_m from those for $\mathcal{M}_{\alpha,\beta}$ via the weighted estimates (2.2.3), as

$$\|T_m\|_{p\to q} \lesssim \|\mathcal{M}_{\alpha,\beta}\|^{1/2}_{(q/2)'\to(p/2)'}.$$

Similarly, one may obtain reverse bounds for $g^*_{\alpha,\beta,\lambda}$ via Theorem 2.4.1.

Corollary 2.5.2. *Let $\alpha, \beta \in \mathbb{R}$ and $2 \leqslant p \leqslant q < \infty$. If $\alpha > 0$ and*

$$\frac{\beta}{d} \geqslant \alpha\Big(\frac{1}{2} - \frac{1}{p}\Big) + \frac{1}{p} - \frac{1}{q},$$

or $\alpha = 0$ and

$$\frac{\beta}{d} = \frac{1}{p} - \frac{1}{q},$$

or $\alpha < 0$ and

$$\frac{\beta}{d} \leqslant \alpha\Big(\frac{1}{2} - \frac{1}{p}\Big) + \frac{1}{p} - \frac{1}{q},$$

and m is a Fourier multiplier satisfying (2.1.4)*, then*

$$\|T_m f\|_q \lesssim \|f\|_p.$$

Also, if $f \in L^q(\mathbb{R}^d)$ is such that $\mathrm{supp}(\widehat{f}) \subseteq \{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$,

$$\|f\|_q \lesssim \|g_{\alpha,\beta}(f)\|_p.$$

Duality allows one to obtain bounds on the multipliers for $1 < p \leqslant q \leqslant 2$. Such bounds recover a number of well-known multiplier theorems since our class (2.1.4) naturally contains those considered by Miyachi [99] – in addition to the classical Hörmander–Mikhlin multipliers and fractional integrals. In particular, as the model multipliers

$|\xi|^{-\beta}e^{i|\xi|^{\alpha}}\chi_{\{|\xi|^{\alpha}\geqslant 1\}}$ are bounded on $L^p(\mathbb{R}^d)$ if and only if $|1/2-1/p|\leqslant \beta/(\alpha d)$, the maximal operators $\mathcal{M}_{\alpha,\beta}$ in (2.2.3) are optimal, in the sense that they cannot be replaced by variants satisfying additional Lebesgue space bounds.

We devote the end of this section to prove Theorem 2.5.1.

Proof of Theorem 2.5.1. We only concern ourselves with the case $\alpha\neq 0$; the case $\alpha=0$ corresponds to the classical fractional Hardy–Littlewood maximal function. Observe that for $\alpha>0$, the possible radii r in the approach region $\Gamma_\alpha(x)$ satisfy $0<r\leqslant 1$ and therefore

$$\mathcal{M}_{\alpha,\beta'}w\leqslant \mathcal{M}_{\alpha,\beta}w$$

for $0<\beta<\beta'$. A similar analysis for the case $\alpha<0$ reveals that it is enough to show that $\mathcal{M}_{\alpha,\beta}$ is bounded from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$, where $1<p\leqslant q\leqslant\infty$, on the line

$$\beta=\frac{\alpha d}{2q}+\frac{d}{2}\Big(\frac{1}{p}-\frac{1}{q}\Big). \tag{2.5.1}$$

We regularise the average in the definition of $\mathcal{M}_{\alpha,\beta}$ and we prove the estimates for the pointwise larger maximal operator (in the case of weights)

$$\widetilde{\mathcal{M}}_{\alpha,\beta}w(x)=\sup_{(y,r)\in\Gamma_\alpha(x)} r^{2\beta}|P_r*w(y)|,$$

where P is a nonnegative compactly supported bump function which is positive on $B(0,1)$ and $P_r(x):=r^{-d}P(x/r)$. Trivially,

$$|\widetilde{\mathcal{M}}_{\alpha,0}w(x)|=\sup_{(y,r)\in\Gamma_\alpha(x)}|P_r*w(y)|\leqslant\|P\|_1\|w\|_\infty$$

and

$$|\widetilde{\mathcal{M}}_{\alpha,\frac{d}{2}}w(x)|=\sup_{(y,r)\in\Gamma_\alpha(x)} r^{d}|P_r*w(y)|\leqslant\|P\|_\infty\|w\|_1,$$

for every $x \in \mathbb{R}^d$. Analytic interpolation between these two estimates gives $\widetilde{\mathcal{M}}_{\alpha,\frac{d}{2p}}$: $L^p(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d)$ for $1 \leqslant p \leqslant \infty$. A further application of analytic interpolation shows that boundedness for $\widetilde{\mathcal{M}}_{\alpha,\beta}$ from $L^p(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ holds for α, β as in (2.5.1) provided $\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} : H^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$. To this end, it suffices to see

$$\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a\|_1 \lesssim 1$$

uniformly in $H^1(\mathbb{R}^d)$-atoms a; recall that an atom a is a function defined on $\mathbb{R}^d$ supported in a cube Q such that $\int_Q a = 0$ and $\|a\|_\infty \leqslant \frac{1}{|Q|}$. By translation-invariance, we may assume that the cube Q is centered at the origin. We have the following standard bounds

$$|P_r * a(y)| \lesssim \begin{cases} 1/|Q| & \text{if } r < |Q|^{1/d}, |y| \lesssim |Q|^{1/d} \\ |Q|^{1/d}/r^{d+1} & \text{if } r > |Q|^{1/d}, |y| \lesssim r \\ 0 & \text{otherwise.} \end{cases}$$

This estimate is a consequence of the following elementary considerations. Of course

$$|P_r * a(y)| = \Big| \int_Q P_r(y-z) a(z) dz \Big|.$$

Observe that $P_r(y - \cdot)$ is supported in a ball of center y and radius r. If r is small, say $r < |Q|^{1/d}$, and $|y| \gtrsim |Q|^{1/d}$, we have that $B(y,r) \cap Q = \varnothing$ and then $P_r * a(y) = 0$. Analogously, if $r > |Q|^{1/d}$ and $|y| \gtrsim r$, we have that $B(y,r) \cap Q = \varnothing$ and then $P_r * a(y) = 0$.

For the remaining cases, one may trivially apply $L^1 - L^\infty$ duality to obtain the bound

$$\Big| \int_Q P_r(y-z) a(z) dz \Big| \leqslant \int_Q |a(z)||P_r(y-z)| dz \leqslant \frac{1}{|Q|} \|P\|_1 \lesssim \frac{1}{|Q|}.$$

This is a good estimate for small r. However, one may do better for large r, as P_r is essentially constant for large r, which would imply that $P_r * a$ tends to $\int a = 0$ as $r \to \infty$. To exploit this, we use the mean value zero of a to obtain the improved bound

$$|P_r * a(y)| = \Big| \int_Q (P_r(y-z) - P_r(y)) a(z) dz \Big| \leqslant \int_Q |\nabla P_r(\xi_z)||z||a(z)| dz \lesssim \frac{|Q|^{1/d}}{r^{d+1}}.$$

This concludes the discussion on the bounds for $|P_r * a|$.

In order to obtain the required bound $\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a\|_1 \lesssim 1$, we need to argue differently depending on the value of α, as the nature of the region Γ_α changes dramatically for $\alpha < 0$, $0 < \alpha \leqslant 1$ and $\alpha > 1$.

Case $\alpha > 1$: we have $0 < r < 1$, so we divide our analysis in $|Q|^{1/d} > 1$ and $|Q|^{1/d} < 1$. Assume $|Q|^{1/d} > 1$. As $r < 1$, we are in the situation $|Q|^{1/d} > r$, so

$$|P_r * a(y)| \lesssim \begin{cases} 1/|Q| & \text{if } |y| \lesssim |Q|^{1/d} \\ 0 & \text{otherwise} \end{cases}$$

If $|x| \lesssim |Q|^{1/d}$, for any $(y,r) \in \Gamma_\alpha(x)$, we have $|P_r * a(y)| \lesssim 1/|Q|$, so

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) = \sup_{(y,r)\in\Gamma_\alpha(x)} r^{\alpha d} |P_r * a(y)| \leqslant \frac{1}{|Q|}.$$

If $|x| \gtrsim |Q|^{1/d}$, we would like to make $r^{\alpha d}$ as big as possible with $|y| \lesssim |Q|^{1/d}$, so the supremum is attained at $r \sim |x|^{\frac{1}{1-\alpha}}$ and then

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} w(x) \lesssim \frac{|x|^{\frac{\alpha d}{1-\alpha}}}{|Q|}.$$

This leads to

$$\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a\|_1 \leqslant \int_{|x|\lesssim|Q|^{1/d}} \frac{1}{|Q|} dx + \int_{|x|\gtrsim|Q|^{1/d}} \frac{|x|^{\frac{\alpha d}{1-\alpha}}}{|Q|} dx \lesssim 1 + |Q|^{\frac{\alpha}{1-\alpha}} \lesssim 1.$$

Now assume $|Q|^{1/d} \leqslant 1$. For any x there is always a small radius r such that $r < |Q|^{1/d}$ and $(y,r) \in \Gamma_\alpha$ with $|y| < |Q|^{1/d}$; recall that the case $\alpha > 1$ allows tangential approach to infinite order. Hence

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) \leqslant |Q|^\alpha \frac{1}{|Q|} = |Q|^{-(1-\alpha)}.$$

The contribution to $\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a\|_1$ of those x such that $|x| \lesssim |Q|^{(1-\alpha)/d}$ is 1. When $|x| \gtrsim |Q|^{(1-\alpha)d}$ one may obtain a better estimate, as to impose $|y| \lesssim |Q|^{1/d}$ one needs to take $r \sim |x|^{\frac{1}{1-\alpha}}$. Then

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) \lesssim |x|^{\frac{\alpha d}{1-\alpha}} \frac{1}{|Q|},$$

which integrates 1 over the region $|x| \gtrsim |Q|^{(1-\alpha)/d}$. This concludes the case $\alpha > 1$.

<u>Case $\alpha < 0$:</u> we have $r > 1$. We split again our analysis in $|Q|^{1/d} > 1$ and $|Q|^{1/d} < 1$. Assume $|Q|^{1/d} < 1$. As $r > 1$, we are in the situation $|Q|^{1/d} < r$, so

$$|P_r * a(y)| \lesssim \begin{cases} |Q|^{1/d}/r^{d+1} & \text{if } |y| \lesssim r, \\ 0 & \text{otherwise.} \end{cases}$$

If $|x| \lesssim 1$, $(0,1) \in \Gamma_\alpha(x)$, so

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) \lesssim |Q|^{1/d} \lesssim 1,$$

which integrates 1 in $|x| \lesssim 1$. If $|x| \gtrsim 1$, a similar reasoning to the one in the previous case tells us that the smallest r such that $|y| \lesssim r$ and $(y,r) \in \Gamma_\alpha$ is given by $r \sim |x|^{\frac{1}{1-\alpha}}$. Thus,

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) \lesssim |x|^{-d} |x|^{\frac{-1}{1-\alpha}} |Q|^{1/d},$$

which integrates 1 in $|x| \gtrsim 1$. Then $\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a\|_1 \lesssim 1$.

Now assume $|Q|^{1/d} > 1$. Taking $r \sim |x|^{\frac{1}{1-\alpha}}$, there is a y such that $(y,r) \in \Gamma_\alpha(x)$ with

$|y| \lesssim r$. If $|x| \leqslant |Q|^{(1-\alpha)/d}$, then $r < |Q|^{1/d}$ and

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) \lesssim |x|^{\frac{\alpha d}{1-\alpha}} \frac{1}{|Q|},$$

which integrates 1 in $|x| \leqslant |Q|^{1/d}$. For the case $|x| \geqslant |Q|^{(1-\alpha)/d}$, we have $r > |Q|^{1/d}$, so

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) \lesssim |x|^{\frac{\alpha d}{1-\alpha}} |Q|^{1/d} |x|^{\frac{-d}{1-\alpha}} |x|^{\frac{-1}{1-\alpha}} = |Q|^{1/d} |x|^{-d} |x|^{\frac{-1}{1-\alpha}},$$

which integrates 1 in the region $|x| > |Q|^{1/d}$. Again, $\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a\|_1 \lesssim 1$ and this completes the case $\alpha < 0$.

Case $\alpha = 1$: the approach region is

$$\Gamma_1(x) = \{(y,r) : 0 < r < 1, |y - x| < 1\}.$$

We make again the distinction between $|Q|^{1/d} > 1$ and $|Q|^{1/d} < 1$. If $|Q|^{1/d} > 1$,

$$|P_r * a(y)| \lesssim \begin{cases} 1/|Q| & \text{if } |y| \lesssim |Q|^{1/d}, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for $|x| \geqslant |Q|^{1/d}$, $\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) = 0$. For $|x| \leqslant |Q|^{1/d}$, the supremum will be attained for $r = 1$, so

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) \lesssim \frac{1}{|Q|},$$

which integrates 1 in the region $|x| \leqslant |Q|^{1/d}$.

Assume $|Q|^{1/d} \leqslant 1$. If $|x| \gtrsim 1$, there is no y such that $(y,r) \in \Gamma_1(x)$ with $|y| < r$ or $|y| < |Q|^{1/d}$. Then $\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}} a(x) = 0$. For $|x| \lesssim 1$, either there is $(y,r) \in \Gamma_1(x)$ with $r < |Q|^{1/d}$ or with $r > |Q|^{1/d}$ or both. In any case the supremum is always controlled by

1,

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a(x) \lesssim 1,$$

which integrates 1 in $|x| \lesssim 1$. Thus, $\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a\|_1 \lesssim 1$ for the case $\alpha = 1$.

<u>Case $0 < \alpha < 1$:</u> we have $0 < r < 1$. If $|Q|^{1/d} \geqslant 1$,

$$|P_r * a(y)| \lesssim \begin{cases} 1/|Q| & \text{if } |y| \lesssim |Q|^{1/d}, \\ 0 & \text{otherwise.} \end{cases}$$

Reasoning as before, if $|x| \gtrsim |Q|^{1/d}$, we have $\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a(x) = 0$. For $|x| \lesssim |Q|^{1/d}$ we have $|y| \lesssim |Q|^{1/d}$ with $(y,r) \in \Gamma_\alpha(x)$, and taking $r < 1$,

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a(x) \lesssim \frac{1}{|Q|},$$

which integrates 1 over that region on x.

Finally, if $|Q|^{1/d} \lesssim 1$, we have again $\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a(x) = 0$ for $|x| \gtrsim 1$. If $|x| \lesssim 1$, we take $(y,r) \in \Gamma_\alpha(x)$ with $r \sim |x|^{1/(1-\alpha)}$ in order to satisfy $|y| \lesssim r$. If $r \sim |x|^{1/(1-\alpha)} \gtrsim |Q|^{1/d}$,

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a(x) \lesssim |x|^{\frac{\alpha d}{1-\alpha}}|Q|^{1/d}|x|^{\frac{-d-1}{1-\alpha}} = |x|^{-d}|x|^{\frac{-1}{1-\alpha}}|Q|^{1/d},$$

which integrates 1 in the region $|Q|^{(1-\alpha)/d} \lesssim |x| \lesssim 1$. In the case $r \sim |x|^{1/(1-\alpha)} \lesssim |Q|^{1/d}$,

$$\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a(x) \lesssim \frac{|x|^{\frac{\alpha d}{(1-\alpha)}}}{|Q|},$$

which integrates 1 in the range $|x| \lesssim |Q|^{(1-\alpha)/d}$. Then $\|\widetilde{\mathcal{M}}_{\alpha,\frac{\alpha d}{2}}a\|_1 \lesssim 1$, and this finishes the case $0 < \alpha < 1$ and the proof of the theorem.

□

2.6 Applications to oscillatory kernels and dispersive PDE

We now discuss some applications of Theorem 2.2.1 and Corollary 2.2.2 in the setting of oscillatory kernels and dispersive partial differential equations.

2.6.1 Oscillatory kernels

An observation of Sjölin [121] using the method of stationary phase allows one to obtain similar pointwise and general-weighted estimates for classes of highly oscillatory convolution kernels. For example we have the following:

Corollary 2.6.1. *For $a > 0$, $a \neq 1$ and $b \geqslant d(1 - \frac{a}{2})$, let $K_{a,b} : \mathbb{R}^d \to \mathbb{C}$ be given by*

$$K_{a,b}(x) = \frac{e^{i|x|^a}}{|x|^b}(1 - \eta(x)),$$

where $\eta \in C_c^\infty(\mathbb{R}^d)$ is such that $\eta(x) = 1$ for all x belonging to a neighbourhood of the origin. Then for any $\lambda > 0$,

$$g_{\alpha,\beta}(K_{a,b} * f)(x) \lesssim g^*_{\alpha,0,\lambda}(f)(x) \tag{2.6.1}$$

and

$$\int_{\mathbb{R}^d} |K_{a,b} * f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\alpha,\beta} M^4 w, \tag{2.6.2}$$

where $\alpha = \frac{a}{a-1}$ and $\beta = \frac{da/2-d+b}{a-1}$.

It is interesting to compare the oscillatory kernels in Corollary 2.6.1 with the kernel associated to the disc multiplier

$$K(x) := \mathcal{F}^{-1}m(x) = c\frac{e^{2\pi i|x|} + e^{-2\pi i|x|} + o(1)}{|x|^{\frac{d+1}{2}}}. \tag{2.6.3}$$

Perhaps remarkably, this kernel takes the form of the missing endpoint case $a = 1$ in Corollary 2.6.1, although it should be noted that the behaviour of these kernels is notoriously discontinuous there.

Proof. In [121] Sjölin establishes that the multiplier $\widehat{K}_{a,b}$ satisfies the Miyachi condition (2.1.1), leading to the conclusion

$$g_{\alpha,\beta}(K_{a,b} * f)(x) \lesssim g^*_{\alpha,0,\lambda}(f)(x)$$

for any $\lambda > 0$, by a direct application of Theorem 2.2.1.

In order to prove (2.6.2), we must force the support condition on the multiplier $\widehat{K}_{a,b}$. We thus choose a function $\varphi \in C^\infty(\mathbb{R}^d)$ such that $\varphi(\xi) = 0$ when $|\xi|^\alpha \leqslant 1$ and $\varphi(\xi) = 1$ when $|\xi|^\alpha \geqslant 2$ and write $\widehat{K}_{a,b} = (1-\varphi)\widehat{K}_{a,b} + \varphi\widehat{K}_{a,b} = m_0 + m_\infty$. The multiplier m_∞ is supported in $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$ and satisfies the Miyachi condition (2.1.1), so Corollary 2.2.2 immediately yields (2.6.2) for T_{m_∞}. The inequality for the portion T_{m_0} follows from a straightforward adaptation of the techniques used in the proof of Theorem 2.4.1. Since $K_0 = \widehat{m}_0$ is a rapidly decreasing function, the Cauchy–Schwarz inequality and Fubini's theorem allow us to write

$$\int_{\mathbb{R}^d} |T_{m_0} f|^2 w \lesssim \|K_0\|_1 \int_{\mathbb{R}^d} |f|^2 |K_0| * w \lesssim \int_{\mathbb{R}^d} |f|^2 A_1^* w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\alpha,\beta} M^4 w,$$

where the last inequality follows from the pointwise bound

$$A_1^* w \lesssim A_1 A_1^* w \lesssim \mathcal{M}_{\alpha,\beta} A_1^* w \lesssim M^2 \mathcal{M}_{\alpha,\beta} M^4 w.$$

$\square$

2.6.2 Dispersive and wave-like equations

The specific multipliers $m_{\alpha,\beta}(\xi) := (1+|\xi|^2)^{-\beta/2}e^{i|\xi|^\alpha}$ yield weighted estimates for the solution $u(x,s) = e^{is(-\Delta)^{\alpha/2}}f(x)$ of the dispersive or wave-like equation

$$\begin{cases} i\partial_s u + (-\Delta)^{\alpha/2}u = 0 \\ u(\cdot,0) = f. \end{cases} \tag{2.6.4}$$

For example, we have the following immediate application.

Corollary 2.6.2. *Let $\alpha \in \mathbb{N}$. Then*

$$\int_{\mathbb{R}^d} |e^{is(-\Delta)^{\alpha/2}}f|^2 w \lesssim \int_{\mathbb{R}^d} |(I - s^{2/\alpha}\Delta)^{\beta/2}f|^2 M^2\mathcal{M}^s_{\alpha,\beta}M^4 w, \tag{2.6.5}$$

where

$$\mathcal{M}^s_{\alpha,\beta}w(x) := \sup_{(y,r)\in\Gamma^s_\alpha(x)} \frac{r^{2\beta}}{|B(y,s^{1/\alpha}r)|}\int_{B(y,s^{1/\alpha}r)} w$$

and

$$\Gamma^s_\alpha(x) = \{(y,r) \in \mathbb{R}^d \times \mathbb{R}_+ : 0 < r \leqslant 1, |x-y| \leqslant s^{1/\alpha}r^{1-\alpha}\}.$$

Of course the case $\alpha = 2$ corresponds to the setting of the free Schrödinger equation. It is interesting to interpret the above weighted estimates in this framework. As it is mentioned at the beginning of this chapter, the maximal operators $\mathcal{M}^s_{\alpha,\beta}$ are highly non-local for $\alpha > 1$, capturing the dispersive nature of the Schrödinger equation.

Corollary 2.6.2 follows from Theorem 2.2.2 via an elementary rescaling argument after noting the scaling identity

$$e^{is(-\Delta)^{\alpha/2}}f(x) = T_{m_{\alpha,\beta}}((I-\Delta)^{\beta/2}f_s)(x/s^{1/\alpha}),$$

where $f_s(x) = f(s^{1/\alpha}x)$. We remark that in order to apply Theorem 2.2.2 to the multiplier

$m_{\alpha,\beta}$ it is necessary to consider its behaviour near the origin separately, as in the proof of Corollary 2.6.1.

One could also obtain weighted estimates for the solution $u(x,s)$ from the corresponding estimates for the Fourier multipliers $\widetilde{m}_{\alpha,\beta} = |\xi|^{-\beta} e^{i|\xi|^\alpha}$, which shall yield estimates in the context of homogeneous Sobolev spaces. Observe that, for $\alpha \in \mathbb{N}$, these multipliers satisfy the estimates (2.1.1) in $\{|\xi|^\alpha \geqslant 1\}$, but $|D^\gamma m(\xi)| \lesssim |\xi|^{-\beta-|\gamma|}$ in $\{|\xi|^\alpha \leqslant 1\}$. We stablish a more general result for multipliers satisfying those differential conditions, that is,

$$|D^\gamma m(\xi)| \lesssim \begin{cases} |\xi|^{-\beta+|\gamma|(\alpha-1)}, & \text{if } \ |\xi|^\alpha \geqslant 1 \\ |\xi|^{-\beta-|\gamma|}, & \text{if } \ |\xi|^\alpha \leqslant 1, \end{cases} \tag{2.6.6}$$

for all $\gamma \in \mathbb{N}^d$ such that $|\gamma| \leqslant \lfloor \frac{d}{2} \rfloor + 1$.

Corollary 2.6.3. *If $m : \mathbb{R}^d \backslash \{0\} \to \mathbb{C}$ satisfies* (2.6.6) *for all $\gamma \in \mathbb{N}^d$ such that $|\gamma| \leqslant \lfloor \frac{d}{2} \rfloor + 1$, then*

$$\int_{\mathbb{R}^d} |T_m f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathfrak{M}_{\alpha,\beta} M^4 w, \tag{2.6.7}$$

where

$$\mathfrak{M}_{\alpha,\beta} w(x) = \sup_{(y,r)\in\Lambda_\alpha(x)} \frac{1}{|B(y,r)|^{1-2\beta/d}} \int_{B(y,r)} w,$$

and

$$\Lambda_\alpha(x) := \{(y,r) \in \mathbb{R}^d \times \mathbb{R}_+ : |x-y| \leqslant r^{1-\alpha}\}.$$

Observe that when $\beta < d/2$, $\mathfrak{M}_{\alpha,\beta}$ satisfies the trivial $L^{d/(2\beta)} \to L^\infty$ bound by a simple application of Hölder's inequality. This observation and Corollary 2.6.3 quickly lead, via the duality argument (1.1.4), to the sharp $L^p \to L^q$ bounds for the class of multipliers satisfying (2.6.6); see Miyachi [98]. Hence for $\alpha \neq 0$, $\mathfrak{M}_{\alpha,\beta}$ necessarily fails to satisfy any other $L^p \to L^q$ inequalities. This is in contrast with the maximal functions $\mathcal{M}_{\alpha,\beta}$ associated with the regions $\Gamma_\alpha(x)$ studied in the previous sections, where $L^p \to L^q$ bounds exist with $q < \infty$ in views of Theorem 2.5.1.

Proof of Corollary 2.6.3. Let $\eta \in C^\infty(\mathbb{R}^d)$ be such that $\eta(\xi) = 0$ when $|\xi|^\alpha \leqslant 1$ and $\eta(\xi) = 1$ when $|\xi|^\alpha \geqslant 2$, and write $m = m_1 + m_2$, with $m_1 = m\eta$ and $m_2 = m(1-\eta)$. As m_1 is supported in $\{|\xi|^\alpha \geqslant 1\}$ and satisfies the Miyachi condition (2.1.1), Theorem 2.2.2 gives

$$\int_{\mathbb{R}^d} |T_{m_1} f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\alpha,\beta} M^4 w.$$

Similarly, the multiplier m_2 satisfies the condition (2.1.1) for $\alpha = 0$, so another application of Theorem 2.2.2 gives

$$\int_{\mathbb{R}^d} |T_{m_2} f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{0,\beta} M^4 w.$$

As the maximal operator $\mathcal{M}_{0,\beta}$ is pointwise comparable to the classical fractional Hardy–Littlewood maximal function of order 2β,

$$M_{2\beta} w(x) = \sup_{r>0} \frac{1}{r^{d-2\beta}} \int_{B(x,r)} w,$$

one trivially has $\mathcal{M}_{0,\beta} \lesssim \mathfrak{M}_{\alpha,\beta}$ for any $\alpha \in \mathbb{R}$. This, together with the obvious $\mathcal{M}_{\alpha,\beta} \leqslant \mathfrak{M}_{\alpha,\beta}$, gives (2.6.7) for m_1 and m_2, from which the result follows. ☐

Of course a straightforward scaling argument leads to the following corollary.

Corollary 2.6.4. *Let $\alpha \in \mathbb{N}$. Then*

$$\int_{\mathbb{R}^d} |e^{is(-\Delta)^{\alpha/2}} f|^2 w \lesssim \int_{\mathbb{R}^d} |(-\Delta)^{\beta/2} f|^2 M^2 \mathfrak{M}^s_{\alpha,\beta} M^4 w, \tag{2.6.8}$$

where

$$\mathfrak{M}^s_{\alpha,\beta} w(x) := \sup_{(y,r)\in\Lambda^s_\alpha(x)} \frac{1}{|B(y,r)|^{1-2\beta/d}} \int_{B(y,r)} w,$$

and

$$\Lambda_\alpha^s(x) := \{(y,r) \in \mathbb{R}^d \times \mathbb{R}_+ : |x-y| \leqslant sr^{1-\alpha}\}.$$

As with the classical fractional maximal operators, $\mathfrak{M}_{\alpha,\beta}$ behaves well on the power weights $w_\gamma(x) := |x|^{-\gamma}$, with $0 \leqslant \gamma < d$. Indeed one may verify that $\mathfrak{M}_{\alpha,\gamma/2} w_\gamma(x) \lesssim 1$ uniformly in $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{R}$, and so Corollary 2.6.4 at $s = 1$ gives

$$\int_{\mathbb{R}^d} |e^{i(-\Delta)^{\alpha/2}} f(x)|^2 |x|^{-\gamma} dx \lesssim \int_{\mathbb{R}^d} |(-\Delta)^{\gamma/4} f(x)|^2 dx. \tag{2.6.9}$$

This special case is somewhat degenerate as the presence of the parameter α is not detected in the estimate. Observe that, alternatively, (2.6.9) may be proved directly by an application of the classical Hardy inequality

$$\int_{\mathbb{R}^d} |h(x)|^2 |x|^{-\gamma} dx \lesssim \int_{\mathbb{R}^d} |(-\Delta)^{\gamma/4} h(x)|^2 dx,$$

followed by the energy conservation identity $\|e^{i(-\Delta)^{\alpha/2}} f\|_2 = \|f\|_2$.

It should be observed that Corollary 2.6.4, combined with the trivial uniform $L^{d/(2\beta)} \to L^\infty$ bound on $\mathfrak{M}_{\alpha,\beta}^s$ allow to recover the elementary sharp homogeneous Strichartz inequality

$$\|e^{is(-\Delta)^{\alpha/2}} f\|_{L_s^\infty L_x^q} \lesssim \|f\|_{\dot{H}^\beta}; \quad \beta = d\left(\frac{1}{2} - \frac{1}{q}\right), \qquad 2 \leqslant q < \infty.$$

A classical prove for the above estimate follows by Sobolev embedding and energy conservation.

Finally, let us interpret Inequality (2.6.8) as a "local energy estimate" that also captures dispersive effects of the propagator $e^{is(-\Delta)^{\alpha/2}}$ via the s-evolution of the region $\Lambda_\alpha^s(x)$. Indeed the sets $\Lambda_\alpha^s(x)$ are increasing in s, so that, in particular

$$\sup_{0<s\leqslant 1} \int_{\mathbb{R}^d} |e^{is(-\Delta)^{\alpha/2}} f|^2 w \lesssim \int_{\mathbb{R}^d} |(-\Delta)^{\beta/2} f|^2 M^2 \mathfrak{M}_{\alpha,\beta} M^4 w, \tag{2.6.10}$$

where $\mathfrak{M}_{\alpha,\beta} := \mathfrak{M}^1_{\alpha,\beta}$. It is interesting to compare this inequality with the weighted maximal estimates in [1] (or [94]) at the interface with geometric measure theory.

It is a very interesting question to determine if, for β beyond some critical threshold, (2.6.10) may be strengthened to

$$\int_{\mathbb{R}^d} \sup_{0<s\leqslant 1} |e^{is(-\Delta)^{\alpha/2}} f|^2 w \lesssim \int_{\mathbb{R}^d} |(-\Delta)^{\beta/2} f|^2 \mathfrak{M}_{\alpha,\beta} w, \tag{2.6.11}$$

modulo suitable factors of M or any other suitable maximal operator $\mathcal{M}$; see for example Rogers and Seeger [116] for related estimates in an unweighted setting. This question seems to be a lot harder due to the nature of the maximal Schrödinger operator, defined by

$$u^*(x) := \sup_{0<s\leqslant 1} |e^{is\Delta} f(x)|.$$

Bounds for this operator are often obtained via Fourier restriction theory; note that $u(x,s) = (\widehat{f} d\mu)\widehat{\ }(x,s)$, where $d\mu$ denotes the parametrised Lebesgue measure on the paraboloid. This question served as a motivation to study certain easier maximal-multiplier operators, which led to the work on the Carleson operator in Chapter 4. Also, we make some remarks on Fourier restriction theory in Chapter 5, in views of attacking the question posed in (2.6.11) in the near future.

Chapter 3

Pseudodifferential operators associated to Hörmander symbol classes

In this chapter we establish general weighted L^2 inequalities for pseudodifferential operators associated to the Hörmander symbol classes $S^m_{\rho,\delta}$. Via such inequalities, we are able to control pseudodifferential operators by maximal functions of the type $\mathcal{M}_{\alpha,\beta}$, previously introduced in Chapter 2. The control by these maximal functions is optimal, as we may recover the sharp $L^p - L^q$ bounds for the symbols classes $S^m_{\rho,\delta}$. Our results apply to the full range of admissible parameters for $S^m_{\rho,\delta}$, that is, $m \in \mathbb{R}$, $0 \leqslant \delta \leqslant \rho \leqslant 1$, $\delta < 1$.

In contrast with the Fourier multiplier case, the weighted inequalities here do not follow from a *pointwise* estimate. The non-translation-invariant nature of the pseudodifferential operators fails to make the g-function approach effective in this case. However, the techniques used still capture the ideas developed in Chapter 2.

The content of this chapter is mostly based on the submitted work [3].

3.1 Weighted control for Hörmander symbol classes

The study of pseudodifferential operators was initiated by Kohn and Nirenberg [77] and Hörmander [69], and it has played a central role in the theory of partial differential equations. Given a smooth function $a \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$, we define the associated pseudodifferential operator T_a by

$$T_a f(x) = \int_{\mathbb{R}^d} e^{ix\cdot\xi} a(x,\xi) \widehat{f}(\xi) d\xi,$$

where $f \in \mathcal{S}(\mathbb{R}^d)$. The smooth function a is typically referred to as the *symbol*. Throughout this chapter, we shall assume that a belongs to the symbol classes $S^m_{\rho,\delta}$, introduced by Hörmander in [69]. Given $m \in \mathbb{R}$ and $0 \leqslant \delta, \rho \leqslant 1$, we say that $a \in S^m_{\rho,\delta}$ if it satisfies the differential inequalities

$$|\partial_x^\nu \partial_\xi^\sigma a(x,\xi)| \lesssim (1+|\xi|)^{m-\rho|\sigma|+\delta|\nu|} \tag{3.1.1}$$

for all multi-indices $\nu, \sigma \in \mathbb{N}^d$.

Of course if a symbol $a(x,\xi)$ is x-independent, T_a is a multiplier operator. Some of the multipliers studied in Chapter 2 may be naturally viewed as symbols. In particular, for $0 \leqslant \alpha \leqslant 1$, if a multiplier m satisfies the differential inequalities (2.1.1) for any multi-index $\gamma \in \mathbb{N}^d$, then $m \in S^{-\beta}_{1-\alpha,0}$. Obvious model examples are the classical multipliers $m_{\alpha,\beta}(\xi) = e^{i|\xi|^\alpha}(1+|\xi|^2)^{-\beta/2}(1-\chi(\xi))$, where χ denotes a smooth cut-off that equals 1 in a neighbourhood of the origin.

In view of the results in Chapter 2 it is natural to ask whether it is possible to obtain analogues for Theorem 2.2.1 and Corollary 2.2.2 in the context of the Hörmander symbol classes $S^m_{\rho,\delta}$. Our main result is a positive answer in the case of the weighted inequalities.

Theorem 3.1.1. *Let $a \in S^m_{\rho,\delta}$, where $m \in \mathbb{R}$, $0 \leqslant \delta \leqslant \rho \leqslant 1$, $\delta < 1$. Then*

$$\int_{\mathbb{R}^d} |T_a f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w \tag{3.1.2}$$

for any weight w, where $\mathcal{M}_{\rho,m} w := \mathcal{M}_{1-\rho,-m} w$.

We introduce the maximal functions $\mathcal{M}_{\rho,m}$ for the ease of notation. Observe that this theorem covers the full range of admissible values for m, ρ, δ except for an endpoint case corresponding to the symbol classes $S^m_{1,1}$. This is to be expected, as it is well known that there are symbols in the class $S^0_{1,1}$ that fail to be bounded on L^2 (see [129]), and thus (3.1.2) would fail on taking $w \equiv 1$.

As discussed at the beginning of this chapter, our approach to proving Theorem 3.1.1 differs from the one adopted for the multiplier case, although some of the main ideas are still present. As is to be expected, the case of pseudodifferential operators adds complexity, and more delicate arguments seem to be required. In particular, appropriate applications of the symbolic calculus and the Cotlar–Stein almost orthogonality principle play important roles. We refer to the end of Section 3.2 for a discussion of the approach taken on the problem, together with an outline of our proof.

The maximal operators $\mathcal{M} = M^2 \mathcal{M}_{\rho,m} M^5$ are *optimal* in (3.1.2). The general mechanism (1.1.4) reveals that if $a \in S^m_{\rho,\delta}$, where $m \in \mathbb{R}$, $0 \leqslant \delta \leqslant \rho \leqslant 1$, $\delta < 1$, the inequality (3.1.2) implies

$$\|T_a\|_{p \to q} \lesssim \|\mathcal{M}_{\rho,m}\|^{1/2}_{(q/2)' \to (p/2)'}.$$

As in the Fourier multiplier case, this allows to transfer $L^p - L^q$ bounds for $\mathcal{M}_{\rho,m}$ to bounds for T_a; in particular the bounds for the maximal operator $\mathcal{M}_{\rho,m}$ obtained in Theorem 2.5.1 allow one to recover the optimal bounds for the symbol classes $S^m_{\rho,\delta}$.

Corollary 3.1.2. *Let $m \in \mathbb{R}$ and $2 \leqslant p \leqslant q < \infty$. Assume $0 \leqslant \delta \leqslant \rho < 1$ and*

$$\frac{-m}{d} \geqslant (1-\rho)\Big(\frac{1}{2}-\frac{1}{p}\Big)+\frac{1}{p}-\frac{1}{q},$$

or $\rho = 1$, $\delta < 1$, and

$$\frac{-m}{d} = \frac{1}{p}-\frac{1}{q}.$$

If $a \in S^m_{\rho,\delta}$, then

$$\|T_a f\|_q \lesssim \|f\|_p.$$

As the class of pseudodifferential operators associated to symbols in a specific class $S^m_{\rho,\delta}$ is closed under adjoints, one also obtains by duality the corresponding Lebesgue space bounds for T_a on the range $1 < q \leqslant p \leqslant 2$.

This corollary is sharp in view of the estimates satisfied by the classical symbol $a_{\rho,m}(\xi) = e^{i|\xi|^{1-\rho}}(1+|\xi|^2)^{m/2}(1-\chi(\xi))$, which fails to be bounded on $L^p(\mathbb{R}^d)$ if $|\frac{1}{p}-\frac{1}{2}| > \frac{m}{d(\rho-1)}$. This recovers well known results on the L^p-boundedness of pseudodifferential operators. Bounds for these operators have been extensively studied, see for instance the work of Calderón and Vaillancourt [18] for the L^2-boundedness of the classes $S^0_{\rho,\rho}$, with $0 \leqslant \rho < 1$, or Hörmander [69], Fefferman [52] or Stein [129] for L^p bounds for the symbol classes $S^m_{\rho,\delta}$. Weighted L^p-boundedness in the context of the A_p Muckenhoupt classes has also been studied, see for example the work of Miller [97], Chanillo and Torchinsky [28], or the most recent work of Michalowski, Rule and Staubach [95, 96]. We note that our Theorem 3.1.1 does not fall beyond the scope of the A_p theory.

We end this discussion with the interesting remark that the maximal operators $\mathcal{M}_{\rho,m}$ are significant improvements of some variants of the Hardy–Littlewood maximal function. In particular, for any $s \geqslant 1$, a crude application of Hölder's inequality reveals that when

$2sm = (\rho - 1)d$,

$$\begin{aligned}
\mathcal{M}_{\rho, \frac{(\rho-1)d}{2s}} w(x) &= \sup_{(y,r) \in \Gamma_{1-\rho}(x)} \frac{1}{r^d r^{\frac{(\rho-1)d}{s}}} \int_{|y-z| \leqslant r} w \\
&\leqslant \sup_{(y,r) \in \Gamma_{1-\rho}(x)} \frac{1}{r^d r^{\frac{(\rho-1)d}{s}}} \Big(\int_{|y-z| \leqslant r} w^s \Big)^{1/s} r^{d(1-\frac{1}{s})} \\
&= \sup_{(y,r) \in \Gamma_{1-\rho}(x)} \Big(\frac{1}{r^{d\rho}} \int_{|y-z| \leqslant r} w^s \Big)^{1/s} \\
&\leqslant \sup_{(y,r) \in \Gamma_{1-\rho}(x)} \Big(\frac{1}{r^{d\rho}} \int_{|y-z| \leqslant r^\rho} w^s \Big)^{1/s} \\
&\leqslant (Mw^s(x))^{1/s}.
\end{aligned}$$

At the level of Lebesgue space bounds, the maximal operators $\mathcal{M}_{\rho,m}$ are bounded on L^s, for $s > 1$, when $2sm = (\rho - 1)d$, a property that the maximal functions $(Mw^s)^{1/s}$ do not enjoy. This allows us to reconcile Theorem 3.1.1 with more classical results in the context of A_p weights. For $s = 1$, we obtain the following.

Corollary 3.1.3. *Let $a \in S^{-d(1-\rho)/2}_{\rho,\delta}$, where $0 \leqslant \delta \leqslant \rho \leqslant 1$, $\delta < 1$. Then*

$$\int_{\mathbb{R}^d} |T_a f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^8 w. \tag{3.1.3}$$

In particular, we may recover the L^2-case of a result of Chanillo and Torchinksy [28], and Michalowski, Rule and Staubach [95], in which it is established that the symbol classes $S^{-d(1-\rho)/2}_{\rho,\delta}$, with $0 < \rho < 1$, are bounded on $L^p(w)$ for $w \in A_{p/2}$ and $2 \leqslant p < \infty$. The inequalities (3.1.3) improve on the existing two-weight inequalities with controlling maximal function $(Mw^s)^{1/s}$, which are implicit in the works [28, 95] from the elementary observation that $(Mw^s)^{1/s} \in A_1$ for any $s > 1$. We remark that in the case of the standard symbol class $S^0 := S^0_{1,0}$ and the classes $S^0_{1,\delta}$, with $\delta < 1$, the inequality (3.1.3) holds with maximal operator M^3; this is a consequence of the inequality (1.1.6) for Calderón–Zygmund operators. We note that the number of compositions of M in (3.1.2) and (3.1.3) does not need to be sharp here; we shall not concern ourselves with such finer points.

3.2 Failure of g-function approach

First of all we observe that the similarity between the differential inequalities (3.1.1) satisfied by the symbols a and those satisfied by the multipliers m (2.1.1) suggests a decomposition of the ξ-space, where ξ corresponds to the frequency variable of f, into $(1-\rho)$-subdyadic balls. However, as T_a is a non-translation-invariant operator, the frequency variables of $T_a f$ and f are not the same. This is manifested, for instance, by the fact that if ψ_B is a bump function adapted to a $(1-\rho)$-subdyadic ball B,

$$T_a(f * \psi_B) \neq T_a f * \psi_B, \tag{3.2.1}$$

in contrast to $T_m(f * \psi_B) = T_m f * \psi_B$. The failure of this property makes the subdyadic square functions $g_{\alpha,\beta}$ not as effective in the setting of pseudodifferential operators, as the decoupling inequality (2.3.4) does not interact well with a subdyadic decomposition at the level of $\widehat{f}$. It is not obvious for us how to adapt the argument in order to make Stein's g-function approach work in this context. Therefore, the weighted estimates (3.1.2) are obtained in a more direct way, and do not follow from a pointwise estimate of the type (1.5.3).

Despite the apparent failure of the g-function approach, it is important to observe the following property from the proof of the decoupling estimate (2.3.4). Let B, B' be subdyadic balls with $r(B) \sim r(B')$ and let f_B, $f_{B'}$ be functions whose Fourier support lies in B and B' respectively. Let $\widetilde{w}$ be a weight function with Fourier support lying in a ball centered at the origin of radius $r(B) \sim r(B')$. Then, Parseval's theorem reveals the orthogonality property

$$\int_{\mathbb{R}^d} f_B \overline{f_{B'}} \widetilde{w} = \int_{\mathbb{R}^d} \widehat{f_B} \widehat{\overline{f_{B'}}} * \widehat{\widetilde{w}} = 0$$

if $\operatorname{dist}(B, B') \gtrsim r(B)$. By translation-invariance, this orthogonality remains valid for $T_m f_B$ and $T_m f_{B'}$, but not for $T_a f_B$ and $T_a f_{B'}$, in view of (3.2.1). In the case of Fourier multipliers only "diagonal" terms contribute to the whole sum, that is,

$$\int_{\mathbb{R}^d} |\sum_{r(B)\sim K} T_m f_B|^2 \widetilde{w} = \int_{\mathbb{R}^d} \sum_{r(B)\sim r(B')\sim K} T_m f_B \overline{T_m f_{B'}} \widetilde{w} \sim \int_{\mathbb{R}^d} \sum_{r(B)\sim K} |T_m f_B|^2 \widetilde{w},$$

and one may thus invoke the elementary Proposition 2.1.2 for the diagonal terms; here K is a suitable fixed scale. The key idea for pseudodifferential operators is that despite $T_a f_B$ and $T_a f_{B'}$ not being orthogonal with respect to the weight $\widetilde{w}$, it is possible to show that

$$\int_{\mathbb{R}^d} T_a f_B \overline{T_a f_{B'}} \widetilde{w} \sim \text{small}$$

if $\operatorname{dist}(B, B') \gtrsim r(B)$, and therefore such "off-diagonal" terms do not significantly contribute to the term

$$\int_{\mathbb{R}^d} |\sum_{r(B)\sim K} T_a f_B|^2 \widetilde{w}.$$

This may be seen as a certain almost orthogonality property between $T_a f_B$ and $T_a f_{B'}$, and for this reason it will be appropriate to make use of the Cotlar–Stein almost orthogonality principle, provided we have a good estimate for the "diagonal terms".

The previous ideas rely on the following observations on the weight $\widetilde{w}$:

- Given a fixed subdyadic ball B, the Fourier support of $\widetilde{w}$ is contained in a ball centered at the origin of radius $r(B) \sim K$. Then, $\widetilde{w}$ is only effective to detect (almost) orthogonality among subdyadic balls B' such that $r(B') \sim r(B) \sim K$.

- Given an arbitrary weight w, we need to find a suitable weight $\widetilde{w}$ satisfying the above properties and controlling the original w.

As the almost orthogonality property depends on the scale K, a first Littlewood–Paley type reduction for the problem seems suitable; observe that if B and B' are subdyadic balls lying on the same dyadic annulus $\{|\xi| \sim 2^k\}$, then $r(B) \sim r(B') \sim 2^{k\rho}$. In contrast to the multiplier case, the weighted Littlewood–Paley theory from Section 1.4 will not suffice for our purposes, and a quantitative version of the *symbolic calculus* will be needed. On each dyadic annulus, we will be able to control w by a suitably band-limited weight $\widetilde{w}$ satisfying the desired properties. Taking a supremum over all dyadic scales will give rise to the maximal operators $\mathcal{M}_{\rho,m}$.

Finally, observe that as $0 \leqslant \rho \leqslant 1$, an $(1-\rho)$-subdyadic decomposition is only suitable in $\{\xi \in \mathbb{R}^d : |\xi| \geqslant 1\}$. This does not represent any obstacle, as the differential inequalities (3.1.1) on $\{|\xi| \leqslant 1\}$ become

$$|\partial_x^\nu \partial_\xi^\sigma a(x,\xi)| \lesssim 1$$

for all multi-indices $\nu, \sigma \in \mathbb{N}^d$. For the portion of a supported in $\{|\xi| \leqslant 1\}$, these differential inequalities will suffice to deduce an appropriate two-weighted inequality for T_a by elementary means.

Outline of the proof

At a very general level, the above ideas may be summarised in the following scheme.

1. Write $a(x,\xi) = a_0(x,\xi) + a_1(x,\xi)$, where a_0 is ξ-supported on $\{|\xi| \lesssim 1\}$ and a_1 is ξ-supported on $\{|\xi| \gtrsim 1\}$, and establish the elementary estimate

$$\int_{\mathbb{R}^d} |T_{a_0} f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w.$$

2. Apply weighted Littlewood–Paley theory to $T_{a_1} f$,

$$\int_{\mathbb{R}^d} |T_{a_1} f|^2 w \lesssim \int_{\mathbb{R}^d} \sum_{k \geqslant 0} |\Delta_k (T_{a_1} f)|^2 M^3 w,$$

where Δ_k is a frequency projection to a dyadic annulus of width 2^k.

3. For every $k \geqslant 0$, majorise the weight $M^3 w$ by a weight $\widetilde{w}_k$ whose Fourier transform is supported in a ball centered at the origin of radius $2^{k\rho}$,

$$\int_{\mathbb{R}^d} |\Delta_k(T_{a_1} f)|^2 M^3 w \lesssim \int_{\mathbb{R}^d} |\Delta_k(T_{a_1} f)|^2 \widetilde{w}_k.$$

4. For every $k \geqslant 0$, use symbolic calculus to "interchange" Δ_k and T_{a_1} provided we introduce some terms of "lower order" and an error term. That is,

$$\Delta_k(T_{a_1} f) = T_{a_1}(\Delta_k f) + \sum_{1 \leqslant |\gamma| < N} T^\gamma(\Delta_k f) + T_{e_k} f,$$

   where T^γ are pseudodifferential operators whose symbols have lower order and T_{e_k} is a pseudodifferential operator associated to a symbol of negative enough order. The decay on e_k allows to easily establish

$$\int_{\mathbb{R}^d} |T_{e_k} f|^2 \widetilde{w}_k \lesssim 2^{-k} \int_{\mathbb{R}^d} |f|^2 M \widetilde{w}_k.$$

   Observe that now T_{a_1} and T^γ are acting on functions f whose Fourier support lies in a dyadic annulus.

5. For every $k \geqslant 0$, establish

$$\int_{\mathbb{R}^d} |T(\Delta_k f)|^2 \widetilde{w}_k \lesssim \int_{\mathbb{R}^d} |\Delta_k f|^2 2^{2km} M \widetilde{w}_k,$$

   for $T = T_{a_1}$ and $T = T^\gamma$. To establish such estimates, we decompose $\Delta_k f = \sum_B f_B$, where B are $(1-\rho)$-subdyadic balls such that $r(B) \sim 2^{k\rho}$, and we establish suitable almost orthogonality estimates for an application of Cotlar–Stein's almost orthogonality principle.

6. Finally, take supremum on $\widetilde{w}_k$ over $k \geqslant 0$ and use weighted Littlewood–Paley theory

to put together the dyadic pieces,

$$\sum_k \int_{\mathbb{R}^d} |\Delta_k f|^2 \widetilde{w}_k \lesssim \int_{\mathbb{R}^d} |f|^2 M(\sup_{k\geqslant 0} 2^{2km} M\widetilde{w}_k).$$

We sum in k the error terms coming from the operators T_{e_k}, which leads to an acceptable term.

Of course the passage of w to $\sup_{k\geqslant 0} 2^{2km} M\widetilde{w}_k$ leads to the maximal operators $\mathcal{M}_{\rho,m} w$, where the approach regions $\Gamma_{1-\rho}$ naturally arise from the subdyadic nature of the operator T_{a_1}. This is reminiscent of the proof of Theorem 2.4.1 in Chapter 2.

We devote the rest of this chapter to make these ideas formal and to provide a proof for Theorem 3.1.1. We start with an auxiliary section that contains several useful lemmas to which we will appeal to through the proof. We note that for the convenience of exposition, the domination of the weight exposed in step 3 has been done in two stages, with the second one incorporated in the inequality in step 5, after a suitable scaling argument.

3.3 Auxiliary results

3.3.1 Symbolic calculus

The composition structure of pseudodifferential operators has been extensively studied; we refer to the work of Hörmander [69] in the case of the symbol classes $S^m_{\rho,\delta}$. We require the following quantitative version when the outermost symbol is a cut-off function on the frequency space adapted to a dyadic annulus.

Theorem 3.3.1. *Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ be such that $\mathrm{supp}(\widehat{\varphi}) \subseteq \{|\xi| \sim 1\}$ and given $R > 1$, let φ_R be defined by $\widehat{\varphi}_R(\xi) := \widehat{\varphi}(R^{-1}\xi)$. Let $a \in S^m_{\rho,\delta}$, where $0 \leqslant \delta \leqslant \rho$ and $\delta < 1$. Then, there exists a symbol $c \in S^m_{\rho,\delta}$ such that*

$$T_c = T_{\widehat{\varphi}_R} \circ T_a.$$

Moreover, for $\epsilon \geqslant 0$ and $\kappa > 0$, the symbol

$$e^N := c - \sum_{|\gamma|<N} \frac{i^{-|\gamma|}}{\gamma!} \partial_\xi^\gamma \widehat{\varphi}_R \partial_x^\gamma a \in S_{\rho,\delta}^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon}$$

for all $N > \frac{d\delta+\kappa\delta+\epsilon}{1-\delta}$, and satisfies

$$|\partial_x^\nu \partial_\xi^\sigma e^N(x,\xi)| \lesssim R^{-\epsilon}(1+|\xi|)^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon-|\sigma|\rho+|\nu|\delta} \tag{3.3.1}$$

for any multi-indices $\nu, \sigma \in \mathbb{N}^d$.

This very specific version of the more general symbolic calculus in [69] allows us to obtain quantitative control for the differential inequalities satisfied by the error term e^N in terms of R, which corresponds to the scale of the frequency projection φ_R. The implicit constants in (3.3.1) depend on finitely many C^k norms of $\widehat{\varphi}$ and on the implicit constants in the differential inequalities (3.1.1) satisfied by a, and they will be acceptable for our purposes for being independent of the parameter R.

We remark that the order of the error symbol e^N in Theorem 3.3.1 is not necessarily sharp here, but one may choose N sufficiently large so that e^N has sufficiently large negative order. Modulo such an error term, we may understand the composition of φ_R with a pseudodifferential operator as the action of the pseudodifferential operator itself, and some other pseudodifferential operators of lower order, on functions with frequency support on the dyadic annulus $\{|\xi| \sim R\}$. We provide the proof of Theorem 3.3.1 in Appendix B for completeness, which consists of a careful modification of the symbolic calculus developed in [129] for the standard symbol classes S^m.

3.3.2 The kernel of a pseudodifferential operator

A pseudodifferential operator with symbol of sufficiently negative order is to all intents and purposes a convolution operator with an integrable kernel. This is an easy consequence of the following observation in Hörmander [69]. Let $a \in S^m_{\rho,\delta}$, $m \in \mathbb{R}, 0 \leqslant \delta, \rho \leqslant 1$, $\delta < 1$ and let $K(x,y)$ denote the distribution kernel of T_a. Then if $\gamma \in \mathbb{N}^d$ satisfies $m - |\gamma|\rho < -d$, the distribution $(x-y)^\gamma K(x,y)$ coincides with a function,

$$(x-y)^\gamma K(x,y) = \int_{\mathbb{R}^d} e^{i(x-y)\cdot\xi}(-iD_\xi)^\gamma a(x,\xi)d\xi. \tag{3.3.2}$$

In view of the differential inequalities (3.1.1), this quickly allows us to deduce that if a symbol $a \in S^m_{\rho,\delta}$ has sufficiently negative order, that is, $m < -d$, then

$$|K(x,y)| \lesssim \frac{1}{(1+|x-y|^2)^{L/2}}$$

for any $L \geqslant 0$. In particular, taking $L > d$, one may control the pseudodifferential operator T_a by a convolution operator with an integrable kernel.

This elementary observation will be very useful to handle the pseudodifferential operator associated with the error symbol e^N obtained after an application of Theorem 3.3.1. Considering the differential inequalities (3.3.1) satisfied by e^N, the identity (3.3.2) reveals that if N is chosen such that $m - N(1-\delta) + d\delta + \kappa\delta + \epsilon < -d$ then the kernel K_{e^N} associated to the symbol e^N satisfies

$$|K_{e^N}(x,y)| \lesssim \frac{R^{-\epsilon}}{(1+|x-y|^2)^{L/2}} \tag{3.3.3}$$

for any $L \geqslant 0$. As in (3.3.1), the implicit constant here is independent of R, and only depends on finitely many C^k norms of $\widehat{\varphi}$ and on the implicit constants in the differential inequalities (3.1.1) satisfied by a. Taking $L > d$, this allows us to bound T_{e^N} by an integrable convolution kernel with a quantitative control of the constant in terms of the scale of the frequency projection φ_R. As we shall see in Section 3.4, such a quantitative control is required for summability purposes in the proof of Theorem 3.1.1.

3.3.3 Almost orthogonality

To obtain a good estimate on each dyadic annulus, we will make use of the Cotlar–Stein almost orthogonality principle.

Lemma 3.3.2 (Cotlar–Stein, [129] p. 280)**.** *Let* $\{T_j\}_{j\in\mathbb{Z}^d}$ *be a family of operators and* $T=\sum_{j\in\mathbb{Z}^d} T_j$*. Let* $\{c(j)\}_{j\in\mathbb{Z}^d}$ *be a family of positive constants such that*

$$A=\sum_{j\in\mathbb{Z}^d} c(j)<\infty$$

and assume that

$$\|T_i^*T_j\|_{2\to 2}\leqslant c(i-j)^2 \quad \text{and} \quad \|T_iT_j^*\|_{2\to 2}\leqslant c(i-j)^2.$$

Then

$$\|T\|_{2\to 2}\leqslant A.$$

3.3.4 L^2-boundedness of integral operators

We also require the following standard version of the Schur test, which is a simple consequence of the Cauchy–Schwarz inequality; see for example Theorem 5.2 in [65].

Lemma 3.3.3 (Schur's test, [65])**.** *Suppose* T *is given by*

$$Tf(x)=\int_{\mathbb{R}^d} K(x,z)f(z)dz$$

and assume there exist measurable functions $h_1,h_2>0$ *and positive constants* C_1 *and* C_2 *such that*

$$\int_{\mathbb{R}^d}|K(x,z)|h_1(z)dz\leqslant C_1h_2(x) \quad \text{and} \quad \int_{\mathbb{R}^d}|K(x,z)|h_2(x)dx\leqslant C_2h_1(z).$$

Then

$$\|T\|_{2\to 2} \leqslant (C_1C_2)^{1/2}.$$

3.4 Proof of Theorem 3.1.1

Let $a \in S^m_{\rho,\delta}$ with $m \in \mathbb{R}$, $0 \leqslant \delta \leqslant \rho \leqslant 1$, $\delta < 1$. By the embeddings of the symbol classes is enough to prove Theorem 3.1.1 for $a \in S^m_{\rho,\rho}$ with $0 \leqslant \rho < 1$, and $a \in S^m_{1,\delta}$ with $\delta < 1$; recall that

$$S^{m_1}_{\rho_1,\delta_1} \subseteq S^{m_2}_{\rho_2,\delta_2} \quad \text{if} \quad m_1 \leqslant m_2,\ \ \rho_1 \geqslant \rho_2,\ \ \delta_1 \leqslant \delta_2.$$

Observe that the upcoming Theorem 3.4.2 is also valid for the symbol classes $S^m_{1,\delta}$ with $\delta < 1$, as they are embedded in $S^m_{1,1}$.

As discussed in Section 3.2, a symbol a satisfying the differential inequalities (3.1.1) behaves differently in the regions $\{|\xi| \leqslant 1\}$ and $\{|\xi| \geqslant 1\}$. Let $\eta \in C^\infty(\mathbb{R}^d)$ be a smooth function supported in $|\xi| \leqslant 2$ and let $a_0(x,\xi) = a(x,\xi)\eta(\xi)$ and a_1 be such that $a = a_0 + a_1$. Theorem 3.1.1 will follow from establishing the required weighted inequalities for both T_{a_0} and T_{a_1}.

In view of (3.1.1), the symbol a_0 satisfies the differential inequalities

$$|\partial_x^\nu \partial_\xi^\sigma a_0(x,\xi)| \lesssim 1$$

for all multi-indices $\nu, \sigma \in \mathbb{N}^d$. Together with the support condition on the variable ξ that we just imposed on $a_0(x,\xi)$, this leads to the following rather elementary weighted inequality.

Proposition 3.4.1.

$$\int_{\mathbb{R}^d} |T_{a_0} f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 A_1^* w,$$

where $A_1^* w = \sup_{t \geqslant 1} A_t$ *and* $A_t w(x) = \frac{1}{|B(x,t)|} \int_{B(x,t)} w$.

We provide a proof of this proposition in Section 3.5. The inequality (3.1.2) for T_{a_0} follows from noting that, as in Chapter 2

$$A_1^* w \lesssim A_1 A_1^* w \lesssim \mathcal{M}_{\rho,m} A_1^* w \lesssim \mathcal{M}_{\rho,m} M w \lesssim M^2 \mathcal{M}_{\rho,m} M^5 w.$$

The difficulty relies thus on understanding the operator T_{a_1}. We reduce the proof of Theorem 3.1.1 to the following theorem, which corresponds to an analogous statement but over the class of functions whose Fourier support lies in a dyadic annulus and whose proof is postponed to Section 3.6.

Theorem 3.4.2. *Let $a \in S^m_{\rho,\rho}$, where $0 \leqslant \rho \leqslant 1$. Let f be a function such that* $\operatorname{supp}(\widehat{f}) \subseteq \{\xi \in \mathbb{R}^d : |\xi| \sim R\}$*, where $R \geqslant 1$. Then*

$$\int_{\mathbb{R}^d} |T_a f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 \mathcal{A}_{\rho,m,R} w$$

uniformly in $R \geqslant 1$, where

$$\mathcal{A}_{\rho,m,R} w(x) := R^{2m} \int_{\mathbb{R}^d} \Big(\sup_{|y-z| \leqslant R^{-\rho}} w(z) \Big) \frac{R^{\rho d}}{(1 + R^{2\rho}|x-y|^2)^{N_0/2}} dy$$

and N_0 is any natural number satisfying $N_0 > d$.

The reduction to Theorem 3.4.2 is done as follows. A first application of Proposition 1.4.2 to the function $T_{a_1} f$ gives

$$\int_{\mathbb{R}^d} |T_{a_1} f|^2 w \lesssim \sum_{k \geqslant 0} \int_{\mathbb{R}^d} |\Delta_k (T_{a_1} f)|^2 M^3 w.$$

Let Φ be a smooth function such that $\widehat{\Phi} = 1$ in $\{\eta \in \mathbb{R}^d : |\eta| \lesssim 1\}$ and define Φ_k by $\widehat{\Phi}_k(\eta) = \widehat{\Phi}(2^{-k}\eta)$ for any $k \geqslant 0$. As $\widehat{\Delta_k g}(\eta) = \widehat{P}(2^{-k}\eta)\widehat{g}(\eta)$ and $\operatorname{supp}(\widehat{P}) \subseteq \{\eta \in \mathbb{R}^d : |\eta| \sim 1\}$, we have $\Delta_k(T_{a_1} f) = \Delta_k(T_{a_1} f) * \Phi_k$, provided the implicit constants are chosen appropriately. An application of the Cauchy–Schwarz inequality and Fubini's theorem gives

$$\int_{\mathbb{R}^d} |\Delta_k(T_{a_1}f)|^2 M^3 w = \int_{\mathbb{R}^d} |\Delta_k(T_{a_1}f) * \Phi_k|^2 M^3 w \lesssim \int_{\mathbb{R}^d} |\Delta_k(T_{a_1}f)|^2 |\Phi_k| * M^3 w \quad (3.4.1)$$

uniformly in $k \geqslant 0$, as the functions Φ_k are normalised on $L^1(\mathbb{R}^d)$.

At this stage, one would like to interchange Δ_k and T_{a_1} in order to apply Theorem 3.4.2. As discussed in Section 3.3.1, this may be done provided we introduce terms of lower order. As $\delta < 1$, fixing $\epsilon > 0$ and $\kappa > 0$, an application of Theorem 3.3.1 for any $k \geqslant 0$ gives

$$\Delta_k(T_{a_1}f) = T_{a_1}(\Delta_k f) + \sum_{1 \leqslant |\gamma| < N} \frac{i^{-|\gamma|}}{\gamma!} T_k^\gamma f + T_{e_k} f,$$

where

$$T_k^\gamma f(x) := \int_{\mathbb{R}^d} e^{ix\cdot\xi} \partial_\xi^\gamma \widehat{P}_k(\xi) \partial_x^\gamma a_1(x,\xi) \widehat{f}(\xi),$$

and e_k is a symbol satisfying

$$|\partial_x^\nu \partial_\xi^\sigma e_k(x,\xi)| \lesssim 2^{-k\epsilon}(1+|\xi|)^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon-|\sigma|\rho+|\nu|\delta}$$

for any multi-indices $\nu, \sigma \in \mathbb{N}^d$. Here $\gamma \in \mathbb{N}^d$, and we choose N to be a positive integer satisfying

$$m - N(1-\delta) + d\delta + \kappa\delta + \epsilon < -d;$$

for ease of notation we remove the dependence of N in the error term e_k, as N is a chosen fixed number independent of k. Such a choice of N allows one to argue as in Section 3.3.2, and the inequality (3.3.3) reads here as

$$|K_{e_k}(x,y)| \lesssim \frac{2^{-k\epsilon}}{(1+|x-y|^2)^{L/2}}$$

for any $L \geqslant 0$. Taking $L > d$, and setting $\Psi^{(L)}(x) := (1+|x|^2)^{-L/2}$, an application of the Cauchy–Schwarz inequality and Fubini's theorem gives

$$\int_{\mathbb{R}^d} |T_{e_k} f|^2 |\Phi_k| * M^3 w \lesssim 2^{-2k\epsilon} \int_{\mathbb{R}^d} |f|^2 \Psi^{(L)} * |\Phi_k| * M^3 w \lesssim 2^{-2k\epsilon} \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w,$$

with implicit constant independent of $k \geqslant 0$; the last inequality follows from the observation that

$$\Psi^{(L)} * |\Phi_k| * M^3 w \lesssim A_1^* M^4 w \lesssim A_1 A_1^* M^4 w \lesssim \mathcal{M}_{\rho,m} A_1^* M^4 w \lesssim M^2 \mathcal{M}_{\rho,m} M^5 w.$$

This is an acceptable bound for each T_{e_k}, as summing over all $k \geqslant 0$ we obtain

$$\sum_{k \geqslant 0} \int_{\mathbb{R}^d} |T_{e_k} f|^2 |\Phi_k| * M^3 w \lesssim \sum_{k \geqslant 0} 2^{-2k\epsilon} \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w$$

for any $\varepsilon > 0$.

For the term corresponding to $T_{a_1}(\Delta_k f)$, we invoke Theorem 3.4.2,

$$\int_{\mathbb{R}^d} |T_{a_1}(\Delta_k f)|^2 |\Phi_k| * M^3 w \lesssim \int_{\mathbb{R}^d} |\Delta_k f|^2 \mathcal{A}_{\rho,m,2^k}(|\Phi_k| * M^3 w) \lesssim \int_{\mathbb{R}^d} |\Delta_k f|^2 M \mathcal{M}_{\rho,m} M^4 w,$$

where the last inequality follows by taking the supremum over all $k \geqslant 0$ on the weight function. Now, one may recouple the dyadic frequency pieces using the standard weighted Littlewood–Paley theory from Proposition 1.4.1,

$$\sum_{k \geqslant 0} \int_{\mathbb{R}^d} |\Delta_k f|^2 M \mathcal{M}_{\rho,m} M^4 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w.$$

Finally, we need to study the terms T_k^γ for $1 \leqslant |\gamma| < N$. Observe that $\partial_\xi^\gamma \widehat{P}_k$ is supported in $\{\xi \in \mathbb{R}^d : |\xi| \sim 2^k\}$ for any $\gamma \in \mathbb{N}^d$, so we are still able to use Theorem 3.4.2 here. To this end, let θ be a smooth function such that $\widehat{\theta}(\xi) = 1$ in $\{\xi \in \mathbb{R}^d : |\xi| \sim 1\}$ and that vanishes outside a slightly enlargement of that set. Let Θ_k be the operator defined by $\widehat{\Theta_k g}(\xi) = \widehat{\theta}_k(\xi)\widehat{g}(\xi)$, where $\widehat{\theta}_k(\xi) = \widehat{\theta}(2^{-k}\xi)$. Then $T_k^\gamma f = T_k^\gamma(\Theta_k f)$, provided the implicit constants are chosen appropriately. Observing that the symbol $\partial_\xi^\gamma \widehat{P}_k(\xi)\partial_x^\gamma a_1(x,\xi) \in S^m_{\rho,\delta}$ uniformly in $k \geqslant 0$ (by embedding of symbol classes), Theorem 3.4.2 leads to

$$\int_{\mathbb{R}^d} |T_k^\gamma f|^2 |\Phi_k| * M^3 w = \int_{\mathbb{R}^d} |T_k^\gamma(\Theta_k f)|^2 |\Phi_k| * M^3 w \lesssim \int_{\mathbb{R}^d} |\Theta_k f|^2 \mathcal{A}_{\rho,m,2^k}(|\Phi_k| * M^3 w)$$

uniformly in $k \geqslant 0$, for every γ such that $1 \leqslant |\gamma| < N$. The sum in γ is not a problem as there is a finite number of terms in that sum, so

$$\sum_{k\geqslant 0} \sum_{1\leqslant|\gamma|<N} \frac{1}{\gamma!} \int_{\mathbb{R}^d} |T_k^\gamma f|^2 |\Phi_k| * M^3 w \lesssim \sum_{k\geqslant 0} \int_{\mathbb{R}^d} |\Theta_k f|^2 \mathcal{A}_{\rho,m,2^k}(|\Phi_k| * M^3 w).$$

For the sum in k we use again standard weighted Littlewood–Paley theory (Proposition 1.4.1) to conclude that

$$\sum_{k\geqslant 0} \int_{\mathbb{R}^d} |\Theta_k f|^2 \mathcal{A}_{\rho,m,2^k}(|\Phi_k| * M^3 w) \leqslant \sum_{k\geqslant 0} \int_{\mathbb{R}^d} |\Theta_k f|^2 M \mathcal{M}_{\rho,m} M^5 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w,$$

where the first inequality follows from taking the supremum in $k \geqslant 0$ in the weight function. Putting the pieces together, we have shown that

$$\int_{\mathbb{R}^d} |T_{a_1} f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 M^2 \mathcal{M}_{\rho,m} M^5 w,$$

and therefore the proof of Theorem 3.1.1 is completed provided we verify the statements of Proposition 3.4.1 and Theorem 3.4.2.

3.5 The part $\{|\xi| \leqslant 1\}$: proof of Proposition 3.4.1

It is crucial to realise that as $a_0(x,\xi)$ has compact support in the ξ variable, we may write

$$T_{a_0} f(x) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i(x-y)\cdot\xi} a_0(x,\xi) f(y) dy d\xi,$$

as the double integral is absolutely convergent. Denoting by K_0 the kernel of T_{a_0},

$$K_0(x,z) = \int_{\mathbb{R}^d} e^{iz\cdot\xi} a_0(x,\xi) d\xi,$$

we may write

$$T_{a_0} f(x) = \int_{\mathbb{R}^d} K_0(x, x-y) f(y) dy,$$

We may interpret T_{a_0} as the convolution of the function $K(x,\cdot)$ with f evaluated at the point x and

$$\int_{\mathbb{R}^d} |T_{a_0} f(x)|^2 w(x) dx \leqslant \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |K_0(x,z)||f(x-z)| dz \Big)^2 w(x) dx.$$

We split the range of integration for the inner integral in two parts, $|z| \leqslant 1$ and $|z| \geqslant 1$. For the first term, the Cauchy-Schwarz inequality, Plancherel's theorem and the estimates on a_0 give

$$\begin{aligned}
\Big(\int_{|z|\leqslant 1} |K_0(x,z)||f(x-z)| dz \Big)^2 &\leqslant \Big(\int_{\mathbb{R}^d} |K_0(x,z)|^2 dz \Big) \Big(\int_{|z|\leqslant 1} |f(x-z)|^2 dz \Big) \\
&\lesssim \Big(\int_{|\xi|\leqslant 2} |a_0(x,\xi)|^2 d\xi \Big) \Big(\int_{|z|\leqslant 1} |f(x-z)|^2 dz \Big) \\
&\lesssim \int_{\mathbb{R}^d} |f(x-z)|^2 \frac{1}{(1+|z|^2)^L} dz.
\end{aligned}$$

Similarly, for the second term,

$$\begin{aligned}\Big(\int_{|z|\geqslant 1}|K_0(x,z)||f(x-z)|dz\Big)^2 &\leqslant \Big(\int_{\mathbb{R}^d}|K_0(x,z)|^2|z|^{2\sigma}dz\Big)\Big(\int_{|z|\geqslant 1}\frac{1}{|z|^{2L}}|f(x-z)|^2dz\Big)\\ &\leqslant \Big(\int_{\mathbb{R}^d}\sum_{|\sigma|=L}|z^\sigma K_0(x,z)|^2dz\Big)\Big(\int_{|z|\geqslant 1}\frac{|f(x-z)|^2}{|z|^{2L}}dz\Big)\\ &\lesssim \Big(\int_{|\xi|\leqslant 2}\sum_{|\sigma|=L}|D_\xi^\sigma a_0(x,\xi)|^2d\xi\Big)\Big(\int_{|z|\geqslant 1}\frac{|f(x-z)|^2}{(1+|z|^2)^L}dz\Big)\\ &\lesssim \int_{\mathbb{R}^d}|f(x-z)|^2\frac{1}{(1+|z|^2)^L}dz,\end{aligned}$$

where $\sigma \in \mathbb{N}^d$ is a multi-index of order L. Putting things together and setting $\Psi^{(2L)}(y) = (1+|y|^2)^{-L}$, Fubini's theorem gives

$$\int_{\mathbb{R}^d}|T_{a_0}f(x)|^2w(x)dx \lesssim \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(x-z)|^2\frac{1}{(1+|z|^2)^L}dzw(x)dx = \int_{\mathbb{R}^d}|f(z)|^2\Psi^{(2L)} * w(z).$$

Proposition 3.4.1 follows from noting that $\Psi^{(2L)} * w \lesssim A_1^* w$ for $L > d/2$.

3.6 The dyadic pieces in $\{|\xi| \geqslant 1\}$

By analogy with the proof provided in [129] for the L^2-boundedness of the symbol classes $S^0_{\rho,\rho}$, with $0 \leqslant \rho < 1$, we reduce Theorem 3.4.2 to a similar statement for the symbol classes $S^0_{0,0}$. As we shall see, this is achieved using Bessel potentials and an elementary scaling argument. For the proof of the weighted inequality for the class $S^0_{0,0}$ we perform an equally spaced decomposition and make an application of the Cotlar–Stein almost orthogonality principle.

3.6.1 Reduction to the symbol classes $S^0_{\rho,\rho}$

It is enough to prove the following version of Theorem 3.4.2 for the symbol classes $S^0_{\rho,\rho}$.

Proposition 3.6.1. *Let $a \in S^0_{\rho,\rho}$, where $0 \leqslant \rho \leqslant 1$. Let f be a function such that*

$\operatorname{supp}(\widehat{f}) \subseteq \{\xi \in \mathbb{R}^d : |\xi| \sim R\}$ *with* $R \geqslant 1$. *Then*

$$\int_{\mathbb{R}^d} |T_a f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 \mathcal{A}_{\rho,0,R} w$$

uniformly in $R \geqslant 1$.

Theorem 3.4.2 follows from the above proposition via the following observation. Let J_m denote the Bessel potential of order m, that is $\widehat{J_m f}(\xi) = (1 + |\xi|^2)^{m/2} \widehat{f}(\xi)$. Then

$$T_a f(x) = \int_{\mathbb{R}^d} e^{ix\cdot\xi} a(x,\xi)(1+|\xi|^2)^{m/2}(1+|\xi|^2)^{-m/2}\widehat{f}(\xi)d\xi = T_{\tilde{a}}(J_m f)(x),$$

where $\widetilde{a}(x,\xi) = a(x,\xi)(1+|\xi|^2)^{-m/2} \in S^0_{\rho,\rho}$. By Proposition 3.6.1

$$\begin{aligned}\int_{\mathbb{R}^d} |T_a f|^2 w \lesssim \int_{\mathbb{R}^d} |T_{\tilde{a}}(J_m f)|^2 w \lesssim \int_{\mathbb{R}^d} |J_m f|^2 \mathcal{A}_{\rho,0,R} w \lesssim \int_{\mathbb{R}^d} |f|^2 R^{2m} \Psi_R^{(L)} * \mathcal{A}_{\rho,0,R} w \\ \lesssim \int_{\mathbb{R}^d} |f|^2 \mathcal{A}_{\rho,m,R} w,\end{aligned}$$

where $\Psi_R^{(L)}(x) := \frac{R^d}{(1+R^2|x|^2)^{L/2}}$ with $L > d$. Here we use that $\Psi_R * \Psi_{R^\rho} \lesssim \Psi_{R^\rho}$; see Lemma A.1 in Appendix A. The penultimate inequality here follows from the following elementary inequality.

Lemma 3.6.2. *Let* f *be such that* $\operatorname{supp}(\widehat{f}) \subseteq \{\xi \in \mathbb{R}^d : |\xi| \sim R\}$ *with* $R \geqslant 1$. *Then*

$$\int_{\mathbb{R}^d} |J_m f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 R^{2m} \Psi_R^{(L)} * w$$

for any $L > d$ *and any weight* w.

The proof of this lemma is very similar to that of Proposition 2.1.2.

Proof. Let φ be a smooth function such that $\widehat{\varphi}(\xi) = 1$ in $\{\xi \in \mathbb{R}^d : |\xi| \sim 1\}$ and that vanishes outside a slightly enlargement of it, and define φ_R by $\widehat{\varphi_R}(\xi) = \widehat{\varphi}(R^{-1}\xi)$. Then, provided the implicit constants are chosen appropriately, $f = f * \varphi_R$. By the Cauchy–Schwarz inequality and Fubini's theorem,

$$\int_{\mathbb{R}^d} |J_m f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 \|J_m \varphi_R\|_1 |J_m \varphi_R| * w. \tag{3.6.1}$$

Observe that

$$\frac{(I-\Delta_\xi)^N}{(1+R^2|x|^2)^N} e^{iRx\cdot\xi} = e^{iRx\cdot\xi},$$

for any $N \geqslant 0$. Using this and integrating by parts,

$$\begin{aligned}
|J_m \varphi_R(x)| &= \Big| \int_{\mathbb{R}^d} e^{ix\cdot\xi} \widehat{\varphi_R}(\xi)(1+|\xi|^2)^{m/2} d\xi \Big| \\
&= \Big| \int_{\mathbb{R}^d} e^{iRx\cdot\xi} \widehat{\varphi}(\xi)(1+R^2|\xi|^2)^{m/2} R^d d\xi \Big| \\
&\leqslant \Big| \int_{\mathbb{R}^d} \frac{e^{iRx\cdot\xi}}{(1+R^2|x|^2)^N} (I-\Delta_\xi)^N [\widehat{\varphi}(\xi)(1+R^2|\xi|^2)^{m/2}] R^d d\xi \Big|
\end{aligned}$$

Now,

$$\begin{aligned}
(I-\Delta_\xi)^N [\widehat{\varphi}(\xi)(1+R^2|\xi|^2)^{m/2}] &= \sum_{k=0}^{N} c_{N,k} (-\Delta_\xi)^k [\widehat{\varphi}(\xi)(1+R^2|\xi|^2)^{m/2}] \qquad (3.6.2) \\
&= \sum_{k=0}^{N} \sum_{k_1+\cdots+k_d=k} c_{N,k} \partial_{\xi_1}^{2k_1} \cdots \partial_{\xi_d}^{2k_d} [\widehat{\varphi}(\xi)(1+R^2|\xi|^2)^{m/2}]
\end{aligned}$$

Given a multiindex $\gamma \in \mathbb{N}^d$,

$$\begin{aligned}
D^\gamma [\widehat{\varphi}(\xi)(1+R^2|\xi|^2)^{m/2}] &= \sum_{|l|\leqslant|\gamma|} c_{\gamma,l} D^{\gamma-l} \widehat{\varphi}(\xi) D^l [(1+R^2|\xi|^2)^{m/2}] \\
&= \sum_{|l|\leqslant|\gamma|} c_{\gamma,l} D^{\gamma-l} \widehat{\varphi}(\xi) (1+R^2|\xi|^2)^{m/2-|l|} R^{2|l|}
\end{aligned}$$

As $|\xi| \sim 1$ and $|D^\gamma \widehat{\varphi}|$ is uniformly bounded for any multiindex $\gamma \in \mathbb{N}^d$,

$$|D^\gamma[\widehat{\varphi}(\xi)(1 + R^2|\xi|^2)^{m/2}]| \lesssim R^m.$$

Using this in (3.6.2),

$$|(I - \Delta_\xi)^N[\widehat{\varphi}(\xi)(1 + R^2|\xi|^2)^{m/2}]| \lesssim R^m,$$

and so

$$|J_m\varphi_R(x)| \lesssim R^m \frac{R^d}{(1 + R^2|x|^2)^N} = R^m \Psi_R^{(L)}(x),$$

setting $L = 2N$. Also,

$$\|J_m\varphi_R\|_1 = \int_{\mathbb{R}^d} |J_m\varphi(x)|dx \lesssim R^m \int_{\mathbb{R}^d} \frac{R^d}{(1 + R^2|x|^2)^{L/2}} dx \lesssim R^m$$

provided $L > d$. Using these estimates in (3.6.1) concludes the proof. □

3.6.2 Reduction to the symbol classes $S^0_{0,0}$

The goal now is to prove Proposition 3.6.1, that is, the special case of Theorem 3.4.2 for the symbol classes $S^0_{\rho,\rho}$. We shall see that, thanks to an elementary scaling argument, this reduces itself to the following specific case for the symbol class $S^0_{0,0}$.

Proposition 3.6.3. *Let $a \in S^0_{0,0}$. Then*

$$\int_{\mathbb{R}^d} |T_a f|^2 w \lesssim \int_{\mathbb{R}^d} |f|^2 \mathcal{A}w,$$

where $\mathcal{A}w := \Psi^{(N_0)} * \widetilde{w}$, $\widetilde{w}(x) := \sup_{|y-x|\leqslant 1} w(y)$ and $\Psi^{(N_0)}(x) = \frac{1}{(1+|x|^2)^{N_0/2}}$ with $N_0 > d$.

To deduce Proposition 3.6.1 from this, let φ be a smooth function such that $\widehat{\varphi}$ equals 1 in $\{\xi \in \mathbb{R}^d : |\xi| \sim 1\}$ and has compact Fourier support in a slightly enlargement of it, and let φ_R be defined by $\widehat{\varphi}_R(\xi) := \widehat{\varphi}(R^{-1}\xi)$. The Fourier support properties of f allows us to write the reproducing formula $\widehat{f} = \widehat{f}\widehat{\varphi}_R$. We may then replace the symbol $a(x,\xi)$ by $a(x,\xi)\widehat{\varphi}_R(\xi)$, which belongs to the class $S^0_{\rho,\rho}$ uniformly in R:

$$
\begin{aligned}
|D_\xi^\sigma[a(x,\xi)\widehat{\varphi_R}(\xi)]| = |\sum_{|l|\leqslant|\sigma|} D^{\sigma-l}a(x,\xi)D^l\widehat{\varphi}(R^{-1}\xi)| &\lesssim \sum_{|l|\leqslant|\sigma|}(1+|\xi|)^{-\rho(|\sigma|-|l|)}R^{-|l|}\\
&\lesssim R^{-\rho(|\sigma|-|l|)}R^{-|l|} = R^{-\rho|\sigma|}R^{-(1-\rho)|l|} \lesssim R^{-\rho|\sigma|}\\
&\sim (1+|\xi|)^{-\rho|\sigma|},
\end{aligned}
$$

as $|\xi| \sim R \gtrsim 1$. For ease of notation, we shall denote the product symbol $a(x,\xi)\widehat{\varphi}_R(\xi)$ by $a(x,\xi)$, but assuming that $a(x,\xi)$ is supported in $\{|\xi| \sim R\}$. Let

$$
\widetilde{a}(x,\xi) := a(R^{-\rho}x, R^\rho\xi).
$$

It is easy to verify from the differential inequalities (3.1.1) and the support property of $a(x,\xi)$ that the new symbol $\widetilde{a}$ belongs to the class $S^0_{0,0}$ uniformly in R:

$$
\begin{aligned}
|D_x^\nu D_\xi^\sigma \widetilde{a}(x,\xi)| &= |D_x^\nu D_\xi^\sigma a(R^{-\rho}x, R^\rho\xi)|\\
&= R^{-\rho|\nu|}R^{\rho|\sigma|}|(D_x^\nu D_\xi^\sigma a)(R^{-\rho}x, R^\rho\xi)|\\
&\lesssim R^{-\rho|\nu|}R^{\rho|\sigma|}(1+R^\rho|\xi|)^{-\rho|\sigma|+\rho|\nu|}\\
&\sim R^{-\rho|\nu|}R^{\rho|\sigma|}R^{-\rho|\sigma|+\rho|\nu|} \sim 1,
\end{aligned}
$$

as $R^\rho|\xi| \sim R$; note that $\widetilde{a}$ is ξ-supported in an annulus of width $O(R^{1-\rho})$. The change of variables $x \mapsto R^{-\rho}x$, $\xi \mapsto R^\rho\xi$ and Proposition 3.6.3 lead to

$$
\int_{\mathbb{R}^d} |T_a f|^2 w = \int_{\mathbb{R}^d} |T_{\widetilde{a}} f_R|^2 w_R \lesssim \int_{\mathbb{R}^d} |f_R|^2 \mathcal{A} w_R
$$

for functions f such that $\mathrm{supp}(\widehat{f}) \subseteq \{|\xi| \sim R\}$, where

$$w_R(x) := w(R^{-\rho}x)R^{-\rho d}$$

and

$$\widehat{f}_R(\xi) := \widehat{f}(R^{\rho}\xi)R^{\rho d}.$$

Proposition 3.6.1 now follows from noting that

$$\mathcal{A}w_R(R^{\rho}x)R^{\rho d} = \mathcal{A}_{\rho,0,R}w(x).$$

This is a consequence of the definitions of $\mathcal{A}$ and $\mathcal{A}_{\rho,0,R}$, along with the following elementary scaling argument. Observe that

$$\begin{aligned}
\widetilde{w_R}(x) &= \sup_{|y-x|\leqslant 1} w_R(y) = \sup_{|R^{\rho}y-x|\leqslant 1} w(y)R^{-\rho d} \\
&= \sup_{|y-R^{-\rho}x|\leqslant R^{-\rho}} w(y)R^{-\rho d} = \widetilde{w}^{R}(R^{-\rho}x)R^{-\rho d} \\
&= (\widetilde{w}^{R})_R(x),
\end{aligned}$$

where $\widetilde{\ }^{R}$ denotes a local supremum at scale $R^{-\rho}$, that is, $\widetilde{w}^{R}(x) = \sup_{|y-x|\leqslant R^{-\rho}} w(y)$. Also,

$$\begin{aligned}
\Psi^{(N)} * w_R(x) &= \int_{\mathbb{R}^d} \Psi^{(N)}(y)w_R(x-y)dy = \int_{\mathbb{R}^d} \Psi^{(N)}(y)w(R^{-\rho d}(x-y))R^{-\rho d}dy \\
&= \int_{\mathbb{R}^d} \Psi^{(N)}(R^{\rho}y)R^{\rho d}w(R^{-\rho d}x-y)R^{-\rho d}dy \\
&= \Psi^{(N)}_{R^{\rho}} * w(R^{-\rho d}x)R^{-\rho d},
\end{aligned}$$

where $\Psi^{(N)}_{R^{\rho}}(x) = \frac{R^{\rho d}}{(1+R^{\rho 2}|x|^2)^{N/2}}$. Then

$$\mathcal{A}w_R(x) = \Psi^{(N)} * \widetilde{w_R}(x) = \Psi^{(N)} * (\widetilde{w}^{R})_R(x) = \Psi^{(N)}_{R^{\rho}} * \widetilde{w}^{R}(R^{-\rho}x)R^{-\rho d},$$

and

$$\mathcal{A}w_R(R^\rho x)R^{\rho d} = \Psi^{(N)}_{R^\rho} * \widetilde{w^R}(x) = \mathcal{A}_{\rho,0,R}w(x),$$

by definition of $\mathcal{A}_{\rho,m,R}$.

3.6.3 The symbol class $S^0_{0,0}$: proof of Proposition 3.6.3

In this section we assume that $a \in S^0_{0,0}$. We first observe that the weight w is pointwise controlled by $\mathcal{A}w$. This is contained in the following lemma, which we borrow from [7]; see [8] for the origins of this. Its short proof is included for completeness.

Lemma 3.6.4 ([7, 8]). $w \lesssim \mathcal{A}w$.

Proof. It is trivial to observe that $w \leqslant \widetilde{w}$, so we only need to show $\widetilde{w} \lesssim \mathcal{A}w$. By translation invariance, it is enough to see that

$$\widetilde{w}(0) \lesssim \mathcal{A}w(0).$$

As $\widetilde{w} \geqslant 0$ and $\Psi^{(N_0)}(y) \gtrsim 1$ for $|y| \leqslant 1$,

$$\mathcal{A}w(0) = \int_{\mathbb{R}^d} \frac{1}{(1+|y|^2)^{N_0/2}}\widetilde{w}(y)dy \gtrsim \int_{|y|\leqslant 1} \widetilde{w}(y)dy.$$

Let $B_1, \ldots, B_{2^d}$ be the intersections of the unit ball with the 2^d coordinate hyperoctants of $\mathbb{R}^d$. It is enough to show that there exists $\ell^* \in \{1, \ldots, 2^d\}$ such that $\widetilde{w}(y) \geqslant \widetilde{w}(0)$ for all $y \in B_{\ell^*}$, as then

$$\mathcal{A}\widetilde{w}(0) \gtrsim \int_{|y|\leqslant 1} \widetilde{w}(y)dy = \int_{B_{\ell^*}} \widetilde{w}(y)dy + \sum_{\ell\neq\ell^*}\int_{B_\ell} \widetilde{w}(y)dy \geqslant |B_{\ell^*}|\widetilde{w}(0) \gtrsim \widetilde{w}(0),$$

which would conclude the proof. We prove our claim by contradiction. Suppose that for each $1 \leqslant \ell \leqslant 2^d$ there exist $y_\ell \in B_\ell$ such that $\widetilde{w}(y_\ell) < \widetilde{w}(0)$. By the definition of $\widetilde{w}$,

$$\sup_{|z-y_\ell|\leqslant 1} w(z) < \widetilde{w}(0) \quad \text{for} \quad 1 \leqslant \ell \leqslant 2^d.$$

As

$$\{|z| \leqslant 1\} \subseteq \bigcup_{\ell=1}^{2^d} \{|z-y_\ell| \leqslant 1\},$$

we have

$$\widetilde{w}(0) = \sup_{|z|\leqslant 1} w(z) \leqslant \sup_{\bigcup_{\ell=1}^{2^d}\{|z-y_\ell|\leqslant 1\}} w(z) = \max_{1\leqslant \ell\leqslant 2^d} \sup_{|z-y_\ell|\leqslant 1} w(z) < \max_{1\leqslant \ell\leqslant 2^d} \widetilde{w}(0) = \widetilde{w}(0),$$

which is of course a contradiction. □

The above lemma reduces the proof of Proposition 3.6.3 to the weighted inequality

$$\int_{\mathbb{R}^d} |T_a f|^2 \mathcal{A}w \lesssim \int_{\mathbb{R}^d} |f|^2 \mathcal{A}w. \tag{3.6.3}$$

Defining the operator $Sf := T_a((\mathcal{A}w)^{-1/2} f)(\mathcal{A}w)^{1/2}$, it is enough to show

$$\int_{\mathbb{R}^d} |Sf|^2 \lesssim \int_{\mathbb{R}^d} |f|^2 \tag{3.6.4}$$

with bounds independent of w; (3.6.3) just follows by taking $f = (\mathcal{A}w)^{1/2} f$ in (3.6.4). Observe first that $(\mathcal{A}w)^\ell$ is a well-defined function for any $\ell \in \mathbb{R}$, as $\mathcal{A}w > 0$. Also, the operator S is well-defined for $f \in \mathcal{S}(\mathbb{R}^d)$; this is due to the fact that any power of $\mathcal{A}w$ has polynomial growth, as well as all its derivatives, see the forthcoming Lemma 3.6.5. Leibniz's formula ensures then that $(\mathcal{A}w)^\ell f \in \mathcal{S}(\mathbb{R}^d)$ for any $\ell \in \mathbb{R}$, and that S maps $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$.

Lemma 3.6.5. *For any $\ell \in \mathbb{R}$ and any $\gamma \in \mathbb{N}^d$,*

$$|D^\gamma (\mathcal{A}w)^\ell(x)| \lesssim (\mathcal{A}w)^\ell(x) \lesssim (1+|x|^2)^{N_0|\ell|/2} (\mathcal{A}w)^\ell(0).$$

Proof. From the trivial fact that $|D^\gamma \Psi^{(N_0)}(x)| \lesssim \Psi^{(N_0)}(x)$ for any $\gamma \in \mathbb{N}^d$, by definition of $\mathcal{A}$ we have

$$|D^\gamma \mathcal{A}w(x)| \leqslant |D^\gamma \Psi^{(N_0)}| * \widetilde{w}(x) \lesssim \Psi^{(N_0)} * \widetilde{w}(x) = \mathcal{A}w(x),$$

as $\widetilde{w} \geqslant 0$. The chain rule quickly reveals

$$|D^\gamma (\mathcal{A}w)^\ell(x)| \lesssim (\mathcal{A}w)^\ell(x).$$

For the second inequality, by Lemma A.2 in Appendix A, one has

$$\mathcal{A}w(0) \frac{1}{(1+|x|^2)^{N_0/2}} \lesssim \mathcal{A}w(x) \lesssim (1+|x|^2)^{N_0/2} \mathcal{A}w(0).$$

Then, if $\ell > 0$, $(\mathcal{A}w)^\ell(x) \lesssim (1+|x|^2)^{N_0\ell/2}(\mathcal{A}w)^\ell(0)$, and if $\ell < 0$, $(\mathcal{A}w)^\ell(x) \lesssim (1+|x|^2)^{N_0|\ell|/2}(\mathcal{A}w)^\ell(0)$, which concludes the proof. □

We shall prove the L^2-boundedness of the operator S from an application of the Cotlar–Stein principle to a suitable family of operators. To construct such a family we introduce the following partition of unity. Let ψ be a smooth, nonnegative function supported in the unit cube $Q_0 = \{x \in \mathbb{R}^d : |x_j| \leqslant 1\}$ and such that

$$\sum_{i \in \mathbb{Z}^d} \psi(x - i) = 1, \tag{3.6.5}$$

and let $a_{\mathbf{i}}(x,\xi) = a(x,\xi)\psi(x-i)\psi(\xi-i')$, where $\mathbf{i} = (i,i') \in \mathbb{Z}^{2d}$. Then

$$a = \sum_{\mathbf{i}\in\mathbb{Z}^{2d}} a_{\mathbf{i}}.$$

This gives a decomposition of the space associated to the ξ variable into balls of radius $O(1)$. Note that in the passage of rescaling the symbol class $S^0_{0,0}$ into $S^0_{\rho,\rho}$, this amounts to a decomposition of the dyadic annulus $\{|\xi| \sim R\}$ into $O(R^{(1-\rho)d})$ balls of radius $O(R^\rho)$; this would correspond to the prototypical Example 2.3.1 of a $(1-\rho)$-subdyadic decomposition. We remark that the decomposition given by ψ was used in the proof of the L^2-boundedness of the class $S^0_{0,0}$ that one may find in [129].

This decomposition allows us to write the operator S as

$$Sf = \sum_{\mathbf{i}\in\mathbb{Z}^{2d}} S_{\mathbf{i}}f,$$

where $S_{\mathbf{i}}f = T_{a_{\mathbf{i}}}((\mathcal{A}w)^{-1/2}f)(\mathcal{A}w)^{1/2}$. We aim to apply Lemma 3.3.2 to the family of operators $\{S_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{Z}^{2d}}$. To this end we need to establish

$$\|S_{\mathbf{i}}^*S_{\mathbf{j}}\|_{2\to 2} \lesssim c(\mathbf{i}-\mathbf{j})^2$$

and

$$\|S_{\mathbf{i}}S_{\mathbf{j}}^*\|_{2\to 2} \lesssim c(\mathbf{i}-\mathbf{j})^2$$

for a family of constants $\{c(\mathbf{i})\}_{\mathbf{i}\in\mathbb{Z}^{2d}}$ such that

$$\sum_{\mathbf{i}\in\mathbb{Z}^{2d}} c(\mathbf{i}) < \infty.$$

Observe that $S_{\mathbf{i}}^* f = (\mathcal{A}w)^{-1/2} T_{a_{\mathbf{i}}}^*((\mathcal{A}w)^{1/2} f)$, where

$$T_{a_{\mathbf{i}}}^* g(y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i\xi\cdot(y-z)} \overline{a}_{\mathbf{i}}(z,\xi) g(z) d\xi dz$$

is a well-defined operator that maps $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$. The decomposition of the x variable via (3.6.5) ensures the kernel of the operator $S_{\mathbf{i}}^* S_{\mathbf{j}}$ to be well defined; also the symmetric role of the x and ξ variables in $a(x,\xi) \in S_{0,0}^0$ suggests such a decomposition in the x variable.

The L^2-boundedness of $S_{\mathbf{i}}^* S_{\mathbf{j}}$

The operator $S_{\mathbf{i}}^* S_{\mathbf{j}}$ may be realised as

$$\begin{aligned} S_{\mathbf{i}}^*(S_{\mathbf{j}} f)(x) &= (\mathcal{A}w)^{-1/2}(x) T_{a_{\mathbf{i}}}^*(\mathcal{A}w\, T_{a_{\mathbf{j}}}((\mathcal{A}w)^{-1/2} f))(x) \\ &= (\mathcal{A}w)^{-1/2}(x) \int_{\mathbb{R}^d} K_{\mathbf{i},\mathbf{j}}(x,z) f(z) (\mathcal{A}w)^{-1/2}(z) dz, \end{aligned}$$

where

$$K_{\mathbf{i},\mathbf{j}}(x,z) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i\xi\cdot(x-y)} e^{i\eta\cdot(y-z)} \overline{a_{\mathbf{i}}}(y,\xi) a_{\mathbf{j}}(y,\eta) \mathcal{A}w(y) dy d\eta d\xi.$$

The kernel $K_{\mathbf{i},\mathbf{j}}$ is well-defined by the support properties of $a_{\mathbf{i}}$ and $a_{\mathbf{j}}$. Note that if $i-j \notin Q_0$, then $K_{\mathbf{i},\mathbf{j}} = 0$.

Integrating by parts in $K_{\mathbf{i},\mathbf{j}}$, after making use of the identities

$$(I - \Delta_y)^{N_1} e^{iy\cdot(\eta-\xi)} = (1 + |\xi - \eta|^2)^{N_1} e^{iy\cdot(\eta-\xi)},$$

$$(I - \Delta_\eta)^{N_2} e^{i\eta\cdot(y-z)} = (1 + |y - z|^2)^{N_2} e^{i\eta\cdot(y-z)}$$

and

$$(I - \Delta_\xi)^{N_3} e^{i\xi\cdot(x-y)} = (1 + |x - y|^2)^{N_3} e^{i\eta\cdot(x-y)},$$

leads to

$$K_{\mathbf{i},\mathbf{j}}(x,z) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\xi\cdot(x-y)} e^{i\eta\cdot(y-z)} \frac{(I-\Delta_\xi)^{N_3}}{(1+|x-y|^2)^{N_3}} \frac{(I-\Delta_\eta)^{N_2}}{(1+|y-z|^2)^{N_2}} \Big[\frac{(I-\Delta_y)^{N_1}}{(1+|\xi-\eta|^2)^{N_1}} (\overline{a_{\mathbf{i}}}(y,\xi) a_{\mathbf{j}}(y,\eta)\mathcal{A}w(y)) \Big] dy d\eta d\xi,$$

for any $N_1, N_2, N_3 \geqslant 0$. Observe that $|D^\gamma \psi(y-k)| \leqslant \|\psi\|_{C^{|\gamma|}} \chi(y-k)$ for any multi-index $\gamma \in \mathbb{N}^d$, where χ is the characteristic function of Q_0. This, Lemma 3.6.5, which ensures that $|D^\gamma(\mathcal{A}w)| \lesssim \mathcal{A}w$, and the differential inequalities $|D_x^\nu D_\xi^\sigma a(x,\xi)| \lesssim 1$ for $a = a_{\mathbf{i}}, a_{\mathbf{j}}$, allows us to deduce, after an application of Leibniz's formula,

$$\begin{aligned} |K_{\mathbf{i},\mathbf{j}}(x,z)| &\lesssim \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\chi(\xi-i')\chi(\eta-j')}{(1+|\xi-\eta|^2)^{N_1}} d\xi d\eta \int_{\mathbb{R}^d} \frac{\mathcal{A}w(y)\chi(y-i)\chi(y-j)}{(1+|y-z|^2)^{N_2}(1+|y-x|^2)^{N_3}} dy \\ &\lesssim \frac{1}{(1+|i'-j'|^2)^{N_1}} \int_{\mathbb{R}^d} \frac{\mathcal{A}w(y)\chi(y-i)\chi(y-j)}{(1+|y-z|^2)^{N_2}(1+|y-x|^2)^{N_3}} dy; \end{aligned} \tag{3.6.6}$$

the implicit constant here depends on finitely many C^k norms of ψ. Now we apply Schur's test to the kernel

$$\widetilde{K_{\mathbf{i},\mathbf{j}}}(x,z) = K_{\mathbf{i},\mathbf{j}}(x,z)(\mathcal{A}w)^{-1/2}(x)(\mathcal{A}w)^{-1/2}(z)$$

with the auxiliary functions $h_1 = h_2 = (\mathcal{A}w)^{1/2}$. We check first that the integral condition with respect to z is satisfied. Observe that from Lemma A.1 in Appendix A, $(\mathcal{A}w)*\Psi^{(N_0)} \lesssim \mathcal{A}w$. Using this, and taking $2N_2 = 2N_3 = N_0 > d$ in (3.6.6), we have

$$\begin{aligned} \int_{\mathbb{R}^d} |\widetilde{K_{\mathbf{i},\mathbf{j}}}(x,z)| h_1(z) dz &\lesssim \frac{(\mathcal{A}w)^{-1/2}(x)}{(1+|i'-j'|^2)^{N_1}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\mathcal{A}w(y)\chi(y-i)\chi(y-j)}{(1+|y-z|^2)^{N_2}(1+|y-x|^2)^{N_3}} dz dy \\ &\lesssim \frac{(\mathcal{A}w)^{-1/2}(x)}{(1+|i'-j'|^2)^{N_1}} \int_{\mathbb{R}^d} \frac{\mathcal{A}w(y)}{(1+|y-x|^2)^{N_3}} dy \\ &\lesssim \frac{(\mathcal{A}w)^{1/2}(x)}{(1+|i'-j'|^2)^{N_1}}, \quad \text{if } i-j \in Q_0, \end{aligned}$$

for any $N_1 \geqslant 0$. On the other hand, $\widetilde{K_{\mathbf{i},\mathbf{j}}} = 0$ if $i - j \notin Q_0$, so combining both cases,

$$\int_{\mathbb{R}^d} |\widetilde{K_{\mathbf{i},\mathbf{j}}}(x,z)| h_1(z) dz \lesssim \frac{(\mathcal{A}w)^{1/2}(x)}{(1 + |\mathbf{i} - \mathbf{j}|^2)^{N_1}},$$

for any $N_1 \geqslant 0$. As the integral condition with respect to the x variable is symmetric, Lemma 3.3.3 yields

$$\|S_{\mathbf{i}}^* S_{\mathbf{j}}\|_{2\to 2} \lesssim \frac{1}{(1 + |\mathbf{i} - \mathbf{j}|^2)^{N_1}} \tag{3.6.7}$$

for any $N_1 \geqslant 0$. The constant $c(\mathbf{i}) = (1 + |\mathbf{i}|^2)^{-N_1/2}$ will be sufficient for an application of the Cotlar–Stein lemma.

The L^2-boundedness of $S_{\mathbf{i}} S_{\mathbf{j}}^*$

Our goal now is to see that $\|S_{\mathbf{i}} S_{\mathbf{j}}^*\|_{2\to 2}$ also satisfies the bound (3.6.7). The operator $S_{\mathbf{i}} S_{\mathbf{j}}^*$ may be realised as

$$\begin{aligned}
S_{\mathbf{i}}(S_{\mathbf{j}}^* f)(x) &= (\mathcal{A}w)^{-1}(y) T_{a_{\mathbf{j}}}^*((\mathcal{A}w)^{1/2} f)(y) \\
&= (\mathcal{A}w)^{1/2}(x) \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i\xi\cdot(x-y)} a_{\mathbf{i}}(x,\xi) (\mathcal{A}w)^{-1}(y) T_{a_{\mathbf{j}}}^*(f(\mathcal{A}w)^{1/2})(y) \psi(y-k) dy d\xi \\
&= (\mathcal{A}w)^{1/2}(x) \int_{\mathbb{R}^d} L_{\mathbf{i},\mathbf{j}}(x,z) f(z) (\mathcal{A}w)^{1/2}(z) dz,
\end{aligned}$$

where $L_{\mathbf{i},\mathbf{j}}$ is taken to be the formal sum

$$L_{\mathbf{i},\mathbf{j}}(x,z) := \sum_{k\in\mathbb{Z}^d} L_{\mathbf{i},\mathbf{j}}^k(x,z) \tag{3.6.8}$$

and

$$L_{\mathbf{i},\mathbf{j}}^k(x,z) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i\xi\cdot(x-y)} e^{i\eta\cdot(y-z)} a_{\mathbf{i}}(x,\xi) \overline{a_{\mathbf{j}}}(z,\eta) (\mathcal{A}w)^{-1}(y) \psi(y-k) dy d\xi d\eta.$$

Observe that, a priori, the formal sum

$$L_{\mathbf{i},\mathbf{j}}(x,z) = \sum_{k\in\mathbb{Z}^d} L^k_{\mathbf{i},\mathbf{j}}(x,z) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\xi\cdot(x-y)} e^{i\eta\cdot(y-z)} a_{\mathbf{i}}(x,\xi)\overline{a_{\mathbf{j}}}(z,\eta)(\mathcal{A}w)^{-1}(y)dyd\xi d\eta,$$

may not be well-defined, as the triple integral in the right hand side does not necessarily converge absolutely. For this reason, we introduce the partition of unity (3.6.5) in the y variable; the integral that defines $L^k_{\mathbf{i},\mathbf{j}}$ is now absolutely convergent. Our analysis below shows, in particular, that such sum is finite.

Again, integration by parts with respect to y,η,ξ gives

$$\begin{aligned} L_{\mathbf{i},\mathbf{j}}(x,z) = \sum_{k\in\mathbb{Z}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\xi\cdot(x-y)} e^{i\eta\cdot(y-z)} \frac{(I-\Delta_\xi)^{N_3}}{(1+|x-y|^2)^{N_3}}\frac{(I-\Delta_\eta)^{N_2}}{(1+|y-z|^2)^{N_2}}\Big[\\ \frac{a_{\mathbf{i}}(x,\xi)\overline{a_{\mathbf{j}}}(z,\eta)}{(1+|\eta-\xi|^2)^{N_1}}\Big](I-\Delta_y)^{N_1}\big((\mathcal{A}w)^{-1}(y)\psi(y-k)\big)dyd\xi d\eta, \end{aligned}$$

for any $N_1,N_2,N_3 \geqslant 0$. The same observations as in the previous case allows us to deduce, after an application of Leibniz's formula,

$$|L_{\mathbf{i},\mathbf{j}}(x,z)| \lesssim \sum_{k\in\mathbb{Z}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\chi(x-i)\chi(\xi-i')}{(1+|x-y|^2)^{N_3}}\frac{\chi(z-j)\chi(\eta-j')}{(1+|y-z|^2)^{N_2}}\frac{(\mathcal{A}w)^{-1}(y)\chi(y-k)}{(1+|\eta-\xi|^2)^{N_1}}dyd\xi d\eta.$$

As the functions $\{\chi(\cdot-k)\}_{k\in\mathbb{Z}^d}$ have bounded overlap, we may sum in the k variable and

$$\begin{aligned} |L_{\mathbf{i},\mathbf{j}}(x,z)| &\lesssim \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{\chi(\xi-i')\chi(\eta-j')}{(1+|\eta-\xi|^2)^{N_1}}d\xi d\eta\int_{\mathbb{R}^d}\frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}}\frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_3}}dy \\ &\lesssim \frac{1}{(1+|i'-j'|^2)^{N_1}}\int_{\mathbb{R}^d}\frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}}\frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_3}}dy. \end{aligned} \tag{3.6.9}$$

The integration in the y variable is finite, so the sum taken in the definition of $L_{\mathbf{i},\mathbf{j}}$ in

(3.6.8) is well defined. In particular, for $N_2 = N_3 > N_0 + d$, it is possible to show that

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_2}} dzdy \lesssim \frac{(\mathcal{A}w)^{-1}(x)}{(1+|i-j|^2)^{N_2/2}}. \tag{3.6.10}$$

As the role of the variables x and z is symmetric here, the same follows with $(\mathcal{A}w)^{-1}(x)$ replaced by $(\mathcal{A}w)^{-1}(z)$ in the right hand side of (3.6.10).

Assuming the estimate (3.6.10) is true, one may successfully apply Schur's test to the kernel

$$\widetilde{L}_{\mathbf{i},\mathbf{j}}(x,z) = L_{\mathbf{i},\mathbf{j}}(x,z)(\mathcal{A}w)^{1/2}(x)(\mathcal{A}w)^{1/2}(z)$$

with the auxiliary functions $h_1 = h_2 = (\mathcal{A}w)^{-1/2}$. Using (3.6.9) and (3.6.10), we have

$$\begin{aligned}\int_{\mathbb{R}^d} |\widetilde{L}_{\mathbf{i},\mathbf{j}}(x,z)|h_1(z)dz &\lesssim \frac{(\mathcal{A}w)^{1/2}(x)}{(1+|i'-j'|^2)^{N_1}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_2}} dydz \\ &\lesssim \frac{(\mathcal{A}w)^{1/2}(x)}{(1+|i'-j'|^2)^{N_1}} \frac{(\mathcal{A}w)^{-1}(x)}{(1+|i-j|)^{N_2/2}} \\ &\lesssim \frac{(\mathcal{A}w)^{-1/2}(x)}{(1+|\mathbf{i}-\mathbf{j}|^2)^{N_2/2}},\end{aligned}$$

for $N_2 > N_0 + d$; the last inequality follows from taking $N_1 = N_2/2$. As the integral condition with respect to the x variable is symmetric, an application of Lemma 3.3.3 yields

$$\|S_{\mathbf{i}}S_{\mathbf{j}}^*\|_{2\to 2} \lesssim \frac{1}{(1+|\mathbf{i}-\mathbf{j}|^2)^{N_2/2}}$$

for any $N_2 > N_0 + d$.

The L^2-boundedness of S

We just saw that the family of operators $\{S_{\mathbf{i}}\}_{i\in\mathbb{Z}^{2d}}$ satisfies the bounds

$$\|S_{\mathbf{i}}^*S_{\mathbf{j}}\|_{2\to 2} \lesssim \frac{1}{(1+|\mathbf{i}-\mathbf{j}|^2)^{N_1}}, \tag{3.6.11}$$

for any $N_1 \geqslant 0$, and

$$\|S_{\mathbf{i}}S_{\mathbf{j}}^*\|_{2\to 2} \lesssim \frac{1}{(1+|\mathbf{i}-\mathbf{j}|^2)^{N_2/2}}$$

for any $N_2 > N_0 + d$. Taking $N_1 = N_2/2$ in (3.6.11) and noting that the series

$$\sum_{\mathbf{i}\in\mathbb{Z}^{2d}} \frac{1}{(1+|\mathbf{i}|^2)^{N_2/4}} < \infty$$

for $N_2/2 > 2d$, an application of the Cotlar–Stein almost orthogonality principle (Lemma 3.3.2) to the family of operators $\{S_{\mathbf{i}}\}_{i\in\mathbb{Z}^{2d}}$ ensures that

$$\|S\|_{2\to 2} \leqslant \sum_{\mathbf{i}\in\mathbb{Z}^{2d}} \frac{1}{(1+|\mathbf{i}|^2)^{N_2/4}} < \infty,$$

provided $N_2 > \max\{N_0 + d, 4d\}$. As we may choose N_2 as large as we please, the estimate (3.6.4) follows. This finishes the proof of Theorem 3.1.1, provided the estimate (3.6.10) is shown to be true.

The validity of the estimate (3.6.10)

At this stage we are only left with proving (3.6.10), that is

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_2}} dzdy \lesssim \frac{(\mathcal{A}w)^{-1}(x)}{(1+|i-j|^2)^{N_2/2}}.$$

To this end, we divide the range for the y-integration into two half-spaces, H_x and H_z, that contain the points x and z respectively and that are the result of splitting $\mathbb{R}^d$ by a hyperplane perpendicular to the line segment joining x and z at its midpoint. Note that for $y \in H_x$, $|y-z| \geqslant \frac{1}{2}|x-z|$, so

$$\frac{1}{(1+|y-z|^2)^{N_2}} \leqslant \frac{2^{2N_2}}{(1+|x-z|^2)^{N_2}}$$

and

$$\int_{\mathbb{R}^d}\int_{H_x} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_2}} dydz$$
$$\lesssim \int_{\mathbb{R}^d} \frac{\chi(z-j)\chi(x-i)}{(1+|x-z|^2)^{N_2}} dz \int_{H_x} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|x-y|^2)^{N_2}} dy$$
$$\lesssim \frac{1}{(1+|i-j|^2)^{N_2/2}} \int_{\mathbb{R}^d} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|x-y|^2)^{N_2/2}} dy.$$

Similarly, for $y \in H_z$, $|x-y| \geqslant \frac{1}{2}|x-z|$, so

$$\frac{1}{(1+|x-y|^2)^{N_2}} \leqslant \frac{2^{2N_2}}{(1+|x-z|^2)^{N_2}}$$

and

$$\int_{\mathbb{R}^d}\int_{H_z} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_2}} dydz$$
$$\lesssim \int_{\mathbb{R}^d}\int_{H_z} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-z|^2)^{N_2}} dydz.$$

By the elementary inequality

$$\frac{1}{(1+|y-z|^2)^{N_2/2}} \frac{1}{(1+|x-z|^2)^{N_2/2}} \lesssim \frac{1}{(1+|x-y|^2)^{N_2/2}},$$

which is a simple consequence of the triangle inequality, we have

$$\int_{\mathbb{R}^d}\int_{H_z} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-z|^2)^{N_2}} dydz$$
$$\lesssim \int_{\mathbb{R}^d} \frac{\chi(z-j)\chi(x-i)}{(1+|x-z|^2)^{N_2/2}} dz \int_{H_z} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|x-y|^2)^{N_2/2}} dy$$
$$\lesssim \frac{1}{(1+|i-j|^2)^{N_2/2}} \int_{\mathbb{R}^d} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|x-y|^2)^{N_2/2}} dy.$$

Putting both estimates together,

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(\mathcal{A}w)^{-1}(y)}{(1+|y-z|^2)^{N_2}} \frac{\chi(z-j)\chi(x-i)}{(1+|x-y|^2)^{N_2}} dzdy \lesssim \frac{(\mathcal{A}w)^{-1} * \Psi^{(N_2)}(x)}{(1+|i-j|^2)^{N_2/2}},$$

so the inequality (3.6.10) is satisfied if

$$(\mathcal{A}w)^{-1} * \Psi^{(N_2)}(x) \lesssim (\mathcal{A}w)^{-1}(x).$$

As $\widetilde{w} \geqslant 0$, by Lemma A.2,

$$\Psi^{(N_0)} * \widetilde{w}(x) \geqslant \frac{1}{(1+|x-y|^2)^{N_0/2}} \Psi^{(N_0)} * \widetilde{w}(y),$$

so by definition of $\mathcal{A}w$,

$$(\mathcal{A}w)^{-1}(x) \leqslant (1+|x-y|^2)^{N_0/2} (\mathcal{A}w)^{-1}(y);$$

in particular

$$(\mathcal{A}w)^{-1}(x-y) \leqslant (1+|y|^2)^{N_0/2} (\mathcal{A}w)^{-1}(x).$$

Thus

$$\begin{aligned}
(\mathcal{A}w)^{-1} * \Psi^{(N_2)}(x) &= \sum_{l\in\mathbb{Z}^d} \int_{l+[0,1]^d} (\mathcal{A}w)^{-1}(x-y)\Psi^{(N_2)}(y)dy \\
&\leqslant \sum_{l\in\mathbb{Z}^d} (\mathcal{A}w)^{-1}(x) \int_{l+[0,1]^d} (1+|y|^2)^{(N_0-N_2)/2} dy \\
&\lesssim (\mathcal{A}w)^{-1}(x) \sum_{l\in\mathbb{Z}^d} (1+|l|^2)^{(N_0-N_2)/2} \\
&\lesssim (\mathcal{A}w)^{-1}(x),
\end{aligned}$$

provided $N_2 > N_0 + d$, and the inequality (3.6.10) follows.

3.7 Towards a pointwise estimate: a sparse approach

As is mentioned at the beginning of this chapter, the weighted inequalities obtained for the pseudodifferential operators T_a, with $a \in S^m_{\rho,\delta}$, do not follow from a *pointwise* estimate of the type (1.5.3). In this final section, we explore if any such pointwise estimates could be obtained. Indeed, for some specific classes $S^m_{\rho,\delta}$ pointwise estimates have been proved through the Fefferman–Stein sharp maximal function. For example, Chanillo and Torchinksy [28] showed that if $a \in S^{d(\rho-1)/2}_{\rho,\delta}$, $0 < \rho < 1$, $\delta < \rho$, and $f \in C_0^\infty(\mathbb{R}^d)$, then

$$M^{\#}(T_a f)(x) \lesssim M_2 f(x),$$

and more recently, Michalowski, Rule and Staubach [96] showed that if $a \in S^{d(\rho-1)}_{\rho,\delta}$, $0 < \rho \leqslant 1$, $0 \leqslant \delta < 1$, and $f \in C_0^\infty(\mathbb{R}^d)$, then, for any $s > 1$,

$$M^{\#}(T_a f)(x) \lesssim M_s f(x).$$

Of course, these pointwise estimates led to weighted results in the context of A_p weights through the corresponding weighted estimates on $M^{\#}$ and M_s.

As discussed in Section 1.3, in recent years there have been refinements of the above type of pointwise estimates when the operator under study is a Calderón–Zygmund operator, with the auxiliary operator on the left entirely absent and the auxiliary operator on the right being a dyadic sparse operator. Our goal here is to explore if any domination by dyadic sparse operators is possible for the symbol classes satisfying the above estimates, that is $S^{d(\rho-1)/2}_{\rho,\delta}$, or $S^{d(\rho-1)}_{\rho,\delta}$, with the respective restrictions on δ and ρ. We have an affirmative result for the latter symbol classes.

Proposition 3.7.1. *Let $a \in S^{d(\rho-1)}_{\rho,\delta}$, $0 < \rho \leqslant 1$, $0 \leqslant \delta < 1$. Then for every $f \in C_0^\infty(\mathbb{R}^d)$*

and any $r > 1$, there exists a sparse family $\mathcal{S}$ such that for a.e. $x \in \mathbb{R}^d$,

$$|T_a f(x)| \lesssim \mathcal{A}_{r,\mathcal{S}} f(x).$$

This allows to recover the $L^p(w)$ boundedness of T_a for $1 < p < \infty$ and $w \in A_p$ established in [96] through the corresponding boundedness of the operators $\mathcal{A}_{r,\mathcal{S}}$. In particular, one may obtain quantitative control on the A_p characteristic of the weight, $[w]_{A_p}$, applying a result of Di Plinio and Lerner [42] on the operators $\mathcal{A}_{r,\mathcal{S}}$. We also note that the forthcoming two-weight inequality in Theorem 4.4.1 applies to this context.

Proposition 3.7.1 is a consequence of Lerner's sparse domination Theorem 1.3.4. In order to apply that theorem we will need some good decay bounds on the kernel associated to T_a, which, as in Section 3.5, is defined by

$$K(x,z) := \int_{\mathbb{R}^d} e^{i\xi\cdot z} a(x,\xi) d\xi.$$

If $\rho > 0$ or $m < -d$, it satisfies

$$|K(x,z)| \lesssim |z|^{-N} \quad \text{for any} \quad N > 0 \quad \text{and} \quad |x-z| \geqslant 1. \tag{3.7.1}$$

For points around the diagonal, K satisfies the following Hörmander-type estimate; see Michalowski, Rule and Staubach [96] or a prior result of Chanillo and Torchinsky [28].[1]

Lemma 3.7.2 ([96])**.** *Let $a \in S^m_{\rho,\delta}$, $0 \leqslant \delta \leqslant 1$, $0 < \rho \leqslant 1$. Then for $|x - x_B| \leqslant r \leqslant 1$, $\theta \in [0,1]$, $p \in [1,2]$, $\frac{m}{\rho} + \frac{d}{p\rho} < l < \frac{m}{\rho} + \frac{d}{p\rho} + \frac{1}{\rho}$, $\frac{1}{2} < c_1 < 2c_2 < \infty$ and $k \geqslant 1$, the following*

[1] We note that the pointwise estimate using the sharp maximal function makes use of these estimates on the kernel.

estimate holds:

$$\Big(\int_{c_1 2^k r^\theta<|y-x_B|<c_2 2^{k+1}r^\theta} |K(x,x-y)-K(x_B,x_B-y)|^{p'}dy\Big)^{1/p'} \lesssim 2^{-kl} r^{l(\rho-\theta)-m-\frac{d}{p}}.$$

We proceed now with the proof of Proposition 3.7.1.

Proof of Proposition 3.7.1. As $a \in S^{d(\rho-1)}_{\rho,\delta}$, we have that T_a is bounded on L^p for $1 < p < \infty$ and is of weak-type $(1,1)$, see for example [129]. In order to apply Theorem 1.3.4, we only need to verify that the grand maximal function $\mathcal{N}_{T_a}$ is of weak-type (r,r) for any $r > 1$; we indeed show that it is bounded on L^r for $r > 1$. Many of the following ideas are quite standard, and may be found, for instance, in [96].

Given a point x and a cube $Q \ni x$, we distinguish two cases, $|Q| \leqslant 1$ and $|Q| > 1$. The latter case is easy to deal with, as we may use the decay of the kernel away from the diagonal, that is (3.7.1). Given $z \in Q$,

$$\begin{aligned}
|T_a(f\chi_{\mathbb{R}^d\setminus 3Q})(z)| &= \Big|\int_{\mathbb{R}^d\setminus 3Q} K(z,z-y)f(y)dy\Big| \\
&\leqslant \sum_{k=0}^{\infty}\int_{2^{k+1}3Q\setminus 2^k(3Q)} |K(z,z-y)||f(y)|dy \\
&\lesssim \sum_{k=0}^{\infty}\int_{2^{k+1}3Q\setminus 2^k(3Q)} |z-y|^{-d-\varepsilon}|f(y)|dy \\
&\lesssim \sum_{k=0}^{\infty}\frac{1}{(2^k\mathrm{diam}(Q))^\varepsilon}\frac{1}{(2^k\mathrm{diam}(Q))^d}\int_{2^{k+1}3Q}|f(y)|dy \\
&\lesssim \sum_{k=0}^{\infty} 2^{-k\varepsilon}Mf(x) \\
&\lesssim Mf(x),
\end{aligned}$$

where we explicitly use that $\mathrm{diam}(Q) > 1$.

If $|Q| \leqslant 1$, one needs to be slightly more subtle. For any $z, x' \in Q$,

$$\begin{aligned}
|T_a(f\chi_{\mathbb{R}^d\backslash 3Q})(z)| &\leqslant |T_a(f\chi_{\mathbb{R}^d\backslash 3Q})(z) - T_a(f\chi_{\mathbb{R}^d\backslash 3Q})(x')| + |T_a(f)(x')| + |T_a(f\chi_{3Q})(x')| \\
&= \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}$$

For the terms II and III we shall use the L^r-boundedness of T_a for $1 < r < \infty$. To deal with I, we use Lemma 3.7.2,

$$\begin{aligned}
I &= \Big| \int_{\mathbb{R}^d\backslash 3Q} (K(z, z-y) - K(x', x'-y))f(y)dy \Big| \\
&\leqslant \sum_{k=0}^{\infty} \int_{2^{k+1}(3Q)\backslash 2^k(3Q)} |K(z, z-y) - K(x', x'-y)||f(y)|dy \\
&\leqslant \sum_{k=0}^{\infty} \Big(\int_{2^{k+1}(3Q)\backslash 2^k(3Q)} |f(y)|^p dy \Big)^{1/p} \Big(\int_{2^{k+1}(3Q)\backslash 2^k(3Q)} |K(z, z-y) - K(x', x'-y)|^{p'} dy \Big)^{1/p'} \\
&\lesssim \sum_{k=0}^{\infty} 2^{-kl} 2^{kd/p} \mathrm{diam}(Q)^{l(\rho-1)-m} \Big(\frac{1}{2^{kd}\mathrm{diam}(Q)^d} \int_{2^{k+1}(3Q)} |f(y)|^p dy \Big)^{1/p} \\
&\lesssim \sum_{k=0}^{\infty} 2^{-kl} 2^{kd/p} \mathrm{diam}(Q)^{l(\rho-1)-m} M_p f(x) \\
&\lesssim M_p f(x),
\end{aligned}$$

provided $l - d/p > 0$ and $l(\rho-1) - m \geqslant 0$. As $m = d(\rho-1)$, this means we require $d/p < l \leqslant d$. In order to apply Lemma 3.7.2 we also require

$$d - \frac{d}{\rho} + \frac{d}{p\rho} < l < d - \frac{d}{\rho} + \frac{d}{p\rho} + \frac{1}{\rho}.$$

So we need to check the admissibility of the following condition

$$\max\Big\{\frac{d}{p}, d - \frac{d}{\rho} + \frac{d}{p\rho}\Big\} < l < \min\Big\{d, d - \frac{d}{\rho} + \frac{d}{p\rho} + \frac{1}{\rho}\Big\}.$$

Clearly $\frac{d}{p} \geqslant d - \frac{d}{\rho} + \frac{d}{p\rho}$ for $0 < \rho \leqslant 1$, and as $\frac{d}{p} < d$ for $p > 1$, we only need to check whether the condition

$$\frac{d}{p} < d - \frac{d}{\rho} + \frac{d}{p\rho} + \frac{1}{\rho}$$

is admissible. This is equivalent to

$$0 < d(1 - \frac{1}{\rho})(1 - \frac{1}{p}) + \frac{1}{\rho},$$

and given a fixed ρ, this is true for a p sufficiently close to 1.

So, in all,

$$|T_a(f\chi_{\mathbb{R}^d\setminus 3Q})(z)| \lesssim Mf(x) + M_pf(x) + |T_af(x')| + |T_a(f\chi_{3Q})(x')|,$$

for any $x' \in Q$.

Raising the above estimate to a power $1 < s < r$, integrating with respect to $x' \in Q$, and raising it again to the power $1/s$,

$$\begin{aligned} |T_a(f\chi_{\mathbb{R}\setminus 3Q})(z)| &\lesssim Mf(x) + M_pf(x) + \Big(\frac{1}{|Q|}\int_Q |T_af(x')|^s dx'\Big)^{1/s} \\ &\quad + \Big(\frac{1}{|Q|}\int_Q |T_a(f\chi_{3Q})(x')|^s dx'\Big)^{1/s} \\ &\lesssim Mf(x) + M_pf(x) + M_s(T_af)(x) + \|T_a\|_s \Big(\frac{1}{|Q|}\int_{3Q} |f(x')|^s dx'\Big)^{1/s} \\ &\lesssim Mf(x) + M_pf(x) + M_s(T_af)(x) + \|T_a\|_s M_sf(x), \end{aligned}$$

where we have taken supremum over all $Q \ni x$ and used the boundedness of T_a in L^s. Thus,

$$\mathcal{N}_{T_a}f(x) \lesssim Mf(x) + M_pf(x) + M_s(T_af)(x) + \|T_a\|_s M_sf(x).$$

Taking L^r-norms, with $r > \max(p, s)$,

$$\|\mathcal{N}_{T_a} f\|_r \lesssim \big(\|M\|_r + \|M\|_{r/p} + \|M\|_{r/s}\|T_a\|_r + \|T_a\|_s\|M\|_{r/s}\big)\|f\|_r.$$

As p and s may be chosen arbitrarily close to 1, $\mathcal{N}_{T_a}$ is bounded on L^r for any $r > 1$. Thus, an application of Theorem 1.3.4 yields

$$|T_a f(x)| \lesssim A_{r,\mathcal{S}} f(x).$$

□

In the case of the symbol classes $S^m_{\rho,\delta}$ with $m < d(\rho - 1)$, we may indeed improve the sparse domination given by Proposition 3.7.1.

Proposition 3.7.3. *Let $a \in S^m_{\rho,\delta}$, with $m < d(\rho - 1)$, $0 < \rho \leqslant 1$, $0 \leqslant \delta < 1$. Then for every $f \in C_0^\infty(\mathbb{R}^d)$, there exists a sparse family $\mathcal{S}$ such that for a.e. $x \in \mathbb{R}^d$,*

$$|T_a f(x)| \lesssim \mathcal{A}_{1,\mathcal{S}} f(x).$$

As T_a is dominated by the sparse operators $\mathcal{A}_{1,\mathcal{S}}$, one may of course also recover the boundedness in $L^p(w)$ for $w \in A_p$ and $1 < p < \infty$ for such symbol classes, and obtain quantitative bounds in terms of the A_p characteristic of the weight, see for instance [84].

It is possible to prove this pointwise control using the local mean oscillation decomposition formula from Theorem 1.3.5. By embedding of the symbol classes, it is enough to prove if for $m = d(\rho - 1) - \varepsilon$, for $0 < \varepsilon \ll 1$ arbitrarily small. We have that

$$\omega_\lambda(T_a f; Q) \lesssim \frac{1}{|\bar{Q}|}\int_{\bar{Q}} |f| + \sum_{k=0}^{\infty} \frac{1}{2^{k\tau}}\Big(\frac{1}{|2^k Q|}\int_{2^k Q} |f|\Big)$$

for some $\tau > 0$. To see this, write $f = f^0 + f^\infty$, where $f^0 = f\chi_{\bar{Q}}$. Denoting by c_Q the center of the cube Q,

$$|T_a f(x) - T_a f^\infty(c_Q)| \leqslant |T_a f^0(x)| + |T_a f^\infty(x) - T_a f^\infty(c_Q)|,$$

and choosing $c = T_a f^\infty(c_Q)$ in the definition of $\omega_\lambda(T_a f; Q)$,

$$\omega_\lambda(T_a f; Q) \leqslant ((T_a f^0)\chi_Q)^*(\lambda|Q|) + \sup_{x \in Q} |T_a f^\infty(x) - T_a f^\infty(c_Q)|.$$

For the second term one may proceed as in the proof of Proposition 3.7.1. In this case we use Lemma 3.7.2 with $p = 1$. As $m = d(\rho - 1) - \varepsilon$, choosing $l = d + \tau$, with $0 < \tau < \min\{\frac{1-\varepsilon}{\rho}, \frac{\varepsilon}{1-\rho}\}$, if $0 < \rho < 1$, $0 < \tau < \varepsilon$ if $\rho = 0$, and $0 < \tau < 1 - \varepsilon$ if $\rho = 1$, which is admissible, it is easy to see that

$$|T_a f^\infty(x) - T_a f^\infty(c_Q)| \leqslant \sum_{k=0}^{\infty} \frac{1}{2^{k\tau}} \Big(\frac{1}{|2^k Q|} \int_{2^k Q} |f| \Big).$$

Then one only needs to check that

$$((T_a f^0)\chi_Q)^*(\lambda|Q|) \lesssim \frac{1}{|\bar{Q}|} \int_{\bar{Q}} |f|.$$

But this follows from the fact that T_a is of weak-type $(1, 1)$, since

$$((T_a f^0)\chi_Q)^*(\lambda|Q|) = \inf \left\{ s > 0 : d_{(T_a f^0)\chi_Q}(s) \leqslant \lambda|Q| \right\},$$

where $d_{(T_a f^0)\chi_Q}$ is the distribution function of $(T_a f^0)\chi_Q$ and

$$d_{(T_a f^0)\chi_Q}(s) = |\{x \in Q : |T_a f^0| > s\}| \leqslant \frac{1}{s} \int_{\bar{Q}} |f|.$$

Hence for

$$s \geqslant \frac{1}{\lambda}\frac{1}{|Q|}\int_{\bar{Q}}|f|$$

we have $d_{(T_a f^0)\chi_Q}(s) \leqslant \lambda|Q|$, and taking s to be the infimum over the above quantities

$$((T_a f^0)\chi_Q)^*(\lambda|Q|) = \frac{1}{\lambda}\frac{1}{|Q|}\int_{\bar{Q}}|f| \lesssim \frac{1}{|\bar{Q}|}\int_{\bar{Q}}|f|.$$

Now, fixing a cube Q_0 by Theorem 1.3.5,

$$|T_a f(x) - m_{T_a f}(Q_0)| \lesssim \sum_{Q\in\mathcal{S}}\Big(\frac{1}{|\bar{Q}|}\int_{\bar{Q}}|f|\Big)\chi_Q(x) + \sum_{Q\in\mathcal{S}}\sum_{k=0}^{\infty}\frac{1}{2^{k\tau}}\Big(\frac{1}{|2^k Q|}\int_{2^k Q}|f|\Big)\chi_Q(x), \quad (3.7.2)$$

for *a.e.* $x \in Q_0$, where $\mathcal{S} \subset \mathcal{D}(Q_0)$ is a sparse family. From (3.7.2) and observing that $|m_{T_a f}(Q_0)| \leqslant \frac{\|T_a f\|_{L^{1,\infty}(Q_0)}}{|Q_0|}$ and T_a is of weak-type $(1,1)$, one may proceed as Conde–Alonso and Rey [32] to deduce Proposition 3.7.3 for *a.e.* $x \in Q_0$. A trick of Lerner in [86] allows the passage from *a.e.* $x \in Q_0$ to *a.e.* $x \in \mathbb{R}^d$. We omit such details here.

Chapter 4

The Carleson operator

Motivated by the study of maximal-multiplier operators, we obtain sharp pointwise and weighted inequalities for the Carleson operator $\mathcal{C}$. In particular, we prove that

$$|\mathcal{C}f(x)| \leqslant C\mathcal{A}_{r,\mathcal{S}}f(x)$$

for any $r > 1$, where $\mathcal{A}_{r,\mathcal{S}}$ is the sparse operator defined in (1.3.3), and

$$\int_{\mathbb{R}} |\mathcal{C}f|^p w \leqslant C \int_{\mathbb{R}} |f|^p M^{\lfloor p \rfloor + 1} w,$$

for any weight w and $1 < p < \infty$. These results are obtained using the sparse operator approach developed by Lerner and others, and presented in Section 1.3, together with the theory of Orlicz maximal functions from Section 1.2. Indeed, we deduce the above results for a broad class of maximally modulated Calderón–Zygmund operators which encompasses the classical Calderón–Zygmund operators and the Carleson operator. The above weighted inequalities are the counterparts to those of Pérez [107] for Calderón–Zygmund operators; see (1.1.6). We also present more general two-weight inequalities in Section 4.6.

Most of the content of this chapter may be found in the work [4], which has been accepted for publication.

4.1 Motivation

On a general level, given a Fourier multiplier m and writing $m_t(\xi) := m(t\xi)$ for any $t > 0$, one may define its associated maximal multiplier operator as

$$T_m^* f(x) = \sup_{t>0} |(m_t \widehat{f})^{\vee}(x)|.$$

For a fixed multiplier m, one may hope to identify a maximal function $\mathcal{M}_m$ so that

$$\int_{\mathbb{R}^d} |T_m^* f|^p w \lesssim \int_{\mathbb{R}^d} |f|^p \mathcal{M}_m w \tag{4.1.1}$$

for some $1 < p < \infty$. Answering this question in the setting of the multipliers $m_{\alpha,\beta}$ would give results for the maximal Schrödinger operator, a central operator in partial differential equations. This question was raised at the end of Chapter 2; see (2.6.11).

In general, this might be quite a difficult problem, as it shall evidence our next example. As discussed previously in this thesis, a precedent for Corollary 2.2.2 is the one-dimensional variation-based result of Bennett [7]. Taking such perspective, multipliers of global bounded variation on the real line may be seen to fall under the class $\alpha = \beta = 0$ of multipliers considered in Section 2.1, and they constitute one of the easiest example in that class. Motivated by establishing an inequality of the type (4.1.1) for the multipliers $m_{\alpha,\beta}$, we consider the analogous question but associated to a multiplier m of global bounded variation on the real line.

The essence of the classical Marcinkiewicz multiplier theorem is the observation that such a multiplier often satisfies the same norm inequalities as the Hilbert transform. In particular, if T_m denotes the associated operator to a multiplier m of global bounded variation, one may deduce

$$\int_{\mathbb{R}} |T_m f|^p w \leqslant C \int_{\mathbb{R}} |f|^p M^{\lfloor p \rfloor + 1} w$$

for any weight w; this follows merely from the analogous result for the Hilbert transform (1.1.6) after a suitable application of Minkowski's inequality. Similarly, one may see that results for T_m^* follow from those for the Carleson operator. Write

$$m(t\xi) = \int_{-\infty}^{t\xi} dm(u) = \int_{\mathbb{R}} \chi_{(-\infty,t\xi)}(u) dm(u) = \int_{\mathbb{R}} \chi_{(u,\infty)}(t\xi) dm(u) = \int_{\mathbb{R}} \chi_{(u/t,\infty)}(\xi) dm(u),$$

where dm denotes the Lebesgue-Stieltjes measure associated to m. Defining $S_{(u/t,\infty)}$ as the operator associated to the multiplier $\chi_{(u/t,\infty)}$,

$$T_m^t f(x) = \int_{\mathbb{R}} S_{(u/t,\infty)} f(x) dm(u),$$

and

$$\begin{aligned} T_m^* f(x) = \sup_{t>0} |T_m^t f(x)| &\leqslant \int_{\mathbb{R}} \sup_{t>0} |S_{(u/t,\infty)} f(x)| |dm|(u) \leqslant \int_{\mathbb{R}} (|f(x)| + \mathcal{C} f(x)) |dm|(u) \\ &\lesssim |f(x)| + \mathcal{C} f(x), \end{aligned}$$

as the integral of $|dm|$ is the total variation of m. Here $\mathcal{C}$ denotes the Carleson operator, defined as

$$\mathcal{C} f(x) = \sup_{\alpha \in \mathbb{R}} \left| \text{p. v.} \int_{\mathbb{R}} \frac{e^{2\pi i \alpha y}}{x-y} f(y) dy \right|. \tag{4.1.2}$$

This elementary example evidences the difficulty of studying maximal multiplier operators, as $\mathcal{C}$ is a much more complicated operator than the Hilbert transform, the underlying operator behind T_m. Of course pointwise and weighted estimates for T_m^* follow from those for $\mathcal{C}$. In particular, we are able to obtain the following.

Theorem 4.1.1. *Let $\mathcal{C}$ be the Carleson operator. Then for any $1 < p < \infty$ there is a constant $C < \infty$ such that for every weight w*

$$\int_{\mathbb{R}} |\mathcal{C}f|^p w \leqslant C \int_{\mathbb{R}} |f|^p M^{\lfloor p \rfloor + 1} w. \tag{4.1.3}$$

We remark that weighted inequalities for the Carleson operator have been previously studied by many authors. Hunt and Young [73] established the $L^p(w)$ boundedness of $\mathcal{C}$ for $1 < p < \infty$ and $w \in A_p$, from which a two weight inequality with controlling maximal operator M_s, with $s > 1$, follows. Later, Grafakos, Martell and Soria [58] gave new weighted inequalities for weights in A_∞, as well as vector-valued inequalities for $\mathcal{C}$. More recently, Do and Lacey [43] gave weighted estimates for a variation norm version of $\mathcal{C}$ in the context of A_p theory that strengthened the results in [73]. Indeed, sparse control and sharp weighted norm inequalities for variational Carleson have been obtained by Di Plinio, Do and Uraltsev [41] only a few months ago. Finally, Di Plinio and Lerner [42] obtained $L^p(w)$ bounds for $\mathcal{C}$ in terms of the $[w]_{A_q}$ constants for $1 \leqslant q \leqslant p$. Note that inequality (4.1.3) does not fall within the scope of the classical A_p theory.

4.2 Maximally modulated Calderón-Zygmund operators

We shall prove a more general version of Theorem 4.1.1 that holds for a broad class of maximally modulated Calderón-Zygmund operators studied previously by Grafakos, Martell and Soria [58], and Di Plinio and Lerner [42]. Let $\Phi = \{\phi_\alpha\}_{\alpha \in A}$ be a family of real-valued mesurable functions indexed by an arbitrary set A and let T be a Calderón–Zygmund operator in $\mathbb{R}^d$. The maximally modulated Calderón-Zygmund operator T^Φ is defined by

$$T^\Phi f(x) = \sup_{\alpha \in A} |T(\mathcal{M}^{\phi_\alpha} f)(x)|, \tag{4.2.1}$$

where $\mathcal{M}^{\phi_\alpha} f(x) = e^{2\pi i \phi_\alpha(x)} f(x)$. We will consider operators T^Φ such that for some $r_0 > 1$ satisfy the *a priori* weak-type inequalities

$$\|T^\Phi f\|_{r,\infty} \lesssim \psi(r)\|f\|_r \tag{4.2.2}$$

for $1 < r \leqslant r_0$, where $\psi(r)$ is a function that captures the dependence of the operator norm on r. This definition is motivated by the Carleson operator, since it may be recovered from (4.2.1) by setting $T = H$ and Φ to be the family of functions given by $\phi_\alpha(x) = \alpha x$ for $\alpha \in \mathbb{R}$. We note that simply by taking $\phi_\alpha \equiv 0$ for all α, one recovers the classical Calderón–Zygmund operators.

Implicit in the work of Di Plinio and Lerner [42] there is the following analogue of the estimate (1.1.3) for maximally modulated Calderón-Zygmund operators.[1]

Theorem 4.2.1. *Let T^Φ be a maximally modulated Calderón-Zygmund operator satisfying (4.2.2). Then for any $s > 1$ and $1 < p < \infty$ there is a constant $C < \infty$ such that for any weight w*

$$\int_{\mathbb{R}^d} |T^\Phi f|^p w \leqslant C \int_{\mathbb{R}^d} |f|^p M_s w. \tag{4.2.3}$$

Following our discussion in Section 1.1 and the remark that yields (1.1.5), for any fixed $1 < p < \infty$ and $1 < s < 2$, the operator M_s is not a sharp controlling maximal operator. One may address the question of obtaining optimal control for T^Φ. Combining the ideas developed by Pérez in [107, 108] with Di Plinio and Lerner's argument [42], we obtain the following, which constitutes the main result of this chapter.

Theorem 4.2.2. *Let T^Φ be a maximally modulated Calderón-Zygmund operator satisfying (4.2.2). Then for any $1 < p < \infty$ there is a constant $C < \infty$ such that for any weight w*

$$\int_{\mathbb{R}^d} |T^\Phi f|^p w \leqslant C \int_{\mathbb{R}^d} |f|^p M^{\lfloor p \rfloor + 1} w. \tag{4.2.4}$$

[1]This result may be seen as a consequence of the A_∞ theory in [58].

This is best possible in the sense that $\lfloor p \rfloor + 1$ *cannot be replaced by* $\lfloor p \rfloor$.

As it is well known that the Carleson operator $\mathcal{C}$ satisfies the condition (4.2.2), Theorem 4.1.1 follows from this more general statement. Of course, Theorem 4.2.2 extends the estimate (1.1.6) for Calderón–Zygmund operators. As observed for (1.1.6), given $1 < p < \infty$, the control given by the maximal operator $M^{\lfloor p \rfloor + 1}$ is optimal here.

Indeed Theorem 4.2.2 may be viewed as a corollary of a more precise statement, that allows one to replace $M^{\lfloor p \rfloor + 1}$ by a sharper class of maximal operators. This strategy builds up on the work of Pérez [107] for the case of unmodulated Calderón-Zygmund operators, involving Young functions A and their associated Orlicz maximal functions M_A.

Theorem 4.2.3. *Let* T^Φ *be a maximally modulated Calderón-Zygmund operator satisfying* (4.2.2) *and* $1 < p < \infty$. *Suppose that* A *is a doubling Young function satisfying*

$$\int_c^\infty \left(\frac{t}{A(t)} \right)^{p'-1} \frac{dt}{t} < \infty \tag{4.2.5}$$

for some $c > 0$. *Then there is a constant* $C < \infty$ *such that for any weight* w

$$\int_{\mathbb{R}^d} |T^\Phi f|^p w \leqslant C \int_{\mathbb{R}^d} |f|^p M_A w. \tag{4.2.6}$$

In the unmodulated setting, Pérez [107] pointed out that condition (4.2.5) is necessary for (4.2.6) to hold for the Riesz transforms. Hence it also becomes a necessary condition for Theorem 4.2.3 to be stated in such a generality, characterizing the class of Young functions for which (4.2.6) holds.

4.3 Control by sparse operators

It was observed in [42] that maximally modulated Calderón–Zygmund operators satisfying the weak-type condition (4.2.2) are controlled, in Banach space norm, by the sparse operators $\mathcal{A}_{r,\mathcal{S}}$. The equivalent to Theorem 1.3.3 in this case is the following.

Proposition 4.3.1 ([42]). *Let X be a Banach function space over $\mathbb{R}^d$ equipped with Lebesgue measure. Let T^Φ be a maximally modulated Calderón-Zygmund operator satisfying* (4.2.2). *Then*

$$\|T^\Phi f\|_X \lesssim \inf_{1<r\leqslant r_0} \left\{ \psi(r) \sup_{\mathcal{D},\mathcal{S}} \|\mathcal{A}_{r,\mathcal{S}} f\|_X \right\},$$

where the supremum is taken over all dyadic grids $\mathcal{D}$ and all sparse families $\mathcal{S} \subset \mathcal{D}$.

As in the case of Calderón–Zygmund operators, this was achieved using the local mean oscillation decomposition formula in Theorem 1.3.5. Observe that Proposition 4.3.1 reduces the proof of Theorem 4.2.3 to its equivalent statement for $\mathcal{A}_{r,\mathcal{S}}$, as long as it is uniform on the sparse families $\mathcal{S}$ and the dyadic grids $\mathcal{D}$.

For those maximally modulated Calderón–Zygmund operators satisfying strong-type estimates, that is,

$$\|T^\Phi f\|_r \lesssim \|f\|_r \tag{4.3.1}$$

for $1 < r \leqslant r_0$, where $r_0 > 1$, one may obtain pointwise control by the sparse operators $\mathcal{A}_{r,\mathcal{S}}$; this might also be possible for those only satisfying the weak-type estimates (4.2.2), although we do not pursue this subtle point here. Of course this is the case of the Carleson operator $\mathcal{C}$, as it is well known that $\mathcal{C}$ is bounded in L^r for $1 < r < \infty$; this is the celebrated Carleson–Hunt theorem, see for instance [25, 72, 53, 78].

Theorem 4.3.2. *Let T^Φ be a maximally modulated Calderón–Zygmund operator satisfying* (4.3.1). *Then, for any $1 < r \leqslant r_0$ and every compactly supported function $f \in L^r$, there exists a sparse family $\mathcal{S}$ such that*

$$|T^\Phi f(x)| \leqslant C\mathcal{A}_{r,\mathcal{S}} f(x). \tag{4.3.2}$$

Proof. This is a corollary of Lerner's general Theorem 1.3.4, and the proof is very similar to that of Proposition 3.7.1. In view of that theorem, it suffices to show that the grand maximal function $\mathcal{N}_{T^\Phi}$ is of weak type (r,r); we indeed prove that is bounded on L^r.

Given $x \in \mathbb{R}^d$, let $Q \ni x$, $z \in Q$, and consider another arbitrary point $x' \in Q$. By the triangle inequality,

$$|T^\Phi(f\chi_{\mathbb{R}^d\setminus 3Q})(z)| \leqslant |T^\Phi(f\chi_{\mathbb{R}^d\setminus 3Q})(z) - T^\Phi(f\chi_{\mathbb{R}^d\setminus 3Q})(x')| + |T^\Phi(f\chi_{\mathbb{R}^d\setminus 3Q})(x')| = \mathrm{I} + \mathrm{II}.$$

An estimate for the term I is standard,

$$\begin{aligned}
&|T^\Phi(f\chi_{\mathbb{R}^d\setminus 3Q})(z) - T^\Phi(f\chi_{\mathbb{R}^d\setminus 3Q})(x')| \\
&\qquad = \Big| \sup_{\alpha\in A} |T(\mathcal{M}^{\phi_\alpha} f\chi_{\mathbb{R}^d\setminus 3Q})(z)| - \sup_{\alpha\in A} |T(\mathcal{M}^{\phi_\alpha} f\chi_{\mathbb{R}^d\setminus 3Q})(x')| \Big| \\
&\qquad \leqslant \sup_{\alpha\in A} |T(\mathcal{M}^{\phi_\alpha} f\chi_{\mathbb{R}^d\setminus 3Q})(z) - T(\mathcal{M}^{\phi_\alpha} f\chi_{\mathbb{R}^d\setminus 3Q})(x')| \\
&\qquad = \sup_{\alpha\in A} \Big| \int_{\mathbb{R}^d\setminus 3Q} f(y) e^{2\pi i\phi_\alpha(y)} (K(z,y) - K(x',y)) dy \Big| \\
&\qquad \leqslant \sum_{k=0}^{\infty} \int_{2^{k+1}(3Q)\setminus 2^k(3Q)} |f(y)||K(z,y) - K(x',y)| dy \\
&\qquad \lesssim \sum_{k=0}^{\infty} \int_{2^{k+1}(3Q)\setminus 2^k(3Q)} |f(y)| \frac{|z-x'|^\delta}{|x'-y|^{d+\delta}} dy \\
&\qquad \lesssim \sum_{k=0}^{\infty} \int_{2^{k+1}(3Q)\setminus 2^k(3Q)} |f(y)| \frac{\ell(Q)^\delta}{(2^k \ell(Q))^{d+\delta}} dy \\
&\qquad \leqslant \sum_{k=0}^{\infty} \frac{1}{2^{k\delta}} \frac{1}{(2^k\ell(Q))^d} \int_{2^{k+1}(3Q)} |f(y)| dy \\
&\qquad \lesssim \sum_{k=0}^{\infty} 2^{-k\delta} Mf(x) \\
&\qquad \lesssim Mf(x).
\end{aligned}$$

For the second term, we crudely estimate

$$\mathrm{II} \leqslant |T^\Phi f(x')| + |T^\Phi(f\chi_{3Q})(x')|.$$

Putting both estimates together,

$$|T^{\Phi}(f\chi_{\mathbb{R}^d\setminus 3Q})(z)| \lesssim Mf(x) + |T^{\Phi}f(x')| + |T^{\Phi}(f\chi_{3Q})(x')|$$

for any $x' \in Q$. Raising the above estimate to a power $1 < s < r$, integrating with respect to $x' \in Q$, and raising it again to the power $1/s$,

$$\begin{aligned}|T^{\Phi}(f\chi_{\mathbb{R}^d\setminus 3Q})(z)| &\lesssim Mf(x) + \Big(\frac{1}{|Q|}\int_Q |T^{\Phi}f(x')|^s dx'\Big)^{1/s} + \Big(\frac{1}{|Q|}\int_Q |T^{\Phi}(f\chi_{3Q})(x')|^s dx'\Big)^{1/s} \\ &\lesssim Mf(x) + M_s(T^{\Phi}f)(x) + \|T^{\Phi}\|_s\Big(\frac{1}{|Q|}\int_{3Q} |f(x')|^s dx'\Big)^{1/s} \\ &\lesssim Mf(x) + M_s(T^{\Phi}f)(x) + \|T^{\Phi}\|_s M_s f(x),\end{aligned}$$

where we have taken supremum over all $Q \ni x$ and used the boundedness of T^{Φ} in L^s. Thus,

$$\mathcal{N}_{T^{\Phi}}f(x) \lesssim Mf(x) + M_s(T^{\Phi}f)(x) + \|T^{\Phi}\|_s M_s f(x),$$

and taking L^r norms, as $1 < s < r$,

$$\|\mathcal{N}_{T^{\Phi}}f\|_r \lesssim \big(\|M\|_r + \|M\|_{r/s}\|T^{\Phi}\|_r + \|T^{\Phi}\|_s\|M\|_{r/s}\big)\|f\|_r.$$

Then $\mathcal{N}_{T^{\Phi}}$ is bounded in L^r and (4.3.2) follows from an application of Theorem 1.3.4. □

We remark that the proofs given in [87, 32, 79] for the pointwise control of Calderón–Zygmund operators in Theorem 2.2.1 do not seem to extend to the case of the Carleson operator; this is in contrast with the most recent proof provided by Lerner.

4.4 Proof of Theorem 4.2.3

In this section we give a proof of Theorem 4.2.3 and we use it, thanks to an observation due to Pérez [107, 108], to deduce Theorem 4.2.2. Our proof follows a similar pattern of a proof of Di Plinio and Lerner in [42].

As seen in Section 4.3, weighted inequalities for T^Φ can be essentially reduced to the uniform weighted inequalities for the sparse operators $\mathcal{A}_{r,\mathcal{S}}$. Observe that this does not require the pointwise bound from Theorem 4.3.2; for this purposes the prior Theorem 4.3.1 suffices. In particular, we have the following estimate.

Theorem 4.4.1. *Let $1 < p < \infty$, $\mathcal{D}$ be a dyadic grid and $\mathcal{S} \subset \mathcal{D}$ a sparse family of cubes. Suppose that A is a doubling Young function satisfying* (4.2.5)*. Then there is a constant $C_{d,p,A} < \infty$ independent of $\mathcal{S}$, $\mathcal{D}$ and the weight w such that*

$$\|\mathcal{A}_{r,\mathcal{S}} f\|_{L^p(w)} \leqslant C_{d,p,A} \left(\left(\frac{p+1}{2r} \right)' \right)^{1/r} \|f\|_{L^p(M_A w)}$$

holds for any $1 < r < \frac{p+1}{2}$.

Proof. We may assume that $f \geqslant 0$. We first linearise the operator $\mathcal{A}_{r,\mathcal{S}}$; recall that

$$\mathcal{A}_{r,\mathcal{S}} f(x) = \sum_{Q \in \mathcal{S}} \frac{1}{|\bar{Q}|} \left(\int_{\bar{Q}} f^r \right)^{1/r} \chi_Q(x).$$

For any Q, by L^p duality, there exists g_Q supported in $\bar{Q}$ such that $\frac{1}{|\bar{Q}|} \int_{\bar{Q}} g_Q^{r'} = 1$ and

$$\left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} f^r \right)^{1/r} = \frac{1}{|\bar{Q}|} \int_{\bar{Q}} f g_Q.$$

Of course the sequence of functions $\{g_Q\}_Q$ depends on the function f. Given such a sequence, we can define a linear operator L_f by

$$L_f h(x) = \sum_{Q \in \mathcal{S}} \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} h g_Q \right) \chi_Q(x).$$

Note that evaluating in f one recovers $\mathcal{A}_{r,\mathcal{S}} f$, that is $L_f(f) = \mathcal{A}_{r,\mathcal{S}} f$. Then, in order to obtain an estimate for $\|\mathcal{A}_{r,\mathcal{S}}\|_{L^p(w)}$ independent of $\mathcal{S}$ and $\mathcal{D}$, it is enough to obtain the corresponding estimate for $\|L_f h\|_{L^p(w)}$ uniformly in the functions g_Q. For ease of notation we remove the dependence of f in L_f. By duality, the estimate

$$\|Lh\|_{L^p(w)} \leqslant C_{d,p,A} \left(\left(\frac{p+1}{2r} \right)' \right)^{1/r} \|h\|_{L^p(M_A w)}$$

is equivalent to

$$\|L^* h\|_{L^{p'}((M_A w)^{1-p'})} \leqslant C_{d,p,A} \left(\left(\frac{p+1}{2r} \right)' \right)^{1/r} \|h\|_{L^{p'}(w^{1-p'})} \tag{4.4.1}$$

where L^* denotes the $L^2(\mathbb{R}^d)$-adjoint operator of L. Since A satisfies (4.2.5), one can apply Theorem 1.2.2 with p replaced by p'. Using (1.2.3) with $u \equiv 1$, the estimate (4.4.1) follows from

$$\|L^* h\|_{L^{p'}((M_A w)^{1-p'})} \leqslant C_d \left(\left(\frac{p+1}{2r} \right)' \right)^{1/r} \|Mh\|_{L^{p'}((M_A w)^{1-p'})}. \tag{4.4.2}$$

We focus then on obtaining (4.4.2). By duality, there exists $\eta \geqslant 0$ such that $\|\eta\|_{L^p(M_A w)} = 1$ and

$$\|L^* h\|_{L^{p'}((M_A w)^{1-p'})} = \int_{\mathbb{R}^d} L^*(h)\eta = \int_{\mathbb{R}^d} hL\eta.$$

By Hölder's inequality and the $L^{r'}$ boundedness of g_Q,

$$\begin{aligned}
\int_{\mathbb{R}^n} hL\eta &= \sum_{Q \in \mathcal{S}} \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} \eta g_Q \right) \int_Q h \leqslant \sum_{Q \in \mathcal{S}} \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} \eta^r \right)^{1/r} \int_Q h \\
&\leqslant \sum_{Q \in \mathcal{S}} \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} \eta^r \right)^{1/r} \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} h \right) C_d |Q| \\
&= C_d \sum_{Q \in \mathcal{S}} \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} \eta^r \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} h \right)^{\frac{r}{p+1}} \right)^{1/r} \left(\frac{1}{|\bar{Q}|} \int_{\bar{Q}} h \right)^{\frac{p}{p+1}} |Q|.
\end{aligned} \tag{4.4.3}$$

Recall that by definition of the Hardy–Littlewood maximal operator

$$\frac{1}{|\bar{Q}|}\int_{\bar{Q}} h(x)dx \leqslant Mh(y) \tag{4.4.4}$$

holds for every $y \in \bar{Q}$. Combining this and the sparseness of $\mathcal{S}$

$$\begin{aligned}(4.4.3) &\leqslant C_d \sum_{Q\in\mathcal{S}} \left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} \left((Mh)^{\frac{1}{p+1}}\eta\right)^r\right)^{1/r} \left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} h\right)^{\frac{p}{p+1}} |E(Q)| \\ &\leqslant C_d \sum_{Q\in\mathcal{S}} \int_{E(Q)} M_r((Mh)^{\frac{1}{p+1}}\eta)(Mh)^{\frac{p}{p+1}} \\ &\leqslant C_d \int_{\mathbb{R}^d} M_r((Mh)^{\frac{1}{p+1}}\eta)(Mh)^{\frac{p}{p+1}}, \end{aligned} \tag{4.4.5}$$

where we have used that $(E(Q))_{Q\in\mathcal{S}}$ are pairwise disjoint and that (4.4.4) also holds for $y \in E(Q) \subseteq Q \subseteq \bar{Q}$. By Hölder's inequality with exponents $\rho = \frac{p+1}{2}$ and $\rho' = \frac{p+1}{p-1}$,

$$\begin{aligned}(4.4.5) &= C_d \int_{\mathbb{R}^d} M_r((Mh)^{\frac{1}{p+1}}\eta)(M_A w)^{\frac{1}{p+1}}(Mh)^{\frac{p}{p+1}}(M_A w)^{-\frac{1}{p+1}} \\ &\leqslant C_d \|M_r((Mh)^{\frac{1}{p+1}}\eta)\|_{L^{\frac{p+1}{2}}((M_A w)^{1/2})} \|Mh\|^{\frac{p}{p+1}}_{L^{p'}((M_A w)^{1-p'})}. \end{aligned} \tag{4.4.6}$$

For $r < \frac{p+1}{2}$, we can apply the classical Fefferman–Stein inequality described in (1.1.2) to the first term in (4.4.6)

$$\begin{aligned}&\|M_r((Mh)^{\frac{1}{p+1}}\eta)\|_{L^{\frac{p+1}{2}}((M_A w)^{1/2})} \\ &\qquad\leqslant C_d \left(\left(\frac{p+1}{2r}\right)'\right)^{1/r} \|(Mh)^{\frac{1}{p+1}}\eta\|_{L^{\frac{p+1}{2}}(M((M_A w)^{1/2}))},\end{aligned}$$

and by Proposition 1.2.3

$$\|(Mh)^{\frac{1}{p+1}}\eta\|_{L^{\frac{p+1}{2}}(M((M_A w)^{1/2}))} \leqslant C_d \|(Mh)^{\frac{1}{p+1}}\eta\|_{L^{\frac{p+1}{2}}((M_A w)^{1/2})}.$$

Finally, by an application of Hölder's inequality with $\rho = 2p'$ and $\rho' = \frac{2p}{p+1}$

$$\begin{aligned}
\|(Mh)^{\frac{1}{p+1}}\eta\|_{L^{\frac{p+1}{2}}((M_A w)^{1/2})} &= \left(\int_{\mathbb{R}^d} \left((Mh)^{\frac{1}{2}}(M_A w)^{-\frac{1}{2p}}\right)\left(\eta^{\frac{p+1}{2}}(M_A w)^{\frac{p+1}{2p}}\right)\right)^{\frac{2}{p+1}} \\
&\leqslant \|Mh\|_{L^{p'}((M_A w)^{1-p'})}^{\frac{1}{p+1}} \|\eta\|_{L^p(M_A w)} \\
&= \|Mh\|_{L^{p'}((M_A w)^{1-p'})}^{\frac{1}{p+1}},
\end{aligned}$$

where the last equality holds since $\|\eta\|_{L^p(M_A w)} = 1$. Altogether,

$$\|L^* h\|_{L^{p'}((M_A w)^{1-p'})} \leqslant C_d \left(\left(\frac{p+1}{2r}\right)'\right)^{1/r} \|Mh\|_{L^{p'}((M_A w)^{1-p'})}.$$

This concludes the proof. $\square$

We are now able to prove Theorem 4.2.3.

Proof of Theorem 4.2.3. By Proposition 4.3.1, it is enough to show that for any $1 < p < \infty$,

$$\inf_{1<r\leqslant r_0} \left\{\psi(r) \sup_{\mathcal{D},\mathcal{S}} \|\mathcal{A}_{r,\mathcal{S}} f\|_{L^p(w)}\right\} \lesssim \|f\|_{L^p(M_A w)}. \tag{4.4.7}$$

By Theorem 4.4.1,

$$\sup_{\mathcal{D},\mathcal{S}} \|\mathcal{A}_{r,\mathcal{S}} f\|_{L^p(w)} \leqslant C_{d,p,A} \left(\left(\frac{p+1}{2r}\right)'\right)^{1/r} \|f\|_{L^p(M_A w)} \tag{4.4.8}$$

for any $1 < r < \frac{p+1}{2}$, since the bound was independent of $\mathcal{D}$, $\mathcal{S}$.

For every $p > 1$, consider

$$r_p = \min\left\{r_0, 1 + \frac{p-1}{3}\right\} = \min\left\{r_0, \frac{p+2}{3}\right\}.$$

We have that $1 < r_p \leqslant r_0$ and $r_p < \frac{p+1}{2}$. Then

$$\|T^{\Phi} f\|_{L^p(w)} \lesssim \psi(r_p) \sup_{\mathcal{D},\mathcal{S}} \|\mathcal{A}_{r_p,\mathcal{S}} f\|_{L^p(w)} \leqslant C_{d,p,A} \left(\left(\frac{p+1}{2r_p} \right)' \right)^{1/r_p} \|f\|_{L^p(M_A w)}.$$

This concludes the proof. □

Observe that this proof of Theorem 4.2.3 could be extended to other operators whose bounds depend on a suitable way on those of $\mathcal{A}_{r,\mathcal{S}}$. This will be the case of the vector-valued extension presented in Section 4.5.1.

Now one may deduce Theorem 4.2.2 from Theorem 4.2.3 via the following observation due to Pérez [107, 108].

Proof of Theorem 4.2.2. Using Theorem 4.2.3, it is enough to prove that there exists a Young function A satisfying (4.2.5) such that

$$M_A w(x) \leqslant C M^{\lfloor p \rfloor + 1} w(x)$$

with C independent of w. Let $A(t) = t \log^{\lfloor p \rfloor}(1 + t)$. It is an elementary computation to show that A satisfies (4.2.5) for any $c > 0$. Then it suffices to prove that there is a constant $C < \infty$ such that for every cube Q

$$\|w\|_{A,Q} \leqslant C \frac{1}{|Q|} \int_Q M^{\lfloor p \rfloor} w(x) dx =: \lambda_Q.$$

This is equivalent to showing that

$$\left\| \frac{w}{\lambda_Q} \right\|_{A,Q} \leqslant 1,$$

which by definition of the Luxemburg norm will follow from

$$\frac{1}{|Q|}\int_Q A\left(\frac{w(x)}{\lambda_Q}\right)dx = \frac{1}{|Q|}\int_Q \frac{w(x)}{\lambda_Q}\log^{\lfloor p\rfloor}\left(1+\frac{w(x)}{\lambda_Q}\right)dx \leqslant 1.$$

Iterating $\lfloor p\rfloor$ times the inequality

$$\int_Q f(x)\log^k(1+f(x))dx \leqslant \tilde{C}\int_Q Mf(x)\log^{k-1}(1+Mf(x))dx$$

from [124], with $f = w/\lambda_Q$, we obtain

$$\frac{1}{|Q|}\int_Q \frac{w(x)}{\lambda_Q}\log^{\lfloor p\rfloor}\left(1+\frac{w(x)}{\lambda_Q}\right)dx \leqslant \frac{\tilde{C}^{\lfloor p\rfloor}}{|Q|}\int_Q M^{\lfloor p\rfloor}\left(\frac{w}{\lambda_Q}\right)(x)dx.$$

By choosing $C = \tilde{C}^{\lfloor p\rfloor} < \infty$, we have

$$\frac{1}{|Q|}\int_Q A\left(\frac{w(x)}{\lambda_Q}\right)dx \leqslant 1.$$

Thus $M_A w(x) \leqslant M^{\lfloor p\rfloor+1}w(x)$, as required.

Finally, it is not possible to replace $\lfloor p\rfloor + 1$ by $\lfloor p\rfloor$ in the statement of Theorem 4.2.2, as the resulting inequality is shown to be false for the (unmodulated) Hilbert transform [107]. □

4.5 Further remarks

4.5.1 Vector-valued extensions

Theorem 4.2.3 has natural vector-valued extensions. Given a sequence of functions $f = (f_j)_{j\in\mathbb{N}}$, consider the vector-valued extension of T^Φ, given by $\bar{T}^\Phi f = (T^\Phi f_j)_{j\in\mathbb{N}}$. For $q \geqslant 1$, we define the function $|f|_q$ by

$$|f(x)|_q = \Big(\sum_{j=1}^{\infty} |f_j(x)|^q\Big)^{1/q}.$$

As in the case of T^Φ, we will assume that the operator $\bar{T}^\Phi$ satisfies the a priori weak type inequalities

$$\|\bar{T}^\Phi f\|_{L^{r,\infty}(\ell^q)} \lesssim \psi(r)\|f\|_{L^r(\ell^q)} \tag{4.5.1}$$

for $1 < r \leqslant r_0$ and some $r_0 > 1$. Theorem 4.2.3 extends naturally for $\bar{T}^\Phi$ in $L^p(\ell^q)$.

Theorem 4.5.1. *For $q \geqslant 1$, let $\bar{T}^\Phi$ be a vector-valued maximally modulated Calderón-Zygmund operator satisfying* (4.5.1) *and $1 < p < \infty$. Suppose that A is a doubling Young function satisfying*

$$\int_c^\infty \left(\frac{t}{A(t)}\right)^{p'-1} \frac{dt}{t} < \infty$$

for some $c > 0$. Then there is a constant $C < \infty$ such that for any weight w

$$\int_{\mathbb{R}^d} |\bar{T}^\Phi f(x)|_q^p w(x)dx \leqslant C \int_{\mathbb{R}^d} |f(x)|_q^p M_A w(x)dx.$$

Theorem 4.5.1 follows from Theorem 4.2.3 by controlling the Banach space norm of $|\bar{T}^\Phi f|_q$ by that of $\mathcal{A}_{r,\mathcal{S}}|f|_q$. This may be done in the same way as for standard Calderón–Zygmund operators; for instance applying Proposition 1.3.5 to $|\bar{T}^\Phi f|_q$. We do not provide any further detail here, and we just note that it relies on the following standard observation.

Proposition 4.5.2. *Let $q \geqslant 1$ and $\bar{T}^\Phi$ be a vector-valued maximally modulated Calderón-Zygmund operator satisfying* (4.5.1). *Then, for any $1 < r \leqslant r_0$,*

$$\omega_\lambda(|\bar{T}^\Phi f|_q; Q) \lesssim \psi(r)\left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} |f|_q^r\right)^{1/r} + \sum_{m=0}^{\infty} \frac{1}{2^{m\delta}}\left(\frac{1}{|2^m Q|}\int_{2^m Q} |f|_q\right). \tag{4.5.2}$$

The proof of this proposition is quite standard and very close to the ones already used in the proofs of Proposition 3.7.1, Proposition 3.7.2 or Theorem 4.3.2; see also [111] for a similar argument in the case of vector-valued Calderón-Zygmund operators.

4.5.2 The Polynomial Carleson operator

Let $D \in \mathbb{N}$. The polynomial Carleson operator is defined as

$$\mathcal{C}_D f(x) := \sup_{\deg(P) \leqslant D} \left| \mathrm{p.\,v.} \int_{\mathbb{R}} \frac{e^{iP(y)}}{y} f(x-y) dy \right|, \tag{4.5.3}$$

where the supremum is taken over all real-coefficient polynomials P of degree at most D. Note that for $D = 1$ one recovers the definition of the Carleson operator.

It was conjectured by Stein that the operator $\mathcal{C}_D$ is bounded in L^p for $1 < p < \infty$. In the case of periodic functions, this conjecture has been recently solved by Lie [90] via time-frequency analysis techniques; see [91] for his previous work for $\mathcal{C}_2$.

One may write $\mathcal{C}_D f(x) = \sup_{\deg(P) \leqslant D} |H^{\mathbb{T}}(\mathcal{M}^P f)(x)|$ for $x \in \mathbb{T}$, where $\mathcal{M}^{\mathcal{P}} f(x) = e^{iP(x)} f(x)$ and $H^{\mathbb{T}}$ denotes the periodic Hilbert transform. Straightforward modifications in the proof of Theorem 4.2.2 yield a similar result for the periodic case and thus, for any $1 < p < \infty$ there is a constant $C < \infty$ such that for any weight w

$$\int_{\mathbb{T}} |\mathcal{C}_D f(x)|^p w(x) dx \leqslant C \int_{\mathbb{T}} |f(x)|^p M^{\lfloor p \rfloor + 1} w(x) dx.$$

4.5.3 Lacunary Carleson operator

Let $\Lambda = \{\lambda_j\}_j$ be a lacunary sequence of integers, that is, $\lambda_{j+1} \geqslant \theta \lambda_j$ for all j and for some $\theta > 1$ and consider the lacunary Carleson maximal operator

$$\mathcal{C}_\Lambda f(x) = \sup_{j \in \mathbb{N}} \left| \mathrm{p.\,v.} \int_{\mathbb{R}} \frac{e^{2\pi i \lambda_j y}}{x-y} f(y) dy \right|.$$

Of course one has the pointwise estimate $\mathcal{C}_\Lambda f(x) \leqslant \mathcal{C}f(x)$, so the weighted inequality (4.2.4) trivially holds for $\mathcal{C}_\Lambda$. This may be reconciled with a similar result for $\mathcal{C}_\Lambda$ obtained by more classical techniques. Consider the classical version of the lacunary Carleson operator in terms of the lacunary partial Fourier integrals. Following the lines of [22],

$$S_\Lambda^* f(x) = \sup_k |S_{\lambda_k} f(x)| \leqslant cMf(x) + \Big(\sum_k |S_{\lambda_k}(f * \psi_k)(x)|^2 \Big)^{1/2},$$

where $\widehat{S_{\lambda_k} f}(\xi) := \chi_{[-\lambda_k, \lambda_k]}(\xi) \widehat{f}(\xi)$, ψ is a suitable Schwartz function, and $\widehat{\psi}_k(\xi) := \widehat{\psi}(\theta^{-k}\xi)$. Since S_{λ_k} satisfies the same Lebesgue space inequalities as the Hilbert transform, from the estimate (1.1.6) and the weighted Littlewood–Paley theory in Section 1.4, one may deduce the inequality (4.2.4) for $\mathcal{C}_\Lambda$ with a higher number of compositions of M.

4.6 More general two-weight inequalities

We conclude this chapter with the study of two-weight inequalities for Carleson-like operators from a different point of view. In this case, we look for sufficient conditions on a pair of weights (u, v) for the inequality

$$\int_{\mathbb{R}^d} |T^\Phi f|^p u \leqslant C_{p,d,u,v} \int_{\mathbb{R}^d} |f|^p v \tag{4.6.1}$$

to hold. Observe that this more general formulation encodes the inequalities (4.1.3), as we have shown that whenever $v = M^{\lfloor p \rfloor + 1} u$, the above inequality holds with $C_{p,d,u,v} = C_{p,d}$. Before stating sufficient conditions on (u, v), we briefly survey the two-weight problem for Calderón–Zygmund operators.

4.6.1 Testing conditions and sufficient conditions

The problem of two-weight inequalities is of considerable more difficulty than the one-weight problem. As mentioned in Section 1.1 and in contrast with the one-weight case, the condition $(u,v) \in A_p$ is necessary but not sufficient to guarantee $M, H : L^p(v) \to L^p(u)$. In the case of the Hardy-Littlewood maximal operator, Sawyer [119] showed that $M : L^p(v) \to L^p(u)$ if and only if the pair of weights (u,v) satisfies, for every cube Q,

$$\int_Q (M(v^{1-p'}\chi_Q))^p u \leqslant C \int_Q v^{1-p'}.$$

Sawyer [120] also characterised those weights (u,v) that give two-weight estimates for fractional integrals. In this case, $I_\alpha : L^p(v) \to L^p(u)$ if and only if

$$\int_Q (I_\alpha(v^{1-p'}\chi_Q))^p u \leqslant C \int_Q v^{1-p'} \tag{4.6.2}$$

and

$$\int_Q (I_\alpha(u\chi_Q))^{p'} v^{1-p'} \leqslant C \int_Q u \tag{4.6.3}$$

for every cube $Q \subset \mathbb{R}^d$. These conditions are typically referred to as *testing* or *Sawyer* conditions. Note that the linearity and self-adjointness of the operator I_α makes appear the dual testing condition (4.6.3).

The above result for fractional integrals leads one to conjecture whether the testing conditions (4.6.2) and (4.6.3) give also a characterisation for a pair of weights in the case of two-weighted estimates for Calderón-Zygmund operators. Partial progress has been done in that direction; we should mention the work of Nazarov, Treil and Volberg [104, 105] and Lacey, Sawyer, Shen and Uriarte-Tuero [80]. In a recent paper, Hytönen [74] characterised those weights that satisfy a two-weighted L^2 inequality for the Hilbert transform; such characterisation is given in terms of the testing conditions and a variant of the two-weight A_2 condition.

There is an alternative approach in the study of two-weight inequalities based on just looking for sufficient conditions on the pair of weights (u,v). Despite not being a characterisation of the weights, those sufficient conditions are given by general conditions on the weights that do not involve the operator itself; note that the testing conditions (4.6.2) and (4.6.3) involve the operator under study. The sufficient conditions are close in spirit to the two-weight A_p condition. Note that (1.1.1) can be rewritten as

$$[u, v]_{A_p} = \sup_{Q \subset \mathbb{R}^d} \|u^{1/p}\|_{p,Q} \|v^{-1/p}\|_{p',Q} < \infty.$$

The idea consists in making the L^p and $L^{p'}$ norms larger, losing the necessity given by the A_p condition but obtaining sufficient conditions instead.

In this direction, Neugebauer [106] showed that if for some $r > 1$

$$\sup_{Q \subset \mathbb{R}^d} \|u^{1/p}\|_{rp,Q} \|v^{-1/p}\|_{rp',Q} < \infty,$$

then $M, T : L^p(v) \to L^p(u)$. Pérez [108] refined Neugebauer's result in the case of the Hardy-Littlewood maximal operator, showing that $M : L^p(v) \to L^p(u)$ if

$$\sup_{Q \subset \mathbb{R}^d} \|u^{1/p}\|_{p,Q} \|v^{-1/p}\|_{B,Q} < \infty$$

for every cube $Q \subset \mathbb{R}^d$, where B is a doubling Young function such that $\bar{B} \in B_p$.

In the case of Calderón-Zygmund operators, Cruz-Uribe and Pérez [37] conjectured that a sufficient condition for $T : L^p(v) \to L^p(u)$ is

$$\sup_{Q \subset \mathbb{R}^d} \|u^{1/p}\|_{A,Q} \|v^{-1/p}\|_{B,Q} < \infty$$

where A, B are doubling Young functions such that $\bar{A} \in B_{p'}$ and $\bar{B} \in B_p$. After some partial results by Cruz-Uribe, Martell and Pérez [36], and Lerner [83], Lerner [84] finally proved that this conjecture is true, reducing its proof to sparse operators. Following this idea, we give a sufficient condition on a pair of weights for a maximally modulated Calderón-Zygmund operator to satisfy a two-weight inequality.

4.6.2 Maximally modulated Calderón-Zygmund operators

One may adapt the proof of Theorem 4.2.3 to obtain a two weighted inequality for maximally modulated Calderón-Zygmund operators provided the weights satisfy a so-called bump condition.

Theorem 4.6.1. *Let T^Φ be a maximally modulated Calderón-Zygmund operator satisfying* (4.2.2) *and $1 < p < \infty$. Let A and B be doubling Young functions such that $\bar{A} \in B_{p'}$ and $\bar{B} \in B_{\frac{p+1}{2r}}$. Assume that (u, v) is a pair of weights satisfying*

$$\sup_{Q \subset \mathbb{R}^d} \|u^{1/p}\|_{A,Q} \|v^{-r/p}\|_{B,Q}^{1/r} < \infty$$

for some $r < \min\{r_0, p\}$. Then there exists a constant $C = C_{d,p,A,B,u,v} < \infty$ such that

$$\int_{\mathbb{R}^d} |T^\Phi f|^p u \leqslant C \int_{\mathbb{R}^d} |f|^p v. \tag{4.6.4}$$

There is an alternative way of proving Theorem 4.6.1 that does not involve any linearisation and adjoint operator argument. This approach follows the ideas of a similar two-weighted inequality for Calderón-Zygmund operators proved by Lerner in [84].

Alternative proof. By Proposition 4.3.1 it is enough to see that

$$\|\mathcal{A}_{r,\mathcal{S}} f\|_{L^p(u)} \leqslant C \|f\|_{L^p(v)}$$

uniformly on the dyadic sparse family $\mathcal{S}$. By duality there exists $g \in L^{p'}$, $\|g\|_{p'} = 1$ such that

$$\left(\int_{\mathbb{R}^d} \mathcal{A}_{r,\mathcal{S}} f(x)^p u(x) dx \right)^{1/p} = \left| \int_{\mathbb{R}^d} \mathcal{A}_{r,\mathcal{S}} f(x) u(x)^{1/p} g(x) dx \right|.$$

Then,

$$
\begin{aligned}
\Big|\int_{\mathbb{R}^d} (\mathcal{A}_{r,\mathcal{S}} f) u^{1/p} g\Big| &\leqslant \sum_{Q\in\mathcal{S}} \left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} |f|^r\right)^{1/r} \int_Q u^{1/p}|g| \\
&\leqslant \sum_{Q\in\mathcal{S}} \left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} |f|^r v^{r/p} v^{-r/p}\right)^{1/r} \left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}} u^{1/p}|g|\right) |\bar{Q}| \\
&\lesssim \sum_{Q\in\mathcal{S}} \|f^r v^{r/p}\|_{\bar{B},\bar{Q}}^{1/r} \|v^{-r/p}\|_{B,\bar{Q}}^{1/r} \|u^{1/p}\|_{A,\bar{Q}} \|g\|_{\bar{A},\bar{Q}} |E(Q)| \\
&\leqslant \sum_{Q\in\mathcal{S}} \int_{E(Q)} (M_{\bar{B}}(f^r v^{r/p}))^{1/r} M_{\bar{A}} g \\
&\leqslant \int_{\mathbb{R}^d} (M_{\bar{B}}(f^r v^{r/p}))^{1/r} M_{\bar{A}} g \\
&\leqslant \|M_{\bar{B}}(f^r v^{r/p})\|_{p/r}^{1/r} \|M_{\bar{A}} g\|_{p'} \\
&\lesssim \|f^r v^{r/p}\|_{p/r}^{1/r} \|g\|_{p'} \\
&= \|f\|_{L^p(v)},
\end{aligned}
$$

where we have used Hölder's inequality for Young functions, the sparseness of the family $\mathcal{S}$ and the boundedness of the operators $M_{\bar{A}}$ and $M_{\bar{B}}$ in $L^{p'}$ and $L^{p/r}$ respectively. □

We should remark that in the above proof it is enough that $\bar{B} \in B_{p/r}$ instead of the stronger condition $\bar{B} \in B_{\frac{p+1}{2r}}$ that one would obtain following the proof of Theorem 4.2.3.

Remark 4.6.2. The obvious vector-valued extensions considered in Section 4.5.1 also hold for this more general two-weighted case.

Remark 4.6.3. One may recover the Fefferman-Stein weighted inequalities (4.2.6) from Theorem 4.6.1 by considering the pair of weights $(w, M_\Gamma w)$, where $\Gamma(t) = A(t^{1/p})$ and the Young function $B(t) = t^{(p/r)'+\varepsilon}$, that satisfies $\bar{B} \in B_{p/r}$. In this case, the constant C in (4.6.4) does not depend on w, since

$$
[w, M_\Gamma w]_{A,B} = \sup_{Q\subset\mathbb{R}^d} \|w^{1/p}\|_{A,Q} \|(M_\Gamma w)^{-r/p}\|_{B,Q}^{1/r}
$$

$$
\begin{aligned}
&= \sup_{Q\subset\mathbb{R}^d} \|w\|_{\Gamma,Q}^{1/p} \left(\frac{1}{|Q|} \int_Q (M_\Gamma w)^{-(r/p)((p/r)'+\varepsilon)} \right)^{\frac{1}{(p/r)'+\varepsilon}\frac{1}{r}} \\
&\leqslant \sup_{Q\subset\mathbb{R}^d} \left(\frac{1}{|Q|} \int_Q (M_\Gamma w)^{-(r/p)((p/r)'+\varepsilon)} (M_\Gamma w)^{(1/p)((p/r)'+\varepsilon)r} \right)^{\frac{1}{(p/r)'+\varepsilon}\frac{1}{r}} \\
&= \sup_{Q\subset\mathbb{R}^d} \left(\frac{1}{|Q|} \int_Q 1 \right)^{\frac{1}{(p/r)'+\varepsilon}\frac{1}{r}} \\
&= 1,
\end{aligned}
$$

where the second equality follows from the definition of Luxemburg norm.

4.6.3 Multilinear weighted inequalities

The alternative proof given for Theorem 4.6.1 in the previous section has the advantage that it does not involve any linear duality, and it can thus be adapted to a multilinear setting. In particular, we are going to see how it applies to multilinear Calderón-Zygmund operators.

The theory of multilinear Calderón-Zygmund operators was formally introduced by Grafakos and Torres in [59]. Given $2 \leqslant k \leqslant d$, we say that a multilinear operator T is a multilinear (or k-linear) Calderón-Zygmund operator if it is bounded from $L^{q_1}\times\cdots\times L^{q_k} \to L^q$ for some $1 \leqslant q_1, \ldots, q_k < \infty$ satisfying $\frac{1}{q} = \frac{1}{q_1} + \cdots \frac{1}{q_k}$ and if it can be represented as

$$
T(f_1, \ldots, f_k)(x) = \int_{(\mathbb{R}^d)^k} K(x, y_1, \ldots, y_k) f_1(y_1) \cdots f_k(y_k) dy_1 \cdots dy_k
$$

for all $x \notin \cap_{j=1}^k \operatorname{supp} f_j$, where the kernel $K : (\mathbb{R}^d)^{k+1} \backslash \Delta \to \mathbb{R}$, with $\Delta = \{(x, y_1, \cdots, y_k) : x = y_1 = \cdots = y_k\}$, satisfies the following size condition

$$
|K(y_0, y_1, \ldots, y_k)| \leqslant \frac{A}{(\sum_{l,m=0}^k |y_l - y_m|)^{kd}},
$$

and the regularity condition

$$|K(y_0,\dots,y_j,\dots,y_k)-K(y_0,\dots,y_j',\dots,y_k)|\leqslant \frac{A|y_j-y_j'|^\delta}{(\sum_{l,m=0}^k|y_l-y_m|)^{kd+\delta}}$$

for some $\delta>0$ and all $0\leqslant j\leqslant k$, whenever $|y_j-y_j'|\leqslant\frac{1}{2}\max_{0\leqslant l\leqslant k}|y_j-y_l|$.

These operators satisfy analogous Lebesgue space bounds to their linear counterparts, that is

$$\|T(f_1,\dots,f_k)\|_{L^p}\leqslant C\prod_{j=1}^k\|f_j\|_{L^{p_j}},$$

if $1<p_j<\infty$ and $\frac{1}{p}=\frac{1}{p_1}+\dots\frac{1}{p_k}$, and

$$\|T(f_1,\dots,f_k)\|_{L^{p,\infty}}\leqslant C\prod_{j=1}^k\|f_j\|_{L^{p_j}}$$

in case there is $p_j=1$. Recently, a weighted theory for these operators has been developed in terms of a multilinear version of the classical A_p theory, see for example [88, 40, 89].

As in Section 4.6.2, it is possible to obtain a sufficient condition on a tuple of weights $(u,v_1,\dots,v_k)$ to have $T:L^{p_1}(v_1)\times\dots\times L^{p_k}(v_k)\to L^p(u)$, with a very similar proof to the one given for Theorem 4.6.1.

Theorem 4.6.4. *Let $1<p_1,\dots,p_k<\infty$ and $p\geqslant 1$ such that $\frac{1}{p}=\frac{1}{p_1}+\dots+\frac{1}{p_k}$. Let A, $B_1,\dots,B_k$ be doubling Young functions such that $\bar{A}\in B_{p'}$ and $\bar{B}_j\in B_{p_j}$ for $j=1,\dots,k$. If $(u,v_1,\dots,v_k)$ are weights such that*

$$\sup_{Q\subset\mathbb{R}^d}\|u^{1/p}\|_{A,Q}\prod_{j=1}^k\|v_j^{-1/p_j}\|_{B_j,Q}<\infty,$$

then

$$\|T(f_1,\dots,f_k)\|_{L^p(u)}\leqslant C\prod_{j=1}^k\|f_j\|_{L^{p_j}(v_j)}.$$

Proof. A multilinear version of Proposition 4.3.1 in [40] allows one to reduce the proof to the multilinear dyadic sparse operators

$$\mathcal{A}_{1,\mathcal{S}}(f_1,\ldots,f_k)(x):=\sum_{Q\in\mathcal{S}}\prod_{j=1}^{k}\left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}}|f_j|\right)\chi_Q(x).$$

By duality there exists $g\in L^{p'}$, $\|g\|_{p'}=1$ such that

$$\left(\int_{\mathbb{R}^d}(\mathcal{A}_{\mathcal{S}}(f_1,\ldots,f_k))^p u\right)^{1/p}=\left|\int_{\mathbb{R}^d}\mathcal{A}_{\mathcal{S}}(f_1,\ldots,f_k)u^{1/p}g\right|.$$

Then,

$$\begin{aligned}\left|\int_{\mathbb{R}^d}\mathcal{A}_{\mathcal{S}}(f_1,\ldots,f_k)u^{1/p}g\right| &\leqslant \sum_{Q\in\mathcal{S}}\prod_{j=1}^{k}\left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}}|f_j|\right)\int_Q u^{1/p}|g|\\ &=\sum_{Q\in\mathcal{S}}\prod_{j=1}^{k}\left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}}|f_j|v_j^{1/p_j}v_j^{-1/p_j}\right)\left(\frac{1}{|\bar{Q}|}\int_{\bar{Q}}u^{1/p}|g|\right)|\bar{Q}|\\ &\lesssim \sum_{Q\in\mathcal{S}}\prod_{j=1}^{k}\|f_jv_j^{1/p}\|_{\bar{B},\bar{Q}}\|v_j^{-1/p_j}\|_{B,\bar{Q}}\|u^{1/p}\|_{A,\bar{Q}}\|g\|_{\bar{A},\bar{Q}}|E(Q)|\\ &\lesssim \sum_{Q\in\mathcal{S}}\int_{E(Q)}M_{\bar{A}}g\prod_{j=1}^{k}M_{\bar{B}}(f_jv_j^{1/p_j})\\ &\leqslant \int_{\mathbb{R}^d}M_{\bar{A}}g\prod_{j=1}^{k}M_{\bar{B}}(f_jv_j^{1/p_j})\\ &\leqslant \|M_{\bar{A}}g\|_{p'}\prod_{j=1}^{k}\|M_{\bar{B}}(fv^{1/p_j})\|_{p_j}\\ &\leqslant \|g\|_{p'}\prod_{j=1}^{k}\|fv^{1/p_j}\|_{p_j}\\ &=\prod_{j=1}^{k}\|f\|_{L^p_j(v_j)},\end{aligned}$$

using sparseness and multilinear Hölder's inequality. □

As a consequence of Theorem 4.6.4 one can get the following weighted Fefferman-Stein inequality for multilinear Calderon-Zygmund operators; recall Remark 4.6.3 in the linear case for the same kind of result.

Corollary 4.6.5. *Let T be a multilinear Calderón-Zygmund operator. Let $p \geqslant 1$ and $1 < p_1, \ldots, p_k < \infty$ satisfying $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}$. Let A and Γ be doubling Young functions satisfying $\overline{\Gamma(t^p)} = \bar{A}(t) \in B_{p'}$. Then there exists a constant $C < \infty$ such that for every weight w,*

$$\int_{\mathbb{R}^d} |T(f_1, \ldots, f_k)|^p w \leqslant C \prod_{j=1}^{k} \left(\int_{\mathbb{R}^d} |f_j|^{p_j} M_\Gamma w \right)^{p/p_j}. \tag{4.6.5}$$

Proof. It is enough to apply Theorem 4.6.4 with the tuple $(w, M_\Gamma w, \ldots, M_\Gamma w)$ and the Young functions $B_j(t) = t^{p_j'+\varepsilon}$ for $j = 1, \ldots, k$. □

Corollary 4.6.5 allows one to recover the result obtained by Hu [70] via different methods; Hu obtained the above result by induction on the level of linearity k and using the linear result (Theorem 1.1.6) as the base case.

Remark 4.6.6. As in Theorem 4.2.2, taking $\Gamma(t) = t \log^{\lfloor p \rfloor}(1+t)$ one obtains (4.6.5) with $M^{\lfloor p \rfloor + 1}$ in the place of M_Γ.

Chapter 5

The Fourier restriction conjecture: a multilinear reduction

The Fourier restriction phenomenon is of central importance in Euclidean harmonic analysis, and it has been a main object of study over the last decades. This phenomenon consists in studying whether the Fourier transform of a function may be meaningfully restricted to a k-dimensional manifold S in $\mathbb{R}^d$; in our discussion we only concern ourselves with the case of S being a hypersurface in $\mathbb{R}^d$.

If a function $f \in L^1(\mathbb{R}^d)$, the Riemann-Lebesgue lemma ensures that $\widehat{f}$ is a continuous function and thus it may be restricted to any subset of $\mathbb{R}^d$. However, if a function $f \in L^p(\mathbb{R}^d)$ for $1 < p \leqslant 2$, the Hausdorff–Young inequality only ensures that $\widehat{f} \in L^{p'}(\mathbb{R}^d)$ and, in general, it may not be well restricted to sets of measure zero. In the late 1960's, Stein made the remarkable observation that under certain appropriate curvature hypothesis on S, there exists $1 < p_0(S) < 2$ such that every $f \in L^p(\mathbb{R}^d)$, with $1 \leqslant p \leqslant p_0(S)$, has a Fourier transform that restricts to S; this is due to $L^p(\mathbb{R}^d) - L^q(d\sigma)$ bounds on the restriction operator $\mathcal{R}_S f = \widehat{f}|_S$, where $d\sigma$ is the induced Lebesgue measure on S.

Establishing the sharp Lebesgue exponents $1 \leqslant p, q \leqslant \infty$ for which the restriction of the Fourier transform to a manifold S defines a bounded map from $L^p(\mathbb{R}^d) - L^q(d\sigma)$ constitutes the so-called *Fourier restriction conjecture*. This conjecture is of crucial importance due to its numerous connections with many other problems and disciplines, such as the Kakeya [51, 11, 145] or Bochner–Riesz conjectures [20, 135], local smoothing [123], Strichartz estimates [130] and almost everywhere convergence questions for dispersive PDE [12], Falconer's distance set problem [93, 46], or problems in incidence geometry [15] and number theory [60, 15, 16]; we do not intend to discuss all these connections here.

The conjecture is still open for $d \geqslant 3$, and the best known partial results have been achieved taking a multilinear perspective on the problem. The goal of this chapter is to obtain a better understanding of the role of multilinear estimates in the original Fourier restriction; in particular, we carefully study the method developed by Bourgain and Guth [17] to obtain linear estimates from their multilinear counterparts.

5.1 The linear and multilinear restriction conjectures

Let $d \geqslant 2$ and S be a smooth compact hypersurface in $\mathbb{R}^d$. For a function $f \in L^p(\mathbb{R}^d)$ we define the *restriction operator* associated to S as $\mathcal{R}_S f = \widehat{f}|_S$. The restriction problem is typically studied in its adjoint form, seeking for $L^p(d\sigma) - L^q(\mathbb{R}^d)$ bounds for the *extension operator* $\mathcal{R}^* g(\xi) = \widehat{gd\sigma}(\xi)$, that is

$$\|\widehat{gd\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|g\|_{L^p(d\sigma)},$$

where $g : S \to \mathbb{C}$ and $d\sigma$ denotes the induced Lebesgue measure on S. We should note that as S is compact, this is equivalent to the estimate

$$\|\widehat{gd\mu}\|_{L^q(\mathbb{R}^d)} \lesssim \|g\|_{L^p(d\mu)},$$

where $d\mu$ denotes the parametrised measure on S; given an open set $U \subset \mathbb{R}^{d-1}$ and a parametrisation $\Sigma : U \to S$, the measure μ is defined by

$$\int_S g(x) d\mu(x) = \int_U g(\Sigma(y)) dy.$$

Due to this equivalence we use $d\sigma$ and $d\mu$ interchangeably in what follows, as we only concern ourselves about norm estimates.

There is a trivial $L^1(d\sigma) - L^\infty(\mathbb{R}^d)$ estimate for the extension operator,

$$\|\widehat{gd\sigma}\|_{L^\infty(\mathbb{R}^d)} \leqslant \|g\|_{L^1(d\sigma)}. \tag{5.1.1}$$

As mentioned above, Stein observed that under appropriate curvature hypothesis on S, other $L^p(d\sigma) - L^q(\mathbb{R}^d)$ estimates may hold besides the trivial one. This is in contrast with the case of absence of curvature hypothesis. For example, let S be a portion of the dth coordinate hyperplane given by the parametrisation $\Sigma : U \to S$, where $\Sigma(x') = (x', 0)$ and U is an open set in $\mathbb{R}^{d-1}$. Then the function

$$\widehat{gd\mu}(\xi) = \int_S g(x) e^{ix\cdot\xi} d\mu(x) = \int_U g(\Sigma(x')) e^{i\Sigma(x')\cdot\xi} dx' = \int_U g(\Sigma(x')) e^{ix'\cdot\xi'} dx' = \widehat{g \circ \Sigma}(\xi')$$

is independent of the ξ_d coordinate. Thus $\widehat{gd\mu} \notin L^q(\mathbb{R}^d)$ for $q < \infty$ unless $g \equiv 0$.

Stein's observation led to set the *restriction conjecture for the Fourier transform*, which in the case of hypersurfaces S with positive Gaussian curvature reads the following.[1]

Conjecture 5.1.1 (Linear restriction conjecture)**.** *If S has everywhere positive Gaussian curvature, $\frac{1}{q} < \frac{d-1}{2d}$ and $\frac{1}{q} \leqslant \frac{d-1}{d+1}\frac{1}{p'}$ then*

$$\|\widehat{gd\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim \|g\|_{L^p(d\sigma)}. \tag{5.1.2}$$

[1]From now on we focus our discussion on the specific case of hypersurfaces with positive Gaussian curvature, for which the prototypical example is a compact piece of the paraboloid.

We denote the estimate (5.1.2) by $\mathcal{R}^*(p \to q)$. The first condition on the exponents corresponds to the integrability of the measure, since $|\widehat{d\sigma}(\xi)| \lesssim (1+|\xi|)^{-(d-1)/2}$ if S has nonvanishing Gaussian curvature; this may be seen via a stationary phase argument, see for example [129]. The second condition follows from testing the estimate in the characteristic function of small caps in S. We refer to the surveys [6] and [136] for more details about the formulation of this conjecture and the forthcoming multilinear analogues. We should remark that the main difficulty in the restriction conjecture is to make the value of q lower; interpolation with the trivial estimate (5.1.1) gives the estimates for bigger values of q and Hölder's inequality and factorisation theory [11] allow to increase and decrease respectively the value of p.

The Fourier restriction conjecture is fully solved for $d = 2$ by Fefferman [49], but it is still open in higher dimensions. Stein and Tomas [139] established $\mathcal{R}^*(2 \to \frac{2(d+1)}{d-1})$, giving a result on the sharp line $\frac{1}{q} = \frac{d-1}{d+1}\frac{1}{p'}$. The striking work of Bourgain [11] led to a new perspective to the problem, linking the Fourier restricion phenomenon with the Kakeya conjecture, and developing the now standard technique *wave-packet decomposition*. Consecutive improvements on the state-of-the-art for the restriction conjecture have been obtained by Wolff [144], Moyua, Vargas and Vega [100, 101], Tao, Vargas and Vega [137], Tao and Vargas [133], Tao [132], Bourgain and Guth [17], Temur [138] and Guth [64, 61].

A fundamental ingredient in the most recent developments in restriction theory is the multilinear approach. This originated with a bilinear formulation of the problem. If S_1 and S_2 are compact hypersurfaces with positive Gaussian curvature, it is obvious that the restriction conjecture induces a bilinear analogue conjecture via Hölder's inequality, that is

$$\|\widehat{g_1 d\sigma_1}\widehat{g_2 d\sigma_2}\|_{q/2} \leqslant \|\widehat{g_1 d\sigma_1}\|_q \|\widehat{g_2 d\sigma_2}\|_q \lesssim \|g_1\|_{L^p(d\sigma_1)}\|g_2\|_{L^p(d\sigma_2)} \tag{5.1.3}$$

with p, q as in Conjecture 5.1.1. However, the range of exponents such that (5.1.3) may hold is wider than the ones given by Conjecture 5.1.1 if we assume that the hypersurfaces S_1 and S_2 are transversal. By transversal we mean that if v_1 and v_2 are unit normal vectors to S_1 and S_2 respectively, then $|v_1 \wedge v_2| > c$ for some constant $c > 0$. This led to the following bilinear conjecture.

Conjecture 5.1.2 (Bilinear restriction conjecture). *Let S_1 and S_2 be smooth compact transversal hypersurfaces with positive Gaussian curvature. If $\frac{1}{q} < \frac{d-1}{2d}$, $\frac{1}{q} \leqslant \frac{d}{d+2}\frac{1}{p'}$ and $\frac{1}{q} \leqslant \frac{d-2}{d+2}\frac{1}{p'} + \frac{1}{d+2}$ then*

$$\|\widehat{g_1 d\sigma_1}\widehat{g_2 d\sigma_2}\|_{L^{q/2}(\mathbb{R}^d)} \lesssim \|g_1\|_{L^p(d\sigma_1)}\|g_2\|_{L^p(d\sigma_2)}. \tag{5.1.4}$$

We denote estimate (5.1.4) by $\mathcal{R}^*(p \times p \to q/2)$. Observe that for functions in $L^2(d\sigma)$, the bilinear conjecture has admissible values for q smaller than the Stein–Tomas exponent. This is of considerable interest, as it permits to exploit, for such values of q, the aforementioned wave-packet decomposition of Bourgain, which fails to work if $g_j \notin L^2(d\sigma)$.

Bilinear estimates became of central importance in the problem, due to a remarkable observation of Tao, Vargas and Vega [137], who showed that the linear and the bilinear restriction conjectures are essentially equivalent. That equivalence is obtained via a Whitney decomposition of the product manifold $S \times S$ around the diagonal $\Delta := \{(\xi, \xi) : \xi \in S\}$; this allows one to decompose $(S \times S)\backslash\Delta$ as a union of sets of the type $S_1 \times S_2$, where S_1 and S_2 are transversal subsets of S, and thus the role of bilinear estimates becomes apparent.

Theorem 5.1.3 ([137]). *Let $1 < p, q < \infty$ be such that $\frac{1}{q} \leqslant \frac{d-1}{2d}$ and $\frac{1}{q} \leqslant \frac{d-1}{d+1}\frac{1}{p'}$. Then $\mathcal{R}^*(p \times p \to q/2) \Leftrightarrow \mathcal{R}^*(p \to q)$.*

The extra transversality assumption on the bilinear estimate makes such estimates more tractable than the linear ones. Thus, the above equivalence together with good bilinear estimates constitutes a way to make progress on the linear restriction conjecture.

The best progress by this method was achieved by Tao [132], who established the bilinear conjecture for functions on $L^2(d\sigma)$, except for the endpoint case.

Theorem 5.1.4 ([132]). *Let S_1, S_2 be any disjoint compact subsets of the paraboloid* [2]. *Then $\mathcal{R}^*(2 \times 2 \to q/2)$ holds for any $q > 2(d+2)/d$. In particular, $\mathcal{R}^*(p \to q)$ holds for $\frac{1}{q} \leqslant \frac{d-1}{d+1}\frac{1}{p'}$ and $q > 2(d+2)/d$.*

One may extend the bilinear setting into a k-linear one, leading naturally to the following conjecture.

Conjecture 5.1.5 (k-linear restriction conjecture). *Suppose that $2 \leqslant k \leqslant d$ and that $S_1, \ldots, S_k$ are transversal* [3] *hypersurfaces with positive Gaussian curvature. If $\frac{1}{q} < \frac{d-1}{2d}$, $\frac{1}{q} \leqslant \frac{d+k-2}{d+k}\frac{1}{p'}$ and $\frac{1}{q} \leqslant \frac{d-k}{d+k}\frac{1}{p'} + \frac{k-1}{k+d}$ then*

$$\Big\| \prod_{j=1}^{k} \widehat{g_j d\sigma_j} \Big\|_{L^{q/k}(\mathbb{R}^d)} \lesssim \prod_{j=1}^{k} \|g_j\|_{L^p(d\sigma_j)}. \tag{5.1.5}$$

We denote the estimate (5.1.5) by $\mathcal{R}^*(p \times \cdots \times p \to q/k)$. The case $k = d$ turns out to be rather special, as the curvature hypothesis does not seem to play any role. In that case one obtains the same range of exponents even without the curvature hypothesis; indeed standard examples allow one to conjecture the validity of the estimate at the missing endpoint $q = 2d/(d-1)$ in Conjecture 5.1.5. By multilinear interpolation, the d-linear conjecture is equivalent to such endpoint case for $p = 2$. The d-linear conjecture is nearly solved; Bennett, Carbery and Tao proved in [9] the following local version, which morally corresponds to the conjecture away from the endpoint.

[2] Observe that the curvature of the paraboloid induces transversallity on any two disjoint compact subsets.

[3] We naturally extend the transversal concept into a multilinear setting; for any $v_1, \ldots, v_k$ unit normal vectors to $S_1, \ldots, S_k$ respectively, the hypersurfaces are *ν-transversal* if $|v_1 \wedge \cdots \wedge v_k| > \nu > 0$, where $\nu > 0$.

Theorem 5.1.6 ([9])**.** *Let $S_1, \dots, S_d$ be transversal hypersurfaces. Then for any $\varepsilon > 0$ there exists $C_\varepsilon < \infty$ such that*

$$\Big\| \prod_{j=1}^{d} \widehat{g_j d\sigma_j} \Big\|_{L^{2/(d-1)}(B(0,R))} \leqslant C_\varepsilon R^\varepsilon \prod_{j=1}^{d} \|g_j\|_{L^2(d\sigma_j)}.$$

Away from the endpoint, it was proved in [9] that the d-linear conjecture is equivalent to a d-linear maximal Kakeya conjecture, strengthening the connections between the original Fourier restriction and the Kakeya conjecture, and highlighting, even more, the strong combinatorial flavour of the Fourier restriction conjecture. We do not intend to discuss the Kakeya maximal conjecture here; its d-linear version was proved, away from the endpoint, by Bennett, Carbery and Tao [9], and the endpoint was first obtained by Guth [62], involving algebraic and topological techniques, and later simplified by Carbery and Valdimarsson [24]. We shall remark that Guth has recently given a short proof for a weaker version of the multilinear Kakeya conjecture away from the endpoint [63], and that Bejenaru [2] has also given an alternative proof for Theorem 5.1.6.

As remarked in [9], the techniques in Theorem 5.1.6 also apply to make partial progress on the k-linear conjecture.

Theorem 5.1.7 ([9])**.** *Let $k \leqslant d$. If $S_1, \dots, S_k$ are transversal hypersurfaces, then for any $\varepsilon > 0$ and $\frac{1}{q} \leqslant \frac{k-1}{2k}$, there exists $C_\varepsilon < \infty$ such that*

$$\Big\| \prod_{j=1}^{k} \widehat{g_j d\sigma_j} \Big\|_{L^{q/k}(B(0,R))} \leqslant C_\varepsilon R^\varepsilon \prod_{j=1}^{k} \|g_j\|_{L^2(d\sigma_j)}. \tag{5.1.6}$$

In the spirit of Proposition 5.1.3, Bourgain and Guth [17] developed a new technique that allows to use these multilinear estimates to make improvement on the linear restriction conjecture. We shall revisit their strategy in Section 5.2, which combined with Theorem 5.1.7 allows to deduce the following.

Theorem 5.1.8 ([17]). *Let $d \geqslant 3$ and S be a compact smooth hypersurface with positive Gaussian curvature. Then $\mathcal{R}^*(\infty \to q)$ holds for*

- $d \equiv 0 \pmod 3$, $q > 2\frac{4d+3}{4d-3}$,
- $d \equiv 1 \pmod 3$, $q > \frac{2d+1}{d-1}$,
- $d \equiv 2 \pmod 3$, $q > \frac{4(d+1)}{2d-1}$.

In the case $d = 3$, Bourgain and Guth refined their argument combining it with a maximal Kakeya estimate of Wolff [144], leading to a small improvement in the value of the exponent q. Similarly, Temur [138] observed that such improvement could be further exploited for any $d \equiv 0 \pmod 3$.

A careful inspection of the method developed by Bourgain and Guth [17], allows one to make the following conjectural theorem. It consists on determining the impact on the linear conjecture of the conjectured k-linear estimates (5.1.5) for $g_j \in L^2(d\sigma)$.[4]

Theorem 5.1.9. *Assume that Conjecture 5.1.5 holds for $k = \lfloor \frac{d+2}{2} \rfloor$, $p = 2$ and $\frac{1}{q} < \frac{d+k-2}{2(d+k)}$. Then $\mathcal{R}^*(\infty \to q)$ holds for*

- $q > 2\frac{3d+1}{3d-3}$ *for d odd,*
- $q > 2\frac{3d+2}{3d-2}$ *for d even.*

This observation, which is a simple consequence of a careful reading of [17], constitutes the main remark of this chapter. Our purpose is to put k-linear estimates in a central scene towards the future developments in restriction theory and related topics. It is interesting to observe that the conjectured optimal multilinear estimates with level of linearity higher than $\sim \frac{d}{2}$ would not lead to any extra benefit on the linear problem using the ideas in [17]; similarly, the known non-optimal multilinear estimates (5.1.6) with level of linearity

[4]We only run the argument assuming that Conjecture 5.1.5 is true for $g_j \in L^2(d\sigma)$, as in that case one may use the *wave-packet decomposition*.

higher than $\sim \frac{2d}{3}$ do not lead to a further improvement on the linear problem. This, which is established in Section 5.3, shows the limitations of the Bourgain–Guth method, as it does not exploit the multilinear estimates for which the level of linearity is close to the dimension. Note that the optimal exponent in the d-linear case corresponds to the conjectured exponent in the Conjecture 5.1.1.

After this analysis was carried out, Guth [61] obtained a "restriction estimate" which amounts to a *weaker* version of Conjecture 5.1.5 for $p = 2$. This weaker version, however, leads to the improvement on the Conjecture 5.1.1 anticipated by Theorem 5.1.9, using a small variant of the method in [17], and so, the exponents in Theorem 5.1.9 correspond to the best current state-of-the-art in the Fourier restriction conjecture for $d \geqslant 4$. In the 3-dimensional case, the best known result is also due to Guth [64], where $q > 3.25$. The main ingredient in both papers is the use of polynomial partitioning in a Fourier restriction setting. We do not discuss this any further in this thesis.

5.2 The Bourgain-Guth method

In this section we recall the Bourgain–Guth method in [17]. As is mentioned above, it permits to deduce linear estimates from their multilinear counterparts. Our aim is to obtain a better understanding of the role of the multilinear estimates in that method. In particular, this allows us to state the following k-linear reduction for the restriction conjecture.

Theorem 5.2.1. *Let $2 \leqslant k \leqslant d$. If $\frac{1}{p} \leqslant \frac{1}{q} < \frac{d-1}{2d}$ and*

$$q > \tilde{q}(k) := 2\min\left(\frac{m}{m-1}, \frac{2d-m+1}{2d-m-1}\right), \quad 2 \leqslant m \leqslant k-1, \tag{5.2.1}$$

then $\mathcal{R}^*(p \to q) \iff \mathcal{R}^*(p \times \cdots \times p \to q/k)$.

Proof. The only relevant content of the theorem is $\mathcal{R}^*(p \times \cdots \times p \to q/k) \Rightarrow \mathcal{R}^*(p \to q)$; the reverse inequality follows from multilinear Hölder's inequality.

We prove indeed a localised version of the restriction estimate $\mathcal{R}^*(p \to q)$. For $R \gg 1$ we will see that

$$\|\widehat{gd\sigma}\|_{L^q(B(0,R))} \leqslant R^\varepsilon C \|g\|_{L^\infty(d\sigma)} \tag{5.2.2}$$

for $q > \tilde{q}(k)$ as in the statement of the theorem. The use of standard "epsilon-removal" lemmas -like the ones in [136] and [133]- allows one to deduce the global estimate $\mathcal{R}^*(p \to q)$ for $q > \tilde{q}(k)$. To lower down $p = \infty$ to $p \geqslant q$ one may use factorization theory; see [11].

Let $\mathcal{C}(R)$ denote the best constant C in the inequality (5.2.2). Our goal is to see that $\mathcal{C}(R) \lesssim 1$. Given a constant K, we denote by $P(K)$ any (positive) power of K. We shall use this notation when the powers of K are irrelevant.

Let $1 \ll K_k \ll R^\varepsilon$ and $\{S^k_\alpha\}_\alpha$ be a partition of S in caps of diameter $1/K_k$ with finite overlapping. For $g \in L^p(d\sigma)$, write $g^k_\alpha(x) = g(x)\chi_{S^k_\alpha}(x)$. Then

$$\widehat{gd\sigma}(\xi) = \sum_\alpha \widehat{g^k_\alpha d\sigma}(\xi).$$

Tile $B(0,R)$ into cubes Q^k of sidelength K_k. By uncertainty principle considerations,[5] we may think of $|\widehat{g^k_\alpha d\sigma}|$ as being essentially constant in the cubes Q^k. Fixing Q^k, either

(I) there exist $\alpha_1, \ldots, \alpha_k$ with $S^k_{\alpha_1}, \ldots, S^k_{\alpha_k}$ being $(K_k)^{-k}$–transversal such that

$$|\widehat{g^k_{\alpha_1} d\sigma}(\xi)|, \ldots, |\widehat{g^k_{\alpha_k} d\sigma}(\xi)| \geqslant K_k^{-(d-1)} \max_\alpha |\widehat{g^k_\alpha d\sigma}(\xi)|$$

for every $\xi \in Q^k$, or

[5]Technically, we should replace $|\widehat{g^k_\alpha d\sigma}|$ by a pointwise majorant satisfying such property, but we refrain from doing that for simplicity of the argument, as it does not contribute to the main ideas in the Bourgain–Guth argument. We develop this further in Appendix D.

(II) there exists a $(k-1)$-dimensional subspace V_{k-1} of $\mathbb{R}^d$ such that for those S_α^k with $\mathrm{dist}(S_\alpha^k, E_{k-1}) \gtrsim 1/K_k$, where E_{k-1} denotes the image of $V_{k-1} \cap \mathbb{S}^{d-1}$ under the Gauss map, then

$$|\widehat{g_\alpha^k d\sigma}(\xi)| < K_k^{-(d-1)} \max_\alpha |\widehat{g_\alpha^k d\sigma}(\xi)|$$

for every $\xi \in Q^k$.

Note that $\alpha_1, \ldots, \alpha_k$ in (I) may be chosen to be the same for all $\xi \in Q^k$, as $|\widehat{g_\alpha^k d\sigma}(\xi)|$ are essentially constant in Q^k. Similarly, the subspace V_{k-1} may be chosen to be the same for all $\xi \in Q^k$. We should also remark that $K_k^{-(d-1)}$ in (I) and (II) may be replaced by any power $K_k^{-\gamma}$ provided $\gamma \geqslant d-1$.

If (I) we run a multilinear argument,

$$\begin{aligned}
|\widehat{g d\sigma}(\xi)| &\leqslant \sum_\alpha |\widehat{g_\alpha^k d\sigma}(\xi)| \\
&\leqslant K_k^{d-1} \max_\alpha |\widehat{g_\alpha^k d\sigma}(\xi)| \\
&\leqslant K_k^{2(d-1)} \prod_{j=1}^k |\widehat{g_{\alpha_j}^k d\sigma}(\xi)|^{1/k} \\
&\leqslant K_k^{2(d-1)} \Big(\sum_{\substack{\alpha_1,\ldots,\alpha_k \\ \text{trans}}} \prod_{j=1}^k |\widehat{g_{\alpha_j}^k d\sigma}(\xi)|^{q/k} \Big)^{1/q}
\end{aligned}$$

for any $\xi \in Q^k$, where the sum is taken over all $\alpha_1, \ldots, \alpha_k$ for which $S_{\alpha_1}^k, \ldots, S_{\alpha_k}^k$ are transversal. We note that such sum has been taken so that the choice of α_j in the right hand side above is independent of the cube Q^k. This allows to sum in Q^k in what follows. Taking the power q, integrating in those Q^k for which (I) holds and using the hypothesis $\mathcal{R}^*(p \times \cdots \times p \to q/k)$, we conclude that

$$\sum_{\substack{Q^k \\ \text{(I) holds}}} \|\widehat{g d\sigma}\|_{L^q(Q^k)}^q \leqslant K_k^{2(d-1)q} \sum_{\substack{\alpha_1,\ldots,\alpha_k \\ \text{trans}}} \sum_{\substack{Q^k \\ \text{(I) holds}}} \Big\| \prod_{j=1}^k \widehat{g_{\alpha_j} d\sigma} \Big\|_{L^{q/k}(Q^k)}^{q/k}$$

$$
\begin{aligned}
&\leqslant K_k^{2(d-1)q} \sum_{\substack{\alpha_1,\dots,\alpha_k\\ \text{trans}}} \Big\| \prod_{j=1}^{k} \widehat{g_{\alpha_j} d\sigma} \Big\|^{q/k}_{L^{q/k}(B(0,R))} \\
&\lesssim R^{\varepsilon} K_k^{2(d-1)q} \sum_{\substack{\alpha_1,\dots,\alpha_k\\ \text{trans}}} \prod_{j=1}^{k} \|g^k_{\alpha_j}\|^{q/k}_{L^p(d\sigma)} \\
&\lesssim R^{\varepsilon} K_k^{2(d-1)q} \sum_{\substack{\alpha_1,\dots,\alpha_k\\ \text{trans}}} \sum_{j=1}^{k} \|g^k_{\alpha_j}\|^{q}_{L^p(d\sigma)} \\
&\lesssim R^{\varepsilon} P(K_k) \sum_{\alpha} \|g^k_{\alpha}\|^{q}_{L^p(d\sigma)} \\
&\lesssim R^{\varepsilon} P(K_k) \Big(\sum_{\alpha} \|g^k_{\alpha}\|^{p}_{L^p(d\sigma)} \Big)^{q/p} \\
&\leqslant R^{\varepsilon} P(K_k) \|g\|^{q}_{L^p(d\sigma)},
\end{aligned}
$$

where in the one to last inequality we have used Hölder's inequality and that $p \geqslant q$. We note that the powers of K_k here are irrelevant, as K_k will be a chosen fixed number independent of R and therefore $\lesssim R^{\varepsilon}$.

If (II) we write

$$
|\widehat{gd\sigma}(\xi)| \leqslant \Big| \int_{\{x:\operatorname{dist}(x,E_{k-1})\lesssim 1/K_k\}} e^{i\xi\cdot x} g(x) d\sigma(x) \Big| + \Big| \sum_{\operatorname{dist}(S^k_\alpha, E_{k-1})\gtrsim 1/K_k} \widehat{g^k_\alpha d\sigma}(\xi) \Big| \tag{5.2.3}
$$

for any $\xi \in Q^k$. For the second term in (5.2.3),

$$
\begin{aligned}
\Big| \sum_{\operatorname{dist}(S^k_\alpha, E_{k-1})\gtrsim 1/K_k} \widehat{g^k_\alpha d\sigma}(\xi) \Big| &< \sum_{\operatorname{dist}(S^k_\alpha, E_{k-1})\gtrsim 1/K_k} K_k^{-(d-1)} \max_\alpha |\widehat{g^k_\alpha d\sigma}(\xi)| \\
&\lesssim \max_\alpha |\widehat{g^k_\alpha d\sigma}(\xi)| \\
&\leqslant \Big(\sum_\alpha |\widehat{g^k_\alpha d\sigma}(\xi)|^q \Big)^{1/q},
\end{aligned}
$$

for any $\xi \in Q^k$, where the sum is taken over all caps S^k_α. As before, such sum has been taken so that the choice of α in the right hand side of the above estimate does not depend on Q^k. Taking the power q, and integrating in those Q^k for which (II) holds,

$$\begin{aligned}\sum_{\substack{Q^k\\ \text{(II) holds}}} \Big\| \sum_{\mathrm{dist}(S^k_\alpha, E_{k-1})\gtrsim 1/K_k} \widehat{g^k_\alpha d\sigma}\Big\|^q_{L^q(Q^k)} &\lesssim \sum_\alpha \sum_{\substack{Q^k\\ \text{(II) holds}}} \|\widehat{g^k_\alpha d\sigma}\|^q_{L^q(Q^k)} \\ &\leqslant \sum_\alpha \|\widehat{g^k_\alpha d\sigma}\|^q_{L^q(B(0,R))} \\ &\lesssim \mathcal{C}(R) K_k^{d+1-\frac{(d-1)q}{p'}} \sum_\alpha \|g^k_\alpha\|^q_{L^p(d\sigma)} \\ &\lesssim \mathcal{C}(R) K_k^{d+1-\frac{(d-1)q}{p'}} K_k^{\frac{(d-1)}{(p/q)'}} \Big(\sum_\alpha \|g^k_\alpha\|^p_{L^p(d\sigma)}\Big)^{q/p} \\ &\lesssim \mathcal{C}(R) K_k^{2d-(d-1)q} \|g\|^q_{L^p(d\sigma)}, \end{aligned} \tag{5.2.4}$$

where we have used the rescaling condition in Appendix E and Hölder's inequality in the sum in α. Here the powers of K_k are relevant, as we are inducting on the size of the caps in our inequality (5.2.2).

This is an acceptable term if $K_k^{2d-(d-1)q} \ll 1$. As by assumption, $q > 2d/(d-1)$ - this is one of the conditions on the exponent in the restriction conjecture -, it is enough to pick K_k sufficiently large and independent of R.

For the first term in (5.2.3), we introduce a new parameter $1 \ll K_{k-1} \ll K_k$ and we consider $\{S^{k-1}_\beta\}_\beta$ to be a partition of S in caps of diameter $1/K_{k-1}$. For $\xi \in Q^k$, we write

$$\begin{aligned}\Big|\int_{\{x:\mathrm{dist}(x,E_{k-1})\lesssim 1/K_k\}} e^{i\xi\cdot x} g(x)d\sigma(x)\Big| &= \Big|\sum_\beta \int_{\{x:\mathrm{dist}(x,E_{k-1})\lesssim 1/K_k\}\cap S^{k-1}_\beta} e^{i\xi\cdot x} g(x)d\sigma(x)\Big| \\ &= \Big|\sum_\beta \widehat{g^{k-1}_\beta d\sigma}(\xi)\Big|.\end{aligned}$$

Again, by the uncertainty principle the quantities $|\widehat{g^{k-1}_\beta d\sigma}|$ are essentially constant in cubes Q^{k-1} of sidelength K_{k-1}. We run a multilinear analysis as before, and for every $Q^{k-1} \subset Q^k$, either

(i) there exist $\beta_1, \dots, \beta_{k-1}$ with $S^{k-1}_{\beta_1}, \dots, S^{k-1}_{\beta_{k-1}}$ being $(K_{k-1})^{-(k-1)}$–transversal such that

$$|\widehat{g^{k-1}_{\beta_1} d\sigma}(\xi)|, \dots, |\widehat{g^{k-1}_{\beta_{k-1}} d\sigma}(\xi)| \geqslant K_{k-1}^{-(k-2)} \max_{\beta} |\widehat{g^{k-1}_{\beta} d\sigma}(\xi)|$$

for every $\xi \in Q^{k-1}$, or

(ii) there exists a $(k-2)$-dimensional subspace $V_{k-2} \subset V_{k-1}$ such that for those S^{k-1}_{β} with $\mathrm{dist}(S^{k-1}_{\beta}, E_{k-2}) \gtrsim 1/K_{k-1}$, where E_{k-2} is the image of $V_{k-2} \cap \mathbb{S}^{d-1}$ under the Gauss map, then

$$|\widehat{g^{k-1}_{\beta} d\sigma}(\xi)| < K_{k-1}^{-(k-2)} \max_{\beta} |\widehat{g^{k-1}_{\beta} d\sigma}(\xi)|$$

for every $\xi \in Q^{k-1}$.

As in the previous case, the caps indexed by $\beta_1, \dots, \beta_{k-1}$ and the subspace V_{k-2} may be chosen to be the same for all $\xi \in Q^{k-1}$, and the power $K_{k-1}^{-(k-2)}$ in (i) and (ii) may be replaced by $K_{k-1}^{-\gamma}$ provided $\gamma \geqslant k-2$. Observe that we may write

$$\widehat{g^{k-1}_{\beta} d\sigma}(\xi) = \sum_{\substack{\alpha : S^k_\alpha \subseteq S^{k-1}_\beta \\ \mathrm{dist}(S_\alpha, E_{k-1}) \lesssim 1/K_k}} \widehat{g^k_\alpha d\sigma}(\xi).$$

We have two possibilities for the case (i). The first one consists in the use of the multilinear estimates from Theorem 5.1.7, and for this reason powers of K_k and K_{k-1} will be irrelevant. In case $q \geqslant \frac{2(k-1)}{k-2}$, for every $\xi \in Q^{k-1}$,

$$\begin{aligned}
\Big|\sum_{\beta} \widehat{g^{k-1}_{\beta} d\sigma}(\xi)\Big| &\leqslant K_{k-1}^{2(k-2)} \prod_{j=1}^{k-1} |\widehat{g^{k-1}_{\beta_j} d\sigma}(\xi)|^{1/(k-1)} \\
&\leqslant K_{k-1}^{2(k-2)} \prod_{j=1}^{k-1} \Big(\sum_{\substack{\alpha_j : S^k_{\alpha_j} \subseteq S^{k-1}_{\beta_j} \\ \mathrm{dist}(S_{\alpha_j}, E_{k-1}) \lesssim 1/K_k}} |\widehat{g^k_{\alpha_j} d\sigma}(\xi)| \Big)^{1/(k-1)} \\
&\leqslant K_{k-1}^{2(k-2)} \Big(\frac{K_k}{K_{k-1}}\Big)^{k-2} \prod_{j=1}^{k-1} \max_{\alpha_j : S^k_{\alpha_j} \subseteq S^{k-1}_{\beta_j}} |\widehat{g^k_{\alpha_j} d\sigma}(\xi)|^{1/(k-1)}
\end{aligned}$$

$$\leqslant P(K_k)P(K_{k-1})\prod_{j=1}^{k-1}\Big(\sum_{\alpha_j:S^k_{\alpha_j}\subseteq S^{k-1}_{\beta_j}}|\widehat{g^k_{\alpha_j}d\sigma}(\xi)|^{q/(k-1)}\Big)^{1/q}$$

$$\lesssim P(K_k)P(K_{k-1})\Big(\sum_{\substack{\alpha_1,\dots,\alpha_{k-1}\\ \alpha_j:S^k_{\alpha_j}\subseteq S^{k-1}_{\beta_j}}}\prod_{j=1}^{k-1}|\widehat{g^k_{\alpha_j}d\sigma}(\xi)|^{q/(k-1)}\Big)^{1/q}$$

$$\leqslant P(K_k)P(K_{k-1})\Big(\sum_{\substack{\beta_1,\dots,\beta_{k-1}\\ \text{trans}}}\sum_{\substack{\alpha_1,\dots,\alpha_{k-1}\\ \alpha_j:S^k_{\alpha_j}\subseteq S^{k-1}_{\beta_j}}}\prod_{j=1}^{k-1}|\widehat{g^k_{\alpha_j}d\sigma}(\xi)|^{q/(k-1)}\Big)^{1/q}$$

$$\leqslant P(K_k)P(K_{k-1})\Big(\sum_{\substack{\alpha_1,\dots,\alpha_{k-1}\\ \text{trans}}}\prod_{j=1}^{k-1}|\widehat{g^k_{\alpha_j}d\sigma}(\xi)|^{q/(k-1)}\Big)^{1/q},$$

where the last sum is taken over all $\alpha_1,\dots,\alpha_k$ such that $S^k_{\alpha_1},\dots,S^k_{\alpha_{k-1}}$ are transversal. Observe that the transversal caps appearing in the right hand side of the above estimate are independent of Q^{k-1} and Q^k, which will allows us to sum both in $Q^{k-1}\subset Q^k$ and Q^k. Taking the q-th power and integrating for every $Q^{k-1}\subset Q^k$ for which (i) holds,

$$\sum_{\substack{Q^k\\ \text{(II) holds}}}\sum_{\substack{Q^{k-1}\subset Q^k\\ \text{(i) holds}}}\Big\|\sum_\beta\widehat{g^{k-1}_\beta d\sigma}\Big\|^q_{L^q(Q^{k-1})}$$

$$\lesssim P(K_k)P(K_{k-1})\sum_{\substack{\alpha_1,\dots,\alpha_{k-1}\\ \text{trans}}}\sum_{\substack{Q^k\\ \text{(II) holds}}}\sum_{\substack{Q^{k-1}\subset Q^k\\ \text{(i) holds}}}\Big\|\prod_{j=1}^{k-1}\widehat{g^k_{\alpha_j}d\sigma}\Big\|^{q/(k-1)}_{L^{q/(k-1)}(Q^{k-1})}$$

$$\leqslant P(K_k)P(K_{k-1})\sum_{\substack{\alpha_1,\dots,\alpha_{k-1}\\ \text{trans}}}\Big\|\prod_{j=1}^{k-1}\widehat{g^k_{\alpha_j}d\sigma}\Big\|^{q/(k-1)}_{L^{q/(k-1)}(B(0,R))}$$

$$\lesssim R^\varepsilon P(K_k)P(K_{k-1})\sum_{\substack{\alpha_1,\dots,\alpha_{k-1}\\ \text{trans}}}\prod_{j=1}^{k-1}\|g^k_{\alpha_j}\|^{q/(k-1)}_{L^p(d\sigma)}$$

$$\lesssim R^\varepsilon P(K_k)P(K_{k-1})\sum_\alpha\|g^k_\alpha\|^q_{L^p(d\sigma)}$$

$$\lesssim R^\varepsilon P(K_k)P(K_{k-1})\|g\|^q_{L^p(d\sigma)},$$

which is an acceptable term for $q\geqslant 2(k-1)/(k-2)$ by the multilinear estimate (5.1.6) in

Theorem 5.1.7. We note that in the one-to-last inequality we have argued as in the end of the case (I).

The second possibility for (i) consists in a slightly different use of the multilinear estimates (5.1.6), that exploits that we are not under the k-transversal case (I). The multilinear estimates are used to obtain a multilinear version of Córdoba's square function estimate [35], see Remark F.2 in Appendix F. We adopt this approach when $q < \frac{2(k-1)}{k-2}$. First, proceeding in the same way as in (I),

$$\begin{aligned}\Big|\sum_{\beta}\widehat{g_\beta^{k-1}d\sigma}(\xi)\Big| &\leqslant K_{k-1}^{2(k-2)}\prod_{j=1}^{k-1}|\widehat{g_{\beta_j}^{k-1}d\sigma}(\xi)|^{1/(k-1)}\\ &\leqslant K_{k-1}^{2(k-2)}\Big(\sum_{\substack{\beta_1,\dots,\beta_{k-1}\\ \text{trans}}}\prod_{j=1}^{k-1}|\widehat{g_{\beta_j}^{k-1}d\sigma}(\xi)|^{q/(k-1)}\Big)^{1/q},\end{aligned}$$

for any $\xi \in Q^{k-1}$, where the sum is taken over all $\beta_1,\dots,\beta_{k-1}$ such that $S_{\beta_1}^{k-1},\dots,S_{\beta_{k-1}}^{k-1}$ are transversal. As in previous cases, such sum is taken so that the transversal caps in right hand side are independent of Q^{k-1}; this will allows us to sum in Q^{k-1}. The contribution of those $Q^{k-1}\subset Q^k$ for which (i) holds is given by

$$\begin{aligned}&\sum_{\substack{Q^k\\ \text{(II) holds}}}\sum_{\substack{Q^{k-1}\subset Q^k\\ \text{(i) holds}}}\Big\|\sum_{\beta}\widehat{g_\beta^{k-1}d\sigma}\Big\|_{L^q(Q^{k-1})}^q\\ &\leqslant P(K_{k-1})\sum_{\substack{Q^k\\ \text{(II) holds}}}\sum_{\substack{\beta_1,\dots,\beta_{k-1}\\ \text{trans}}}\sum_{\substack{Q^{k-1}\subset Q^k\\ \text{(i) holds}}}\Big\|\prod_{j=1}^{k-1}\widehat{g_{\beta_j}^{k-1}d\sigma}\Big\|_{L^{q/(k-1)}(Q^{k-1})}^{q/(k-1)}\\ &\leqslant P(K_{k-1})\sum_{\substack{Q^k\\ \text{(II) holds}}}\sum_{\substack{\beta_1,\dots,\beta_{k-1}\\ \text{trans}}}\Big\|\prod_{j=1}^{k-1}\Big(\sum_{\substack{\alpha_j:S_{\alpha_j}^k\subseteq S_{\beta_j}^{k-1}\\ \operatorname{dist}(S_{\alpha_j}^k,E_{k-1})\lesssim 1/K_k}}\widehat{g_{\alpha_j}^k d\sigma}\Big)\Big\|_{L^{q/(k-1)}(Q^k)}^{q/(k-1)}\\ &\lesssim P(K_{k-1})K_k^{\varepsilon}\sum_{\substack{Q^k\\ \text{(II) holds}}}\sum_{\substack{\beta_1,\dots,\beta_{k-1}\\ \text{trans}}}\Big\|\prod_{j=1}^{k-1}\Big(\sum_{\substack{\alpha_j:S_{\alpha_j}^k\subseteq S_{\beta_j}^{k-1}\\ \operatorname{dist}(S_{\alpha_j}^k,E_{k-1})\lesssim 1/K_k}}|\widehat{g_{\alpha_j}^k d\sigma}|^2\Big)^{1/2}\Big\|_{L^{q/(k-1)}(Q^k)}^{q/(k-1)},\end{aligned}$$

where the last inequality follows from Remark F.2 for any $2 \leqslant q \leqslant \frac{2k}{k-1}$; this may be seen as a multilinear square function estimate. By Hölder's inequality, and using the information that the caps $S^k_{\alpha_j}$ concentrate among E_{k-1},

$$
\begin{aligned}
&\lesssim P(K_{k-1})K_k^{\varepsilon}\Big(\frac{K_k}{K_{k-1}}\Big)^{(k-2)(\frac{q}{2}-1)} \sum_{\substack{Q^k \\ \text{(II) holds}}} \sum_{\substack{\beta_1,\dots,\beta_{k-1} \\ \text{trans}}} \Big\| \prod_{j=1}^{k-1} \Big(\sum_{\substack{\alpha_j : S^k_{\alpha_j} \subseteq S^{k-1}_{\beta_j} \\ \operatorname{dist}(S^k_{\alpha_j}, E_{k-1}) \lesssim 1/K_k}} |\widehat{g^k_{\alpha_j} d\sigma}|^q \Big)^{1/q} \Big\|^{q/(k-1)}_{L^{q/(k-1)}(Q^k)} \\
&\lesssim P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)} \sum_{\substack{Q^k \\ \text{(II) holds}}} \sum_{\beta} \Big\| \Big(\sum_{\substack{\alpha : S^k_{\alpha} \subseteq S^{k-1}_{\beta} \\ \operatorname{dist}(S^k_{\alpha}, E_{k-1}) \lesssim 1/K_k}} |\widehat{g^k_{\alpha} d\sigma}|^q \Big)^{1/q} \Big\|^{q}_{L^{q}(Q^k)} \\
&\lesssim P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)} \sum_{\substack{Q^k \\ \text{(II) holds}}} \sum_{\beta} \sum_{\substack{\alpha : S^k_{\alpha} \subseteq S^{k-1}_{\beta} \\ \operatorname{dist}(S^k_{\alpha}, E_{k-1}) \lesssim 1/K_k}} \|\widehat{g^k_{\alpha} d\sigma}\|^{q}_{L^{q}(Q^k)} \\
&\leqslant P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)} \sum_{\beta} \sum_{\alpha : S^k_{\alpha} \subseteq S^{k-1}_{\beta}} \sum_{Q^k} \|\widehat{g^k_{\alpha} d\sigma}\|^{q}_{L^{q}(Q^k)} \\
&\leqslant P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)} \sum_{\alpha} \|\widehat{g^k_{\alpha} d\sigma}\|^{q}_{L^{q}(B(0,R))} \\
&\leqslant P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)} K_k^{d+1-\frac{(d-1)q}{p'}} \sum_{\alpha} \|g^k_{\alpha}\|^{q}_{L^{p}(d\sigma)} \\
&\leqslant P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)+d+1-\frac{(d-1)q}{p'}} K_k^{\frac{d-1}{(p/q)'}} \Big(\sum_{\alpha} \|g^k_{\alpha}\|^{p}_{L^{p}(d\sigma)} \Big)^{q/p} \\
&\leqslant P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)+2d-(d-1)q} \mathcal{C}(R) \|g\|^{q}_{L^{p}(d\sigma)},
\end{aligned}
$$

where we have used the parabolic rescaling condition in Appendix E. This use of induction hypothesis makes powers of K_k to be relevant again in our argument. In order for the above estimate to be an acceptable term we ask $P(K_{k-1})K_k^{\varepsilon+(k-2)(\frac{q}{2}-1)+2d-(d-1)q} \ll R^{\varepsilon}$, which will be true if

$$
(k-2)\left(\frac{q}{2}-1\right) + 2d - (d-1)q < 0
$$

and we pick K_k sufficiently large and independent of R; note that the choice of K_k will depend on the K_{k-1} chosen for an equivalent condition to (5.2.4) to hold for K_{k-1} in the next step. The above condition may be written as

$$q > \frac{2(2d-(k-2))}{2d-k}.$$

So we choose the second possibility for (i) whenever

$$\frac{2(2d-(k-2))}{2d-k} < \frac{2(k-1)}{k-2}.$$

Hence for the case (i) we choose the option that gives us a wider range of values for q, that is

$$q > 2\min\left(\frac{k-1}{k-2}, \frac{2d-k+2}{2d-k}\right).$$

For (ii) we do a similar analysis to case (II), replacing $k-1$ by $k-2$. One ultimately obtains that if

$$q > 2\min\left(\frac{m}{m-1}, \frac{2d-m+1}{2d-m-1}\right), \quad 2 \leqslant m \leqslant k-1,$$

$q \geqslant \frac{2d}{d-1}$ and $\mathcal{R}^*(p \times \cdots \times p \to q/k)$ holds, then the restriction estimate $\mathcal{R}^*(p \to q)$ follows for $p \geqslant q$. $\square$

5.3 Analysis of the exponents

The Bourgain-Guth method described in Section 5.2 suggests than better linear restrictions estimates should be obtained in case one uses better multilinear restriction estimates than the ones given by Theorem 5.1.7. The aim of this section is to discuss this issue, together with yielding a proof for the conditional Theorem 5.1.9.

It is reasonable to expect that the k-linear restriction conjectured estimates that might be proven in the future are the ones corresponding to functions on $L^2(d\sigma)$, due to the wave packet decomposition and orthogonality considerations. Thus, under the hypothesis of Conjecture 5.1.5, one hopes the local estimates

$$\Big\| \prod_{j=1}^{k} \widehat{g_j d\sigma_j} \Big\|_{L^{\frac{2(d+k)}{k(d+k-2)}}(B(0,R))} \lesssim R^{\varepsilon} \prod_{j=1}^{k} \|g_j\|_{L^2(d\sigma_j)} \tag{5.3.1}$$

to hold for any $\varepsilon > 0$. Observe that (5.3.1) corresponds to a k-linear version of Tao's bilinear estimate in [132]. Note that in contrast with the known k-linear L^2-estimates from Theorem 5.1.7, the estimates (5.3.1) involve the curvature hypothesis.

A natural question is to understand how the Bourgain-Guth method would improve the current state-of-the-art on the linear restriction conjecture in case we knew the above conjectured estimates (5.3.1) to be true.

Before proceeding with our analysis, we should introduce the following notation. For any given $2 \leqslant k \leqslant d$, we denote

- $q_{cs}(k) := \frac{2(d+k)}{d+k-2}$ the exponent corresponding to the conjectured multilinear estimate (5.3.1), where the subscript in q_{cs} refers to *curvature sensitive.*

- $q_{ci}(k) := \frac{2k}{k-1}$ the exponent corresponding to the known multilinear estimates (5.1.6), where the subscript in q_{ci} refers to *curvature insensitive.*

- $q_{sf}(k) := \frac{2(2d-k+1)}{2d-k-1}$ the exponent obtained throughout the proof of Theorem 5.2.1 after the use of the multilinear theory via a multilinear square function estimate, where the subscript in q_{sf} refers to *square function.*

5.3.1 The trilinear case

We study first if we would obtain any improvement for the linear restriction conjecture via Theorem 5.2.1, that is the Bourgain-Guth method, in case we knew the conjectured estimate (5.3.1) for $k = 3$, that is $\mathcal{R}^*(p \times p \times p \to q/3)$ for $q > q_{cs}(3) = \frac{2(d+3)}{d+1}$. Observe that the conditions imposed on q by Theorem 5.2.1 for $k = 3$ are $\frac{1}{p} \leqslant \frac{1}{q} < \frac{d-1}{2d}$ and

$q > (4d-2)/(2d-3)$. The conjectured exponent $q > q_{cs}(3) = 2(d+3)/(d+1)$ is an admissible exponent for $d \geqslant 4$, since

$$\frac{2(d+3)}{d+1} \geqslant \frac{4d-2}{2d-3} \Leftrightarrow \frac{1}{d+1} \geqslant \frac{1}{2d-3} \Leftrightarrow d \geqslant 4.$$

However the use of the conjectured trilinear estimate with $q > q_{cs}(3)$ would only make improvement on the Bourgain–Guth state-of-the-art for the linear restriction conjecture in the case $d = 4$. This may be easily checked comparing $q_{cs}(3)$ with the exponents in Theorem 5.1.8.

- For $d \equiv 0 \pmod 3$,

$$\frac{2(d+3)}{d+1} < 2\frac{4d+3}{4d-3} \Leftrightarrow \frac{2}{d+1} < \frac{6}{4d-3} \Leftrightarrow 4d-3 < 3d+3 \Leftrightarrow d < 6,$$

  so no improvement would be obtained in this case.

- For $d \equiv 1 \pmod 3$,

$$\frac{2(d+3)}{d+1} < \frac{2d+1}{d-1} \Leftrightarrow \frac{4}{d+1} < \frac{3}{d-1} \Leftrightarrow 4d-4 < 3d+3 \Leftrightarrow d < 7,$$

  so in the case $d = 4$, the conjectured trilinear estimate would improve on the Bourgain–Guth results for the linear restriction conjecture.

- For $d \equiv 2 \pmod 3$,

$$\frac{2(d+3)}{d+1} < \frac{4(d+1)}{2d-1} \Leftrightarrow \frac{2}{d+1} < \frac{3}{2d-1} \Leftrightarrow 4d-2 < 3d+3 \Leftrightarrow d < 5,$$

  so no improvement would be obtained in this case.

The above observation suggests that for higher dimensions, a “good” trilinear estimate is "less efficient" than a "worse" but higher level of linearity estimate, where that "higher" level of linearity is "close" to the dimension. We make that informal comment more precise in the coming subsections.

5.3.2 The k-linear case

Here we study how to use the conjectured estimate (5.3.1) for a fixed k to obtain improvement for the linear restriction conjecture in $\mathbb{R}^d$; in particular we deduce for which dimensions $d = d(k)$ the conjectured k-linear estimate would provide improvement. Observe that $q_{sf}(m) < q_{ci}(m)$ if and only if

$$\frac{2d-m+1}{2d-m-1} < \frac{m}{m-1} \Leftrightarrow \frac{2}{2d-m-1} < \frac{1}{m-1} \Leftrightarrow 3m < 2d+1 \Leftrightarrow m < \frac{2d+1}{3}.$$

The condition (5.2.1) in Theorem 5.2.1 implies, in particular that

$$\begin{aligned} q > q_{sf}(k-1) \quad &\text{if} \quad k < \frac{2d+4}{3}, \\ q > q_{ci}(k-1) \quad &\text{if} \quad k \geqslant \frac{2d+4}{3}. \end{aligned}$$

As $q_{ci}(k-1) > q_{ci}(k) > q_{cs}(k)$, the conjectured exponent $q_{cs}(k)$ is not an admissible exponent for Theorem 5.2.1 when $k \geqslant \frac{2d+4}{3}$; in other words, the conjectured exponent $q_{cs}(k)$ would not lead to any improvement (via the Bourgain–Guth argument) on the linear restriction conjecture in $\mathbb{R}^d$ if $k \geqslant \frac{2d+4}{3}$.

Thus we only consider those d such that $k < \frac{2d+4}{3}$. Since $q_{sf}(m)$ is increasing as a function of m, the condition (5.2.1) on the exponent q becomes $q > q_{sf}(k-1)$. In view of Theorem 5.2.1, an admissible value for q is given by

$$q > \max(q_{cs}(k), q_{sf}(k-1)).$$

Observe that $q_{cs}(k) \geqslant q_{sf}(k-1)$ if and only if

$$\frac{d+k}{d+k-2} \geqslant \frac{2d-k+2}{2d-k} \Leftrightarrow \frac{1}{d+k-2} \geqslant \frac{1}{2d-k} \Leftrightarrow d \geqslant 2k-2 \Leftrightarrow k \leqslant \frac{d+2}{2}.$$

Then

$$q > \max(q_{cs}(k), q_{sf}(k-1)) = \begin{cases} q_{cs}(k) & \text{if } k \leqslant (d+2)/2, \\ q_{sf}(k-1) & \text{if } (d+2)/2 \leqslant k < (2d+4)/3. \end{cases}$$

We compare this value of q with the Bourgain–Guth state-of-the-art for the linear restriction conjecture (Theorem 5.1.8), to detect when the conjectured inequality (5.3.1) would lead to an improvement.

We distinguish two cases. For $k \leqslant (d+2)/2$, the condition on the exponent q is given by $q > q_{cs}(k)$. We compare this with the exponents in Theorem 5.1.8.

- For the case $d \equiv 0 \pmod 3$,

$$\frac{2(d+k)}{d+k-2} < \frac{2(4d+3)}{4d-3} \Leftrightarrow \frac{1}{d+k-2} < \frac{3}{4d-3} \Leftrightarrow k > \frac{d+3}{3},$$

  so we would get improvement when $k > (d+3)/3$.

- For the case $d \equiv 1 \pmod 3$,

$$\frac{2(d+k)}{d+k-2} < \frac{2d+1}{d-1} \Leftrightarrow \frac{4}{d+k-2} < \frac{3}{d-1} \Leftrightarrow k > \frac{d+2}{3},$$

  so we would get improvement when $k > (d+2)/3$.

- For the case $d \equiv 2 \pmod 3$,

$$\frac{2(d+k)}{d+k-2} < \frac{4(d+1)}{2d-1} \Leftrightarrow \frac{2}{d+k-2} < \frac{3}{2d-1} \Leftrightarrow k > \frac{d+4}{3},$$

so we would get improvement when $k > (d+4)/3$.

On the other hand, for $\frac{(d+2)}{2} \leqslant k < \frac{2d+4}{3}$, the condition on q is given by $q > q_{sf}(k-1)$. A similar case analysis as before tells us that,

- For the case $d \equiv 0 \pmod 3$,

$$\frac{2d-k+2}{2d-k} < \frac{4d+3}{4d-3} \Leftrightarrow \frac{1}{2d-k} < \frac{3}{4d-3} \Leftrightarrow k < \frac{2d+3}{3},$$

  so we would get improvement when $k < (2d+3)/3$.

- For the case $d \equiv 1 \pmod 3$,

$$\frac{2(2d-k+2)}{2d-k} < \frac{2d+1}{d-1} \Leftrightarrow \frac{4}{2d-k} < \frac{3}{d-1} \Leftrightarrow k < \frac{2d+4}{3},$$

  so we would get improvement when $k < (2d+4)/3$.

- For the case $d \equiv 2 \pmod 3$,

$$\frac{2(2d-k+2)}{2d-k} < \frac{4(d+1)}{2d-1} \Leftrightarrow \frac{2}{2d-k} < \frac{3}{2d-1} \Leftrightarrow k < \frac{2d+2}{3},$$

  so we would get improvement when $k < (2d+2)/3$.

Hence, given k we obtain progress on the linear restriction conjecture for those d such that $\frac{d}{3} \lesssim k \lesssim \frac{2d}{3}$, that is for those d such that $\frac{3k}{2} \lesssim d \lesssim 3k$.

5.3.3 Optimal level of linearity for a given dimension d

Here we study, for a fixed dimension d, which is the level of linearity $k = k(d)$ that gives the best improvement on the linear restriction conjecture in $\mathbb{R}^d$ via the Bourgain–Guth method. We refer to such level as the *optimal* level of linearity, in the sense that by using the estimates (5.3.1) in higher levels of linearity we no longer obtain any improvement on the linear restriction conjecture.

From Section 5.3.2 we know that improvement on the linear problem requires $k < \frac{2d+4}{3}$. Recall that an admissible q for Theorem 5.2.1 needs to satisfy

$$q > \max(q_{cs}(k), q_{sf}(k-1)) = \begin{cases} q_{cs}(k) & \text{if } k \leqslant (d+2)/2, \\ q_{sf}(k-1) & \text{if } (d+2)/2 \leqslant k < (2d+4)/3. \end{cases}$$

One should observe the following:

- as $q_{sf}(k-1)$ is increasing as a function of k, we would like k to be the smallest integer in the range $[\frac{d+2}{2}, \frac{2d+4}{3})$, i.e., $k \leqslant \lfloor \frac{d+3}{2} \rfloor$.

- as $q_{cs}(k)$ is decreasing as a function of k, we would like k to be the biggest integer in $[2, \frac{d+2}{2}]$, i.e., $k \geqslant \lfloor \frac{d+1}{2} \rfloor$.

This tells us that the optimal level of linearity satisfies $\lfloor \frac{d+1}{2} \rfloor \leqslant k \leqslant \lfloor \frac{d+3}{2} \rfloor$. Observe that

- if $(d+2)/2 \in \mathbb{N}$, that is d even, the value $k = \frac{d+2}{2}$ is the best level of linearity; in this case $q_{cs}(k) = q_{sf}(k-1)$.

- if $(d+2)/2 \notin \mathbb{N}$, that is d odd, we need to compare $q_{cs}(k)$ for $k = (d+1)/2$ and $q_{sf}(k-1)$ for $k = (d+3)/2$ and choose the smallest q. But it turns out that

$$q_{cs}\Big(\frac{d+1}{2}\Big) = 2\frac{3d+1}{3d-3} = 2\frac{3d-1}{3d-3} = q_{sf}\Big(\frac{d+3}{2} - 1\Big),$$

  so we may choose $k = (d+1)/2$ as the optimal level.

Thus the level of linearity that leads to the biggest improvement given a fixed dimension d is $k = \lfloor \frac{d+2}{2} \rfloor$. In particular the above analysis tells us that the conjectured estimate (5.3.1) with $k = \lfloor \frac{d+2}{2} \rfloor$ would establish the restriction conjecture for

- $q > 2\frac{3d+1}{3d-3}$ for d odd,
- $q > 2\frac{3d+2}{3d-2}$ for d even,

which is the statement of Theorem 5.1.9.

Connections with the Schrödinger propagator

As it is mentioned at the end of Chapter 2, the solution to the free linear Schrödinger equation,

$$i\partial_s u - \Delta u = 0, \qquad u(0,x) = f(x),$$

where $(s,x) \in \mathbb{R}^{1+d}$, satisfies

$$\mathcal{F}_d u(s,\xi) = e^{is|\xi|^2}\mathcal{F}_d f(\xi),$$

where $\mathcal{F}_d$ denotes the spatial Fourier transform (in $\mathbb{R}^d$). Then one may write the solution as

$$u(s,x) = e^{-is\Delta}f(x) = \int_{\mathbb{R}^d} \mathcal{F}_d f(\xi) e^{i(s|\xi|^2 + x\cdot\xi)} d\xi.$$

Observe that this corresponds to the extension operator associated to the paraboloid on $\mathbb{R}^{d+1}$. Let S denote the paraboloid, and let $\Sigma : \mathbb{R}^d \to \mathbb{R}^{d+1}$ be the parametrisation given by $\Sigma(x) = (x, |x|^2)$. Then for a function g defined on S, we have that

$$\widehat{gd\mu}(\xi) = \int_S g(y) e^{iy\cdot\xi} d\mu(y) = \int_{\mathbb{R}^d} g(\Sigma(x)) e^{i\Sigma(x)\cdot\xi} dx = \int_{\mathbb{R}^d} g(x,|x|^2) e^{i(x\cdot\xi' + |x|^2\cdot\xi_{d+1})} dx,$$

where $\xi' = (\xi_1, \ldots, \xi_d)$; here $\widehat{\ }$ denotes the $\mathbb{R}^{d+1}$ Fourier transform. Setting $g \circ \Sigma = \mathcal{F}_d f$, it is obvious that the expressions for u and $\widehat{gd\mu}$ coincide.

This trivial observation emphasises the importance of the Fourier restriction phenomenon in dispersive PDE. In particular, the theory of Strichartz estimates has been intimately related to that of the Fourier restriction. By a Strichartz estimate we mean control of the full norm of the solution u, integrating in both time and space, in terms of the size of the initial data f. For example, it is known that

$$\|e^{-is\Delta} f\|_{L^q_s L^r_x(\mathbb{R}^{1+d})} \lesssim \|f\|_{L^2(\mathbb{R}^d)}$$

for

$$\frac{2}{q} + \frac{d}{r} = \frac{d}{2}, \quad (q, r, d) \neq (2, \infty, 2), \;\; q, r \geqslant 2.$$

This is in contrast with the weighted estimates for the operator $e^{-is\Delta}$ obtained in Section 2.6.2, where only integration in the space-variable is taken.

The case $q = \infty$ corresponds to estimates for the maximal Schrödinger operator u^*, whose boundedness implies almost everywhere convergence of the solution $u(\cdot, s)$ to the initial data $f \in L^2$ as s approaches 0. More generally one may formulate Strichartz estimates for initial data f in the homogeneous and inhomogeneous Sobolev spaces $\dot{H}^\sigma$ and H^σ respectively. In the context of the maximal Schrödinger operator, determining the optimal Sobolev space H^σ of initial data for which there is a.e. convergence is known as the *Carleson problem* [27]. Many authors have contributed over the last decades to this question [39, 122, 140, 12, 100, 133, 134, 81, 13, 14], which is still open for $d \geqslant 3$. Most of the progress has been obtained via a Fourier restriction approach; in particular, only a few months ago, Du, Guth and Li [44] have established the 2-dimensional case except for the endpoint case, using Fourier restriction theory and polynomial partitioning. These connections with Fourier restriction theory suggest that an inequality of the type (2.6.11) could perhaps be obtained via weighted Fourier restriction estimates rather than via the techniques used in Chapter 2.

Appendix A

Smooth averages

Here we briefly recall some elementary properties of the smoothing functions

$$\Psi_R^{(N)}(x) := \frac{R^d}{(1 + R^2|x|^2)^{N/2}}.$$

Lemma A.1. *Let $N > d$. Let $R \geqslant K$ denote two different scales. Then*

$$\Psi_R^{(N)} * \Psi_K^{(N)} \lesssim \Psi_K^{(N)}.$$

Proof. We need to show

$$\int_{\mathbb{R}^d} \frac{R^d}{(1 + R^2|y - x|^2)^{N/2}} \frac{K^d}{(1 + K^2|y|^2)^{N/2}} dy \lesssim \frac{K^d}{(1 + K^2|x|^2)^{N/2}}$$

for any $x \in \mathbb{R}^d$. Observe first that if $K|x| \leqslant 1$, the estimate is trivial, as

$$\frac{K^d}{(1 + K^2|y|^2)^{N/2}} \leqslant K^d \leqslant \frac{2^{N/2} K^d}{(1 + K^2|x|^2)^{N/2}}$$

and the integral

$$\int_{\mathbb{R}^d} \frac{R^d}{(1 + R^2|y - x|^2)^{N/2}} dy < \infty$$

provided $N > d$.

If $K|x| \geqslant 1$, we divide $\mathbb{R}^d$ into two half-spaces H_x and H_0, that contain the points x and 0 respectively and that are the result of splitting $\mathbb{R}^d$ by a hyperplane perpendicular to the line segment joining x and the origin 0 at its midpoint. If $y \in H_x$, then $|y| \geqslant |x|/2$ and

$$\frac{K^d}{(1+K^2|y|^2)^{N/2}} \leqslant \frac{2^N K^d}{(1+K^2|x|^2)^{N/2}}.$$

Thus

$$\begin{aligned}\int_{H_x} \frac{R^d}{(1+R^2|y-x|^2)^{N/2}} \frac{K^d}{(1+K^2|y|^2)^{N/2}} dy &\leqslant \frac{2^N K^d}{(1+K^2|x|^2)^{N/2}} \int_{H_x} \frac{R^d}{(1+R^2|y-x|^2)^{N/2}} dy \\ &\lesssim \frac{K^d}{(1+K^2|x|^2)^{N/2}}.\end{aligned}$$

If $y \in H_0$, we have $|y-x| \geqslant |x|/2$. Similarly,

$$\frac{R^d}{(1+R^2|y-x|^2)^{N/2}} \leqslant \frac{2^N R^d}{(1+R^2|x|^2)^{N/2}} \leqslant \frac{2^N R^d}{R^N |x|^N} = \frac{2^N R^{d-N}}{|x|^N}.$$

As $R > K$, $N > d$ and $K|x| \geqslant 1$,

$$\frac{2^N R^{d-N}}{|x|^N} \leqslant \frac{2^N K^{d-N}}{|x|^N} = \frac{2^N 2^{N/2} K^d}{(2K^2|x|^2)^{N/2}} \lesssim \frac{K^d}{(1+K^2|x|^2)^{N/2}},$$

and arguing as in the previous case, this concludes the proof. □

For the case $R = 1$, we simply denote $\Psi^{(N)}(x) := \frac{1}{(1+|x|^2)^{N/2}}$. We have the following Harnack-type property.

Lemma A.2. *For $w \geqslant 0$,*

$$w * \Psi^{(N)}(x) \gtrsim \frac{1}{(1+|x-y|^2)^{N/2}} w * \Psi^{(N)}(y).$$

Proof. The triangle inequality quickly reveals that $(1+|x|^2)^{-N/2} \gtrsim (1+|x-y|^2)^{-N/2}(1+|y|^2)^{-N/2}$ for any $N \geqslant 0$. Then, as $w \geqslant 0$,

$$\frac{w(z)}{(1+|x-z|^2)^{N/2}} \gtrsim \frac{1}{(1+|x-y|^2)^{N/2}} \frac{w(z)}{(1+|y-z|^2)^{N/2}},$$

just by replacing $x \mapsto x-z$, $y \mapsto y-z$. The result follows from integrating with respect to the z variable. $\square$

Appendix B

Symbolic Calculus

This appendix is devoted to providing a proof of Theorem 3.3.1, which is a very specific quantitative version of the symbolic calculus in Hörmander [69]. Recall the statement.

Theorem B.1. *Let* $\varphi \in \mathcal{S}$ *be such that* $\operatorname{supp}(\widehat{\varphi}) \subseteq \{|\xi| \sim 1\}$ *and given* $R > 1$*, let* φ_R *be defined by* $\widehat{\varphi}_R(\xi) := \widehat{\varphi}(R^{-1}\xi)$*. Let* $a \in S^m_{\rho,\delta}$*, where* $0 \leqslant \delta \leqslant \rho$ *and* $\delta < 1$*. Then, there exists a symbol* $c \in S^m_{\rho,\delta}$ *such that*

$$T_c = T_{\widehat{\varphi}_R} \circ T_a.$$

Moreover, for $\epsilon \geqslant 0$ *and* $\kappa > 0$*, the symbol*

$$e^N := c - \sum_{|\gamma|<N} \frac{i^{-|\gamma|}}{\gamma!} \partial_\xi^\gamma \widehat{\varphi}_R \partial_x^\gamma a \in S^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon}_{\rho,\delta}$$

for all $N > \frac{d\delta+\kappa\delta+\epsilon}{1-\delta}$*, and satisfies*

$$|\partial_x^\nu \partial_\xi^\sigma e^N(x,\xi)| \lesssim R^{-\epsilon}(1+|\xi|)^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon-|\sigma|\rho+|\nu|\delta} \tag{B.1}$$

for any multi-indices $\nu, \sigma \in \mathbb{N}^d$.

As it is mentioned in Section 3.3.1, the order of the error symbol $e^N \in S^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon}_{\rho,\delta}$ is not necessarily sharp here, but it naturally arises from our proof. Nevertheless, such an order is admissible for our purposes, as one may choose N large enough so that e^N is of sufficiently large negative order. Our proof follows the same structure as that given in Stein [129] for the standard symbol classes S^m.

To justify our computations, we technically should replace a by a_ε, where $a_\varepsilon(x,\xi) = a(x,\xi)\psi(\varepsilon x, \varepsilon\xi)$ and $\psi \in C_0^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ with $\psi(0,0) = 1$. The symbol a_ε, which has compact support, satisfies the same differential inequalities as a uniformly in $0 < \varepsilon \leqslant 1$. As our estimates will be independent of ε, the passage to the limit when $\varepsilon \to 0$ gives the desired result; we refer to [129] for these standard details. Such considerations allow us to suppress the dependence on ε in what follows.

Proof. Observe that we may write

$$T_{\widehat{\varphi}_R}(T_a f)(x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} c(x,\xi)e^{i(x-z)\cdot\xi} f(z)\,dz\,d\xi,$$

where

$$c(x,\xi) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \widehat{\varphi}_R(\eta)a(y,\xi)e^{i(x-y)\cdot(\eta-\xi)}\,dy\,d\eta = \int_{\mathbb{R}^d} \widehat{\varphi}_R(\xi+\eta)\widehat{a}(\eta,\xi)e^{ix\cdot\eta}\,d\eta,$$

and $\widehat{a}$ denotes Fourier transform with respect to the x variable. We first obtain an estimate depending on the size of the support of a; that dependence will be later removed in the second part of the proof.

B.1 Assuming $a(x,\xi)$ has compact support in the x-variable

Integrating by parts,

$$\widehat{a}(\eta,\xi) = \int_{\mathbb{R}^d} \frac{e^{ix\cdot\eta}}{(1+|\eta|^2)^M}(I-\Delta_x)^M a(x,\xi)\,dx,$$

so

$$|\widehat{a}(\eta,\xi)| \lesssim (1+|\eta|)^{-2M}(1+|\xi|)^{m+2M\delta}, \tag{B.2}$$

for any $M \geqslant 0$; the implicit constant above depends on the size of the support of a in the x variable. For $\widehat{\varphi}_R(\xi+\eta)$ we use Taylor's formula around the point ξ,

$$\widehat{\varphi}_R(\xi+\eta) = \sum_{|\gamma|<N} \frac{1}{\gamma!}\partial_\xi^\gamma \widehat{\varphi}_R(\xi)\eta^\gamma + \mathfrak{R}_N(\xi,\eta),$$

where $\mathfrak{R}_N$ is the remainder in Taylor's theorem and is bounded by

$$|\mathfrak{R}_N(\xi,\eta)| \lesssim \max_{|\gamma|=N}\max_{\zeta} |\partial_\xi^\gamma \widehat{\varphi}_R(\zeta)||\eta|^N,$$

where the maximum in ζ is taken on the line segment joining ξ to $\xi+\eta$. Thus

$$\begin{aligned} c(x,\xi) &= \sum_{|\gamma|<N} \frac{1}{\gamma!}\int_{\mathbb{R}^d} \partial_\xi^\gamma \widehat{\varphi}_R(\xi)\eta^\gamma \widehat{a}(\eta,\xi)e^{ix\cdot\eta}d\eta + \int_{\mathbb{R}^d} \mathfrak{R}_N(\xi,\eta)\widehat{a}(\eta,\xi)e^{ix\cdot\eta}d\eta \\ &= \sum_{|\gamma|<N} \frac{i^{-|\gamma|}}{\gamma!}\partial_\xi^\gamma \widehat{\varphi}_R(\xi)\partial_x^\gamma a(x,\xi) + \int_{\mathbb{R}^d} \mathfrak{R}_N(\xi,\eta)\widehat{a}(\eta,\xi)e^{ix\cdot\eta}d\eta \end{aligned}$$

and

$$e^N(x,\xi) = \int_{\mathbb{R}^d} \mathfrak{R}_N(\xi,\eta)\widehat{a}(\eta,\xi)e^{ix\cdot\eta}d\eta.$$

We need to show that the $e^N \in S^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon}_{\rho,\delta}$ and satisfies the differential inequalities (B.1).

Observe that, for γ such that $|\gamma| = N$,

$$\partial_\xi^\gamma \widehat{\varphi}_R(\zeta) = R^{-N}(\partial_\xi^\gamma \widehat{\varphi})(R^{-1}\zeta) \lesssim R^{-\epsilon}(1+|\zeta|)^{-N+\epsilon},$$

as the support condition on $\widehat{\varphi}$ ensures $|\zeta| \sim R \sim |\zeta|+1$, since $R > 1$. This leads to the following estimates for the remainder,

$$|\mathfrak{R}_N(\xi,\eta)| \lesssim R^{-\epsilon}|\eta|^N(1+|\xi|)^{-N+\epsilon} \quad \text{for} \quad |\xi| \geqslant 2|\eta|,$$

and

$$|\mathfrak{R}_N(\xi,\eta)| \lesssim R^{-\epsilon}|\eta|^N \quad \text{for} \quad |\xi| \leqslant 2|\eta|,$$

as $N \geqslant \epsilon$. Using the estimate (B.2) in the form

$$|\widehat{a}(\eta,\xi)| \lesssim (1+|\eta|)^{-2M_1}(1+|\xi|)^{m+2M_1\delta} \quad \text{for} \quad |\xi| \geqslant 2|\eta|,$$

and

$$|\widehat{a}(\eta,\xi)| \lesssim (1+|\eta|)^{-2M_2}(1+|\xi|)^{m+2M_2\delta} \quad \text{for} \quad |\xi| \leqslant 2|\eta|,$$

where $M_1, M_2 \geqslant 0$, we have

$$\begin{aligned}|e^N(x,\xi)| &\lesssim R^{-\epsilon}(1+|\xi|)^{m+2M_1\delta-N+\epsilon}\int_{|\xi|\geqslant 2|\eta|}(1+|\eta|)^{-2M_1}|\eta|^N d\eta \\ &\quad + R^{-\epsilon}(1+|\xi|)^{m+2M_2\delta}\int_{|\xi|\leqslant 2|\eta|}(1+|\eta|)^{-2M_2}|\eta|^N d\eta \\ &\lesssim R^{-\epsilon}(1+|\xi|)^{m+2M_1\delta-N+\epsilon} + R^{-\epsilon}(1+|\xi|)^{m+2M_2\delta-2M_2+N+d}\end{aligned}$$

provided $-2M_1+N+d<0$ and $-2M_2+N+d<0$. Choosing

$$M_1 = (N+d+\kappa)/2$$

and

$$M_2 = \frac{2N+d(1-\delta)-\kappa\delta-\epsilon-N\delta}{2(1-\delta)},$$

which clearly satisfies the condition $-2M_2 + N + d < 0$, as $N > \epsilon + \kappa\delta$, one has

$$|e^N(x,\xi)| \lesssim R^{-\epsilon}(1+|\xi|)^{m-N(1-\delta)+d\delta+\kappa\delta+\epsilon}.$$

In view of the definitions of the symbols c and e^N, the use of the Leibniz formula and the condition $\rho \leqslant 1$ allows one, by the same arguments as above, to deduce the differential inequalities (B.1) for all multi-indices $\nu, \sigma \in \mathbb{N}^d$.

B.2 The case of general $a(x,\xi)$

It suffices to prove the differential inequalities (B.1) for x near an arbitrary but fixed point x_0; in particular we prove them for x such that $|x-x_0| \leqslant 1/2$, with bounds independent of x_0. To this end, let θ be a smooth function which equals 1 on $|y-x_0| \leqslant 1$ and supported in $|y-x_0| \leqslant 2$, and write $a = \theta a + (1-\theta)a = a_1 + a_2$. For a_1, one may argue as before and write

$$c_1(x,\xi) = \sum_{|\gamma|<N} \frac{i^{-|\gamma|}}{\gamma!}(\partial_\xi^\gamma \widehat{\varphi}_R(\xi))(\partial_x^\gamma a_1(x,\xi)) + \int_{\mathbb{R}^d} \mathfrak{R}_N(\xi,\eta)\widehat{a}_1(\eta,\xi)e^{ix\cdot\eta}d\eta.$$

As $a_1 = a$ for $|x-x_0| \leqslant 1/2$ and the size of the support of a_1 in the x variable is constant and independent of x_0, the previous argument reveals that the symbol

$$e_1^N(x,\xi) := c_1(x,\xi) - \sum_{|\gamma|<N} \frac{i^{-|\gamma|}}{\gamma!}(\partial_\xi^\gamma \widehat{\varphi}_R(\xi))(\partial_x^\gamma a(x,\xi))$$

satisfies the differential inequalities (B.1) for $|x-x_0| \leqslant 1/2$, with bounds independent of x_0. As

$$|e^N(x,\xi)| \leqslant |e_1^N(x,\xi)| + |c_2(x,\xi)|,$$

where c_2 is the symbol of $T_{\widehat{\varphi}_R} \circ T_{a_2}$, it is enough to show that c_2 satisfies the same estimates as e_1^N. Indeed, we will show that for $|x - x_0| \leqslant 1/2$,

$$|c_2(x,\xi)| \lesssim R^{-\epsilon}(1+|\xi|)^{m-\bar{N}}$$

for any $\bar{N} \geqslant 0$; the proof then follows by taking

$$\bar{N} = N(1-\delta) - d\delta - \kappa\delta - \epsilon,$$

which is nonnegative for $N > \frac{d\delta+\kappa\delta+\epsilon}{1-\delta}$. Recall that

$$c_2(x,\xi) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \widehat{\varphi}_R(\eta) a_2(y,\xi) e^{i(x-y)\cdot(\eta-\xi)} dy d\eta.$$

Integrating by parts with respect to the η variable,

$$c_2(x,\xi) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\Delta_\eta^{N_1}\widehat{\varphi}_R(\eta)}{|x-y|^{2N_1}} a_2(y,\xi) e^{i(x-y)\cdot(\eta-\xi)} dy d\eta,$$

which is a convergent integral, as for $|x - x_0| \leqslant 1/2$ and $|y - x_0| \geqslant 1$, we have $|x-y| \geqslant 1/2$. Integrating by parts with respect to the y variable,

$$c_2(x,\xi) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\Delta_\eta^{N_1}\widehat{\varphi}_R(\eta)}{(1+|\eta-\xi|^2)^{N_2}} (I-\Delta_y)^{N_2}\Big(\frac{a_2(y,\xi)}{|x-y|^{2N_1}}\Big) e^{i(x-y)\cdot(\eta-\xi)} dy d\eta.$$

In view of the differential inequalities satisfied by $\widehat{\varphi}_R$ and a_2,

$$|c_2(x,\xi)| \lesssim R^{-\epsilon} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{(1+|\eta|)^{-2N_1+\epsilon}(1+|\xi|)^{m+2N_2\delta}}{(1+|\eta-\xi|)^{2N_2}(1+|x-y|)^{2N_1}} dy d\eta.$$

The integration in y is finite if we choose $N_1 > d/2$. The triangle inequality trivially reveals

$$\frac{1}{(1+|\eta-\xi|)^{2N_2}} \leqslant \frac{(1+|\eta|)^{2N_2}}{(1+|\xi|)^{2N_2}},$$

for any $N_2 \geqslant 0$, so

$$|c_2(x,\xi)| \lesssim R^{-\epsilon}(1+|\xi|)^{m-2N_2(1-\delta)} \int_{\mathbb{R}^d} (1+|\eta|)^{-2N_1+2N_2+\epsilon} d\eta \lesssim R^{-\epsilon}(1+|\xi|)^{m-\bar{N}},$$

provided we take $2N_2(1-\delta) = \bar{N}$ and N_1 satisfying $2N_1 - 2N_2 - \epsilon > d$, that is

$$N_1 > \frac{d+\epsilon+\bar{N}/(1-\delta)}{2}.$$

Observe that any such N_1 also satisfies the required condition $N_1 > d/2$, as $\bar{N} \geqslant 0$.

In view of the definition of c_2, the use of the Leibniz formula and of similar arguments to the ones exposed above leads one to deduce that

$$|\partial_x^\nu \partial_\xi^\sigma c_2(x,\xi)| \lesssim R^{-\epsilon}(1+|\xi|)^{m-\rho|\sigma|+\delta|\nu|-\bar{N}}$$

for any $\bar{N} \geqslant 0$ and all multi-indices $\nu, \sigma \in \mathbb{N}^d$, so we may conclude that e^N satisfies the required differential inequalities (B.1).

$\square$

Appendix C

Coifman–Rochberg

In this appendix we provide a proof of Proposition 1.2.3. Recall the statement.

Proposition C.1. *Let A be a Young function. If $0 < \delta < 1$, then $(M_A w)^\delta \in A_1$ with A_1 constant independent of w. In particular,*

$$M\left((M_A w)^\delta\right)(x) \leqslant C_d \frac{1}{1-\delta} (M_A w)^\delta (x)$$

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

Our proof is an alternative to the one given in Proposition 5.32 in [38]. We follow the same method that Coifman and Rochberg [31] used to prove the classical result $(Mw)^\delta \in A_1$ for any $0 \leqslant \delta < 1$. In contrast to the Hardy-Littlewood maximal operator M, the maximal operator M_A is not in general of weak-type $(1,1)$. However, it will be enough to use the following weaker estimate.

Proposition C.2 ([38])**.** *Let A be a Young function. For all function f satisfying $\|f\|_{A,Q} \to 0$ as $|Q| \to \infty$, and all $t > 0$,*

$$|\{x \in \mathbb{R}^d : M_A f(x) > t\}| \leqslant 3^d \int_{\{x\in\mathbb{R}^d : |f(x)|>t/2\}} A\left(\frac{2\cdot 4^d |f(x)|}{t}\right) dx.$$

Proof of Proposition 1.2.3. Following the ideas in [31], we need to show that

$$\frac{1}{|Q|}\int_Q (M_A w)^\delta \leqslant C(M_A w)^\delta(x)$$

holds for every $x \in Q$ and C independent of Q and w. Let $2Q$ denote the cube whose center is the same as Q and whose sidelength is twice that of Q. Write $w = w_1 + w_2$, where $w_1 = w\chi_{2Q}$. By definition of M_A, $M_A w(x) \leqslant M_A w_1(x) + M_A w_2(x)$, so for $0 \leqslant \delta < 1$,

$$(M_A w)^\delta(x) \leqslant (M_A w_1)^\delta(x) + (M_A w_2)^\delta(x).$$

Hence, it suffices to show

$$\frac{1}{|Q|}\int_Q (M_A w_i)^\delta \leqslant C(M_A w)^\delta(x) \qquad i = 1, 2, \tag{C.1}$$

for every $x \in Q$ and C independent of Q and w. As w_1 has compact support, we have that $\|w_1\|_{A,R} \to 0$ as $|R| \to \infty$. Using Proposition C.2,

$$\begin{aligned}\frac{1}{|Q|}\int_Q (M_A w_1)^\delta(y)dy &= \frac{\delta}{|Q|}\int_0^\infty t^{\delta-1}|\{x \in Q : M_A w_1(y) > t\}|dt \\ &\leqslant \frac{\delta}{|Q|}\int_0^\infty t^{\delta-1}\min\left(|Q|, 3^d\int_{\{y\in 2Q: w_1(y)>t/2\}} A\left(\frac{2\cdot 4^d w_1(y)}{t}\right)dy\right)dt.\end{aligned} \tag{C.2}$$

As A is a convex and increasing function, by definition of the Luxemburg norm, we have that for $t > 2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}$,

$$\begin{aligned}3^d\int_{\{y\in 2Q: w_1(y)>t/2\}} A\left(\frac{2\cdot 4^d w_1(y)}{t}\right)dy &\leqslant \frac{1}{2^d}\int_{\{y\in 2Q: w_1(y)>t/2\}} A\left(\frac{2\cdot 4^d\cdot 3^d\cdot 2^d w_1(y)}{t}\right)dy \\ &\leqslant \frac{1}{2^d}\int_{2Q} A\left(\frac{w(y)}{\|w\|_{A,2Q}}\right)dy \leqslant |Q|,\end{aligned}$$

so we can bound (C.2) by

$$\begin{aligned}
\text{(C.2)} \leqslant & \frac{\delta}{|Q|} \int_0^{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}} t^{\delta-1}|Q|dt \\
& + \frac{3^d\delta}{|Q|} \int_{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}}^{\infty} t^{\delta-1} \int_{\{y\in 2Q: w_1(y)>t/2\}} A\left(\frac{2\cdot 4^d w_1(y)}{t}\right) dydt.
\end{aligned}$$

The first term in the right hand side is equal to $(2\cdot 4^d\cdot 3^d\cdot 2^d)^\delta \|w\|_{A,2Q}^\delta$. For the second term, by convexity of A, we have

$$\begin{aligned}
&\frac{3^d\delta}{|Q|} \int_{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}}^{\infty} t^{\delta-1} \int_{\{y\in 2Q: w_1(y)>t/2\}} A\left(\frac{2\cdot 4^d w_1(y)}{t}\right) dydt \\
&\quad \leqslant \frac{\delta}{2^d|Q|} \int_{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}}^{\infty} t^{\delta-1} \int_{\{y\in 2Q: w_1(y)>t/2\}} A\left(\frac{2\cdot 4^d\cdot 3^d\cdot 2^d w_1(y)}{t}\right) dydt \\
&\quad \leqslant \frac{\delta}{2^d|Q|} \int_{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}}^{\infty} t^{\delta-1} \int_{\{y\in 2Q: w_1(y)>t/2\}} \frac{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}}{t} A\left(\frac{w_1(y)}{\|w\|_{A,2Q}}\right) dydt \\
&\quad = \frac{\delta}{2^d|Q|} \int_{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}}^{\infty} t^{\delta-2} 2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q} \int_{\{y\in 2Q: w_1(y)>t/2\}} A\left(\frac{w(y)}{\|w\|_{A,2Q}}\right) dydt \\
&\quad \leqslant \delta \int_{2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q}}^{\infty} t^{\delta-2} 2\cdot 4^d\cdot 3^d\cdot 2^d\|w\|_{A,2Q} dt \\
&\quad = \frac{\delta}{1-\delta}(2\cdot 4^d\cdot 3^d\cdot 2^d)^\delta \|w\|_{A,2Q}^\delta.
\end{aligned}$$

Then

$$\frac{1}{|Q|}\int_Q (M_A w_1)^\delta \leqslant (2\cdot 24^d)^\delta \frac{1}{1-\delta}\|w\|_{A,2Q}^\delta \leqslant (2\cdot 24^d)^\delta (M_A w)^\delta(x)$$

for all $x \in Q$, and (C.1) is proved for w_1.

To prove (C.1) for $i = 2$, we can assume $M_A w_2(x) > 0$. Let $y \in Q$ and R be another cube such that $y \in R$ and $\|w_2\|_{A,R} > 0$. Then $R \nsubseteq 2Q$ and $\ell(R) > \frac{1}{2}\ell(Q)$, where $\ell(\cdot)$ denotes the sidelenght of a cube. This ensures $Q \subset 3R$. We claim that

$$\|w_2\|_{A,R} \leqslant 3^d \|w\|_{A,3R} \leqslant 3^d M_A w(x) \tag{C.3}$$

for every $R \ni y$. Then for every $y \in Q$, we have

$$M_A w_2(y) \leqslant 3^d M_A w(x),$$

so $(M_A w_2)^\delta(y) \leqslant 3^{d\delta}(M_A w)^\delta(x)$ for every $y \in Q$. Thus

$$\frac{1}{|Q|}\int_Q (M_A w_2)^\delta(y)dy \leqslant 3^{d\delta}(M_A w)^\delta(x),$$

as required.

To conclude the proof we still need to show (C.3). But this follows by convexity and monotonicity of A, since

$$\frac{1}{|R|}\int_R A\left(\frac{w_2(z)}{3^d\|w\|_{A,3R}}\right)dz \leqslant \frac{1}{3^d}\frac{1}{|R|}\int_{3R} A\left(\frac{w(z)}{\|w\|_{A,3R}}\right)dz \leqslant 1,$$

and this implies, by definition of Luxemburg norm, $\|w_2\|_{A,R} \leqslant 3^d\|w\|_{A,3R}$. This completes the proof. $\square$

Appendix D

Uncertainty principle

In the proof of Theorem 5.2.1, we have ensured that the quantities $|\widehat{g_\alpha d\sigma}(\xi)|$ are essentially constant at scale K, where g_α denotes a cap in S of radius $1/K$. Technically speaking, this is incorrect; however, the quantities $|\widehat{g_\alpha d\sigma}(\xi)|$ are pointwise controlled by a quantity satisfying such a property, and it is to that other quantity to the one that we should apply the dichotomies in the Bourgain–Guth argument. In what follows we make more formal those uncertainty principle considerations.

Let $\eta \in \mathcal{S}(\mathbb{R}^d)$ such that $\widehat{\eta} = 1$ in $B(0,1)$ and $\widehat{\eta} = 0$ outside $B(0,2)$, and for any $K > 0$, define $\eta_K(\xi) := K^{-d}\eta(K^{-1}\xi)$. Now, fix $\xi \in B_R$ and write

$$\widehat{gd\sigma}(\xi) = \sum_\alpha \widehat{g_\alpha d\sigma}(\xi) = \sum_\alpha e^{-ix_\alpha \cdot \xi} \int_{S_\alpha} e^{-i(x-x_\alpha)\cdot\xi} g(x) d\sigma(x) =: \sum_\alpha e^{-ix_\alpha\cdot\xi} T_\alpha g(\xi),$$

where x_α denotes the center of the cap S_α. It is a straightforward computation to check that $\widehat{T_\alpha g}(y)$ is supported in $B(0, \frac{1}{K})$. Then

$$T_\alpha g(\xi) = T_\alpha g * \eta_K(\xi).$$

As $|\eta(\xi)| \lesssim (1+|\xi|)^{-N}$ for any $N>0$,

$$|T_\alpha g(\xi)| \leqslant \int_{\mathbb{R}^d} |T_\alpha g(\theta)||\eta_K(\xi-\theta)|d\theta \lesssim \int_{\mathbb{R}^d} |T_\alpha g(\theta)|K^{-d}\Big(1+\frac{|\xi-\theta|}{K}\Big)^{-N} d\theta := c_\alpha(\xi),$$

and from the inherent properties of the function $\Psi^{(N)}(\xi) := (1+|\xi|)^{-N}$ discussed in Appendix A, it is clear that $c_\alpha(\xi_1) \sim c_\alpha(\xi_2)$ if $|\xi_1-\xi_2| \lesssim K$.

One should then apply the dichotomies in the proof of Theorem 5.2.1 to the quantities $c_\alpha(\xi)$, which are pointwise majorants of $|\widehat{g_\alpha d\sigma}(\xi)|$. This still gives the result for the extension operator, as long as we choose $N>d$ to guarantee the integrability of $\Psi^{(N)}$, and as we make applications of Fubini's theorem and Hölder's inequality when necessary along the argument. We avoid such computations here.

Appendix E

Parabolic Rescaling

We provide the parabolic scaling for the restriction conjecture used in the proof of Theorem 5.2.1. Let $\mathcal{C}(R)$ denote the smallest constant in the estimate

$$\|\widehat{gd\sigma}\|_{L^q(B(0,R))} \leqslant \mathcal{C}(R)\|g\|_{L^p(d\sigma)}$$

over all $d\sigma$ associated to any quadratic surface. Obviously $\mathcal{C}(R) < \infty$ since we have localised the estimate into a ball of radius R.

Proposition E.1. *Let $g_\alpha = g\chi_{S_\alpha}$, where S_α is a cap in S of diameter $1/K$. Then*

$$\|\widehat{g_\alpha d\sigma}\|_{L^q(B(0,R))} \lesssim \mathcal{C}(R)K^{\frac{d+1}{q}-\frac{d-1}{p'}}\|g_\alpha\|_{L^p(d\sigma)}.$$

Proof. In this case it is more convenient to work with the measure $d\mu$ and the parametrisation $\Sigma : U_\alpha \to S_\alpha$, where U_α is an open set in $\mathbb{R}^{d-1}$, given by $\Sigma(x') = (x', \psi(x'))$; here $\psi : \mathbb{R}^{d-1} \to \mathbb{R}$ is a quadratic function.

Observe that one may write

$$\begin{aligned}\widehat{g_\alpha d\mu}(\xi) &= \int_{\mathbb{R}^d} e^{i\xi\cdot x} g_\alpha(x)d\mu(x)\\ &= \int_{U_\alpha} e^{i\xi\cdot\Sigma(x')} g_\alpha(\Sigma(x'))dx'\end{aligned}$$

$$
\begin{aligned}
&= \int_{U_\alpha} e^{i(\xi'\cdot x' + \xi_d \psi(x'))} g(\Sigma(x'))dx' \\
&= K^{-(d-1)} \int_{\widetilde{U_\alpha}} e^{i(\xi'\cdot y/K + \xi_d \psi(y/K))} \tilde{g}(y)dy,
\end{aligned}
$$

where $\widetilde{U_\alpha} = KU_\alpha$ and $\tilde{g}(y) = g(\Sigma(y/K))$. Then, integrating over $\xi \in B(0,R)$ and doing a parabolic rescaling

$$
\begin{aligned}
\|\widehat{g_\alpha d\mu}\|^q_{L^q(B(0,R))} &= K^{-(d-1)q} \int_{B(0,R)} \left| \int_{\widetilde{U_\alpha}} e^{i(\xi'\cdot y/K + \xi_d \psi(y/K))} \tilde{g}(y)dy \right|^q d\xi \\
&= K^{-(d-1)q} K^{d+1} \int_{\substack{|\tilde{\xi}'| < R/K \\ |\tilde{\xi}_d| < R/K^2}} \left| \int_{\widetilde{U_\alpha}} e^{i(\tilde{\xi}'\cdot y + \tilde{\xi}_d \tilde{\psi}(y))} \tilde{g}(y)dy \right|^q d\tilde{\xi} \\
&\leqslant K^{-(d-1)q + (d+1)} \mathcal{C}(R/K, R/K^2)^q \|\tilde{g}\chi_{\widetilde{U_\alpha}}\|^q_{L^p(d\mu)} \\
&= K^{-(d-1)q + (d+1)} K^{(d-1)\frac{q}{p}} \mathcal{C}(R/K, R/K^2)^q \|g_\alpha\|^q_{L^p(d\mu)} \\
&= K^{-\frac{(d-1)q}{p'} + (d+1)} \mathcal{C}(R/K, R/K^2)^q \|g_\alpha\|^q_{L^p(d\mu)} \\
&\leqslant K^{-\frac{(d-1)q}{p'} + (d+1)} \mathcal{C}(R)^q \|g_\alpha\|^q_{L^p(d\mu)},
\end{aligned}
$$

where $\tilde{\psi}$ denotes another quadratic function. $\square$

Appendix F

Multilinear square function estimate

The multilinear theory from [9] yields certain multilinear square function estimates. This was first observed in [17], to which we refer for proofs.

Lemma F.1 ([17]). *Let $2 \leqslant m \leqslant d$ and V be a subspace of $\mathbb{R}^d$ of dimension m. Let $P_1, \ldots, P_m \in S$ be points that satisfy $n(P_i) \in V$ for all $1 \leqslant i \leqslant m$ and $|n(P_1) \wedge \cdots \wedge n(P_m)| > c$, where $n(P)$ denotes the unit normal to S at the point P. Let $U_1, \ldots, U_m \subset S$ be small neighbourhoods of $P_1, \ldots, P_m$. Let M be large and $D_i \subset U_i$ be subsets of $1/M$ separated points ξ that obey the condition $\operatorname{dist}(n(\xi), V) < c/M$. Then for $f_i \in L^\infty(U_i)$, we have*

$$\frac{1}{|B_M|}\int_{B_M}\prod_{i=1}^{m}\Big|\sum_{\xi\in D_i}\int_{|\eta-\xi|<\frac{c}{M}} f_i(\eta)e^{-ix\cdot\eta}d\sigma(\eta)\Big|^{\frac{2}{m-1}}dx$$
$$\lesssim M^{\varepsilon}\Big(\frac{1}{|B_M|}\int_{B_M}\prod_{i=1}^{m}\Big(\sum_{\xi\in D_i}\Big|\int_{|\eta-\xi|<\frac{c}{M}} f_i(\eta)e^{-ix\cdot\eta}d\sigma(\eta)\Big|^2\Big)^{\frac{1}{2m}}dx\Big)^{\frac{2m}{m-1}}.$$

Remark F.2. Two application of Hölder's inequality, together with Lemma F.1, lead to a multilinear version of the Córdoba square function estimate [35] for $q \leqslant \frac{2m}{m-1}$

$$\begin{aligned}
&\frac{1}{|B_M|}\int_{B_M}\prod_{i=1}^{m}\Big|\sum_{\xi\in D_i}\int_{|\eta-\xi|<\frac{c}{M}} f_i(\eta)e^{-ix\cdot\eta}d\sigma(\eta)\Big|^{\frac{q}{m}}dx\\
&\qquad\leqslant\Big(\frac{1}{|B_M|}\int_{B_M}\prod_{i=1}^{m}\Big|\sum_{\xi\in D_i}\int_{|\eta-\xi|<\frac{c}{M}} f_i(\eta)e^{-ix\cdot\eta}d\sigma(\eta)\Big|^{\frac{2}{m-1}}dx\Big)^{\frac{q(m-1)}{2m}}\\
&\qquad\lesssim M^{\varepsilon}\Big(\frac{1}{|B_M|}\int_{B_M}\prod_{i=1}^{m}\Big(\sum_{\xi\in D_i}\big|\int_{|\eta-\xi|<\frac{c}{M}} f_i(\eta)e^{-ix\cdot\eta}d\sigma(\eta)\big|^{2}\Big)^{\frac{1}{2m}}dx\Big)^{q}\\
&\qquad\leqslant M^{\varepsilon}\frac{1}{|B_M|}\int_{B_M}\prod_{i=1}^{m}\Big(\sum_{\xi\in D_i}\big|\int_{|\eta-\xi|<\frac{c}{M}} f_i(\eta)e^{-ix\cdot\eta}d\sigma(\eta)\big|^{2}\Big)^{\frac{1}{2}\frac{q}{m}}dx,
\end{aligned}$$

so

$$\Big\|\prod_{j=1}^{m}\Big(\sum_{\xi\in D_i}\widehat{f_i^{\xi}d\sigma}\Big)\Big\|_{L^{q/m}(B_M)}\lesssim M^{\varepsilon}\Big\|\prod_{j=1}^{m}\Big(\sum_{\xi\in D_i}|\widehat{f_i^{\xi}d\sigma}|^2\Big)^{1/2}\Big\|_{L^{q/m}(B_M)},$$

where $f_i^{\xi}=f\chi_{S_{\xi}^{M}}$ and S_{ξ}^{M} denotes a cap in S of radius $1/M$ centered at ξ.

List of References

[1] Juan Antonio Barceló, Jonathan Bennett, Anthony Carbery, and Keith M. Rogers. On the dimension of divergence sets of dispersive equations. *Math. Ann.*, 349(3):599–622, 2011.

[2] Ioan Bejenaru. The multilinear restriction estimate: a short proof and a refinement. To appear in *Math. Res. Letters*, Preprint available at arXiv:1601.03336.

[3] David Beltran. Control of pseudo-differential operators by maximal functions via weighted inequalities. Preprint available at arXiv:1608.06571.

[4] David Beltran. A Fefferman-Stein inequality for the Carleson operator. To appear in *Rev. Mat. Iber.* Preprint available at arXiv:1410.6085.

[5] David Beltran and Jonathan Bennett. Subdyadic square functions and applications to weighted harmonic analysis. *Adv. Math.*, 307:72–99, 2017.

[6] Jonathan Bennett. Aspects of multilinear harmonic analysis related to transversality. In *Harmonic analysis and partial differential equations*, volume 612 of *Contemp. Math.*, pages 1–28. Amer. Math. Soc., Providence, RI, 2014.

[7] Jonathan Bennett. Optimal control of singular Fourier multipliers by maximal operators. *Anal. PDE*, 7(6):1317–1338, 2014.

[8] Jonathan Bennett, Anthony Carbery, Fernando Soria, and Ana Vargas. A Stein conjecture for the circle. *Math. Ann.*, 336(3):671–695, 2006.

[9] Jonathan Bennett, Anthony Carbery, and Terence Tao. On the multilinear restriction and Kakeya conjectures. *Acta Math.*, 196(2):261–302, 2006.

[10] Jonathan Bennett and Samuel Harrison. Weighted norm inequalities for oscillatory integrals with finite type phases on the line. *Adv. Math.*, 229(4):2159–2183, 2012.

[11] J. Bourgain. Besicovitch type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.*, 1(2):147–187, 1991.

[12] J. Bourgain. A remark on Schrödinger operators. *Israel J. Math.*, 77(1-2):1–16, 1992.

[13] J. Bourgain. On the Schrödinger maximal function in higher dimension. *Tr. Mat. Inst. Steklova*, 280(Ortogonalnye Ryady, Teoriya Priblizhenii i Smezhnye Voprosy):53–66, 2013.

[14] J. Bourgain. A note on the Schrödinger maximal function. *J. Anal. Math.*, 130:393–396, 2016.

[15] Jean Bourgain and Ciprian Demeter. The proof of the l^2 decoupling conjecture. *Ann. of Math. (2)*, 182(1):351–389, 2015.

[16] Jean Bourgain, Ciprian Demeter, and Larry Guth. Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three. *Ann. of Math. (2)*, 184(2):633–682, 2016.

[17] Jean Bourgain and Larry Guth. Bounds on oscillatory integral operators based on multilinear estimates. *Geom. Funct. Anal.*, 21(6):1239–1295, 2011.

[18] Alberto-P. Calderón and Rémi Vaillancourt. A class of bounded pseudo-differential operators. *Proc. Nat. Acad. Sci. U.S.A.*, 69:1185–1187, 1972.

[19] Anthony Carbery. A weighted inequality for the maximal Bochner-Riesz operator on $\mathbb{R}^2$. *Trans. Amer. Math. Soc.*, 287(2):673–680, 1985.

[20] Anthony Carbery. Restriction implies Bochner-Riesz for paraboloids. *Math. Proc. Cambridge Philos. Soc.*, 111(3):525–529, 1992.

[21] Anthony Carbery, Elena Romera, and Fernando Soria. Radial weights and mixed norm inequalities for the disc multiplier. *J. Funct. Anal.*, 109(1):52–75, 1992.

[22] Anthony Carbery, José L. Rubio de Francia, and Luis Vega. Almost everywhere summability of Fourier integrals. *J. London Math. Soc. (2)*, 38(3):513–524, 1988.

[23] Anthony Carbery and Andreas Seeger. Weighted inequalities for Bochner-Riesz means in the plane. *Q. J. Math.*, 51(2):155–167, 2000.

[24] Anthony Carbery and Stefán Ingi Valdimarsson. The endpoint multilinear Kakeya theorem via the Borsuk-Ulam theorem. *J. Funct. Anal.*, 264(7):1643–1663, 2013.

[25] Lennart Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966.

[26] Lennart Carleson. Two remarks on H^1 and BMO. *Advances in Math.*, 22(3):269–277, 1976.

[27] Lennart Carleson. Some analytic problems related to statistical mechanics. In *Euclidean harmonic analysis (Proc. Sem., Univ. Maryland, College Park, Md., 1979)*, volume 779 of *Lecture Notes in Math.*, pages 5–45. Springer, Berlin, 1980.

[28] Sagun Chanillo and Alberto Torchinsky. Sharp function and weighted L^p estimates for a class of pseudodifferential operators. *Ark. Mat.*, 24(1):1–25, 1986.

[29] Michael Christ. On almost everywhere convergence of Bochner-Riesz means in higher dimensions. *Proc. Amer. Math. Soc.*, 95(1):16–20, 1985.

[30] R. R. Coifman and C. Fefferman. Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.*, 51:241–250, 1974.

[31] R. R. Coifman and R. Rochberg. Another characterization of BMO. *Proc. Amer. Math. Soc.*, 79(2):249–254, 1980.

[32] José M. Conde-Alonso and Guillermo Rey. A pointwise estimate for positive dyadic shifts and some applications. *Math. Ann.*, 365(3-4):1111–1135, 2016.

[33] A. Córdoba and C. Fefferman. A weighted norm inequality for singular integrals. *Studia Math.*, 57(1):97–101, 1976.

[34] Antonio Córdoba. The Kakeya maximal function and the spherical summation multipliers. *Amer. J. Math.*, 99(1):1–22, 1977.

[35] Antonio Córdoba. Geometric Fourier analysis. *Ann. Inst. Fourier (Grenoble)*, 32(3):vii, 215–226, 1982.

[36] D. Cruz-Uribe, J. M. Martell, and C. Pérez. Sharp two-weight inequalities for singular integrals, with applications to the Hilbert transform and the Sarason conjecture. *Adv. Math.*, 216(2):647–676, 2007.

[37] David Cruz-Uribe and Carlos Pérez. On the two-weight problem for singular integral operators. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 1(4):821–849, 2002.

[38] David V. Cruz-Uribe, José Maria Martell, and Carlos Pérez. *Weights, extrapolation and the theory of Rubio de Francia*, volume 215 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer Basel AG, Basel, 2011.

[39] Björn E. J. Dahlberg and Carlos E. Kenig. A note on the almost everywhere behavior of solutions to the Schrödinger equation. In *Harmonic analysis (Minneapolis, Minn., 1981)*, volume 908 of *Lecture Notes in Math.*, pages 205–209. Springer, Berlin-New York, 1982.

[40] Wendolín Damián, Andrei K. Lerner, and Carlos Pérez. Sharp Weighted Bounds for Multilinear Maximal Functions and Calderón–Zygmund Operators. *J. Fourier Anal. Appl.*, 21(1):161–181, 2015.

[41] F. Di Plinio, Y. Do, and G. Uraltsev. Positive sparse domination of variational Carleson operators. To appear in *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, Preprint available at arXiv:1612.03028.

[42] Francesco Di Plinio and Andrei K. Lerner. On weighted norm inequalities for the Carleson and Walsh-Carleson operator. *J. Lond. Math. Soc. (2)*, 90(3):654–674, 2014.

[43] Yen Do and Michael Lacey. Weighted bounds for variational Fourier series. *Studia Math.*, 211(2):153–190, 2012.

[44] Xiumin Du, Larry Guth, and Xiaochun Li. A sharp Schrödinger maximal estimate in $\mathbb{R}^2$. Preprint available at arXiv:1612.08946.

[45] 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.

[46] M. Burak Erdoğan. A bilinear Fourier extension theorem and applications to the distance set problem. *Int. Math. Res. Not.*, (23):1411–1425, 2005.

[47] C. Fefferman and E. M. Stein. Some maximal inequalities. *Amer. J. Math.*, 93:107–115, 1971.

[48] C. Fefferman and E. M. Stein. H^p spaces of several variables. *Acta Math.*, 129(3-4):137–193, 1972.

[49] Charles Fefferman. Inequalities for strongly singular convolution operators. *Acta Math.*, 124:9–36, 1970.

[50] Charles Fefferman. The multiplier problem for the ball. *Ann. of Math. (2)*, 94:330–336, 1971.

[51] Charles Fefferman. On the convergence of multiple Fourier series. *Bull. Amer. Math. Soc.*, 77:744–745, 1971.

[52] Charles Fefferman. L^p bounds for pseudo-differential operators. *Israel J. Math.*, 14:413–417, 1973.

[53] Charles Fefferman. Pointwise convergence of Fourier series. *Ann. of Math. (2)*, 98:551–571, 1973.

[54] Nobuhiko Fujii. A proof of the Fefferman-Stein-Strömberg inequality for the sharp maximal functions. *Proc. Amer. Math. Soc.*, 106(2):371–377, 1989.

[55] José García-Cuerva and José L. Rubio de Francia. *Weighted norm inequalities and related topics*, volume 116 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1985. Notas de Matemática [Mathematical Notes], 104.

[56] John B. Garnett and Peter W. Jones. BMO from dyadic BMO. *Pacific J. Math.*, 99(2):351–371, 1982.

[57] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2009.

[58] Loukas Grafakos, José María Martell, and Fernando Soria. Weighted norm inequalities for maximally modulated singular integral operators. *Math. Ann.*, 331(2):359–394, 2005.

[59] Loukas Grafakos and Rodolfo H. Torres. Multilinear Calderón-Zygmund theory. *Adv. Math.*, 165(1):124–164, 2002.

[60] Ben Green. Roth's theorem in the primes. *Ann. of Math. (2)*, 161(3):1609–1636, 2005.

[61] Larry Guth. Restriction estimates using polynomial partitioning II. Preprint available at arXiv:1603.04250.

[62] Larry Guth. The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture. *Acta Math.*, 205(2):263–286, 2010.

[63] Larry Guth. A short proof of the multilinear Kakeya inequality. *Math. Proc. Cambridge Philos. Soc.*, 158(1):147–153, 2015.

[64] Larry Guth. A restriction estimate using polynomial partitioning. *J. Amer. Math. Soc.*, 29(2):371–413, 2016.

[65] Paul Richard Halmos and Viakalathur Shankar Sunder. *Bounded integral operators on L^2 spaces*, volume 96 of *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]*. Springer-Verlag, Berlin-New York, 1978.

[66] Samuel Harrison. *Control of oscillatory convolution operators via maximal functions in weighted L^2 Inequalities.* PhD thesis, University of Birmingham, 2010.

[67] Henry Helson and Gabor Szegö. A problem in prediction theory. *Ann. Mat. Pura Appl. (4)*, 51:107–138, 1960.

[68] I. I. Hirschman, Jr. On multiplier transformations. *Duke Math. J*, 26:221–242, 1959.

[69] Lars Hörmander. Pseudo-differential operators and hypoelliptic equations. In *Singular integrals (Proc. Sympos. Pure Math., Vol. X, Chicago, Ill., 1966)*, pages 138–183. Amer. Math. Soc., Providence, R.I., 1967.

[70] GuoEn Hu. Weighted norm inequalities for the multilinear Calderón-Zygmund operators. *Sci. China Math.*, 53(7):1863–1876, 2010.

[71] Richard Hunt, Benjamin Muckenhoupt, and Richard Wheeden. Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.*, 176:227–251, 1973.

[72] Richard A. Hunt. On the convergence of Fourier series. In *Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967)*, pages 235–255. Southern Illinois Univ. Press, Carbondale, Ill., 1968.

[73] Richard A. Hunt and Wo Sang Young. A weighted norm inequality for Fourier series. *Bull. Amer. Math. Soc.*, 80:274–277, 1974.

[74] Tuomas P. Hytönen. The two-weight inequality for the Hilbert transform with general measures. Preprint available at arXiv:1312.0833.

[75] Tuomas P. Hytönen. The sharp weighted bound for general Calderón-Zygmund operators. *Ann. of Math. (2)*, 175(3):1473–1506, 2012.

[76] Tuomas P. Hytönen. The A_2 theorem: remarks and complements. In *Harmonic analysis and partial differential equations*, volume 612 of *Contemp. Math.*, pages 91–106. Amer. Math. Soc., Providence, RI, 2014.

[77] J. J. Kohn and L. Nirenberg. An algebra of pseudo-differential operators. *Comm. Pure Appl. Math.*, 18:269–305, 1965.

[78] Michael Lacey and Christoph Thiele. A proof of boundedness of the Carleson operator. *Math. Res. Lett.*, 7(4):361–370, 2000.

[79] Michael T. Lacey. An elementary proof of the A_2 bound. *Israel Journal of Mathematics*, 217(1):181–195, 2017.

[80] Michael T. Lacey, Eric T. Sawyer, Chun-Yen Shen, and Ignacio Uriarte-Tuero. Two-weight inequality for the Hilbert transform: A real variable characterization, I. *Duke Math. J.*, 163(15):2795–2820, 2014.

[81] Sanghyuk Lee. On pointwise convergence of the solutions to Schrödinger equations in $\mathbb{R}^2$. *Int. Math. Res. Not.*, pages Art. ID 32597, 21, 2006.

[82] Sanghyuk Lee, Keith M. Rogers, and Andreas Seeger. Improved bounds for Stein's square functions. *Proc. Lond. Math. Soc. (3)*, 104(6):1198–1234, 2012.

[83] Andrei K. Lerner. A pointwise estimate for the local sharp maximal function with applications to singular integrals. *Bull. Lond. Math. Soc.*, 42(5):843–856, 2010.

[84] Andrei K. Lerner. On an estimate of Calderón-Zygmund operators by dyadic positive operators. *J. Anal. Math.*, 121:141–161, 2013.

[85] Andrei K. Lerner. A simple proof of the A_2 conjecture. *Int. Math. Res. Not. IMRN*, (14):3159–3170, 2013.

[86] Andrei K. Lerner. On pointwise estimates involving sparse operators. *New York J. Math.*, 22:341–349, 2016.

[87] Andrei K. Lerner and Fedor Nazarov. Intuitive dyadic calculus: the basics. Preprint available at arXiv:1508.05639.

[88] Andrei K. Lerner, Sheldy Ombrosi, Carlos Pérez, Rodolfo H. Torres, and Rodrigo Trujillo-González. New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory. *Adv. Math.*, 220(4):1222–1264, 2009.

[89] Kangwei Li, Kabe Moen, and Wenchang Sun. The sharp weighted bound for multilinear maximal functions and Calderón-Zygmund operators. *J. Fourier Anal. Appl.*, 20(4):751–765, 2014.

[90] Victor Lie. The polynomial Carleson operator. Preprint available at arXiv:1105.4504.

[91] Victor Lie. The (weak-L^2) boundedness of the quadratic Carleson operator. *Geom. Funct. Anal.*, 19(2):457–497, 2009.

[92] J. E. Littlewood and R. E. A. C. Paley. Theorems on Fourier Series and Power Series. *J. London Math. Soc.*, S1-6(3):230.

[93] Pertti Mattila. Spherical averages of Fourier transforms of measures with finite energy; dimension of intersections and distance sets. *Mathematika*, 34(2):207–228, 1987.

[94] Pertti Mattila. *Fourier Analysis and Hausdorff Dimension*, volume 150 of *Cambridge studies in advanced mathematics.* Cambridge University Press, 2015.

[95] Nicholas Michalowski, David J. Rule, and Wolfgang Staubach. Weighted norm inequalities for pseudo-pseudodifferential operators defined by amplitudes. *J. Funct. Anal.*, 258(12):4183–4209, 2010.

[96] Nicholas Michalowski, David J. Rule, and Wolfgang Staubach. Weighted L^p boundedness of pseudodifferential operators and applications. *Canad. Math. Bull.*, 55(3):555–570, 2012.

[97] Nicholas Miller. Weighted Sobolev spaces and pseudodifferential operators with smooth symbols. *Trans. Amer. Math. Soc.*, 269(1):91–109, 1982.

[98] Akihiko Miyachi. On some Fourier multipliers for $H^p(\mathbb{R}^n)$. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 27(1):157–179, 1980.

[99] Akihiko Miyachi. On some singular Fourier multipliers. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 28(2):267–315, 1981.

[100] A. Moyua, A. Vargas, and L. Vega. Schrödinger maximal function and restriction properties of the Fourier transform. *Internat. Math. Res. Notices*, (16):793–815, 1996.

[101] A. Moyua, A. Vargas, and L. Vega. Restriction theorems and maximal operators related to oscillatory integrals in $\mathbf{R}^3$. *Duke Math. J.*, 96(3):547–574, 1999.

[102] Benjamin Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.

[103] Alexander Nagel and Elias M. Stein. On certain maximal functions and approach regions. *Adv. in Math.*, 54(1):83–106, 1984.

[104] F. Nazarov, S. Treil, and A. Volberg. The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.*, 12(4):909–928, 1999.

[105] F. Nazarov, S. Treil, and A. Volberg. Two weight inequalities for individual Haar multipliers and other well localized operators. *Math. Res. Lett.*, 15(3):583–597, 2008.

[106] C. J. Neugebauer. Inserting A_p-weights. *Proc. Amer. Math. Soc.*, 87(4):644–648, 1983.

[107] C. Pérez. Weighted norm inequalities for singular integral operators. *J. London Math. Soc. (2)*, 49(2):296–308, 1994.

[108] C. Pérez. On sufficient conditions for the boundedness of the Hardy-Littlewood maximal operator between weighted L^p-spaces with different weights. *Proc. London Math. Soc. (3)*, 71(1):135–157, 1995.

[109] Carlos Pérez. Sharp L^p-weighted Sobolev inequalities. *Ann. Inst. Fourier (Grenoble)*, 45(3):809–824, 1995.

[110] Carlos Pérez. Sharp estimates for commutators of singular integrals via iterations of the Hardy-Littlewood maximal function. *J. Fourier Anal. Appl.*, 3(6):743–756, 1997.

[111] Carlos Pérez and Rodrigo Trujillo-González. Sharp weighted estimates for vector-valued singular integral operators and commutators. *Tohoku Math. J. (2)*, 55(1):109–129, 2003.

[112] S. Petermichl. The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. *Amer. J. Math.*, 129(5):1355–1375, 2007.

[113] Stefanie Petermichl. The sharp weighted bound for the Riesz transforms. *Proc. Amer. Math. Soc.*, 136(4):1237–1249, 2008.

[114] Maria Carmen Reguera. On Muckenhoupt-Wheeden conjecture. *Adv. Math.*, 227(4):1436–1450, 2011.

[115] Maria Carmen Reguera and Christoph Thiele. The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$. *Math. Res. Lett.*, 19(1):1–7, 2012.

[116] Keith M. Rogers and Andreas Seeger. Endpoint maximal and smoothing estimates for Schrödinger equations. *J. Reine Angew. Math.*, 640:47–66, 2010.

[117] Marvin Rosenblum. Summability of Fourier series in $L^p(d\mu)$. *Trans. Amer. Math. Soc.*, 105:32–42, 1962.

[118] José L. Rubio de Francia, Francisco J. Ruiz, and José L. Torrea. Calderón-Zygmund theory for operator-valued kernels. *Adv. in Math.*, 62(1):7–48, 1986.

[119] Eric T. Sawyer. A characterization of a two-weight norm inequality for maximal operators. *Studia Math.*, 75(1):1–11, 1982.

[120] Eric T. Sawyer. A characterization of two weight norm inequalities for fractional and Poisson integrals. *Trans. Amer. Math. Soc.*, 308(2):533–545, 1988.

[121] Per Sjölin. Convolution with oscillating kernels. *Indiana Univ. Math. J.*, 30(1):47–55, 1981.

[122] Per Sjölin. Regularity of solutions to the Schrödinger equation. *Duke Math. J.*, 55(3):699–715, 1987.

[123] Christopher D. Sogge. Propagation of singularities and maximal functions in the plane. *Invent. Math.*, 104(2):349–376, 1991.

[124] E. M. Stein. Note on the class L log L. *Studia Math.*, 32:305–310, 1969.

[125] E. M. Stein. Some problems in harmonic analysis. In *Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1*, Proc. Sympos. Pure Math., XXXV, Part, pages 3–20. Amer. Math. Soc., Providence, R.I., 1979.

[126] E. M. Stein. The development of square functions in the work of A. Zygmund. *Bull. Amer. Math. Soc. (N.S.)*, 7(2):359–376, 1982.

[127] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[128] Elias M. Stein. *Topics in harmonic analysis related to the Littlewood-Paley theory.* Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1970.

[129] 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.

[130] Robert S. Strichartz. Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Math. J.*, 44(3):705–714, 1977.

[131] Jan-Olov Strömberg. Bounded mean oscillation with Orlicz norms and duality of Hardy spaces. *Indiana Univ. Math. J.*, 28(3):511–544, 1979.

[132] T. Tao. A sharp bilinear restrictions estimate for paraboloids. *Geom. Funct. Anal.*, 13(6):1359–1384, 2003.

[133] T. Tao and A. Vargas. A bilinear approach to cone multipliers. I. Restriction estimates. *Geom. Funct. Anal.*, 10(1):185–215, 2000.

[134] T. Tao and A. Vargas. A bilinear approach to cone multipliers. II. Applications. *Geom. Funct. Anal.*, 10(1):216–258, 2000.

[135] Terence Tao. The Bochner-Riesz conjecture implies the restriction conjecture. *Duke Math. J.*, 96(2):363–375, 1999.

[136] Terence Tao. Some recent progress on the restriction conjecture. In *Fourier analysis and convexity*, Appl. Numer. Harmon. Anal., pages 217–243. Birkhäuser Boston, Boston, MA, 2004.

[137] Terence Tao, Ana Vargas, and Luis Vega. A bilinear approach to the restriction and Kakeya conjectures. *J. Amer. Math. Soc.*, 11(4):967–1000, 1998.

[138] Faruk Temur. A Fourier restriction estimate for surfaces of positive curvature in $\mathbb{R}^6$. *Rev. Mat. Iberoam.*, 30(3):1015–1036, 2014.

[139] Peter A. Tomas. A restriction theorem for the Fourier transform. *Bull. Amer. Math. Soc.*, 81:477–478, 1975.

[140] Luis Vega. Schrödinger equations: pointwise convergence to the initial data. *Proc. Amer. Math. Soc.*, 102(4):874–878, 1988.

[141] Stephen Wainger. Special trigonometric series in k-dimensions. *Mem. Amer. Math. Soc. No.*, 59:102, 1965.

[142] J. Michael Wilson. Weighted norm inequalities for the continuous square function. *Trans. Amer. Math. Soc.*, 314(2):661–692, 1989.

[143] Michael Wilson. The intrinsic square function. *Rev. Mat. Iberoam.*, 23(3):771–791, 2007.

[144] Thomas Wolff. An improved bound for Kakeya type maximal functions. *Rev. Mat. Iberoamericana*, 11(3):651–674, 1995.

[145] Thomas Wolff. Recent work connected with the Kakeya problem. In *Prospects in mathematics (Princeton, NJ, 1996)*, pages 129–162. Amer. Math. Soc., Providence, RI, 1999.