On certain geometric maximal functions in Harmonic analysis

Aswin Govindan Sheri

Doctor of Philosophy
University of Edinburgh
2023

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Aswin Govindan Sheri*)

Acknowledgements

I wish to express my sincere gratitude to Dr. Jonathan Hickman, my thesis advisor, for his unwavering support, guidance, and expertise throughout my PhD journey. His insights, energy, and encouragement have been invaluable in keeping me motivated and hopeful, especially during my formative years as an early graduate student, and have helped shape my research and other academic interests in my later years.

I am also deeply grateful to Prof. Jim Wright, my second supervisor, for teaching me a lot of beautiful mathematics in my first year as a MIGSAA student and suggesting some of the problems to work on.

During my time in Edinburgh, I have been fortunate to make many friends who have been a constant source of love, encouragement, and motivation. I must highlight the names of Aparna, Adarsh, Swetha, Sooraj, and Neelakshi, my early co-travellers in this journey, without whom my life in Edinburgh would have been dull and meaningless. Aparna and Vidhya have been crucial in my life by constantly being with me during my periods of struggle. I am also thankful for the love and affection shown by Awanish, Anisha, Dayal, Nikhil, Junaina, and Gayathri, whose presence made my last year in Edinburgh fantastic.

My fellow graduate students, including Reuben, Hrit, Maria, Alina, Guopeng, and other members of our MIGSAA family, have been my friends, colleagues, and a source of inspiration. Their presence certainly has made my PhD journey enjoyable and enriching.

Lastly, I want to acknowledge the support of my family, Ida, and friends in India, who have been a constant source of love and kindness throughout these years.

Lay summary

This thesis investigates specific questions around objects known as maximal functions, essential tools in Harmonic analysis used to study a function's local and global behaviour. To understand what maximal functions do, we first consider an averaging (or smoothing) process, which takes a function f and finds its average over a region in space. This region can be a ball, a sphere, or a line. A maximal function $\mathcal{M}f$ then looks for the worst values such averages can attain by varying the regions where the functions are averaged. The fundamental question we ask is the following.

If we start with a function f that has a tall and thin distribution, does the output function of this process, $\mathcal{M}f$, also carry a similar distribution?

To answer this question meaningfully, we need a tool to measure the distribution of functions, which is done using mathematical objects known as L^p-norms. Here, we consider two types of maximal functions.

In the first type, the averages are taken on singular surfaces (such as curves or two-dimensional surfaces in three-dimensional space). In this case, it is shown that the answer to our question depends on the geometry of the surfaces. In particular, the maximal function preserves distributions (in the sense of L^p norm) if the surface is not entirely flat in the space.

In the second type, averages are taken on long thin tubes (or boxes) in the space, and the maximal function looks at the worst averaged value after varying the directions of the tubes along a curve in a lower-dimensional space. Here, it is shown that the maximal function preserves distributions (in the sense of L^p norm) if the curve bends and twists in the space.

Abstract

The broad theme of the thesis is of geometric maximal functions associated to curved surfaces. We produce novel results about two maximal functions of different types, presented in two parts of the thesis.

In the first part (Chapter 2), we study the $L^p \to L^p$ boundedness of a lacunary maximal function on a graded homogeneous group. The main theorem of this part generalises the existing maximal results in specific homogeneous groups, such as the Euclidean space and the Heisenberg group. Using an iteration scheme, we estimate the maximal function, assuming that the measure associated to the maximal function satisfies a curvature condition.

This second part of this thesis (Chapters 3 and 4) deals with the problem of $L^p \to L^p$ boundedness of a Nikodym maximal function in the Euclidean space. The maximal function is defined using a one-parameter family of tubes in $\mathbb{R}^{d+1}$, whose directions are determined by a non-degenerate curve in $\mathbb{R}^d$. These operators naturally arise in the analysis of maximal averages over space curves. The main theorem generalises the known results for $d = 2$ and $d = 3$ to general dimensions.

Notation

- $(G, \mathrm{d}x)$ always denotes a Lie group with the associated Haar measure.
- $C_c(G)$ denotes the space of continuous functions on G with compact support.
- $C_c^\infty(G)$ denotes the space of infinitely differentiable functions on G with compact supported.
- $L^p(G)$ denotes the space of measurable functions f such that
$$\int_G |f(x)|^p \mathrm{d}x < \infty.$$
- $\mathcal{S}(G)$ denotes the Schwartz space of smooth rapidly decreasing functions in G.
- $W^{k,p}(\mathbb{R}^d)$ denotes the Sobolev space given by
$$W^{k,p}(\mathbb{R}^d) := \{f \in L^p(\mathbb{R}^d) : D^\alpha f \in L^p(\mathbb{R}^d) \text{ for all } 0 \leq |\alpha| \leq k\}$$
where $D^\alpha f$ denotes the weak derivative of f with order α.
- $\mathbb{N}_0$ denotes the set $\mathbb{N} \cup \{0\}$.
- For a set $E \subseteq \mathbb{R}^n$, we denote its characteristic function by χ_E.
- Given $f \in L^1(\mathbb{R}^n)$ we let either $\hat{f}$ or $\mathcal{F}(f)$ denote its Fourier transform and $\check{f}$ or $\mathcal{F}^{-1}(f)$ denote its inverse Fourier transform, which are normalised as follows:
$$\hat{f}(\xi) := \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\,\mathrm{d}x, \qquad \check{f}(\xi) := \int_{\mathbb{R}^n} e^{ix\cdot\xi} f(x)\,\mathrm{d}x.$$
- For $m \in L^\infty(\mathbb{R}^n)$, we denote by $m(\frac{1}{i}\partial_x)$ the Fourier multiplier operator defined by its action on $g \in \mathcal{S}(\mathbb{R}^n)$ as
$$\mathcal{F}(m(\tfrac{1}{i}\partial_x)g)(\xi) := m(\xi)\mathcal{F}(g)(\xi) \qquad \text{for } \xi \in \mathbb{R}^n.$$
- Given two numbers $A, B \geq 0$ and a list of parameters $M_1, \ldots, M_n$, the notation $A \lesssim_{M_1,\ldots,M_n} B$ or $A = O_{M_1,\ldots,M_n}(B)$ signifies that $A \leq CB$ for some constant $C = C_{M_1,\ldots,M_n} > 0$ depending only on the parameters $M_1, \ldots, M_n$.

In addition, $A \sim_{M_1,\ldots,M_n} B$ is used to signify that both $A \lesssim_{M_1,\ldots,M_n} B$ and $B \lesssim_{M_1,\ldots,M_n} A$ hold.

Contents

Chapter 1

Introduction

In the extensive field of harmonic analysis, maximal functions can take various forms and appear in different contexts. In this thesis, we investigate maximal functions linked to 'curved' geometric objects in $\mathbb{R}^n$. In order to provide a unified framework for the questions we address, we present the following general formulation. The discussion here is given in a somewhat imprecise manner to simplify the statements.

Let $n, m \geq 1$ and $E \subset \mathbb{R}$. For each $t \in E$ (the set E will be referred to as the parameter set or the index set), let $\sigma_{(t)}$ be a finite Borel measure supported on a compact set $S_{(t)} \subseteq \mathbb{R}^n$. Consider a smooth function $\Phi : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ and consider an averaging operator

$$A_t f(x) := \int_{\mathbb{R}^n} |f \circ \Phi(x, y)| \, \mathrm{d}\sigma_{(t)}(y) \qquad \text{for } x \in \mathbb{R}^m \text{ and } t \in E, \tag{1.1}$$

originally defined over a 'nice' class of functions so that we can ignore the concerns about the definability of the operator. On this class of functions, define the maximal function

$$\mathcal{M} f(x) := \sup_{t \in E} |A_t f(x)| \qquad \text{for } \; x \in \mathbb{R}^m. \tag{1.2}$$

The general nature of problems studied in the thesis is to seek if $\mathcal{M}$ can be extended as a bounded operator between L^p spaces, and what are the best bounds we can obtain for $\|\mathcal{M}\|_{L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^m)}$ for $p \geq 1$.

While a broad formulation of this type helps present the questions at hand, it is essential to note that it may be deceptive, as a comprehensive theory of these operators falls outside the scope of this thesis. Instead, our analysis focuses on two particular branches of the framework, which will be examined in separate parts of the thesis.

In the first part of our thesis, we consider a specialized group structure (either commutative or non-commutative) on $\mathbb{R}^n$. Specifically, we examine a *homogeneous group* (denoted by $(\mathbb{R}^n, \cdot)$) that possesses a one-parameter family of *dilations* $(\delta_t)_{t>0}$ which are smooth maps on $\mathbb{R}^n$ that are compatible with the group law. These objects are rigorously defined in Chapter 2, and we refer the reader to the content of §2.1.1 for further details. Essentially, a homogeneous group can be thought of as a space that extends many of the fundamental aspects of $\mathbb{R}^n$, such as its abelian group law, manifold structure, Euclidean dilation structure, and the presence of a Haar measure (which is the Lebesgue measure).

Recalling the general setup described in (1.1) – (1.2), we let $m = n$ and $E := \{2^k : k \in \mathbb{Z}\}$. Begin with a compactly supported finite Borel measure σ and define $\sigma_{(t)}$ as a dilate of σ by the action $\langle \sigma_{(t)}, f\rangle = \langle \sigma, f \circ \delta_t\rangle$ on test functions. If σ is supported on a set $S \subset \mathbb{R}^n$, note that the dilated measure $\sigma_{(t)}$ is supported on $S_{(t)} := \delta_t(S)$. Define

$$\Phi(x, y) := x \cdot y^{-1} \qquad \text{for } x, y \in \mathbb{R}^n,$$

where $\cdot$ denotes the new group law on $\mathbb{R}^n$ and y^{-1} denotes the inverse of $y \in \mathbb{R}^n$ with respect to this group law. With these assumptions, we can re-write (1.1) as

$$A_t f(x) := \int_{\mathbb{R}^n} f(x \cdot y^{-1})\, \mathrm{d}\sigma_{(t)}(y).$$

In other words, A_t maps a function f to a convolution product (using the group law) between f and the t-dilate of σ. As the index set contains all the lacunary numbers, the maximal function defined by (1.2) will be called a lacunary maximal function and be denoted by the symbol $\mathcal{M}_{\mathrm{lac}}$. In Chapter 2, we study sufficient conditions for $\mathcal{M}_{\mathrm{lac}}$ to be extended as a bounded operator between L^p spaces for $p > 1$. The main theorem of this chapter (Theorem 2.2.4) generalises many known maximal results in some of the special homogeneous groups, such as the Euclidean space and the Heisenberg group. Our key assumption in this theorem is a 'curvature assumption' on the underlying measure σ (see the discussion in §2.2.2). It is known that without any curvature conditions, non-trivial maximal inequalities may fail.

In the second part of this thesis, we begin with a smooth non-degenerate curve $\gamma : [-1, 1] \to \mathbb{R}^m$. In particular,

$$|\det\left(\gamma^{(1)}(t) \cdots \gamma^{(m)}(t)\right)| > 0 \qquad \text{for all } t \in [-1, 1].$$

The curve γ determines a one-parameter family of directions in $\mathbb{R}^{m+1}$, and these directions will constitute the index set E for the maximal function. Let $E := [-1, 1]$. For each $t \in E$, we consider a long thin box $T_\gamma(t) \subset \mathbb{R}^{m+1}$ with the long axis (of unit length) along the direction $\begin{pmatrix}\gamma(t) & 1\end{pmatrix}^t$ and the other axes are determined by the derivatives of γ (see §4.1 for a precise definition). Define

$$S_{(t)} := T_\gamma(t) \quad \text{and} \quad \mathrm{d}\sigma_{(t)}(y) := |T_\gamma(t)|^{-1}\chi_{T_\gamma(t)}(y)\mathrm{d}y.$$

Fix $n = m + 1$. For $y \in \mathbb{R}^n$, we write $y = (y_1, y_2)$ where $y_1 \in \mathbb{R}^m$ and $y_2 \in \mathbb{R}$ and define $\Phi(x, y) := (x - y_1, y_2)$ for $x \in \mathbb{R}^m$. With these definitions, (1.1) now reads as

$$A_t f(x) := |T_\gamma(t)|^{-1} \int_{T_\gamma(t)} |f(x - y_1, y_2)|\mathrm{d}y_1\mathrm{d}y_2, \qquad \text{for } x \in \mathbb{R}^m.$$

By (1.2), the associated maximal function then seeks the largest average of f over all boxes whose long directions vary in a one-parameter family. Consequently, we refer to the resulting maximal function as a Nikodym maximal function (denoted by $\mathcal{M}_{\mathrm{nik}}$), keeping in line with the existing literature [14, 37, 3]. Maximal functions of this form arise in the study of certain local smoothing problems and the related (more complex) maximal functions. Via Theorem 4.1.2, we establish $L^p(\mathbb{R}^{m+1}) \to L^p(\mathbb{R}^m)$ estimates for $\mathcal{M}_{\mathrm{nik}}$ for dimensions $m \geq 2$ and for exponent $p \geq 2$, and improves on known results for $m = 2$ and $m = 3$. See the discussion in §4.1 after the statement of Theorem 4.1.2 for further details.

An important distinction between the study of $\mathcal{M}_{\mathrm{lac}}$ and $\mathcal{M}_{\mathrm{nik}}$ is seen to be in the role played by 'curvature'. In the $\mathcal{M}_{\mathrm{lac}}$ analysis, a curvature assumption is placed on the underlying singular measures to obtain non-trivial maximal estimates. In the case of the L^p estimates for $\mathcal{M}_{\mathrm{nik}}$, however, the non-degeneracy assumption is placed on the curve that determines the class of directions upon which we take the supremum.

The structure of the thesis is the following.

Part 1:

In Chapter 2, we study the lacunary maximal function $\mathcal{M}_{\mathrm{lac}}$ in the setting of a graded homogeneous group. To familiarise the reader with the theory of homogeneous groups and to keep it self-contained to an extent, we detail many fundamental notions and results in an abstract setup before stating our first main result and its proof. The same chapter discusses many examples of measures satisfying the required curvature assumption.

Part 2:

In Chapter 3, we conduct a literature review to examine how maximal inequalities relate to local smoothing problems. The primary goal of this chapter is to contextualise our main finding (Theorem 4.1.2) within the broader framework of local smoothing estimates. We begin this chapter by discussing a technique from [37] for solving the local smoothing problem for the wave equation in the plane, which leads to the proof of the Bourgain circular maximal theorem. Later, we explore the potential for generalising these methods to higher dimensional local smoothing problems, which involves studying a higher dimensional Nikodym maximal function.

In Chapter 4, we establish Lebesgue estimates for a Nikodym maximal function, a generalisation of the maximal function encountered in Chapter 3. In the final section of Chapter 4, we examine the geometric nature of the maximal function and discuss the challenges involved in devising a geometric proof for the Nikodym maximal estimate.

Chapter 2

L^p estimates for lacunary maximal functions on homogeneous groups

The content of this chapter is based on a joint work [50] with Jonathan Hickman and Jim Wright.

2.1 Background on homogeneous groups

The setting where we place our discussion is that of a *homogeneous group* which is a natural platform to build theories of operators originally arising from the Euclidean harmonic analysis[1]. In this section, we survey the background material on the theory of homogeneous groups. The material is standard and can be found, for instance, in [17] or [18].

2.1.1 Basic definitions

Lie groups and Lie algebras

Definition 2.1.1. A Lie group G is a smooth real manifold endowed with the smooth maps $(x, y) \mapsto x \cdot_G y$ and $x \mapsto x^{-1}$ satisfying the properties

i) $x \cdot_G (y \cdot_G z) = (x \cdot_G y) \cdot_G z$,

ii) $e \cdot_G x = x \cdot_G e = x$ and

iii) $x \cdot_G x^{-1} = x^{-1} \cdot_G x = e$

for all $x, y, z \in G$, where $e \in G$ is an element of the group called the identity element.

Definition 2.1.2. A Lie algebra is a real vector space V endowed with a bilinear map $[\,\cdot\,,\,\cdot\,] : V \times V \to V$ that maps $(a, b) \mapsto [a, b]$, called the Lie bracket of a and b, such that

[1]See, for example, in the introduction to [18], where the authors list several reasons why it is interesting to study operators in such generality.

i) $[a, a] = 0$ for all $a \in V$ and

ii) the Jacobi identity

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$$

holds for all $a, b, c \in V$. By writing $[a + b, a + b] = [a, a] + [a, b] + [b, a] + [b, b]$ we see that the first property is equivalent to

$$[a, b] = -[b, a] \qquad \text{for all a}, b \in V.$$

There is a natural association between Lie groups and Lie algebras. To see this, we begin with a Lie group G and derive a Lie algebra structure using some of the fundamental features of the group.

Definition 2.1.3. Let $a \in G$. A mapping $X_a : C^\infty(G) \to \mathbb{R}$ is called a tangent vector to G at a if

i) $X_a(f + g) = X_a f + X_a g$ and

ii) $X_a(fg) = X_a(f)g(a) + f(a)X_a(g)$

for all $f, g \in C^\infty(G)$. The collection of all tangent vectors at a forms a vector space denoted by T_aG.

T_aG is a finite dimensional[2] vector space with $\dim T_aG = \dim G$ for any $a \in G$. The disjoint union

$$TG := \bigsqcup_{a \in G} T_aG,$$

called the tangent bundle, assumes[3] a natural manifold structure from that of G.

Definition 2.1.4. A vector field on G is a smooth map $X : G \to TG$ such that $X_x := X(x) \in T_xG$ for any $x \in G$. It can act on $C^\infty(G)$ by

$$(Xf)(x) = X_x f \qquad \text{for all } f \in C^\infty(G).$$

We can introduce a bracket structure, called the commutator, on the space of vector fields by

$$[X, Y](x) := X_x Y - Y_x X \qquad \text{for all } x \in G. \tag{2.1}$$

It is easy to verify that $[X, Y]$ is also a vector field and that the commutator satisfies the properties of a Lie bracket. Thus, the collection of all vector fields on G forms a Lie algebra (albeit an infinite dimensional one). We are interested in one of its Lie sub-algebra, which can be introduced by taking the group structure of G also into consideration.

[2] See [17, §1.2] for a proof.
[3] See [17, §1.2] for further details.

For $a \in G$, consider the left and right translation maps $\mathcal{L}_a, \mathcal{R}_a : G \to G$ defined by

$$\mathcal{L}_a(x) := a \cdot_G x, \qquad \mathcal{R}_a(x) := x \cdot_G a, \tag{2.2}$$

and the associated differential maps $d\mathcal{L}_a, d\mathcal{R}_a : TG \to TG$ defined by

$$d\mathcal{L}_a(X)f := X(f \circ \mathcal{L}_a) \circ \mathcal{L}_{a^{-1}} \qquad \text{and} \qquad d\mathcal{R}_a(X)f := X(f \circ \mathcal{R}_a) \circ \mathcal{R}_{a^{-1}}$$

for all $X \in TG$ and $f \in C^\infty(G)$.

Definition 2.1.5. A vector field $X : G \to TG$ is called *left-invariant* if

$$X \circ \mathcal{L}_a = d\mathcal{L}_a \circ X \qquad \text{for all } a \in G. \tag{2.3}$$

Similarly, the vector fields $X : G \to TG$ is called *right-invariant* if

$$X \circ \mathcal{R}_a = d\mathcal{R}_a \circ X \qquad \text{for all } a \in G.$$

The vector space of all left-invariant vector fields is closed under taking Lie brackets. Indeed, using (2.1) and (2.2), the identity

$$[d\mathcal{L}_a(X), d\mathcal{L}_a(Y)] = d\mathcal{L}_a([X,Y])$$

can be verified for any two vector fields X and Y and $a \in G$. Consequently, if X and Y are left-invariant vector fields, then

$$[X,Y] \circ \mathcal{L}_a = [X \circ \mathcal{L}_a, Y \circ \mathcal{L}_a] = [d\mathcal{L}_a(X), d\mathcal{L}_a(Y)] = d\mathcal{L}_a([X,Y])$$

which implies that $[X,Y]$ is also left-invariant. Therefore, the space of all left-invariant vector fields forms a Lie algebra equipped with the commutator bracket.

Definition 2.1.6. The Lie algebra $\mathfrak{g}$ of a Lie group G is the space of all left-invariant vector fields on G equipped with the commutator bracket of vector fields.

One important observation to be noted here is that by Definition 2.1.5, a left-invariant vector field is uniquely determined at all points in G, once it is determined at the identity (or at any other point for that matter). The mapping $X \mapsto X(e)$ is therefore seen to be a vector space isomorphism between $\mathfrak{g}$ and the tangent space T_eG; for a vector $X_e \in T_eG$, we use (2.3) to define the unique vector field that identifies with X_e at the identity, the smoothness of which can be easily checked[4]. We can use this isomorphism to equip T_eG with a Lie bracket induced from the commutator Lie bracket of $\mathfrak{g}$. Then, T_eG becomes a Lie algebra that is isomorphic to $\mathfrak{g}$. In view of this, we use the symbol $X \in \mathfrak{g}$ to represent both a left-invariant vector field and its value at the identity, as long as the notational inconsistency does not lead to any ambiguity in the statement.

[4]See [17, §1.2] for further details.

The exponential map

We can now introduce the *exponential map*, which plays an important role in establishing the Lie algebra - Lie group correspondence[5]. For $X \in \mathfrak{g}$, consider the initial value problem for a function $\gamma : [0, \epsilon) \to G, \epsilon > 0$, satisfying

$$\gamma'(t) = X(\gamma(t)),\ \gamma(0) = e.$$

From the theory of ordinary differential equations, we know that this equation has a unique solution (which is also called a *flow* associated to X) on an interval $[0, \epsilon)$ for a small positive value ϵ, and the solution depends smoothly on X_e. Furthermore, the interval of existence can be extended by taking smaller and smaller vectors X_e around the origin, in particular, so that the solution exists on $[0, 1]$. In this case we define

$$\exp(X) := \gamma(1).$$

By applying the inverse function theorem[6], the exponential map can be shown to be a local diffeomorphism from some open neighborhood of $0 \in \mathfrak{g}$ to some open neighborhood of $e \in G$.

The interaction between the group law, the Lie bracket, and the exponential map is made explicit by the Baker–Campbell–Hausdorff formula. In particular, for any $X, Y \in \mathfrak{g}$ sufficiently close to the origin, we have

$$\exp X \cdot \exp Y = \exp P(X, Y)$$

whenever $P(X, Y)$ converges, where[7]

$$P(X, Y) := X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] + [Y, [Y, X]]) + \text{higher order terms.} \tag{2.4}$$

In what follows, we restrict our attention to Lie algebras where P reduces to a finite degree polynomial.

Definition 2.1.7. A Lie algebra $\mathfrak{g}$ is called *nilpotent* if there exists $m \in \mathbb{N}$ such that the lower central series

$$[X_1, [X_2, [\ldots, [X_m, X_{m+1}]] \ldots]] = 0 \quad \text{for any } X_1, \ldots, X_{m+1} \in \mathfrak{g}. \tag{2.5}$$

The *step* of $\mathfrak{g}$ (and the step of the associated Lie group) is the smallest natural number m satisfying (2.5).

If G is a connected and simply connected nilpotent group (with step m), it is possible to show that[8] the exponential map is a global diffeomorphism from $\mathfrak{g}$ to G. By fixing a basis for $\mathfrak{g}$, one can identify the Lie algebra with $\mathbb{R}^n$ where $n = \dim \mathfrak{g}$. Using the exponential map, we can then identify G with $\mathbb{R}^n$. In view of (2.4), the group law of G is expressed as a polynomial with degree m.

[5] We do not provide a full picture of the Lie algebra - Lie group correspondence here. The interested reader can refer to standard textbooks on Lie theory such as [5] for further details.

[6] See [17, §1.2] for details.

[7] The exact formula may be found in [17, Theorem 1.3.2].

[8] see [18, Proposition 1.2] for details.

Let $\mathrm{d}x$ denote the Lebesgue measure on $\mathfrak{g}$. The push-forward of $\mathrm{d}x$ under the exponential map, a bi-invariant (invariant under the left and right translations) Haar measure on G, will also be denoted by $\mathrm{d}x$. In other words, G is a *unimodular* group, a group where the left-invariant Haar measure coincides with the right-invariant Haar measure.

Homogeneous groups

Generalising the Euclidean dilation structure, one can equip $\mathfrak{g}$ with a one parameter family of automorphisms under the same name.

Definition 2.1.8. For $t > 0$, the *t-dilation* on $\mathfrak{g}$, denoted by δ_t is an algebra automorphism of $\mathfrak{g}$ defined by

$$\delta_t := e^{(A \log t)} = \sum_{l \geq 0} \frac{((\log t)A)^l}{l!}, \tag{2.6}$$

where the dilation matrix A is a diagonalisable linear operator on $\mathfrak{g}$ with positive eigenvalues. In particular, the identity $\delta_{st} = \delta_s \circ \delta_t$ holds for any $s, t \in \mathbb{R}_+$.

Remark 2.1.9. For any $t > 0$, the map $\exp \circ \delta_t \circ \exp^{-1}$ is a group automorphism of G. These maps are called the dilations of G and, by an abuse of notation, are also denoted by δ_t. In particular,

$$\delta_t \exp(X) = \exp(\delta_t X) \qquad \text{for} \quad X \in \mathfrak{g} \text{ and } t > 0. \tag{2.7}$$

Consider a basis $\{X_j \in \mathfrak{g} : 1 \leq j \leq n\}$ for $\mathfrak{g}$ that consists of eigenvectors of A. Once and for all, we fix such a basis for $\mathfrak{g}$ so that the dilation structure on G takes a neat form in exponential coordinates. In particular, if λ_i is the eigenvalue associated to X_i, then

$$\delta_r((x_1, \ldots, x_n)) = (r^{\lambda_1} x_1, \ldots, r^{\lambda_n} x_n) \qquad \text{for all } (x_1, \ldots, x_n) \in G.$$

The following lemma shows that the admission of a dilation structure in a Lie algebra is a stronger condition than the nilpotency.

Lemma 2.1.10. *[18, Proposition 1.3] If $\mathfrak{g}$ admits a family of dilations, it is nilpotent.*

Proof. It suffices to verify (2.5) for some $m \in \mathbb{N}$ with X_i's chosen from a basis of $\mathfrak{g}$. As $\mathfrak{g}$ admits a family of dilations, we may fix a basis formed by eigenvectors of the dilation matrix A (recall Definition 2.1.8).

Suppose $X, Y \in \mathfrak{g}$ are two eigenvectors of A with eigenvalues $a \in \mathbb{R}_+$ and $b \in \mathbb{R}_+$ respectively. Equivalently, for any $t > 0$, they are eigenvectors of δ_t with eigenvalues t^a and t^b, respectively. Since δ_t is an algebra automorphism,

$$\delta_t([X, Y]) = [\delta_t X, \delta_t Y] = t^{a+b}[X, Y].$$

In particular, $[X, Y]$ is an eigenvector of δ_t with a new eigenvalue t^{a+b} for any $t > 0$. Thus, if X_i is an eigenvector of δ_t for $1 \leq i \leq m$, the iterated bracket product

$$[X_1, [X_2, [\dots, [X_m, X_{m+1}]\dots]] \tag{2.8}$$

is also an eigenvector of δ_t with an eigenvalue different from any of X_i. As the dilation map, and equivalently, the dilation matrix A can only have finitely many eigenvalues, there must exist an $m \in \mathbb{N}$ such that the iterated product (2.8) gives 0 for any choice of X_i from the basis. The proof ends here. $\square$

With all these definitions, we can introduce the notion of homogeneous groups.

Definition 2.1.11 (Homogeneous group)**.** A *homogeneous group* G is a connected, simply connected Lie group associated with a Lie algebra $\mathfrak{g}$ which admits a one-parameter family of dilations $\{\delta_t\}_{t>0}$.

Example 2.1.12. The prototypical example of a non-commutative homogeneous group is the Heisenberg group. To introduce its group structure, consider a homogeneous group G with a Lie algebra $\mathfrak{g}$ equipped with a one-parameter family of dilations $(\delta_t)_{t>0}$ such that $\{1, 2\}$ forms the set of eigenvalues for the dilation matrix A. As A is diagonalisable, we write $\mathfrak{g} = V_1 \oplus V_2$, where V_j denotes eigenspace of A associated to the eigenvalue j. As in the proof of Lemma 2.1.10, the inclusions

$$[V_1, V_1] \subseteq V_2 \qquad \text{and} \qquad [V_1, V_2] = [V_2, V_2] = \{0\}$$

follow as easy consequences of the properties of the dilation mapping. In other words, $\mathfrak{g}$ is nilpotent with step 2. Let $\dim V_1 = m$ and $\dim V_2 = d$. By Baker–Campbell–Hausdorff formula, the associated connected and simply connected Lie group G has a group operation, expressed in exponential coordinates as

$$(\vec{x}, \vec{u}) \cdot_G (\vec{y}, \vec{v}) := (\vec{x} + \vec{y}, u_1 + v_1 + \frac{1}{2}\vec{x}^\top J_1 \vec{y}, \dots, u_d + v_d + \frac{1}{2}\vec{x}^\top J_d \vec{y})$$

where $\vec{x}, \vec{y} \in \mathbb{R}^m$, $\vec{u} = (u_1, \dots, u_d), \vec{v} = (v_1, \dots, v_d) \in \mathbb{R}^d$ and each J_i is a skew-symmetric $m \times m$ matrix. Furthermore, for $t > 0$, the dilation map $\delta_t : G \to G$ maps

$$(\vec{x}, \vec{u}) \mapsto (t\vec{x}, t^2\vec{u}),$$

which is again expressed in exponential coordinates. Now, consider the special case when $m = 2k$ even, $d = 1$ and $J_1 = J$ where

$$J := \begin{bmatrix} 0 & I_{k\times k} \\ -I_{k\times k} & 0 \end{bmatrix}.$$

In this case, the Lie algebra $\mathfrak{g}$ (which is identified with $\mathbb{R}^{2k} \oplus \mathbb{R}$) is called the kth Heisenberg Lie algebra (denoted by $\mathfrak{h}^k$) the corresponding group is the kth Heisenberg group (denoted by $\mathbb{H}^k$).

Definition 2.1.13 (Homogeneous dimension)**.** Let G admit a family of dilations. If $\lambda_1, \dots, \lambda_n > 0$ are the eigenvalues (with repetitions) of the dilation matrix A,

then
$$Q(G) := \sum_{1\le i\le n} \lambda_i = \mathrm{tr}(A)$$
is called the homogeneous dimension of G.

Example 2.1.14. It is immediate to see that $Q(\mathbb{R}^n) = n$ and $Q(\mathbb{H}^k) = 2(k+1)$.

Remark 2.1.15. For $\psi \in L^1(G)$ and $t > 0$, by a simple change of variables,
$$\|\psi_t\|_{L^1(G)} = \|\psi\|_{L^1(G)} \tag{2.9}$$
where
$$\psi_t := t^{-Q}\psi \circ \delta_{t^{-1}}. \tag{2.10}$$

Definition 2.1.16 (Homogeneous norm)**.** A *homogeneous norm* on G is a function $|\cdot|_G : G \to [0,\infty)$, that is C^∞ away from the identity, satisfying the following:

(a) $|0|_G = 0$ and $|x|_G > 0$ for any $0 \neq x \in G$.

(b) $|\delta_t(x)|_G = t|x|_G$ and $|x^{-1}|_G = |x|_G$, for any $x \in G$, $t > 0$.

(c) There exists $C > 0$ such that $|x \cdot y|_G \le C(|x|_G + |y|_G)$ for all $x, y \in G$.

To give an example, the Korányi norm on $\mathbb{H}^k$, defined by
$$|\vec{w}|_{\mathbb{H}^k} := (\sum_{i=1}^{2k} x_i^4 + u^2)^{1/4} \qquad \text{whenever } \vec{w} = (x_1, \dots, x_{2k}, u) \in \mathbb{H}^k,$$
is a homogeneous norm.

At this point, it is worth noting that the Euclidean norm G assumes from its coordinate identification with a Euclidean space does not result in a homogeneous norm. Nevertheless, as the following lemma shows, a homogeneous group can always be equipped with a homogeneous norm. The result is taken from [18, §1. A].

Lemma 2.1.17. *At least one homogeneous norm exists on any homogeneous group G.*

Proof. For $x \in G \setminus \{0\}$, the Euclidean norm $|\delta_t x|$ is a monotonically increasing function in t and has the range $(0,\infty)$. Thus, there exists $t = t(x) \in (0,\infty)$ such that $|\delta_t x| = 1$. We define
$$|x|_G := \begin{cases} [t(x)]^{-1} & \text{when } x \neq 0, \\ 0 & \text{when } x = 0. \end{cases}$$
It is easy to verify the conditions (a) and (b) in Definition 2.1.16 for $|\cdot|_G$. By setting
$$C := \sup_{|x|_G+|y|_G=1} |x \cdot y|_G,$$
we can also verify (c) when $|x|_G + |y|_G = 1$ and this can be extended to any $x, y \in G$ via homogeneity assumption in (b). □

By a ball in G, we always refer to a ball that is defined using the homogeneous norm:

Definition 2.1.18. For $x \in G$ and $r \in \mathbb{R}_+$, the symbol $B(x,r)$ represents a ball centered at x with radius r defined by

$$B(x,r) := \{y \in G : |x^{-1}y|_G \leq r\}.$$

Another group feature that is important to our analysis is grading.

Definition 2.1.19. A Lie algebra $\mathfrak{g}$ is *graded* if it is endowed with a vector space decomposition of the form

$$\mathfrak{g} = \bigoplus_{j=1}^{\infty} V_j \quad \text{such that} \quad [V_i, V_j] \subseteq V_{i+j}, \tag{2.11}$$

where all but finitely many of the V_j's are trivial vector spaces. A Lie group is *graded* if it is a connected and simply connected Lie group whose Lie algebra is graded.

Example 2.1.20. The Lie algebra $\mathfrak{h}^k$ associated to the Heisenberg group $\mathbb{H}^k$ is graded and takes the form $\mathfrak{h}^k = \mathbb{R}^{2k} \oplus \mathbb{R}$.

Using the following lemma, we see that the admission of a graded structure is stronger property than the existence of dilations.

Lemma 2.1.21. *If $\mathfrak{g}$ is a graded Lie algebra, it admits a one-parameter family of dilations.*

Proof. Recalling Definition 2.1.8, it suffices to construct a diagonalisable matrix A with positive eigenvalues; for each $t > 0$, the dilation operator δ_t can then be defined by (2.6). Recalling (2.11), we may choose A to be such that

$$AX = jX \qquad \text{whenever } X \in V_j \text{ and } j \in \mathbb{N}.$$

To complete the proof of the lemma, we must show that the dilation operators defined by (2.6) are algebra automorphisms. To this end, take $t \in \mathbb{R}_+$, and by definition, the vector space V_j is the eigenspace of δ_t corresponding to the eigenvalue t^j for each $j \in \mathbb{N}$. Thus, δ_t is a bijective linear map on $\mathfrak{g}$. Furthemore, for $X \in V_j$ and $Y \in V_k$, (2.11) implies that $[X,Y] \in V_{j+k}$. Thus, using (2.6), we have

$$\delta_t([X,Y]) = t^{j+k}[X,Y] = [t^j X, t^k Y] = [\delta_t(X), \delta_t(Y)].$$

As $\mathfrak{g}$ is graded, the linearity of δ_t can now be used to conclude that

$$\delta_t([X,Y]) = [\delta_t X, \delta_t Y] \qquad \text{for any } X, Y \in \mathfrak{g}.$$

The proof ends here. $\square$

Definition 2.1.22. a Lie algebra $\mathfrak{g}$ is *stratified* when it is graded, $\mathfrak{g} = \bigoplus_{j=1}^{\infty} V_j$, and V_1 generates $\mathfrak{g}$ as an algebra. In other words, every element of $\mathfrak{g}$ can be written as a linear combination of iterated Lie brackets of elements of V_1, or equivalently,

$$[V_1, V_j] = V_{1+j}, \qquad \text{for all } 1 \leq j \leq m-1.$$

A Lie group is stratified if it is a connected and simply connected Lie group whose Lie algebra is stratified.

In the next subsection, we discuss some of the actions on the class of smooth functions on the group.

2.1.2 Analysis on homogeneous groups

Definition 2.1.23. For a function $f : G \to \mathbb{C}$, the symbol $\tilde{f}$ denotes its reflection, defined by

$$\tilde{f}(x) := f(x^{-1}) \qquad \text{for } x \in G.$$

Definition 2.1.24. Let $\mathcal{M}(G)$ denote the collection of complex Borel measures on G with finite total variation norm. For $\mu \in \mathcal{M}(G)$, we use the notation $\|\mu\|_{\mathcal{M}(G)}$ to represent its total variation norm.

Action by vector fields

In view of the definition of the exponential map, a vector field $X \in \mathfrak{g}$ can also be interpreted as a left-invariant differential operator on $C^{\infty}(G)$ defined by

$$Xf(x) := \left(\frac{d}{dt}\right)_{t=0} f(x \cdot_G \exp(tX)). \tag{2.12}$$

Let $X \in \mathfrak{g}$ is an eigenvector of the dilation matrix A (recall (2.6)) with eigenvalue λ. Using (2.12) and (2.7), we deduce that

$$\begin{aligned} X[f \circ \delta_s(x)] &= \left(\frac{\mathrm{d}}{\mathrm{dt}}\right)_{t=0} f(\delta_s(x)\delta_s \exp(tX)) \\ &= \left(\frac{\mathrm{d}}{\mathrm{dt}}\right)_{t=0} f(\delta_s(x) \exp(t\delta_s(X))) \\ &= \left(\frac{\mathrm{d}}{\mathrm{dt}}\right)_{t=0} f(\delta_s(x) \exp(ts^{\lambda}X)) \\ &= s^{\lambda}(Xf) \circ \delta_s(x). \end{aligned}$$

The above computation will be extremely useful later in the chapter.

Convolutions

The convolution product in G is always assumed to be the right convolution. In particular, for measurable functions f and g, define

$$f * g(x) := \int_G f(xy^{-1})g(y)\mathrm{d}y = \int_G f(y)g(y^{-1}x)\mathrm{d}y \qquad \text{for } x \in G,$$

if the integrals converge. In the below, we state Young's inequality over homogeneous groups, which in particular, ensures absolute convergence of the integrals almost everywhere when f and g belong to some Lebesgue spaces.

Proposition 2.1.25 ([18], Proposition 1.18[9]). *Suppose $1 \leq p, q, r \leq \infty$ and $1/r + 1 = 1/p + 1/q$. If $f \in L^p(G)$ and $g \in L^q(G)$, then $f * g \in L^r(G)$, and*

$$\|f * g\|_{L^r(G)} \leq \|f\|_{L^p(G)} \|g\|_{L^q(G)}.$$

Furthermore, if $\mu \in \mathcal{M}(G)$, then

$$\|f * \mu\|_{L^p(G)} \leq \|\mu\|_{\mathcal{M}(G)} \|f\|_{L^p(G)} \qquad \text{for any } 1 \leq p \leq \infty. \tag{2.13}$$

The following convolution identities are verified by elementary computations. We omit their proofs here.

(i) For $f, g \in L^1(G)$, the reflection of $f * g$ is same as $\tilde{g} * \tilde{f}$.

(ii) For any $f, g \in L^1(G)$ and $t > 0$, we have $(f * g)_t = f_t * g_t$, where the t-dilate of a function is defined by (2.10).

(iii) Suppose X is a left-invariant and $\tilde{X}$ is a right-invariant vector field. Let $f, g \in C^2(G) \cap L^1(G)$. Then, by differentiating under the integral sign, we have
$$X(f * g) = f * Xg \quad \text{and} \quad \tilde{X}(f * g) = \tilde{X} f * g.$$

(iv) For $X \in \mathfrak{g}$, let $\tilde{X}$ represent the (unique) right-invariant vector field that coincides with X at the tangent space at the identity. Let the functions $f, g \in C^1(G) \cap L^1(G)$ be chosen such that $\tilde{X} f \in L^1(G)$ and $Xg \in L^1(G)$. Then, by integration-by-parts,
$$f * Xg = \tilde{X} f * g.$$

As we have acquired sufficient knowledge of homogeneous groups, we can proceed to our main section, where a specific maximal question will be framed in this setting.

2.2 The lacunary maximal problem

2.2.1 The maximal function

For a compactly supported finite Borel measure σ on G, define the averaging operator $\mathcal{A}[\sigma]$ by

$$\mathcal{A}[\sigma] f(x) := f * \sigma(x) = \int_G f(xy^{-1}) \, \mathrm{d}\sigma(y) \qquad \text{for } x \in G, \tag{2.14}$$

[9] The reference only treats the functional case, but the version of Young's inequality for measures follows in a similar manner

whenever $f \in C_c(G)$. We are interested in a maximal averaging operator defined over G, where the averages are taken with respect to dilations of a fixed measure. For $t > 0$, we define the t-dilate of σ by the action

$$\langle \sigma_t, f \rangle := \int_G f \circ \delta_t(x)\, \mathrm{d}\sigma(x) \qquad \text{for all } f \in C_c(G).$$

Define the lacunary maximal function

$$\mathcal{M}f(x) := \sup_{k \in \mathbb{Z}} |\mathcal{A}[\sigma_{2^k}]f(x)| \qquad \text{for } x \in G,\ f \in C_c(G).$$

We shall investigate the $L^p(G)$ mapping properties of $\mathcal{M}$ under certain 'curvature' conditions on the underlying measure.

In the Euclidean setting (when $G = \mathbb{R}^n$), the first major result on this problem comes from [8], where σ is taken to be the surface measure on the unit sphere $\mathbb{S}^{n-1}$. In this case, [8] proves that $\mathcal{M}$ is strong type (p,p) whenever $1 < p \leq \infty$. A more general result was presented in [16] where maximal functions associated with a larger collection of measures are considered, and the only assumption placed on the class of measures is in terms of the decay of their Fourier transforms. In particular, if $\hat{\sigma}(\xi) = O(|\xi|^{-\epsilon})$ for some $\epsilon > 0$ and for large values of $\xi \in \widehat{\mathbb{R}}^n$, then it was shown that $\mathcal{M}$ is strong type (p,p) for $p \in (1,\infty)$. At the endpoint near L^1, although strong type estimates fail, weak type estimates such as $H^1 \to L^{1,\infty}$ or $L \log L \to L^{1,\infty}$ bounds have been established under various assumptions related to either the Fourier decay of the measure or the curvature features of its singular support (see [27, 39, 48, 12]).

Coming to the non-euclidean regime, the maximal problem has been studied over Heisenberg groups or Métivier groups. The case when σ is the surface measure on the unit sphere under the Korányi norm in the Heisenberg group has been considered in [20, 43]. A more singular case was considered in [2], where the measure is supported on a co-dimension 2 sphere in $\mathbb{H}^k$, $k \geq 2$. A more detailed discussion of these results will be carried out later in the chapter.

2.2.2 Curvature assumption

For all known estimates of the lacunary maximal function, some assumption about the 'curvature' of the singular support of σ is used. It is well established that no non-trivial maximal estimates are available when the averages are taken along lines or other surfaces of zero curvature. From [16], we see that the Fourier decay on the measure σ can be used as the basic assumption in order to obtain operator bounds for $\mathcal{M}$ when $G = \mathbb{R}^n$. When the singular support of the measure is a compact submanifold in $\mathbb{R}^n$, the Fourier decay assumption can also be related to the standard notions of Gaussian or principle curvatures of the manifold.[10] However, at the level of generality of homogeneous groups, the machinery of the Fourier transform becomes extremely difficult to work with. Instead, we work with a different property of a measure well connected to the Fourier decay or the curvature properties of the associated singular support in the Euclidean case,

[10] See [54, VIII. B] for further details.

which easily extends to the abstract setting of a homogeneous group. An ideal candidate is a property of the iterated convolution products of the measure.

Let σ be a compactly supported finite Borel measure on G and recall the definition of $\mathcal{A}[\sigma]$ from (2.14). Observe that its adjoint is given by the map $\mathcal{A}[\tilde{\sigma}]$, where $\tilde{\sigma}$ is defined by the action

$$\langle \tilde{\sigma}, f \rangle := \int_G f(x^{-1})\, \mathrm{d}\sigma(x) \qquad \text{for } f \in C_c(G).$$

Now, for $N \in \mathbb{N}_0$, we define the Nth convolution product $\sigma^{(N)}$ recursively; starting with $\sigma^{(0)} = \sigma$, define

$$\sigma^{(N)} := \begin{cases} \sigma^{(N-1)} * \tilde{\sigma} & \text{when } N \text{ is odd,} \\ \sigma^{(N-1)} * \sigma & N \geq 2 \text{ is even.} \end{cases} \tag{2.15}$$

We frame our 'curvature assumption' around the regularity properties of such iterated products. To motivate the assumption, take $G = \mathbb{R}^n$. Suppose the decay estimate $\hat{\sigma}(\xi) = O(|\xi|^{-\epsilon})$ holds for any $\xi \neq 0$ and for a fixed $\epsilon > 0$. Let $N_\epsilon := 10\lceil n\epsilon^{-1}\rceil$ and $\mu := \sigma^{(N_\epsilon)}$. Observe that

$$\hat{\mu}(\xi) = O_n(|\xi|^{-10n}) \qquad \text{for } \xi \in \hat{\mathbb{R}}^n \setminus \{0\}\ .$$

Since ν is compactly supported with a finite mass, one can improve this further to $|\hat{\nu}(\xi)| \lesssim_n (1 + |\xi|)^{-10n}$ for any $\xi \in \hat{\mathbb{R}}^n$. Consequently, $\hat{\nu}$ lies in $L^2(\hat{\mathbb{R}}^n) \cap L^1(\hat{\mathbb{R}}^n)$. Using Plancherel's theorem, we deduce that ν is absolutely continuous with respect to the Lebesgue measure with an L^2 density, say h. Furthermore,

$$\begin{aligned} |h(x-y) - h(x)| &\lesssim_n \left| \int_{\hat{\mathbb{R}}^n} \hat{h}(\xi) e^{i\langle x,\xi\rangle} \left(e^{-i\langle y,\xi\rangle} - 1\right) \mathrm{d}\xi \right| \\ &\leq \int_{\hat{\mathbb{R}}^n} |\hat{h}(\xi)| |\langle y, \xi\rangle|\, \mathrm{d}\xi \leq |y| \int_{\hat{\mathbb{R}}^n} \min\{|\xi|, |\xi|^{-10n+1}\} \mathrm{d}\xi \\ &\lesssim |y|, \end{aligned}$$

whenever $x, y \in \mathbb{R}^n$. In other words, one can amplify the Fourier decay of the measure by taking iterated convolution products with itself (or its adjoint), ultimately reaching an absolutely continuous measure with a Lipschitz density. In situations where the singular support of σ is (a piece of) a submanifold embedded in $\mathbb{R}^n$, non-trivial Fourier decay is related to non-zero curvature. In these situations, we can also give a geometric interpretation to the first part of the above observation, that the process of iterated self-sums of the submanifold will culminate in a set of positive measure, provided the manifold *curves*.

In [42], while studying the problem of norm-estimating convolution operators where the kernels are supported on smooth low-dimensional varieties in G, the authors studied the regularity properties of the iterated products. Motivated by their results, we consider the following definition.

Definition 2.2.1 (Curvature assumption)**.** Let σ be a Borel measure in G. We say that σ satisfies *curvature assumption* (C) if the following hold.

(i) There exists $N > 0$ such that $\mu := \sigma^{(N)}$ is absolutely continuous with respect to the Haar measure on G.

(ii) If h denotes the density of μ, then h satisfies an L^1-Hölder type condition: there exists $\delta > 0$ and $C_\delta > 0$ such that

$$\int_G |h(xy^{-1}) - h(x)| \mathrm{d}x + \int_G |h(y^{-1}x) - h(x)| \mathrm{d}x \leq C_\delta |y|^\delta \quad \text{for } y \in G. \tag{2.16}$$

Definition 2.2.2. Let $\delta > 0$. We use the notation $L^1_\delta(G)$ to represent space of all $L^1(G)$ functions satisfying (2.16), equipped with the norm

$$\|h\|_{L^1_\delta(G)} := \|h\|_{L^1(G)} + \bar{C}_\delta(h),$$

where $\bar{C}_\delta(h)$ is the smallest constant for which (2.16) holds.

Going back to the discussion on the relation between the Fourier decay of a measure and (C) in the Euclidean setting, we note the equivalence between the two notions:

Lemma 2.2.3. *Let σ be a compactly supported finite Borel measure on $\mathbb{R}^n$. σ satisfies* (C) *if and only if there exists $\epsilon > 0$ such that $\hat{\sigma}(\xi) = O(|\xi|^{-\epsilon})$ for all $\xi \in \hat{\mathbb{R}}^n \setminus \{0\}$.*

Proof. The "if" part has already been verified by the discussion above. To see how the reverse implication holds, we first note that by Plancherel's theorem, Fourier decay of σ can be deduced from that of $\sigma^{(l)}$ for some $l \geq 0$. Fix $N \geq 0$ such that $\sigma^{(N)}(x) = h(x)\mathrm{d}x$ satisfies the regularity conditions (i) and (ii) in Definition 2.2.1. Now,

$$\begin{aligned}
|\left(e^{-i\langle y,\xi\rangle} - 1\right)\hat{h}(\xi)| &= \left|\int_{\mathbb{R}^n} (h(x-y) - h(x))\mathrm{d}x\right| \\
&\leq \int_{\mathbb{R}^n} |h(x-y) - h(x)|\mathrm{d}x \lesssim \|h\|_{L^1_\delta(\mathbb{R}^n)}|y|^\delta.
\end{aligned}$$

Choosing y to be $\pi\xi/|\xi|^2$, we conclude the proof. □

2.2.3 Maximal theorem and applications

We are now in a position to state our maximal theorem.

Theorem 2.2.4. *Let G be a homogeneous group and σ be a finite compactly supported Borel measure on G. If σ satisfies (C), then $\mathcal{M}$ is bounded on $L^p(G)$ for all $p > 1$.*

Before turning to the proof of Theorem 2.2.4, we discuss some applications.

Euclidean case

For $G = \mathbb{R}^n$, in view of Lemma 2.2.3, our theorem recovers a classical result of Duoandikoetxea and Rubio de Francia [16]. In particular, if σ is a finite, compactly supported Borel measure on $\mathbb{R}^n$ such that for some $\epsilon > 0$ the Fourier decay condition

$$|\hat{\sigma}(\xi)| \lesssim |\xi|^{-\epsilon} \qquad \text{for all } \xi \in \widehat{\mathbb{R}}^n \setminus \{0\} \tag{2.17}$$

holds, then the associated lacunary maximal function $\mathcal{M}$ is bounded on $L^p(G)$ for all $1 < p \leq \infty$. Examples of measures satisfying the Fourier decay assumption (2.17) include

(a) surface measures on finite type surfaces in $\mathbb{R}^n$ (see [54, §3.2] for the Fourier decay estimates),

(b) special fractal measures; see [34, Theorem 1.4] for the construction of measures in a torus with varying Fourier and Hausdorff dimensions, and [19] for explicit Salem sets in $\mathbb{R}^n$.

Surfaces in the Heisenberg group

For $G = \mathbb{H}^k$, we list two important applications of Theorem 2.2.4. Verifying the curvature assumption for both cases is postponed until the next subsection.

(i) Let σ be the surface measure on the unit Korányi sphere

$$\mathbb{S}_{\mathrm{kor}} := \{w \in \mathbb{H}^k : |w|_{\mathbb{H}^k} = 1\}.$$

For this case, Theorem 2.2.4 recovers the previously known maximal bounds; the L^p estimates for the Korányi lacunary maximal function obtained in [20, Theorem 1.2] for the range $k \geq 2$. Theorem 2.2.4 also removes the dimensional constraint and proves that the associated maximal function is bounded on $L^p(\mathbb{H}^k)$ for all $1 < p \leq \infty$ and $k \geq 1$.

(ii) For the second example, we let σ be the surface measure on the co-dimension two sphere

$$\mathbb{S}^{2k-1} := \{(z, 0) \in \mathbb{R}^{2k+1} : |z|_{\mathbb{R}^{2k}} = 1\}. \tag{2.18}$$

Theorem 2.2.4 recovers [2, Theorem 1.1] when $k \geq 2$ (see also [43], where the $k = 1$ case and extensions to Métivier groups are considered) although, as remarked in [43], such bounds can be directly deduced from earlier work such as [38].

Both examples are special cases of results on more general groups (see Lemma 2.2.5 and 2.2.6).

Analytic surfaces in general homogeneous groups

In this section, we consider classes of analytic surfaces that satisfy the curvature assumption and, as a consequence, come under the purview of Theorem 2.2.4.[11]

[11] The findings of this section are proved in collaboration with J. Hickman.

Let G be a homogeneous group and S be a smooth submanifold G. We say a Borel measure σ on G is a C_c^∞*-density on* S if it is of the form $\eta \mathrm{d}\sigma_S$ where $\eta \in C_c^\infty(S)$ is a smooth, compactly supported function on S and σ_S is the natural surface measure on S induced by the Haar measure on G.

Lemma 2.2.5. *Let G be a graded homogeneous group with $\dim G \geq 2$. Suppose Ω is an open convex domain in $\mathfrak{g}$ with an analytic boundary $\Sigma := \partial\Omega$. It follows that any C_c^∞-density σ on $\exp(\Sigma)$ satisfies* (C).

For any graded homogeneous group G and a homogeneous norm $\|\cdot\|$ on G, the unit sphere in G with respect to $\|\cdot\|$ can be written as $\exp(\Sigma)$ such that Σ is an analytic boundary to a convex domain as in Lemma 2.2.5. In particular, we recover the maximal estimates associated to the Korányi sphere in $\mathbb{H}^k$ as listed in the previous subsection.

Under a stronger assumption on the group, we can verify the curvature assumption on a larger class of analytic surfaces.

Lemma 2.2.6. *Let G be a stratified group with $\dim G \geq 2$ with the Lie algebra $\mathfrak{g} = \bigoplus_{i=1}^m V_i$. Let $\Sigma \subseteq \mathfrak{g}$ be an analytic submanifold of $\mathfrak{g}$ such that $\Pi_1(\Sigma)$ generates V_1 (in terms of vector addition) where $\Pi_\ell \colon \mathfrak{g} \to V_\ell$ denote the subspace projection onto V_ℓ for $1 \leq \ell \leq m$. It follows that any C_c^∞-density σ on $\exp(\Sigma)$ satisfies* (C).

As the first application of the lemma, we see that the surface measure on the co-dimension two sphere in the Heisenberg group as defined by (2.18) satisfies the curvature assumption. By glancing through the literature, we list a few interesting examples that satisfy the assumptions of Lemma 2.2.6.

(a) *'Non-degenerate' surfaces in* V_1: Let $\Sigma \subseteq V_1$ be a surface that cannot be contained in any proper subspace of V_1. It is clear that Σ satisfies the assumptions of the lemma. A concrete example in this class is the co-dimension two sphere in the Heisenberg group defined by (2.18).

(b) *Tilted 'non-degenerate' surfaces of* V_1: Consider the map

$$\mathfrak{M} : V_1 \to \mathfrak{g}, \qquad X \mapsto \mathfrak{M}(X) := (X, \Lambda(X)),$$

where Λ is an $(n - d_1) \times d_1$ matrix. Let $\Sigma \subseteq V_1$ be a surface that cannot be contained in any proper subspace of V_1 and let Σ^Λ be the image of Σ under this map. If μ denotes the surface measure on Σ, we use μ^Λ to denote the push-forward of μ under $\mathfrak{M}$. Averaging or maximal operators associated to such measures in the Heisenberg group, along with natural extensions to the class of Métivier groups, have been considered in several works [38, 1, 43].

(c) *Analytic curves*: any curve in $\gamma \in \mathfrak{g}$ come under the purview of the lemma, provided $\gamma_1 := \Pi_1(\gamma)$ generates V_1. In particular, the lemma applies moment curves in $\mathfrak{g}$. Furthermore, it applies to a curve $\gamma_{\alpha,\beta} : [0, 1] \to \mathbb{H}^1$ defined by

$$\gamma(s) := (s, s^2, \alpha s^3 + \beta s^\sigma)$$

for any $\alpha, \beta \in \mathbb{R}$ and $\sigma \geq 0$. Averaging operators associated to such curves are studied previously in [46].

Both Lemma 2.2.5 and Lemma 2.2.6 follow from a testing condition provided by [42].

Proposition 2.2.7 (Corollary 2.3, [42]). *Let S be a connected analytic submanifold of a homogeneous group G. If S generates the group G, then any C_c^∞-density σ on S satisfies* (C).

Here we say a set $S \subseteq G$ *generates* G if $G = \langle S \rangle$ where

$$\langle S \rangle := \{s_1 \cdot \ldots \cdot s_N : s_1, \ldots, s_N \in S \cup \tilde{S}\}$$

for $\tilde{S} := \{s^{-1} : s \in S\}$. Ricci–Stein [42] work with the ostensibly weaker condition that $G = \mathrm{clos}(\langle S \rangle)$; however, in all cases we consider (that is, for S a connected analytic submanifold) these conditions turn out to be equivalent.[12]

We remark that the result in [42, Corollary 2.3] is somewhat more general. There, the authors consider a family of connected analytic submanifolds S_j for $1 \leq j \leq N$ such that the iterated product set $S_1 \cdot \ldots \cdot S_N$ contains a non-trivial open subset of G. For each j, one fixes σ_j a smooth density on S_j and considers the convolution product $\sigma_1 * \cdots * \sigma_N$. To recover Proposition 2.2.7, we choose the S_j to alternate between S and the reflection $\tilde{S}$ and, accordingly, the σ_j to alternate between σ and $\tilde{\sigma}$. Using [42, Proposition 1.1], the hypothesis that S generates G implies the existence of some N such that $S_1 \cdot \ldots \cdot S_N$ contains a non-trivial open subset of G, and so [42, Corollary 2.3] applies.

In view of Proposition 2.2.7, it is clear that Lemma 2.2.5 follow from the following result.

Lemma 2.2.8. *Let G be a graded homogeneous group with $\dim G \geq 2$. Suppose Ω is an open convex domain in $\mathfrak{g}$ with an analytic boundary $\Sigma := \partial\Omega$. It follows that $\exp(\Sigma)$ generates G.*

Similarly, Proposition 2.2.7 immediately reduces the proof of Lemma 2.2.6 to the following lemma.

Lemma 2.2.9. *Let G be a stratified m-step homogeneous group and let $\Sigma \subseteq \mathfrak{g}$. $\exp(\Sigma)$ generates the group G if and only if $\Pi_1(\Sigma)$ generates V_1 (in terms of vector addition).*

We will now proceed to prove the lemmas one by one.

Proof of Lemma 2.2.8. Suppose $\mathfrak{g}$ denotes the corresponding graded Lie algebra of the form $\mathfrak{g} = \bigoplus_{i=1}^m V_i$ with $\dim V_j = d_j$ and

$$[V_i, V_j] \subseteq V_{i+j} \qquad \text{for all } 1 \leq i \leq m-1,\ 1 \leq j \leq m-i.$$

For $X \in \mathfrak{g}$, define the commutator function

$$\Phi_X : \mathfrak{g} \to \mathfrak{g}, \qquad Y \mapsto \Phi_X(Y) := [X, Y].$$

Note that Φ_X is linear, and as $\mathfrak{g}$ is graded, the range of Φ_X is contained in $\bigoplus_{j=2}^m V_j$. First, we prove an elementary property about the kernel of this map.

[12] See the cited reference for further details.

Claim A. *For any $X \in \mathfrak{g}$, the nullity of Φ_X is at least 2.*

Proof: We divide the proof into two parts; first, let

$$\dim(V_1) \geq 2.$$

As the range of Φ_X is a subspace of $\bigoplus_{j=2}^m V_j$, the rank–nullity theorem can be applied here to deduce that $\dim \ker(\Phi_X) \geq 2$. This concludes the argument for the first case.

In the second case, $\dim(V_1) = 1$. As the Lie bracket is skew-symmetric, it follows that $[V_1, V_1] = \{0\}$. Therefore, the range of Φ_X is contained in $\bigoplus_{j=3}^m V_j$. By applying the rank-nullity theorem again,

$$\dim \ker(\Phi_X) \geq \dim(V_1) + \dim(V_2) \geq 2,$$

where the final inequality is justified by the assumption that the dimension of the group is at least two. The proof ends here. ■

Assume, without loss of generality, that

$$B(0,1) := \{x \in G : |x|_G \leq 1\} \subseteq \exp(\Omega).$$

Choose $x \in B(0,1)$ and let $X = \exp^{-1}(x)$. Let $\mathbb{S}_\mathfrak{g}$ denote the unit sphere in $\mathfrak{g}$ with respect to the euclidean norm. By Claim A, we can choose a subspace H_X of $\ker \Phi_X$ with dimension 2. The set $D_x := \mathbb{S}_\mathfrak{g} \cap H_X$ is connected with dimension one in $\mathfrak{g}$. By the convexity assumption on Ω, we can construct unique continuous mappings $t, s : D_x \to \mathbb{R}_+$ such that for any $W \in D_x$, we have

$$X + t(W)W,\ X - s(W)W \in \Sigma.$$

Consider the continuous function $F : D_x \to \mathbb{R}$ mapping

$$W \mapsto F(W) := t(W) - s(W).$$

Clearly, $t(-W) = s(W)$ and $s(-W) = t(W)$, so that $F(-W) = -F(W)$. Applying the intermediate value theorem, we deduce the existence of $W_x \in D_x$ such that

$$F(W_x) = 0 \qquad \text{or, equivalently,} \qquad t(W_x) = s(W_x).$$

Let $w_x := \exp(W_x)$. Since $W_x \in H_X$, by the Baker–Campbell–Hausdorff formula and the bilinearity of the Lie bracket,

$$x \cdot_G x = x \cdot_G (t(W_x)w_x) \cdot_G (t(W_x)w_x)^{-1}x \in \exp(\Sigma) \cdot_G \exp(\Sigma).$$

As x was chosen arbitrarily, we can conclude that $\exp(\Sigma) \cdot_G \exp(\Sigma)$ contains an open ball in G. This implies that $\exp(\Sigma)$ generates G. □

Before we begin the proof of Lemma 2.2.9, we must define a few auxiliary mappings and verify some of their basic properties. In what follows, we use capital letters to represent elements of the Lie algebra and small letters to represent elements of the Lie group, related by the exponential map (for instance, $X =$

$\exp^{-1}(x))$.

For $1 \le \ell \le m$, let $\Phi_\ell \colon \mathfrak{g}^\ell \to \mathfrak{g}$ be the mapping

$$\Phi_\ell \colon (X_1, \dots, X_\ell) \mapsto [X_1, [X_2, \dots [X_{\ell-1}, X_\ell] \dots]]$$

which takes a nested sequence of commutators of ℓ elements. If $\mathfrak{g}$ is stratified, the restricted mapping

$$\Phi_\ell|_{V_1^\ell} \colon V_1^\ell \mapsto V_\ell$$

is a surjection. Furthermore, for any $X_1, \dots, X_\ell \in \mathfrak{g}$, we have

$$\Pi_i \circ \Phi_\ell(X_1, \dots, X_\ell) = 0 \qquad \text{for } 1 \le i \le \ell - 1 \tag{2.19}$$

and

$$\Pi_\ell \circ \Phi_\ell(X_1, \dots, X_\ell) = \Phi_\ell\big(\Pi_1(X_1), \dots, \Pi_1(X_\ell)\big). \tag{2.20}$$

Let $\tilde{\Phi}_\ell \colon G^\ell \to G$ be a mapping defined iteratively; $\tilde{\Phi}_\ell(x_1) := x_1$ and

$$\tilde{\Phi}_\ell(x_1, \dots, x_\ell) := x_1 \cdot_G \tilde{\Phi}_{\ell-1}(x_2, \dots, x_{\ell-1}) \cdot_G x_1^{-1} \cdot_G \tilde{\Phi}_{\ell-1}(x_2, \dots, x_{\ell-1})^{-1}.$$

Claim B. *For any $\ell \ge 1$, the mappings Φ_ℓ and $\tilde{\Phi}_\ell$ are related by the identity*

$$\exp^{-1}(\tilde{\Phi}_\ell(x_1, \dots, x_\ell)) = \Phi_\ell(X_1, \dots, X_\ell) + E_{\ell+1}(x_1, \dots, x_\ell) \tag{2.21}$$

for all $x_1, \dots, x_\ell \in G$, where $E_{\ell+1}(x_1, \dots, x_\ell) \in \mathfrak{g}$ is a linear combination of commutators of order at least $\ell + 1$ formed by the vector fields $X_1, \dots, X_\ell$.

Proof: The proof is based on induction. When $\ell = 1$, the statement is obvious. Let $\mathbf{x} := (x_1, \dots, x_\ell)$. Assuming (2.21) for $\ell = k - 1$, take

$$\tilde{\Phi}_k(\mathbf{x}) = x_1 \cdot_G \tilde{\Phi}_{k-1}(\mathbf{x}') \cdot_G x_1^{-1} \cdot_G \tilde{\Phi}_{k-1}(\mathbf{x}')^{-1}$$

where $\mathbf{x}' := (x_2, \dots, x_k)$. By Baker–Campbell–Hausdorff formula,

$$\exp^{-1} \tilde{\Phi}_k(\mathbf{x}) = [X_1, \exp^{-1} \tilde{\Phi}_{k-1}(\mathbf{x}')] + E_3(x_1, \tilde{\Phi}_{k-1}(\mathbf{x}')),$$

where E_3 is a linear combination of Lie bracket products of order at least 3. As $\exp^{-1} \tilde{\Phi}_{k-1}(\mathbf{x}')$ is a linear combination of commutators of order at least $k-1$ by the induction assumption, the minimum order of a commutator in $E_3(x_1, \tilde{\Phi}_{k-1}(\mathbf{x}'))$ is $k + 1$. On the other hand, by the induction assumption again,

$$[X_1, \exp^{-1} \tilde{\Phi}_{k-1}(\mathbf{x}')] = [X_1, \Phi_{k-1}(X_2, \dots, X_\ell)] + [X_1, E_k(\mathbf{x}')].$$

Therefore,

$$\begin{aligned} \exp^{-1} \tilde{\Phi}_k(\mathbf{x}) &= [X_1, \Phi_{k-1}(X_2, \dots, X_\ell)] + E_{k+1}(x_1, \mathbf{x}') \\ &= \Phi_k(X_1, \dots, X_\ell) + E_{k+1}(\mathbf{x}) \end{aligned}$$

where E_{k+1} is a linear combination of commutators of order at least $k+1$. Thus, we can close the induction and complete the proof of (2.21). ∎

For $1 \leq \ell \leq m$, define the group mappings

$$\pi_\ell : G \to G, \qquad \pi_\ell : x \mapsto \exp \circ \Pi_\ell \circ \exp^{-1}(x).$$

Observe that the projection π_1 satisfies

$$\exp^{-1} \circ \pi_1(x_1 \cdot_G x_2) = \exp^{-1} \circ \pi_1(x_1) + \exp^{-1} \circ \pi_1(x_2) \tag{2.22}$$

for all $x_1, x_2 \in G$.

As a consequence of (2.21), (2.19) and (2.20), we record the identities

$$\pi_i \circ \tilde{\Phi}_\ell(x_1, \dots, x_\ell) = e \qquad \text{for } 1 \leq i \leq \ell - 1 \tag{2.23}$$

and

$$\pi_\ell \circ \tilde{\Phi}_\ell(x_1, \dots, x_\ell) = \exp \circ \Phi_\ell\big(\exp^{-1} \circ \pi_1(x_1), \dots, \exp^{-1} \circ \pi_1(x_\ell)\big). \tag{2.24}$$

Proof of Lemma 2.2.9. It suffices to look at the 'if' part, as the other implication trivially holds.

For any $x \in \exp(V_1)$ there exists some $\mathbf{g}(x) \in G$ such that

$$\mathbf{g}(x) \in \langle \exp \Sigma \rangle \quad \text{and} \quad \pi_1(\mathbf{g}(x)) = x. \tag{2.25}$$

Indeed, from our hypothesis on Σ there exists a finite sequence of elements

$$S_1, \dots, S_k \in \Sigma \qquad \text{such that } X = \Pi_1(S_1) + \cdots + \Pi_1(S_k),$$

where $X = \exp^{-1}(x)$. If we define

$$\mathbf{g}(x) := s_1 \cdot_G \cdots \cdot_G s_k$$

for $s_i = \exp(S_i)$, then clearly $\mathbf{g}(x) \in \langle \exp \Sigma \rangle$ whilst, by (2.22), we also have

$$\exp^{-1} \pi_1(\mathbf{g}(x)) = \exp^{-1} \pi_1(s_1) + \cdots + \exp^{-1} \pi_1(s_k) = X.$$

We therefore obtain a function $\mathbf{g} \colon \exp V_1 \to G$. This function is not uniquely defined, but for our purposes it suffices to work with *some* $\mathbf{g}$ satisfying (2.25).[13]

We now use induction to prove that

$$W_\ell := \left\{ \exp\left(\sum_{i=\ell}^{m} Y_i \right) : Y_i \in V_i \text{ for } \ell \leq i \leq m \right\} \subseteq \langle \exp \Sigma \rangle \tag{2.26}$$

for all $1 \leq \ell \leq m+1$, where W_{m+1} is interpreted as $\{0\}$. For $\ell = 1$, the above statement becomes $G = \langle \exp \Sigma \rangle$, which is precisely the content of the lemma.

We take $\ell = m+1$ as the base of the induction, in which case (2.26) is trivial. Let $2 \leq \ell \leq m+1$ and suppose, by way of induction hypothesis, that $W_\ell \subseteq \langle \exp \Sigma \rangle$. To complete the argument, it suffices to show $W_{\ell-1} \subseteq \langle \exp \Sigma \rangle$.

[13] We could easily stipulate additional conditions to ensure $\mathbf{g}$ is uniquely defined and thus avoid arbitrary choices in the definition.

Fix $Y_i \in V_i$ for $\ell - 1 \leq i \leq m$. Since G is stratified, $X_1, \dots, X_{\ell-1} \in V_1$ can be chosen such that

$$\Phi_{\ell-1}(X_1, \dots, X_{\ell-1}) = Y_{\ell-1}.$$

It follows from (2.24) and (2.25) that

$$\begin{aligned}\pi_{\ell-1}\big(\tilde{\Phi}_{\ell-1}(\mathbf{g}(x_1), \dots, \mathbf{g}(x_{\ell-1}))\big) &= \exp\circ\Phi_{\ell-1}\big(\exp^{-1}\circ\pi_1(\mathbf{g}(x_1)), \dots, \exp^{-1}\circ\pi_1(\mathbf{g}(x_{\ell-1}))\big)\\ &= \Phi_{\ell-1}(X_1, \dots, X_{\ell-1}).\end{aligned}$$

On the other hand, from (2.23) we have

$$\pi_i\big(\tilde{\Phi}_{\ell-1}(\mathbf{g}(x_1), \dots, \mathbf{g}(x_{\ell-1}))\big) = e \qquad \text{for } 1 \leq i \leq \ell - 2.$$

Consequently, we may write

$$z := \tilde{\Phi}_{\ell-1}(\mathbf{g}(x_1), \dots, \mathbf{g}(x_{\ell-1})) = \exp\Big(Y_{\ell-1} + \sum_{i=\ell}^{m} Z_i\Big)$$

for some $Z_i \in V_i$ for $\ell \leq i \leq m$. In view of (2.25), we have $z \in \langle \exp \Sigma \rangle$.

By the basic properties of the group operation, there exist polynomial mappings

$$P_i \colon V_{\ell+1} \times \cdots \times V_{i-1} \mapsto V_i$$

such that if $u = \exp(U) := \exp(\sum_{i=\ell}^{m} U_i) \in G$ with $U_i \in V_i$, then

$$\pi_i(u \cdot_G z) = \exp(U_i + P_i(U)) \qquad \text{for } \ell \leq i \leq m$$

where P_i depends only $U_\ell, \dots, U_{i-1}$ and z. In particular, $P_i(U)$ is independent of $U_i, \dots, U_m$ and so the polynomial P_ℓ is constant as a function of U (in fact, $P_\ell(U) = Z_\ell$). On the other hand, the remaining projections are given by

$$\pi_i(u \cdot_G z) = e \quad \text{for } 1 \leq i \leq \ell - 2 \quad \text{and} \quad \pi_{\ell-1}(u \cdot_G z) = \exp Y_{\ell-1}$$

Let $u = \exp(\sum_{i=\ell}^{m} U_i) \in W_\ell$ so that, by our induction hypothesis, $u \cdot_G z \in \langle \exp \Sigma \rangle$. In view of the dependence properties of the P_i, it is possible to inductively choose the u_i so that

$$\begin{aligned} Y_\ell &= U_\ell + P_\ell(U),\\ Y_{\ell+1} &= U_{\ell+1} + P_{\ell+1}(U),\\ &\vdots\\ Y_m &= U_m + P_m(U).\end{aligned}$$

Thus, from the preceding observations,

$$u \cdot_G z = \exp(\sum_{i=\ell-1}^{m} Y_i).$$

Since the right-hand side is an arbitrary element of $W_{\ell-1}$, we conclude that

$$W_{\ell-1} \subseteq \langle \exp \Sigma \rangle.$$

This closes the induction and completes the proof. □

Let us now turn to the proof of Theorem 2.2.4.

2.3 Proof of Theorem 2.2.4

The proof is motivated by an argument presented by Ricci and Stein [42] (with the key ideas tracing back to [11]). In their work [42], the authors investigate the boundedness of classical singular operators and maximal functions along submanifolds, so it is reasonable to expect their strategy to be applicable to our problem as well. The outline of the proof can be summarised as follows.

Through initial reductions, we observe that the maximal estimates can be derived from L^p estimates of *frequency-localised* singular integral operators (such reductions are standard; see, for instance, [16]). To estimate these operators, we employ the Calderón–Zygmund theory adapted to homogeneous groups. In particular, the key steps involve obtaining L^2 estimates for one endpoint and verifying a Hörmander-type condition for the other endpoint near L^1. To prove the L^2 estimate, we utilise a method from [42] in which iterated applications of T^*T allow us to make use of the curvature assumption. On the other hand, the Hörmander-type condition is verified by applying a mean-value theorem for homogeneous groups.

Before commencing the proof, we must survey some aspects of the Littlewood–Paley theory adapted to the setting of homogeneous groups.

2.3.1 A glimpse into Littlewood–Paley theory

First, we discuss the Littlewood–Paley decomposition of a function in this setting. As in the Euclidean case, the frequency localisation of a function is defined as the convolution with a smooth function satisfying large moment conditions. Consider a function $\psi \in C_c^\infty(G)$ which is *mean zero* in the sense that

$$\int_G \psi = 0$$

and, for some large $M \in \mathbb{N}$, satisfies the *higher moment conditions*

$$\int_G \psi(y) y^\alpha \, \mathrm{d}y = 0, \qquad \text{for any } \alpha \in \mathbb{N}_0^n \text{ and } 0 < |\alpha| \leq M. \tag{2.27}$$

Here we recall the standard notation $y^\alpha = \prod_{i=1}^n y_i^{\alpha_i}$ whenever $y = (y_1, \dots, y_n) \in G$ and $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$. To familiarise ourselves with this function, let us try to make sense of (2.27) when $G = \mathbb{R}^n$. In this case, (2.27) implies that

$$\partial_\xi^\alpha \hat{\psi}(0) = 0 \qquad \text{for all } \alpha \in \mathbb{N}_0^n, \ |\alpha| \leq M.$$

By the Taylor series approximation, we deduce that $\hat{\psi}(\xi) = O(|\xi|^M)$ for all $|\xi| \leq 1$. On the other hand, when $|\xi| > 1$, we use the fact that $\hat{\psi} \in \mathcal{S}(\mathbb{R}^n)$ to deduce that $\hat{\psi}(\xi) = O_N(|\xi|^{-N})$ for any $N \in \mathbb{N}_0$. Therefore, it can be concluded that the essential support of $\hat{\psi}$ is contained inside the unit strip

$$D_1 := \{\xi \in \mathbb{R}^n : |\xi| \approx 1\}$$

and as a consequence, convolution with the function ψ amounts to essentially localising the frequency support to D_1.

The following result dictates the degree of orthogonality between two convolution operators whose kernels are different dilates for a function satisfying (2.27). Although the proposition will not be used in the upcoming arguments, it verifies the essential orthogonality properties one expects to be manifested between these 'frequency projection operators'.

Proposition 2.3.1 ([13], Lemma 4.2)**.** *Suppose $\psi, \phi \in C_c^\infty(G)$ satisfy (2.27). For $t, s \in (0, \infty)$, define the operators Q_t, Q'_s by*

$$Q_t f := f * \psi_t \quad \textit{and} \quad Q'_s f := f * \phi_s \qquad \textit{for } f \in L^1(G).$$

For any constant $M \geq 1$, we have

$$\|(Q_t)^* Q'_s\|_{L^2 \to L^2} + \|Q_t (Q'_s)^*\|_{L^2 \to L^2} \lesssim_{\phi,\psi,M} (\min\{t/s, s/t\})^{(M+1)},$$

where Q^* denotes the adjoint of Q.

Proof. We will restrict our attention to estimating $Q_t^* Q'_s$ when $s \geq t$. The argument for the other cases is similar.

By Young's inequality, it suffices to obtain an L^1 norm bound on the kernel of $Q_t^* Q'_s$. The kernel can be expressed as $\phi_s * \tilde{\psi}_t$, where $\tilde{\psi}$ denote the reflection of ψ as given by Definition 2.1.23. Using (2.9), it suffices to estimate the L^1 norm of $\phi * \tilde{\psi}_{t/s}$. Since $s \geq t$, the function $\tilde{\psi}_{t/s}$ and therefore $\phi * \tilde{\psi}_{t/s}$ is supported inside an ball of bounded radius. In view of this, it suffices to estimate the L^∞ norm of $\phi * \tilde{\psi}_{t/s}$.

Temporarily fixing $x \in G$, we define the function ϕ^x by

$$\phi^x(y) := \phi(xy^{-1}) \qquad \text{for } y \in G.$$

Suppose p^x denotes the Taylor polynomial of degree M associated with ϕ^x. Now, using (2.27), we have

$$\begin{aligned} \phi * \tilde{\psi}_{t/s}(x) &= \int_G \phi(xy^{-1}) \tilde{\psi}_{t/s}(y) \mathrm{d}y \\ &= \int_G (\phi^x(y) - p^x(y))\, \tilde{\psi}_{t/s}(y^{-1}) \mathrm{d}y. \end{aligned}$$

Now, $\tilde{\psi}_{t/s}$ is supported inside a Euclidean ball $B_{\text{eucl}}(0, \tilde{C}(t/s))$ for some $\tilde{C} \geq 1$. Therefore,

$$|\phi^x(y) - p^x(y)| \lesssim |y|^{M+1} \lesssim (t/s)^{M+1} \qquad \text{for any } y \in \operatorname{supp} \tilde{\psi}_{t/s},$$

where $|\cdot|$ denotes the Euclidean norm. Thus,

$$\left|\phi * \tilde{\psi}_{t/s}(x)\right| = \left|\int_G (\phi^x(y) - p^x(y))\, \tilde{\psi}_{t/s}(y)\mathrm{d}y\right| \lesssim (t/s)^{(M+1)}\, \|\tilde{\psi}\|_{L^1(G)},$$

completing the proof. □

Finally, we show the existence of a Littlewood–Paley decomposition for any smooth function on G, which is a key tool in our argument.

Proposition 2.3.2. *Let G be a homogeneous group. There exists $\psi \in C_c^\infty(G)$ of mean zero satisfying* (2.27) *such that*

$$f = \sum_{k\in\mathbb{Z}} f * \psi_{2^k} \qquad \text{for all } f \in C_c^1(G), \tag{2.28}$$

where the convergence holds uniformly.

This result is well-known (for instance, it can be deduced from [13, Proposition 3.4]); however, for completeness, we present the straightforward proof.

Proposition 2.3.2 follows from a basic result on L^2 approximate identities. Consider $\phi \in C_c^\infty(G)$ satisfying

$$\int_G \phi = 1. \tag{2.29}$$

Given any $f \in C_c^1(G)$, it follows that

$$\|f * \phi_t - f\|_{L^\infty(G)} \to 0 \quad \text{as } t \to 0_+ \qquad \text{and} \qquad \|f * \phi_t\|_{L^\infty(G)} \to 0 \quad \text{as } t \to \infty; \tag{2.30}$$

the standard proofs are left to the reader.

Proof (Proposition 2.3.2). Suppose $\phi \in C_c^\infty(G)$ satisfies (2.29) as above and also (2.27). By (2.30), we have

$$f = \lim_{K\to\infty} f * \phi_{2^{-K}} - f * \phi_{2^K}$$

and so, by the fundamental theorem of calculus,

$$f = -\lim_{K\to\infty} \int_{2^{-K}}^{2^K} f * \Big(\frac{\partial \phi_t}{\partial t}\Big)\,\mathrm{d}t = -\sum_{k\in\mathbb{Z}} f * \Big(\int_{2^k}^{2^{k+1}} \frac{\partial \phi_t}{\partial t}\,\mathrm{d}t\Big), \tag{2.31}$$

where in each case the convergence holds uniformly over G. A computation shows

$$\frac{\partial \phi_t}{\partial t}(x) = -t^{-1} h_t(x) \qquad \text{for some } h \in C_c^\infty(G).$$

Moreover, if we define

$$\psi(x) := \int_1^2 h_t(x)\,\frac{\mathrm{d}t}{t}, \tag{2.32}$$

then, by a simple change of variables,

$$-\int_{2^k}^{2^{k+1}} \frac{\partial \phi_t}{\partial t}(x)\,\mathrm{d}t = \int_{2^k}^{2^{k+1}} h_t(x)\,\frac{\mathrm{d}t}{t} = \psi_{2^k}(x). \tag{2.33}$$

Combining (2.31) and (2.33), we see that (2.28) holds for ψ as defined in (2.32).

It remains to show ψ is of mean zero and satisfies the moment condition (2.27). Clearly, it suffices to show the same properties hold for the function h. However, since for all $\alpha \in \mathbb{N}_0^n$ we have

$$\int_G h(x)x^\alpha\,\mathrm{d}x = -t\frac{\partial}{\partial t}\int_G \phi_t(x)x^\alpha\,\mathrm{d}x\Big|_{t=1},$$

the mean zero property for h is an immediate consequence of (2.29) whilst the moment condition (2.27) for h is inherited directly from ϕ. □

The proof of Theorem 2.2.4 is divided into three parts; in the first part, we linearise the maximal function, and in the last two parts, we obtain two endpoint estimates for the linearised operator.

2.3.2 ℓ^2- domination and randomisation

Let us begin by choosing a non-negative function $\phi \in C_c^\infty(G)$ such that

$$\int_G \phi = \sigma(G).$$

Introducing the measure $\nu := \sigma - \phi$, we note that ν is a compactly supported Borel measure on G such that $\|\nu\|_{\mathcal{M}(G)} < \infty$ and

$$\nu(G) = 0. \tag{2.34}$$

Using this new measure, we define the maximal function.

$$\widetilde{\mathcal{M}}f := \sup_{k\in\mathbb{Z}} |f * \nu_{2^k}| \qquad \text{for } \ f \in C_c(G)\ .$$

Using Proposition 2.3.2, we can choose a function $\psi \in C_c^\infty(G)$ which satisfies (2.27). Now, consider the maximal function

$$\widetilde{\mathcal{M}}_l f := \sup_{k\in\mathbb{Z}} |f * \psi_{2^{k+l}} * \nu_{2^k}| \qquad \text{for } \ f \in C_c(G) \text{ and } l \in \mathbb{Z}\ .$$

By standard reductions, we deduce L^2 estimates for $\mathcal{M}$ from those for $\widetilde{\mathcal{M}}_l$ (details to follow). After replacing the supremum in k with an ℓ^2 sum, we consider the square function

$$S_l f := \left[\sum_{k\in\mathbb{Z}} |f * \psi_{2^{k+l}} * \nu_{2^k}|^2\right]^{\frac{1}{2}} \qquad \text{for } f \in C_c(G) \text{ and } l \in \mathbb{Z}.$$

The proof of Theorem 2.2.4 is reduced to proving norm estimates for S_l with a decay in the l parameter.

Proposition 2.3.3. *For any $p \in (1, 2]$, there exists $\epsilon(p) > 0$ such that*

$$\|S_l\|_{L^p(G)\to L^p(G)} \lesssim_p 2^{-\epsilon(p)|l|}, \tag{2.35}$$

where the implicit constant is independent of l.

Proposition 2.3.3 $\implies$ Theorem 2.2.4. Observing the trivial case $p = \infty$ in Theorem 2.2.4, it suffices to estimate the maximal function when p lies in the interval $(1, 2]$. The L^p estimates for the maximal function when $p \in (2, \infty)$ then follows by Marcinkiewicz's interpolation theorem [52, §4].

Fix $p \in (1, 2]$ and $f \in C_c^\infty(G)$. Using the definition of $\widetilde{\mathcal{M}}$, the inequality

$$\mathcal{M}f(x) \lesssim_\phi \widetilde{\mathcal{M}}f(x) + \mathcal{M}_\phi f(x) \qquad \text{for } x \in G$$

directly follows, where

$$\mathcal{M}_\phi f := \sup_{k\in\mathbb{Z}} |f * \phi_k|$$

is a variant of the Hardy-Littlewood maximal function on G. It is well-known that $\mathcal{M}_\phi$ is bounded on $L^p(G)$ for all $p > 1$ (see, for instance, [18, §2] for proof). Therefore, it suffices to estimate the maximal function $\widetilde{\mathcal{M}}f$. Using Proposition 2.3.2, we may write

$$f * \nu_{2^k} = \sum_{l\in\mathbb{Z}} f * \psi_{2^{k+l}} * \nu_{2^k}.$$

By triangle inequality and ℓ^2 domination,

$$\widetilde{\mathcal{M}}f(x) \leq \sum_{l\in\mathbb{Z}} \widetilde{\mathcal{M}}_l f(x) \leq \sum_{l\in\mathbb{Z}} S_l f(x) \qquad \text{for } x \in G.$$

Now, Proposition 2.3.3 yields

$$\|\widetilde{\mathcal{M}}f\|_{L^p(G)} \leq \sum_{l\in\mathbb{Z}} \|S_l f\|_{L^p(G)} \lesssim_p \sum_{l\in\mathbb{Z}} 2^{-\epsilon(p)|l|}\|f\|_{L^p(G)} \lesssim \|f\|_{L^p(G)},$$

concluding the proof of Theorem 2.2.4. □

In order to prove Proposition 2.3.3, we randomise/linearise the operator S_l. This step is carried out using the Rademacher system. Let $\vec{r} = (r_k)_{k\in\mathbb{Z}}$ be a sequence of the Rademacher functions. In particular, the functions $r_k : [0, 1] \to \{-1, 1\}$ form a collection of independent and identically distributed random variables.

For $t \in [0, 1]$, consider the function

$$T_{l,\vec{r}(t)} f := \sum_{k\in\mathbb{Z}} r_k(t) T_l^k f$$

where

$$T_l^k f := f * \psi_{2^{k+l}} * \nu_{2^k}.$$

Now, Khintchine's inequality[14] gives

$$|S_l f(x)|^p \approx_p \mathbb{E}\big[|T_{l,\vec{r}} f(x)|^p\big] \qquad \text{for any } x \in G. \tag{2.36}$$

In view of this, we restrict our attention to studying norm estimates of the linearised operator $T_{l,\vec{r}(t)}$ for each $t \in [0,1]$. We prove two endpoint norm estimates to this end.

Lemma 2.3.4. *There exists $\rho_1, \rho_2 > 0$ such that*

$$\|(T_l^k)^* T_l^j\|_{L^2(G)\to L^2(G)} + \|T_l^k (T_l^j)^*\|_{L^2(G)\to L^2(G)} \lesssim \min\{2^{-\rho_1|l|}, 2^{-\rho_2|j-k|}\} \tag{2.37}$$

holds for all $j, k, l \in \mathbb{Z}$.

Lemma 2.3.5. *Let $\vec{r} \in \{-1,1\}^{\mathbb{Z}}$. For any $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that*

$$\|T_{l,\vec{r}}\|_{L^1(G)\to L^{1,\infty}(G)} \le C_\varepsilon 2^{\varepsilon|l|}, \tag{2.38}$$

where the implicit constant is independent of $\vec{r}$ and $l \in \mathbb{Z}$.

To obtain favourable L^2 estimates for $T_{l,\vec{r}}$, we must combine Lemma 2.3.4 with the Cotlar–Stein almost orthogonality lemma. For completeness, we include the result here.[15]

Lemma 2.3.6 (Almost orthogonality lemma)**.** *Let $T^i : L^2(G) \to L^2(G)$ be a collection of operators satisfying the almost-orthogonality condition:*

$$\|T^j (T^k)^*\|_{L^2(G)\to L^2(G)} + \|(T^j)^* T^k\|_{L^2(G)\to L^2(G)} \le [\gamma(j-k)]^2 \qquad \text{for any } j,k \in \mathbb{Z},$$

such that $\sum_{i\in\mathbb{Z}} |\gamma(i)| = A < \infty$. Then, $\sum_{k\in\mathbb{Z}} T^k$ converges in the strong topology and

$$\|\sum_{k\in\mathbb{Z}} T^k\|_{L^2(G)\to L^2(G)} \le A.$$

Proposition 2.3.3 can be proved by combining these lemmas.

Proof of Proposition 2.3.3. Fix $l \in \mathbb{Z}$. From Lemma 2.3.4, we see that Lemma 2.3.6 can be applied for $T^i := T_l^i$ and $\gamma(i) := C 2^{-(\rho_1|l|+\rho_2(i))/2}$ for a uniform constant C. In particular, there exists $\rho > 0$ such that

$$\|T_{l,\vec{r}}\|_{L^2(G)\to L^2(G)} \lesssim 2^{-\rho_1|l|/2}, \tag{2.39}$$

where the implicit constant is independent of $\vec{r}$ and $l \in \mathbb{Z}$.

Interpolating between (2.39) and the weak L^1 estimate from Lemma 2.3.5, we further deduce that for any $p \in (1,2]$, there exists $\epsilon(p) > 0$ such that

$$\|T_{l,\vec{r}}\|_{L^p(G)\to L^p(G)} \lesssim_p 2^{-\epsilon(p)|l|}. \tag{2.40}$$

By duality, the range of exponents allowed in (2.40) extends to $(1,\infty)$. By (2.36), (2.40) implies (2.35) for any p in this range, completing the proof. $\square$

[14] see, for reference, [21, Chapter 12]

[15] a proof of the lemma can be found in [30].

Proofs for Lemma 2.3.4 and Lemma 2.3.5 are given in the next two subsections.

2.3.3 Obtaining the L^2 bounds (proof of Lemma 2.3.4)

We begin by introducing further notation:

- Let $\mathcal{C}_1$ denote the sub-collection of $\mathcal{M}(G)$, consisting of measures μ such that $\mu(G) = 0$.
- Let $\mathcal{C}_2$ be a sub-collection of $\mathcal{M}(G)$ consisting of compactly supported Borel measures on G of the form $\varrho_1 + \cdots + \varrho_q$ for some $q \in \mathbb{N}$, where each ϱ_i lies in $\mathcal{M}(G)$ and satisfies (C).

The main result in this subsection is the following.

Lemma 2.3.7. *Let $\mu, \vartheta \in \mathcal{C}_1 \cap \mathcal{C}_2$. Then, there exists $\rho > 0$ such that*

$$\|\mathcal{A}[\mu_s * \vartheta_t]\|_{L^2(G)\to L^2(G)} \lesssim_{\mu,\vartheta} \big(\min\{s/t,\ t/s\}\big)^{\rho},$$

for any $s, t \in (0, \infty)$.

Assuming the lemma, we proceed to the proofs of the L^2 bounds.

Lemma 2.3.7 $\implies$ Lemma 2.3.4. By unwinding definitions, we can write

$$(T_l^k)^* T_l^j = \mathcal{A}[\psi_{2^{j+l}} * \nu_{2^j} * \tilde{\nu}_{2^k} * \tilde{\psi}_{2^{k+l}}] \tag{2.41}$$

and

$$(T_l^k)(T_l^j)^* = \mathcal{A}[\tilde{\nu}_{2^j} * \tilde{\psi}_{2^{j+l}} * \psi_{2^{k+l}} * \nu_{2^k}]. \tag{2.42}$$

By (2.27) and (2.34), we see that $\psi, \nu, \tilde{\psi}, \tilde{\nu} \in \mathcal{C}_1$. Furthermore, as σ, ϕ and ψ satisfy (C) and $\nu = \sigma - \phi$, we may further deduce that

$$\psi, \nu, \tilde{\psi}, \tilde{\nu} \in \mathcal{C}_1 \cap \mathcal{C}_2.$$

Therefore, Lemma 2.3.7 can be applied to an averaging operator whose kernel is a convolution product of scaled copies of any two among these four measures. Considering the form of the operators in (2.41) and (2.42), we apply the lemma only to the cases listed in the following table. By cases, we mean the different measures and scales we feed into Lemma 2.3.7. In the final column of the table, we list the values which, after applying the lemma, come up in the operator norm estimates for the corresponding averaging operator.

Operator	μ	ϑ	s	t	$\min\{\frac{s}{t}, \frac{t}{s}\}$
$\mathcal{A}[\psi_{2^{j+l}} * \nu_{2^j}]$	ψ	ν	2^{j+l}	2^j	$2^{-\|l\|}$
$\mathcal{A}[\nu_{2^j} * \tilde{\nu}_{2^k}]$	ν	$\tilde{\nu}$	2^j	2^k	$2^{-\|k-j\|}$
$\mathcal{A}[\tilde{\nu}_{2^j} * \tilde{\psi}_{2^{j+l}}]$	$\tilde{\nu}$	$\tilde{\psi}$	2^j	2^{j+l}	$2^{-\|l\|}$
$\mathcal{A}[\tilde{\psi}_{2^{j+l}} * \psi_{2^{k+l}}]$	$\tilde{\psi}$	ψ	2^{j+l}	2^{k+l}	$2^{-\|k-j\|}$

By applying Lemma 2.3.7, we can find $\rho_1, \rho_2 > 0$ such that

$$\|\mathcal{A}[\psi_{2^{j+l}} * \nu_{2^j}]\|_{L^2(G)\to L^2(G)} + \|\mathcal{A}[\tilde{\nu}_{2^j} * \tilde{\psi}_{2^{j+l}}]\|_{L^2(G)\to L^2(G)} \lesssim_{\nu,\psi} 2^{-\rho_1|l|}, \tag{2.43}$$

and

$$\|\mathcal{A}[\nu_{2^j} * \tilde{\nu}_{2^k}]\|_{L^2(G)\to L^2(G)} + \|\mathcal{A}[\tilde{\psi}_{2^{j+l}} * \psi_{2^{k+l}}]\|_{L^2(G)\to L^2(G)} \lesssim_{\nu,\psi} 2^{-\rho_2|k-j|}. \tag{2.44}$$

To deduce the statement of Lemma 2.3.4 from the above estimates, we use Young's inequality. In particular, by repeated applications of (2.13), we obtain the inequalities

$$\begin{aligned}&\|(T_l^k)^* T_l^j\|_{L^2(G)\to L^2(G)} \\ &\qquad \lesssim_{\nu,\psi} \min\left\{\|\mathcal{A}[\psi_{2^{j+l}} * \nu_{2^j}]\|_{L^2(G)\to L^2(G)}, \|\mathcal{A}[\nu_{2^j} * \tilde{\nu}_{2^k}]\|_{L^2(G)\to L^2(G)}\right\}\end{aligned}$$

and

$$\begin{aligned}&\|(T_l^k)(T_l^j)^*\|_{L^2(G)\to L^2(G)} \\ &\qquad \lesssim_{\nu,\psi} \min\left\{\|\mathcal{A}[\tilde{\nu}_{2^j} * \tilde{\psi}_{2^{j+l}}]\|_{L^2(G)\to L^2(G)}, \|\mathcal{A}[\tilde{\psi}_{2^{j+l}} * \psi_{2^{k+l}}]\|_{L^2(G)\to L^2(G)}\right\}.\end{aligned}$$

Combining these with (2.43) and (2.44), we obtain (2.37), completing the proof of Lemma 2.3.4. □

Proof of Lemma 2.3.7. The proof begins with two elementary reductions.

First, notice that the adjoint of $\mathcal{A}[\mu_s * \vartheta_t]$ is given by the operator $\mathcal{A}[\tilde{\vartheta}_t * \tilde{\mu}_s]$. From the Hilbert space theory, we recall that the L^2 operator norm of an operator and its adjoint are the same. Therefore, it suffices to prove the existence of a $\rho > 0$ such that

$$\|\mathcal{A}[\mu_s * \vartheta_t]\|_{L^2(G)\to L^2(G)} \lesssim_{\mu,\vartheta} (s/t)^\rho,$$

whenever $s \le t$, $\mu \in \mathcal{C}_1$ and $\vartheta \in \mathcal{C}_2$.

Second, if $\vartheta \in \mathcal{C}_2$, then we have a decomposition $\vartheta = \varrho_1 + \cdots + \varrho_q$ with each $\varrho_i \in \mathcal{M}(G)$ satisfying (C). This induces a decomposition for the operator $\mathcal{A} := \mathcal{A}[\mu_s * \vartheta_t]$, given by

$$\mathcal{A} = \sum_{i=1}^{q} \mathcal{A}[\mu_s * (\varrho_i)_t].$$

To estimate $\mathcal{A}$, it suffices to estimate each summand individually. Because of this, we may further assume that ϑ satisfies (C).

Let $u := \frac{s}{t} \le 1$. After recalling (2.15), we consider the measure

$$\varpi_{(n)} := \mu_u * \vartheta^{(n)} \qquad \text{for } n \in \mathbb{N}_0.$$

By a simple rescaling argument, we see that

$$\|\mathcal{A}\|_{L^2(G)\to L^2(G)} = \|\mathcal{A}[\varpi_{(0)}]\|_{L^2(G)\to L^2(G)}. \tag{2.45}$$

We claim that for any $n \in \mathbb{N}_0$, the following inequality holds:

$$\|\mathcal{A}[\varpi_{(n)}]\|_{L^2(G)\to L^2(G)} \leq \|\mu\|_{\mathcal{M}(G)}^{1/2}\|\vartheta\|_{\mathcal{M}(G)}^{n/2}\|\mathcal{A}[\varpi_{(n+1)}]\|_{L^2(G)\to L^2(G)}^{1/2}. \tag{2.46}$$

To prove the claim, we use the Hilbert space identity

$$\|\mathcal{A}[\varpi_{(n)}]\|_{L^2(G)\to L^2(G)} = \|(\mathcal{A}[\varpi_{(n)}])^* \circ \mathcal{A}[\varpi_{(n)}]\|_{L^2(G)\to L^2(G)}^{1/2}.$$

However,

$$(\mathcal{A}[\varpi_{(n)}])^* \circ \mathcal{A}[\varpi_{(n)}] = \mathcal{A}[\varpi_{(n)} * \mathcal{R}(\vartheta^{(n)}) * \widetilde{\mu}_u],$$

where $\mathcal{R}$ maps a measure ϱ to $\tilde{\varrho}$, its reflection. In view of (2.15), we deduce that

$$\begin{aligned}\varpi_{(n)} * \mathcal{R}(\vartheta^{(n)}) * \tilde{\mu}_u &= \mu_u * \vartheta^{(n)} * \mathcal{R}(\vartheta^{(n)}) * \tilde{\mu}_u \\ &= \mu_u * \vartheta^{(n+1)} * \mathcal{R}(\vartheta^{(n-1)}) * \tilde{\mu}_u \\ &= \varpi_{(n+1)} * \mathcal{R}(\vartheta^{(n-1)}) * \tilde{\mu}_u.\end{aligned}$$

Thus, by repeated applications of (2.13), we obtain the claim (2.46) (note that the $(n-1)$th convolution product as defined by (2.15), involves the convolution of n measures).

By combining (2.46) and (2.45), we deduce that

$$\|\mathcal{A}\|_{L^2(G)\to L^2(G)} \leq C_{\mu,\vartheta}(n)\|\mathcal{A}[\varpi_{(n)}]\|_{L^2(G)\to L^2(G)}^{\frac{1}{2^n}} \quad \text{for } n \in \mathbb{N}_0, \tag{2.47}$$

where

$$C_{\mu,\vartheta}(n) = \begin{cases} 1 & \text{if } n = 0, \\ \|\mu\|_{\mathcal{M}(G)}^{\sum_{k=1}^n \frac{1}{2^k}} \|\vartheta\|_{\mathcal{M}(G)}^{\sum_{k=1}^n \frac{k-1}{2^k}} & \text{if } n \geq 1. \end{cases}$$

Recall the assumption that ϑ satisfies (C). Thus, there exists $N \in \mathbb{N}_0$ such that $\vartheta^{(n)}$ is absolutely continuous with density function h which lies in L^1_δ for some $\delta > 0$. We fix $n = N$, and in view of (2.47), it suffices to estimate the operator norm of $\mathcal{A}[\varpi_{(N)}]$.

By Young's inequality, the operator bound for $\mathcal{A}[\varpi_{(N)}]$ follows from an estimate on $\|\varpi_{(N)}\|_{L^1(G)}$. Since $\mu \in \mathcal{C}_1$, it follows that

$$\begin{aligned}\|\varpi_{(N)}\|_{L^1(G)} &= \|\mu_u * h\|_{L^1(G)} \\ &= \int_G \left|\int_G \left[h(y^{-1}x) - h(x)\right] \mathrm{d}\mu_u(y)\right| \mathrm{d}x \\ &\leq \int_G \left(\int_G \left|h(y^{-1}x) - h(x)\right| \mathrm{d}x\right) \mathrm{d}|\mu_u|(y) \\ &\leq \|h\|_{L^1_\delta(G)} \int_G |\delta_u y|^\delta \mathrm{d}|\mu|(y) \\ &\leq u^{\lambda_{\min}\delta}\|h\|_{L^1_\delta(G)} \int_G |y|^\delta \mathrm{d}|\mu|(y),\end{aligned}$$

where $\lambda_{\min} := \min_{1\leq i\leq n} \lambda_i$, the smallest eigenvalue of the dilation matrix A. Combining this deduction with (2.47) and Young's inequality, we obtain the desired estimate

$$\|\mathcal{A}\|_{L^2(G)\to L^2(G)} \le C_{\mu,\vartheta}(N)\|\varpi_{(N)}\|_{L^1(G)}^{\frac{1}{2^N}} \lesssim C'_{\mu,\vartheta}\left(\frac{s}{t}\right)^{\rho},$$

where $\rho = 2^{-N}\lambda_{\min}\delta$ and $C'_{\mu,\vartheta} = C_{\mu,\vartheta}(N)\|h\|_{L^1_\delta(G)}^{\frac{1}{2^N}}$. This concludes the proof. $\square$

2.3.4 Proof of Lemma 2.3.5

Fix a vector $\vec{r} = (r_k)_{k\in\mathbb{Z}} \in \{-1,1\}^{\mathbb{Z}}$ and let K_l denote the kernel of $T_{l,\vec{r}}$. We may write

$$K_l = \sum_{k\in\mathbb{Z}} r_k K_l^k, \tag{2.48}$$

where $K_l^k := (\psi_{2^l} * \nu)_{2^k}$ is the kernel of T_l^k.

Using Calderón–Zygmund theory adapted to the homogeneous setting[16], we see that to prove (2.38), we must verify the Hörmander-condition

$$\sup_{y\in G}\int_{|x|_G\ge C_0|y|_G} |K_l(y^{-1}x) - K_l(x)|\mathrm{d}x \lesssim_\varepsilon 2^{\varepsilon|l|} \qquad \text{for any} \quad \varepsilon > 0, \tag{2.49}$$

where C_0 is some fixed constant. In our attempts to prove the above inequality, we will fix the value of C_0 depending on C, the constant appearing in item (c) of Definition 2.1.16. In particular, we define $C_0 := 2C$.

In view of (2.48), it is clear that (2.49) follows from the estimate

$$\sup_{y\in G}\sum_{k\in\mathbb{Z}} I_l^k(y) \lesssim_\varepsilon 2^{\varepsilon|l|} \qquad \text{for any} \quad \varepsilon > 0, \tag{2.50}$$

where

$$I_l^k(y) := \int_{|x|_G\ge C_0|y|_G} |K_l^k(y^{-1}x) - K_l^k(x)|\mathrm{d}x.$$

Our first step towards (2.50) is to identify the region where I_l^k vanishes. By unwinding the definition of K_l^k, we write

$$I_l^k(y) = \int_{|x|_G\ge C_0|y|_G} |(\psi_{2^l} * \nu)_{2^k}(y^{-1}x) - (\psi_{2^l} * \nu)_{2^k}(x)|\mathrm{d}x \tag{2.51}$$

$$= \int_{|x|_G\ge C_0 2^{-k}|y|_G} |(\psi_{2^l} * \nu)((\delta_{2^{-k}}y)^{-1}x) - (\psi_{2^l} * \nu)(x)|\mathrm{d}x. \tag{2.52}$$

As both ν and ψ are assumed to be compactly supported, we can find a constant $C_1 > 1$, depending only on ν and ψ, such that the support of $\psi_{2^l} * \nu$ is contained inside the ball $B(0, C_1 2^{\max\{l,0\}})$. Set $C_2 := 2C_1$. We claim that

$$I_l^k(y) = 0 \qquad \text{whenever } |y|_G \ge C_2 2^{k+\max\{l,0\}}\ . \tag{2.53}$$

[16] see, for instance, [54, Theorem 3 on p. 19] for a reference.

To see this, we fix a $y \in G$ such that $|y|_G \geq C_2 2^{k+\max\{l,0\}}$. In view of (2.51), it suffices to check that

$$|(\delta_{2^{-k}} y)^{-1} x|_G, |x|_G \geq C_1 2^{\max\{l,0\}} \qquad \text{whenever } |x|_G \geq C_0 |\delta_{2^{-k}} y|_G.$$

The lowerbound on $|x|_G$ is clear, as $C_0 \geq 1$ and $|x|_G \geq C_0 |\delta_{2^{-k}} y|_G \geq C_0 C_2 2^{\max\{l,0\}}$. On the other hand, by item (c) of Definition 2.1.16, we may deduce the chain of inequalities

$$C_0 |\delta_{2^{-k}} y|_G \leq |x|_G \leq C(|(\delta_{2^{-k}} y)^{-1} x|_G + |(\delta_{2^{-k}} y)^{-1}|_G).$$

Therefore,

$$C_1 C 2^{\max\{l,0\}} < C_2 C 2^{\max\{l,0\}} = C_2 (C_0 - C) 2^{\max\{l,0\}} \leq C |(\delta_{2^{-k}} y)^{-1} x|_G.$$

As a consequence, the required lower bound on $|(\delta_{2^{-k}} y)^{-1} x|_G$ is achieved, completing the proof of the claim (2.53).

Now, we can attempt to estimate the non-zero values of I_l^k. Applying Young's inequality shows that the kernel K_l^k is L^1 normalised. Consequently, we have the uniform estimates for I_l^k, given by

$$I_l^k(y) \lesssim 1 \qquad \text{for any } k, l \in \mathbb{Z}. \tag{2.54}$$

However, to sum different I_l^k's and obtain (2.50), one must prove non-trivial decay estimates for I_l^k in the k variable. The decay is proved using a variant of the mean value theorem in G, which we have expressed as a lemma below.

Before stating the lemma, we recall a few features of our setup from Remark 2.1.9. In particular, the set $\{X_i\}_{1 \leq i \leq n}$ represents a basis for $\mathfrak{g}$, consisting of the eigenvectors of the dilation matrix A. Furthermore, the eigenvalue associated to X_i, denoted by λ_i, is positive for $1 \leq i \leq n$. We also recall that for any $X \in \mathfrak{g}$, the symbol $\tilde{X}$ represents the (unique) right invariant vector field that coincides with X in the tangent space at the identity.

The lemma we present below is a variant of [18, Theorem 1.33].

Lemma 2.3.8. *Let $g \in C^1(G)$. For any $z \in G$, we have*

$$\int_G |g(zx) - g(x)| \mathrm{d}x \lesssim \sum_{j=1}^{n} |z|_G^{\lambda_j} \|\tilde{X}_j g\|_{L^1(G)}. \tag{2.55}$$

The proof of the lemma is postponed till the end of this subsection. Assuming the lemma, we resume the proof of (2.50).

Fix $k, l \in \mathbb{Z}$. After dropping the restriction over the region of integration in

(2.52), we see that

$$\begin{aligned}I_l^k(y) &\leq \int_G |(\psi_{2^l} * \nu)((\delta_{2^{-k}} y)^{-1} x) - (\psi_{2^l} * \nu)(x)| \mathrm{d}x \\ &= \int_G \left| \int_G \psi_{2^l}((\delta_{2^{-k}} y)^{-1} x z^{-1}) - \psi_{2^l}(x z^{-1}) \mathrm{d}\nu(z) \right| \mathrm{d}x \\ &\leq \|\nu\|_{\mathcal{M}(G)} \int_G |\psi_{2^l}((\delta_{2^{-k}} y)^{-1} x) - \psi_{2^l}(x)| \mathrm{d}x.\end{aligned}$$

At this point, apply Lemma 2.3.8 for $g = \psi_{2^l}$ and $z = (\delta_{2^{-k}} y)^{-1}$; since $\tilde{X}_j(\psi_{2^l}) = 2^{-l\lambda_j}(\tilde{X}_j\psi)_{2^l}$, we deduce that

$$I_l^k(y) \lesssim_\nu \sum_{j=1}^n (2^{-k}|y|_G)^{\lambda_j} \|\tilde{X}_j(\psi_{2^l})\|_{L^1(G)} = \sum_{j=1}^n (2^{-(k+l)}|y|_G)^{\lambda_j} \|\tilde{X}_j\psi\|_{L^1(G)}.$$

Combining this with the trivial estimate (2.54), we write

$$I_l^k(y) \lesssim_{\psi,\nu} \min\{1, \sum_{j=1}^n (2^{-(k+l)}|y|_G)^{\lambda_j}\}. \tag{2.56}$$

Now, the estimation of the sum in (2.50) is carried out in two steps:

Case 1 ($2^l \geq 1$): By (2.53) and (2.56), we deduce that

$$\begin{aligned}\sum_{k\in\mathbb{Z}} I_l^k(y) &= \sum_{k\in\mathbb{Z}:|y|_G\leq C_2 2^{(l+k)}} I_l^k(y) \\ &\lesssim_{\psi,\nu} \sum_{j=1}^n (2^{-l}|y|_G)^{\lambda_j} \Big(\sum_{k\in\mathbb{Z}:C_2^{-1}2^{-l}|y|_G\leq 2^k} 2^{-k\lambda_j} \Big) \\ &\lesssim_{\psi,\nu} 1.\end{aligned}$$

Case 2 ($2^l < 1$): Using (2.53) and (2.56) again,

$$\begin{aligned}\sum_{k\in\mathbb{Z}} I_l^k(y) &= \sum_{k\in\mathbb{Z}:\ |y|_G\leq C_2 2^k} I_l^k(y) \\ &= \sum_{k\in\mathbb{Z}:\ 2^{-k}|y|_G\leq C_2 2^l} I_l^k(y) + \sum_{C_2 2^l < 2^{-k}|y|_G \leq C_2} I_l^k(y) \\ &\lesssim_{\psi,\nu} \sum_{j=1}^n (2^{-l}|y|_G)^{\lambda_j} \Big(\sum_{k\in\mathbb{Z}:\ C_2^{-1}2^{-l}|y|_G\leq 2^k} 2^{-k\lambda_j} \Big) + \sum_{C_1 2^l\leq 2^{-k}|y|_G\leq C_1} 1 \\ &\lesssim_{\psi,\nu,\lambda} 1 + |l|.\end{aligned}$$

Combining both cases, we obtain (2.50), completing the proof of Lemma 2.3.5.

Proof of Lemma 2.3.8. The proof has two parts. First, assume that z can be written as $\exp(t\tilde{X}_j)$ for some $1 \leq j \leq n$. In this case, we aim for a sharper estimate than (2.55) by replacing the sum on the left-hand side with its jth summand.

By the Fundamental Theorem of Calculus and the identity (2.3), we write

$$\begin{aligned} g(zx) - g(x) &= \int_0^t \left(\frac{\mathrm{d}}{\mathrm{d}s'}\right)_{s'=s} [g(\exp(s'\tilde{X}_j)x)]\mathrm{d}s \\ &= \int_0^t \tilde{X}_j g(\exp(s\tilde{X}_j)x)\mathrm{d}s. \end{aligned}$$

Therefore,

$$\begin{aligned} \int_G |g(zx) - g(x)|\mathrm{d}x &\leq \int_0^t \int_G |\tilde{X}_j g(\exp(s\tilde{X}_j)x)|\mathrm{d}x\mathrm{d}s \\ &= |t| \|\tilde{X}_j g\|_{L^1(G)}. \end{aligned} \tag{2.57}$$

Since $|z|_G = |\exp(t\tilde{X}_j)|_G = |t|^{\frac{1}{\lambda_j}} |\exp(\tilde{X}_j)|_G$, we can conclude that

$$\int_G |g(zx) - g(x)|\mathrm{d}x \leq C_1 |z|_G{}^{\lambda_j} \|\tilde{X}_j g\|_{L^1(G)}$$

where $C_1 := \max_{1\leq j\leq n} |\exp(\tilde{X}_j)|^{-\lambda_j}$.

In the second part of the proof, z can be chosen to be any arbitrary member of G. Here, we may express z as z_n, where

$$z_j := \exp(t_j\tilde{X}_j) \cdot_G \cdots \cdot_G \exp(t_1\tilde{X}_1) \qquad \text{for } 1 \leq j \leq n \text{ and } t_1, \ldots, t_n \in \mathbb{R},$$

and $g(zx) - g(x)$ as the telescopic sum

$$\sum_{j=1}^{n} (g(z_j x) - g(z_{j-1}x)),$$

where $z_0 x = x$. By repeated applications of the first case, specifically the inequality (2.57), we have

$$\int_G |g(zx) - g(x)|\mathrm{d}x \leq \sum_{j=1}^{n} |t_j| \|\tilde{X}_j g\|_{L^1(G)} \lesssim \sum_{j=1}^{n} |z|_G^{\lambda_j} \|\tilde{X}_j g\|_{L^1(G)}, \tag{2.58}$$

as required. The final step in (2.58) follows from the inequality

$$|z|_G^{-1} \sum_{j=1}^{n} |t_j|^{\frac{1}{\lambda_j}} \leq \sup\left\{\sum_{j=1}^{n} |s_j|^{\frac{1}{\lambda_j}} : |\exp(s_n\tilde{X}_n) \cdot_G \cdots \cdot_G \exp(s_1\tilde{X}_1)|_G = 1\right\} \lesssim 1$$

which becomes obvious when we notice that $|\delta_{|z|_G^{-1}}(z)|_G = 1$ and that we can write

$$\delta_{|z|_G^{-1}}(z) = \exp(t_n|z|_G^{-\lambda_n}\tilde{X}_n) \cdot_G \cdots \cdot_G \exp(t_1|z|_G^{-\lambda_1}\tilde{X}_1),$$

using the basic properties of the group dilations. □

Chapter 3

Maximal estimates and Local smoothing problems

The content of this chapter is broadly based on the works of Mockenhaupt–Seeger–Sogge [37], Guth–Wang–Zhang [26], and Beltran–Guo–Hickman–Seeger [3]. The discussion is included to introduce and contextualise the result in the next chapter.

3.1 Introduction: Spherical and circular maximal theorems

In this chapter, we discuss some of the well-known methods in Harmonic analysis used to study maximal operators. To illustrate these methods, we use the spherical maximal operator in $\mathbb{R}^d$ as a model operator, with a particular emphasis on the case $d = 2$. The first part of the chapter (till the end of §3.4) demonstrates known results in this area, whereas the second part (§3.5) addresses the question of extending these results to higher dimensions.

For $d \in \mathbb{N}$, let σ^d denote the surface measure on the sphere $\mathbb{S}^{d-1} \subset \mathbb{R}^d$, induced by Lebesgue measure on $\mathbb{R}^{d+1}$, normalised to have total mass one. For each $t > 0$, let σ_t^d denote the measure on the dilated sphere $t\mathbb{S}^{d-1}$ defined by the action

$$\langle \sigma_t^d, f \rangle := \int_{\mathbb{S}^{d-1}} f(tx) \mathrm{d}\sigma^d(x) \qquad \text{for any } f \in L^1_{\mathrm{loc}}(\mathbb{R}^d).$$

The Stein spherical maximal function $\mathcal{M}_{\mathbf{s}}^d$ (which is named after E. Stein, who formulated the maximal problem) is defined by the formula

$$\mathcal{M}_{\mathbf{s}}^d g(x) := \sup_{0<t<\infty} |\sigma_t^d * g(x)| \qquad \text{for } g \in \mathcal{S}(\mathbb{R}^d) \text{ and } x \in \mathbb{R}^d. \tag{3.1}$$

As in the previous chapter, we are interested in the problem of L^p boundedness for the maximal function $\mathcal{M}_{\mathbf{s}}^d$. Naturally, the first step towards formulating this problem is to find the necessary conditions on the range of Lebesgue spaces where the $\mathcal{M}_{\mathbf{s}}^d$ can be extended as a bounded operator. To this end, we define

$$g(x) := \frac{1}{|x|^{d-1}\log(|x|^{-1})}\chi_{B(0,1)}(x) \qquad \text{for } x \in \mathbb{R}^d.$$

Fix $d \geq 2$. It is easy to see that $|\mathcal{M}_{\mathbf{s}}^d g(x)| = \infty$ for all $x \in \mathbb{R}^d$. However,

$$\|g\|_{L^p(\mathbb{R}^d)}^2 = \int_{B(0,1)} \frac{1}{(|x|^{d-1}\log(|x|^{-1}))^p}\mathrm{d}x < \infty \iff p \leq d/(d-1).$$

In other words, when $p \leq d/(d-1)$ and $d \geq 2$, the maximal function $\mathcal{M}_{\mathbf{s}}^d$ cannot be L^p bounded (Note that we reach the same conclusion even when we replace $(0,\infty)$ is replaced with $[1,2]$ as the index set for the supremum in (3.1)). When $d = 1$, it is easy to see that $\mathcal{M}_{\mathbf{s}}^d$ cannot be L^p bounded when $p < \infty$, and L^∞ boundedness always holds in all dimensions. Therefore, the right maximal problem to be framed is the following: can $\mathcal{M}_{\mathbf{s}}^d$ be extended as an L^p bounded operator when $d \geq 2$ and $p > d/(d-1)$?

The first result in this direction came from E. Stein, who proved sharp spherical maximal estimates in all dimensions except for the plane.

Theorem 3.1.1 (Stein [53]). *Let $d \geq 3$. For $p > d/(d-1)$, there exists $C_{p,d} > 0$ such that*

$$\|\mathcal{M}_{\boldsymbol{s}}^d\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \leq C_{p,d}.$$

Stein's argument fell short of obtaining any non-trivial maximal estimates when $d = 2$. We will shortly describe the difficulties in the argument that prevented it from being extended to the planar setting. It took almost another ten years before a different approach to the maximal problem was developed by J. Bourgain, and he used it to tackle the L^p boundedness for the maximal function in the plane (which is often called the circular maximal function). To distinguish these two arguments, we use the notation $\mathcal{M}_{\mathrm{b}}$ to represent the maximal operator $\mathcal{M}_{\mathrm{s}}^d$ when $d = 2$. Sharp L^p estimates for $\mathcal{M}_{\mathrm{b}}$ were established by Bourgain.

Theorem 3.1.2 (Bourgain [6]). *For $p \in (2,\infty]$, there exists $C_p > 0$ such that*

$$\|\mathcal{M}_{\mathrm{b}}\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \leq C_p. \tag{3.2}$$

The failure of L^2 boundedness for $\mathcal{M}_{\mathrm{b}}$ is crucial in understanding how Bourgain's proof of the circular maximal theorem (Theorem 3.1.2) contrasts with Stein's proof of the spherical maximal theorem (Theorem 3.1.1). We plan to first sketch the proof of a simplified version of Theorem 3.1.1 and highlight why the argument fails to provide any maximal estimate when $d = 2$. Afterward, we will describe a proof strategy that overcomes these issues and prove Theorem 3.1.2, following an argument of [37].

Proof of a weaker version of Theorem 3.1.1

Fix $d \geq 3$ and drop the superscript in σ_t^d. Simplifying the setup, we restrict our attention to norm-estimating the local maximal operator $\mathcal{M}_{\mathrm{s,loc}}$ defined by

$$\mathcal{M}_{\mathrm{s,loc}} g := \sup_{t \in [1,2]} |\mathcal{A}_t g| \qquad \text{where} \qquad \mathcal{A}_t g := \sigma_t * g$$

for any $g \in \mathcal{S}(\mathbb{R}^d)$. Furthermore, we plan to address the question of only the L^2 boundedness for $\mathcal{M}_{\mathrm{s,loc}}$. Although we are working in such a simplified setup, the core ideas in the full proof of Theorem 3.1.1 are not lost[1].

We begin by introducing a dyadic frequency localisation to the maximal operator. Suppose $\eta, \beta \in C_c^\infty(\mathbb{R})$ are the classical Littlewood–Paley functions such that

$$\operatorname{supp} \eta \subseteq \{r \in \mathbb{R} : |r| \leq 2\}, \qquad \operatorname{supp} \beta \subseteq \{r \in \mathbb{R} : 1/2 \leq |r| \leq 2\} \tag{3.3}$$

and

$$\eta(r) + \sum_{j \in \mathbb{N}} \beta(r/2^j) = 1 \qquad \text{for all } r \in \mathbb{R}. \tag{3.4}$$

Let $g \in \mathcal{S}(\mathbb{R}^d)$. Using Plancherel's theorem,

$$\mathcal{A}_t g(x) = \sum_{j=0}^{\infty} \mathcal{A}_t^j g(x)$$

where

$$\mathcal{A}_t^j(g)(x) := (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle x, \xi \rangle} \hat{\sigma}(t\xi) \beta_{(j)}(t\xi) \hat{g}(\xi) \mathrm{d}\xi \qquad \text{for } x \in \mathbb{R}^d, \tag{3.5}$$

and

$$\beta_{(j)}(\xi) := \begin{cases} \eta(|\xi|) & \text{if } j = 0, \\ \beta(2^{-j}|\xi|) & \text{if } j \in \mathbb{N}, \end{cases}$$

for $\xi \in \mathbb{R}^d$. By the triangle inequality,

$$\mathcal{M}_{\mathrm{s,loc}} g(x) \leq \sum_{j=0}^{\infty} \mathcal{M}_{\mathrm{s,loc}}^j g(x) \qquad \text{where} \qquad \mathcal{M}_{\mathrm{s,loc}}^j g(x) := \sup_{t \in [1,2]} |\mathcal{A}_t^j g(x)|$$

for any $x \in \mathbb{R}^d$. Thus, the L^2 norm of $\mathcal{M}_{\mathrm{s,loc}} g$ can be dominated by the sum of L^2 norms of $\mathcal{M}_{\mathrm{s,loc}}^j g$ over all $j \in \mathbb{N}$. Therefore, to show that $\mathcal{M}_{\mathrm{loc}}$ is strong-type $(2,2)$, it suffices to show that the L^2 operator norm of $\mathcal{M}_{\mathrm{s,loc}}^j$ has a decay in the j parameter.

[1] A full proof of the theorem can be found in [54, Page 510].

Proposition 3.1.3. *Let $d \geq 3$. For all $j \in \mathbb{N}_0$, we have*

$$\|\mathcal{M}^j_{\mathrm{s,loc}}\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim 2^{-j(d-2)/2}. \tag{3.6}$$

The proof of (3.6) contains two key ingredients. The first is a result that will be referred to as the Sobolev embedding lemma hereafter.

Lemma 3.1.4. *Let $F \in C^1(\mathbb{R})$ and $p > 1$. Then it follows that*

$$\sup_{1\leq t\leq 2} |F(t)|^p \leq |F(1)|^p + p\left(\int_1^2 |F(t)|^p \mathrm{d}t\right)^{(p-1)/p} \cdot \left(\int_1^2 |F'(t)|^p \mathrm{d}t\right)^{1/p}$$

Proof. The proof is a simple application of the Fundamental Theorem of Calculus. For $s \in [1, 2]$, we begin with the identity

$$|F(s)|^p = |F(1)|^p + p\int_1^s |F(t)|^{p-1} F'(t)\mathrm{d}t.$$

Applying Hölder's inequality now completes the proof. □

The Sobolev embedding lemma allows us to dominate the maximal function by Fourier integral operators. This brings us to the second ingredient in the proof, which is the estimation of the Fourier decay of the spherical measure. By standard stationary phase methods, it is well known that

$$|\hat{\sigma}(\xi)| + |\nabla_\xi \hat{\sigma}(\xi)| \lesssim (1 + |\xi|)^{-(d-1)/2} \qquad \text{for } \xi \in \mathbb{R}^d. \tag{3.7}$$

The proof of (3.7) can be found in [54, Chapter 7]. With these two ingredients, we can now prove the above proposition.

Proof of Proposition 3.1.3. Since $\sigma * \check{\beta}_{(0)} \in \mathcal{S}(\mathbb{R}^d)$, we can dominate $\mathcal{M}^0_{\mathrm{s,loc}}$ by a variant of the Hardy–Littlewood operator[2], which is bounded for any $p > 1$. Thus, it suffices to look at the case of $j \neq 0$.

Fix $j \in \mathbb{N}$ and $g \in \mathcal{S}(\mathbb{R}^d)$. To highlight that t plays the role of a variable rather than a parameter, define

$$\mathcal{A}^j g(x,t) := \mathcal{A}^j_t g(x) \qquad \text{for } (x,t) \in \mathbb{R}^{d+1}. \tag{3.8}$$

Temporarily fixing $x \in \mathbb{R}^d$, set

$$F(t) := \mathcal{A}^j g(x,t) \qquad \text{for } t \in [1,2].$$

Combining Fubini's theorem with Lemma 3.1.4, we see that (3.6) follows from the norm estimates

$$\|\mathcal{A}^j_1 g\|_{L^2(\mathbb{R}^d)} \lesssim 2^{-j(d-2)/2}\|g\|_{L^2(\mathbb{R}^d)} \tag{3.9}$$

[2]see [54, Chapter 2, §2.1] for more details.

and

$$\|\mathcal{A}^j g\|_{L^2(\mathbb{R}^d\times[1,2])}^{1/2}\|\partial_t\mathcal{A}^j g\|_{L^2(\mathbb{R}^d\times[1,2])}^{1/2} \lesssim 2^{-j(d-2)/2}\|g\|_{L^2(\mathbb{R}^d)}. \tag{3.10}$$

By combining Plancherel's theorem with (3.7), we deduce that

$$\|\mathcal{A}_1^j\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \lesssim 2^{-j(d-1)/2},$$

which is a stronger estimate than (3.9). On the other hand, by combining (3.7) with Fubini's theorem and Plancherel's theorem,

$$\|\mathcal{A}^j\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d\times[1,2])} \lesssim 2^{-j(d-1)/2}. \tag{3.11}$$

Because of (3.5), $\partial_t\mathcal{A}^j g$ can be expressed as a Fourier multiplier operator with the multiplier function $\partial_t[\hat{\sigma}(t\xi)\beta_{(j)}(t\xi)]$. By Leibnitz rule and (3.7), we can deduce that this function is uniformly bounded by $2^j\cdot 2^{-j(d-1)/2}$. Therefore,

$$\|\partial_t\mathcal{A}^j\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d\times[1,2])} \lesssim 2^{-j(d-3)/2}.$$

Combining the above inequality with (3.11), we obtain (3.10), and that concludes the proof. □

When $d = 2$, Stein's argument does not give any decay in the L^2 operator norms of $\mathcal{M}^j_{\mathrm{s,loc}}$. Since we have earlier verified that $\mathcal{M}_{\mathrm{b}}$ (and its localised version) is unbounded on L^2, this was already expected. We must significantly modify these arguments to get decay in their L^p norms for $p > 2$, a requirement to prove the circular maximal theorem. First, note that in the process of obtaining (3.10), we have not used the extra t-integration in the L^2 norms in the left side of the inequality. To improve the argument, we must utilise the averaging in the t variable and gain a non-trivial factor in the operator norm of $\mathcal{A}^j$.

Bourgain's original argument in [6] for the circular maximal theorem relied on a combination of Fourier analytic tools and geometric considerations. An alternate proof of (3.2) was given in [44] purely using geometric tools. For the thesis, however, we wish to focus on a third line of argument that relates maximal estimates with the so-called *local smoothing* estimates for the wave equation. This approach was laid out through the results of Sogge [51] and Mockenhaupt–Seeger–Sogge [37]. Following their strategy, a wide variety of maximal functions were estimated in subsequent years by several other authors (see, for instance, [3], [9], [29]).

3.2 Proof of the circular maximal theorem

As in the proof for the spherical maximal theorem, we introduce frequency localised maximal operators.

Let $g \in \mathcal{S}(\mathbb{R}^2)$. For $j \in \mathbb{N}_0$ and $t \in \mathbb{R}$, recall the definition of the frequency localised averaging operator $\mathcal{A}_t^j g$ from (3.5), where we set $d = 2$ and σ now denotes the normalised surface measure on the unit circle. Introducing the maximal function

$$\mathcal{M}_{\mathrm{b}}^{j} g(x) := \sup_{t\in(0,\infty)} |\mathcal{A}_t^j g(x)| \qquad \text{for } x \in \mathbb{R}^2 \text{ and } g \in \mathcal{S}(\mathbb{R}^2),$$

we see that

$$\|\mathcal{M}_{\mathrm{b}}\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \leq \sum_{j=0}^{\infty} \|\mathcal{M}_{\mathrm{b}}^j\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \qquad \text{for } p \geq 1.$$

Therefore, the proof of Theorem 3.1.2 reduces to the following result.

Proposition 3.2.1. *For $p > 2$, there exists $\varepsilon(p) > 0$ such that*

$$\|\mathcal{M}_{\mathrm{b}}^j\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \lesssim 2^{-j\varepsilon(p)} \qquad \textit{for any } j \in \mathbb{N}_0. \tag{3.12}$$

Compared with $\mathcal{M}_{\mathrm{s,loc}}^j$, the maximal function $\mathcal{M}_{\mathrm{b}}^j$ is a *global* maximal function. In particular, in its definition, the supremum of averages is taken over the entire positive real line. Because of this, one cannot directly apply Lemma 3.1.4 to relate the maximal operator with a Fourier integral operator. To use the lemma, we must consider the following result, which reduces global maximal estimates to local maximal estimates. The lemma and its proof have been reproduced from [4, Lemma 3.4], but it can be traced back to at least Bourgain's result [6].

Lemma 3.2.2 (Lemma 3.4, [4])**.** *Consider a function $m \in L^\infty(\mathbb{R}^d)$ such that*

$$\operatorname{supp} m \subseteq \{\xi \in \mathbb{R}^d : \lambda_0/2 \leq |\xi| \leq 2\lambda_0\}$$

for a fixed $\lambda_0 \in \mathbb{R} \setminus \{0\}$. For $t \in \mathbb{R}$, consider the Fourier multiplier $\tilde{\mathcal{A}}_t$ defined as

$$\tilde{\mathcal{A}}_t g(x) := [m(t \cdot \tfrac{1}{i}\partial_x) g](x), \qquad \textit{whenever } g \in \mathcal{S}(\mathbb{R}^d).$$

Let $p \geq 2$. Suppose there exists $M_p(\lambda_0) > 0$ such that the local maximal estimate

$$\|\sup_{t\in[1,2]} |\tilde{\mathcal{A}}_t|\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \leq M_p(\lambda_0)$$

holds. Then, there exists $C_p > 0$ such that

$$\|\sup_{t\in(0,\infty)} |\tilde{\mathcal{A}}_t|\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \leq C_p M_p(\lambda_0).$$

Proof. Fix $g \in \mathcal{S}(\mathbb{R}^d)$. Let $k_0 \in \mathbb{Z}$ be chosen such that $\lambda_0 \in [2^{k_0}, 2^{k_0+1})$. Recall the definition of $\beta \in C_c^\infty(\mathbb{R}^d)$ from (3.3). For $l \in \mathbb{Z}$, let P_l denote the classical smooth Littlewood–Paley projection operator defined by $P_l := \beta\big(2^{-l}|\tfrac{1}{i}\partial_x|\big)$, so that

$$\sum_{l\in\mathbb{Z}} P_l g(x) = g(x) \qquad \text{for a.e. } x \in \mathbb{R}^d.$$

To be used later in the proof, we recall the Littlewood–Paley square function estimate[3]

$$\|(\sum_{l\in\mathbb{Z}}|P_l g|^2)^{1/2}\|_{L^p(\mathbb{R}^d)} \lesssim_p \|g\|_{L^p(\mathbb{R}^d)} \qquad \text{for any } p\in(1,\infty). \tag{3.13}$$

In view of the support properties of m, we deduce that

$$\tilde{\mathcal{A}}_t g = \sum_{l=-\infty}^{\infty} \tilde{\mathcal{A}}_t P_l g = \sum_{|l-(k+k_0)|\leq 10} \tilde{\mathcal{A}}_t P_l g, \qquad \text{if } t\in[2^{-k},2^{-k+1}].$$

By a simple rescaling, it is easy to show that

$$\|\sup_{t\in[1,2]}|\tilde{\mathcal{A}}_t|\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} = \|\sup_{t\in[2^k,2^{k+1}]}|\tilde{\mathcal{A}}_t|\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \qquad \text{for any } k\in\mathbb{Z}.$$

Fixing $k\in\mathbb{Z}$ for now, we combine these observations with the assumption of the lemma.

$$\begin{aligned}\int_{\mathbb{R}^d}\sup_{t\in[2^k,2^{k+1}]}|\tilde{\mathcal{A}}_t g(x)|^p\mathrm{d}x &= \int_{\mathbb{R}^d}\sup_{t\in[2^k,2^{k+1}]}|\sum_{|l-(k+k_0)|\leq 10}\tilde{\mathcal{A}}_t P_l g(x)|^p\mathrm{d}x\\ &\leq (M_p(\lambda_0))^p\int_{\mathbb{R}^d}|\sum_{|l-(k+k_0)|\leq 10}P_l g(x)|^p\mathrm{d}x\\ &\lesssim (M_p(\lambda_0))^p\int_{\mathbb{R}^d}\sum_{|l-(k+k_0)|\leq 10}|P_l g(x)|^p\mathrm{d}x.\end{aligned}$$

Therefore,

$$\begin{aligned}\int_{\mathbb{R}^d}\sup_{t\in(0,\infty)}|\tilde{\mathcal{A}}_t g(x)|^p\mathrm{d}x &\leq \sum_{k\in\mathbb{Z}}\int_{\mathbb{R}^d}\sup_{t\in[2^k,2^{k+1}]}|\tilde{\mathcal{A}}_t g(x)|^p\mathrm{d}x\\ &\lesssim (M_p(\lambda_0))^p\int_{\mathbb{R}^d}\sum_{k\in\mathbb{Z}}\sum_{|l-(k+k_0)|\leq 10}|P_l g(x)|^p\mathrm{d}x\\ &\lesssim (M_p(\lambda_0))^p\int_{\mathbb{R}^d}\sum_{l\in\mathbb{Z}}|P_l g(x)|^p\mathrm{d}x\\ &\lesssim (M_p(\lambda_0))^p\int_{\mathbb{R}^d}(\sum_{l\in\mathbb{Z}}|P_l g(x)|^2)^{p/2}\mathrm{d}x\\ &\lesssim_p (M_p(\lambda_0))^p\|g\|^p_{L^p(\mathbb{R}^d)},\end{aligned}$$

where we made use of the inclusion $\ell^2\subseteq\ell^p$ for $p>2$ in the penultimate step, and (3.13) in the last step. □

Because of the above lemma, we see that Proposition 3.2.1 follows from its localised version:

[3]See, for a reference, [54, p. 267].

Proposition 3.2.3. *For $p > 2$, there exists $\varepsilon(p) > 0$ such that*

$$\|\mathcal{M}^j_{\mathrm{b,loc}}\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \lesssim 2^{-j\varepsilon(p)} \qquad \textit{for any } j \in \mathbb{N}. \tag{3.14}$$

Proposition 3.2.3 $\implies$ *Proposition 3.2.1.* As in the proof of Proposition 3.1.3, the $j = 0$ case in Proposition 3.2.1 is almost immediate. Thus, it suffices to look at the cases when $j \neq 0$.

Fix $j \in \mathbb{N}$ and consider $m := \hat{\sigma} \cdot \beta_{(j)}$. Observe that the definition of $\tilde{\mathcal{A}}_t$ now coincides with that of $\mathcal{A}^j_t$. By applying Lemma 3.2.2 for $d = 2$, we see that the operator estimates in (3.12) and (3.14) are equivalent, concluding the proof. □

Arguing along the lines of the proof of Proposition 3.1.3, the required decay in the L^p operator norms of the frequency localised maximal operators from certain related Fourier integral estimates.

Proposition 3.2.4. *For $j \in \mathbb{N}$, recall the definition of $\mathcal{A}^j$ from (3.8). There exists $p_0 \in (2, \infty)$ and $\varepsilon_1, \varepsilon_2 > 0$ such that for all $g \in \mathcal{S}(\mathbb{R}^2)$ and $j \in \mathbb{N}$, we have*

$$\|\mathcal{A}^j g(\cdot, 1)\|_{L^{p_0}(\mathbb{R}^2)} \lesssim 2^{-j\varepsilon_1}\|g\|_{L^{p_0}(\mathbb{R}^2)} \tag{3.15}$$

and

$$\|\mathcal{A}^j g\|^{1-1/p_0}_{L^{p_0}(\mathbb{R}^2\times[1,2])}\|\partial_t\mathcal{A}^j g\|^{1/p_0}_{L^{p_0}(\mathbb{R}^2\times[1,2])} \lesssim 2^{-j\varepsilon_2}\|g\|_{L^{p_0}(\mathbb{R}^2)}. \tag{3.16}$$

Proposition 3.2.4 $\implies$ *Proposition 3.2.3.* Fix $j \in \mathbb{N}$. The proposition is proved by interpolating operator norm estimates for $\mathcal{M}^j_{\mathrm{b,loc}}$ between three cases: $p = 2, p_0$ and $p = \infty$.

Case $p = p_0$: Let $g \in \mathcal{S}(\mathbb{R}^d)$. The key result to use here is the Sobolev embedding lemma (Lemma 3.1.4). Temporarily fixing $x \in \mathbb{R}^d$, set

$$F(t) := \mathcal{A}^j g(x, t) \qquad \text{for } t \in [1, 2].$$

Combining (3.15) and (3.16) with Lemma 3.1.4 and applying Fubini's theorem,

$$\|\mathcal{M}^j_{\mathrm{b,loc}}\|_{L^{p_0}(\mathbb{R}^2)\to L^{p_0}(\mathbb{R}^2)} \lesssim 2^{-j\min\{\varepsilon_1,\varepsilon_2\}}.$$

Case $p = 2$: This case has already been investigated in the proof of Proposition 3.1.3. In particular, the Fourier decay of the circular measure yields

$$\|\mathcal{M}^j_{\mathrm{b,loc}}\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \lesssim 1.$$

Case $p = \infty$: An application of young's inequality gives the kernel estimate $\|\sigma * \check{\beta}_{(j)}\|_{L^1(\mathbb{R}^2)} \lesssim 1$. As an immediate consequence,

$$\|\mathcal{M}^j_{\mathrm{b,loc}}\|_{L^\infty(\mathbb{R}^2)\to L^\infty(\mathbb{R}^2)} \lesssim 1.$$

By interpolating between these estimates, we can conclude the proof of Proposition 3.2.3. □

The proof of Proposition 3.2.4 relies on what is known as the local smoothing phenomenon. In the next section, we take a brief detour from the main line of argument to introduce the local smoothing problem; and afterwards, we will see how it relates to the proposition.

3.3 The local smoothing problem in the plane

Let $\rho, \tilde{\beta} \in \mathcal{S}(\mathbb{R})$ be real-valued and chosen such that

$$\operatorname{supp} \hat{\rho} \subseteq [-1,1] \qquad \text{and} \qquad \rho(t) \gtrsim 1 \quad \text{for } t \in [1,2], \tag{3.17}$$

and

$$\operatorname{supp} \tilde{\beta} \subseteq [1/4,4] \qquad \text{and} \qquad \tilde{\beta}(t) = 1 \quad \text{for } t \in [1/2,2]. \tag{3.18}$$

For $\lambda \in \mathbb{N}$, consider the Fourier integral operator defined by

$$\mathfrak{T}^\lambda(g)(x,t) := (2\pi)^{-2} \int_{\mathbb{R}^2} e^{i(\langle x,\xi\rangle + t|\xi|)} \tilde{\beta}(\lambda^{-1}|\xi|)\rho(t)\hat{g}(\xi)\mathrm{d}\xi \tag{3.19}$$

whenever $g \in \mathcal{S}(\mathbb{R}^2)$. The operator $\mathfrak{T}^\lambda$ is closely related to the Euclidean half-wave propagator $e^{it\sqrt{-\Delta}}$ which is used to construct solutions to the Cauchy problem for the wave equation (a detailed discussion on this topic can be found in [28]).

We are interested in the sharp $L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2 \times [1,2])$ estimates for $\mathfrak{T}^\lambda$. One way to obtain this is by freezing the time variable and individually investigating the operators of the form $\mathfrak{T}^\lambda_t(g) := \mathfrak{T}^\lambda(g)(\,\cdot\,,t)$ for each $t \in [1,2]$. To this end, we record the fixed time estimates from Peral [40] and Miyachi [36] (see [47] for a more general version).

Proposition 3.3.1. *Let $\lambda \in 2^{\mathbb{N}}$ and $p \in [2,\infty]$. For any $t > 0$, we have*

$$\|\mathfrak{T}^\lambda_t\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2)} \lesssim_p \lambda^{-(1/p-1/2)}. \tag{3.20}$$

By acting on certain test functions (see, for instance, [4, §2.1]), we can also see that (3.20) is sharp. Now, by Fubini's theorem, Proposition 3.3.1 implies that

$$\|\mathfrak{T}^\lambda\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} \lesssim_p \lambda^{-(1/p-1/2)}. \tag{3.21}$$

In the *local smoothing* problem (for the wave equation in the plane), we seek the possibility of proving stronger estimates for $\mathfrak{T}^\lambda$, compared to what is obtained as the output of the best fixed time estimates.

By *sharp local smoothing estimates*, we refer to the operator norm estimates

$$\|\mathfrak{T}^\lambda\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} \lesssim_{\alpha,p} \lambda^{-\alpha} \qquad \text{for } \alpha < \alpha_{\text{crit}}(p) - 1/2, \tag{3.22}$$

where

$$\alpha_{\text{crit}}(p) = \begin{cases} 1/2 & \text{for } 2 \le p \le 4, \\ 2/p & \text{for } p > 4. \end{cases}$$

Observe that (3.22) is a stronger estimate than (3.21) whenever $p > 2$. In this range, the non-trivial improvement (by a power of λ) from the fixed time estimates can be interpreted as a manifestation of the smoothing effect of averaging in the t variable.

The connection between the local smoothing problem and maximal inequalities was laid out by the works of Sogge [51] and Mockenhaupt–Seeger–Sogge [37]. In the former article, Sogge produced local smoothing estimates for $p > 2$, reformulating Bourgain's proof to the circular maximal theorem in terms of the local smoothing phenomenon (in fact, the main theorem in [51] is a variable-coefficient generalisation of Theorem 3.1.2). In the same article, Sogge conjectured the sharp local smoothing estimates (3.22).

We list out some of the major milestones in the direction of solving the sharp local smoothing estimates:

- Wolff [56] proved (3.22) for all $p > 74$. His argument used the ℓ^p-decoupling phenomenon for the light cone in $\mathbb{R}^3$, kick-starting the theory of Fourier decoupling. Building upon Wolff's argument, the range was later extended to $p > 190/3$ by Garrigós and Seeger [23], and to $p > 20$ by Garrigós—Seeger—Schlag [22].

- By proving L^3 estimates for a square function associated to the light cone in $\mathbb{R}^3$, Lee and Vargas [35] obtained (3.22) whenever $p \in [2, 3]$.

- Bourgain and Demeter [7] established sharp decoupling inequalities for the cone in 2015. Using Wolff's strategy, this implies (3.22) whenever p lies in the range $[6, \infty)$. Thus, the only regime where local smoothing conjecture remained unsolved was the interval $(3, 6)$.

- Guth–Wang–Zhang filled this gap by proving the sharp L^4 estimate for the cone square function in [26] (square functions estimates are, in general, harder to prove compared to the decoupling estimates). Their result completely resolved the local smoothing conjecture in the plane.

By following the arguments from [37] and assuming the square function estimate from [26], we reproduce the sharp local smoothing estimates.

Theorem 3.3.2. *For any $\lambda \in 2^{\mathbb{N}}$ and $p \in [2, \infty]$, the estimate* (3.22) *holds.*

Postponing the proof of Theorem 3.3.2 until later sections, we will investigate how this theorem relates to the circular maximal theorem.

3.3.1 Local smoothing to maximal estimates

Following the discussion in §3.2, the proof of the circular maximal theorem reduces to Proposition 3.2.4. Before commencing the proof of the proposition, we introduce the following definition.

Definition 3.3.3. A function $a \in C^\infty(\mathbb{R}^d \times \mathbb{R})$ is said to be a symbol of order μ if it is smooth away from the origin and for any $\alpha \in \mathbb{N}^d$ and $\beta \in \mathbb{N}$, we have

$$|\partial_\xi^\alpha \partial_t^\beta a(\xi, t)| \lesssim_{\alpha,\beta} (1 + |\xi|)^{\mu - |\alpha|} \qquad \text{for } (\xi, t) \in (\mathbb{R}^d \setminus \{0\}) \times \mathbb{R}.$$

The collection of all symbols with order μ is denoted as S^ν.

To relate the circular maximal function with the local smoothing problem, we write circular averages in the form of Fourier integral operators. Recall that the method of stationary phase yields the formula (for a reference, see [54, chapter 8])

$$\hat{\sigma}(\xi) = (2\pi)^2 \sum_{\pm} a_\pm(|\xi|)e^{\pm i|\xi|}, \tag{3.23}$$

where $a_+, a_- \in S^{-1/2}$ (the $(2\pi)^2$ factor is included in (3.23) for convenience). Using (3.5) and the above formula, we write

$$\mathcal{A}^j g(x,t)\rho(t) = \sum_{\pm} \mathfrak{A}^j_\pm(g)(x,t) \qquad \text{for } g \in \mathcal{S}(\mathbb{R}^2) \tag{3.24}$$

where ρ is recalled from (3.17) and

$$\mathfrak{A}^j_\pm(g)(x,t) := \rho(t)\int_{\mathbb{R}^2} e^{i(\langle x,\xi\rangle \pm t|\xi|)} a_\pm(t|\xi|)\beta(2^{-j}t|\xi|)\hat{g}(\xi)\mathrm{d}\xi$$

for $j \in \mathbb{N}$ and $(x,t) \in \mathbb{R}^3$. Observe that the operators $\mathfrak{A}^j_+$ and $\mathfrak{T}^\lambda$ in (3.19) are similar Fourier integral operators, except for the order of the associated symbol.

Theorem 3.3.2 $\implies$ *Proposition 3.2.4.* Among the two inequalities to be proved in Proposition 3.2.4, let us first focus on the easier one: the fixed time estimate (3.15). We claim that

$$\|\mathcal{A}^j g(\,\cdot\,,t)\|_{L^p(\mathbb{R}^2)} \lesssim 2^{-j/p}\|g\|_{L^p(\mathbb{R}^2)} \qquad \text{for any } t>0 \text{ and } p \in [2,\infty].$$

When $p=2$, Plancherel's theorem and the decay of $\hat{\sigma}$ are sufficient to yield the claim. On the other hand, when $p=\infty$, the claim follows by Young's inequality. By interpolating between these two cases, we deduce the claim for any p in the mentioned range, completing the proof of (3.15) with $\varepsilon_1(p) = 1/p$. We may proceed to the proof of (3.16).

In view of (3.24), it suffices to estimate the operators $\mathfrak{A}^j_\pm$. After setting $\lambda = 2^j$, we can use the support properties of β and $\tilde{\beta}$ to write

$$\lambda^{1/2}|\mathfrak{A}^j_\pm(g)(x,t)| \lesssim |\chi_{1,\pm}(\tfrac{1}{i}\partial_x, t) \circ \mathfrak{T}^\lambda(g)(x,t)|$$

where $\chi_{1,\pm}(\xi,t) := \lambda^{1/2}a_\pm(t|\xi|)\beta(\lambda^{-1}t|\xi|)\rho(t)$.

Direct computations show that for $t \in [1,2]$, we also have

$$\partial_t \mathfrak{A}^j_\pm(g)(x,t) = \int_{\hat{\mathbb{R}}^2} e^{i(\langle x,\xi\rangle + t|\xi|)} b_\pm(\xi,t)\beta(\lambda^{-1}t|\xi|)\hat{g}(\xi)\mathrm{d}\xi,$$

where $b_+, b_- \in S^{1/2}(\mathbb{R}^{2+1})$. Therefore, as above, we write

$$\lambda^{-1/2}|\partial_t\mathfrak{A}^j_\pm(g)(x,t)| \lesssim |\chi_{2,\pm}(\tfrac{1}{i}\partial_x, t) \circ \mathfrak{T}^\lambda(g)(x,t)|,$$

where $\chi_{2,\pm}(\xi,t) := \lambda^{-1/2}b_\pm(\xi,t)\beta(\lambda^{-1}t|\xi|)\rho(t)$. From the order properties of $a_\pm$

and $b_\pm$, we deduce that $\chi_{1,\pm}, \chi_{2,\pm} \in S^0$. Therefore, the Hörmander–Mihilin multiplier theorem[4] can be applied to conclude that

$$\lambda^{1/2}\|\mathfrak{A}^j_\pm(g)\|_{L^p(\mathbb{R}^2\times[1,2])} + \lambda^{-1/2}\|\partial_t\mathfrak{A}^j_\pm(g)\|_{L^p(\mathbb{R}^2\times[1,2])} \lesssim \|\mathfrak{T}^\lambda g\|_{L^p(\mathbb{R}^2\times[1,2])},$$

when $\lambda = 2^j$. Applying Theorem 3.3.2, we obtain

$$\|\mathfrak{A}^j_\pm\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} \lesssim_{\alpha,p} 2^{-j(\alpha+1/2)}$$

and

$$\|\partial_t\mathfrak{A}^j_\pm(g)\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} \lesssim_{\alpha,p} 2^{-j(\alpha-1/2)},$$

for $\alpha < \alpha_{\text{crit}}(p) - 1/2$ and $p \in (2,\infty)$. By combining these norm inequalities with (3.24), we obtain

$$\|\mathcal{A}^j\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} + 2^{-j}\|\partial_t\mathcal{A}^j\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} \lesssim_{\alpha,p} 2^{-j(\alpha+1/2)},$$

for any $\alpha < \alpha_{\text{crit}}(p) - 1/2$ and $p \in (2,\infty)$. Consequently,

$$\|\mathcal{A}^j g\|^{1-1/p}_{L^p(\mathbb{R}^2\times[1,2])}\|\partial_t\mathcal{A}^j g\|^{1/p}_{L^p(\mathbb{R}^2\times[1,2])} \lesssim_{\varepsilon_2,p} 2^{-j\varepsilon_2}\|g\|_{L^p(\mathbb{R}^2)},$$

provided $\varepsilon_2 < \varepsilon_{\text{crit}}(p) := \alpha_{\text{crit}}(p) - 1/p$. Since $\varepsilon_{\text{crit}}(p) > 0$ for any $p \in (2,\infty)$, we obtain (3.16) after choosing any p_0 in this range. Thus, we conclude the proof of Proposition 3.2.4 and thereby establish Theorem 3.1.2. $\square$

3.4 Proving the sharp local smoothing theorem

Fix $\lambda \in 2^{\mathbb{N}}$. We begin by noting an easy but crucial observation about the support properties of the space-time Fourier transform of $\mathfrak{T}^\lambda$.

For $g \in \mathcal{S}(\mathbb{R}^2)$, we use the integral expression (3.19) to deduce that

$$\mathcal{F}_{x,t}\mathfrak{T}^\lambda(g)(\xi,\tau) = \tilde{\beta}(\lambda^{-1}|\xi|)\hat{\rho}(\tau - |\xi|)\hat{g}(\xi).$$

Recalling (3.17) and (3.18), the above identity tells us that the space-time Fourier transform of $\mathfrak{T}^\lambda(g)$ lies in an $O(1)$-neighborhood of the truncated light cone

$$\Lambda_\lambda := \big\{(\xi,\tau) \in \mathbb{R}^3 : 1/4\lambda \le |\xi| \le 4\lambda \text{ and } \tau = |\xi|\big\}.$$

Because of this, we can relate the local smoothing problem with a bunch of questions in Fourier restriction theory surrounding functions frequency-supported in a tiny neighborhood of a compact *curved* sub-manifold (see the survey articles [4] and [25] for a detailed discussion on these topics). In these studies, it has become a standard practice to decompose the neighborhood of the submanifold into many essentially convex boxes. This division induces a natural decomposition of the function, such that we can expect cross-cancellations to occur between different parts while summing them back. Typically, the cancellations are manifested as decoupling or square function inequalities. On the other hand, the individual

[4]for a reference, see [54, Proposition 2 on p.245]

pieces are much easier to study. For instance, the uncertainty principle gives us efficient heuristics about the kernel associated with a Fourier multiplier operator whose multiplier function is a bump function supported inside a convex region.

For our argument, considering the direction of the non-vanishing principle curvature of the light cone, we realise that the right decomposition should be done in the angular variable.

Introducing the angular decomposition in $\mathbb{R}^2$

Let $\{\xi_\nu\}_{\nu\in\mathcal{W}(\lambda)}$ be a maximal $\lambda^{-1/2}$ separated subset of $\mathbb{S}^1$, so that the index set $\mathcal{W}(\lambda)$ is of cardinality $O(\lambda^{1/2})$. For each $\nu \in \mathcal{W}(\lambda)$, let $\chi_\nu \in C^\infty(\mathbb{R}^2)$ be a homogeneous function of degree 0 satisfying the following properties:

(i) $|\partial_\xi^\alpha \chi_\nu(\xi)| \lesssim \lambda^{-|\alpha|/2} \quad$ for $\alpha \in \mathbb{Z}^2$,

(ii) $\operatorname{supp} \chi_\nu \subseteq \{\xi \in \mathbb{R}^2 : |\xi/|\xi| - \xi_\nu| \lesssim \lambda^{-1/2}\}$,

(iii) $\sum_{\nu\in\mathcal{W}(\lambda)} \chi_\nu(\xi) = 1$ for all $\xi \in \mathbb{R}^2$.

Recalling (3.19), write $\mathfrak{T}^\lambda(g) = \sum_{\nu\in\mathcal{W}(\lambda)} \mathfrak{T}_\nu^\lambda(g)$, where

$$\mathfrak{T}_\nu^\lambda(g)(x,t) := \int_{\mathbb{R}^2} e^{i(\langle x,\xi\rangle + t|\xi|)} \tilde{\beta}(\lambda^{-1}|\xi|)\chi_\nu(\xi)\rho(t)\hat{g}(\xi)\mathrm{d}\xi.$$

To underline the importance of the specific decomposition here, notice the non-linearity (in ξ) of the phase function $\phi(x,t,\xi) := \langle x,\xi\rangle + t|\xi|$. By decomposing the $\mathfrak{T}^\lambda$, we have *linearised* the phase function; if $K_\nu^\lambda(\,\cdot\,,t)$ denotes kernel associated to $\mathfrak{T}_\nu^\lambda(\,\cdot\,,t)$ for each $t > 0$, then we can write

$$K_\nu^\lambda(x,t) = (\mathcal{F}_x^{-1} e^{it\langle \xi_\nu,\cdot\rangle} m_\nu^\lambda(\cdot,t))\rho(t),$$

where

$$m_\nu^\lambda(\xi,t) := e^{it(|\xi| - \langle \xi_\nu,\xi\rangle)}\chi_\nu(\xi)\tilde{\beta}(\lambda^{-1}|\xi|)$$

behaves like a bump function. Indeed, using the properties of χ_ν, it is easy to deduce that

$$|\partial_{\xi_\nu}^\alpha m_\nu^\lambda(\xi,t)| \lesssim_\alpha \lambda^{-|\alpha|}, \qquad |\partial_{\xi_\nu^\perp}^\alpha m_\nu^\lambda(\xi,t)| \lesssim_\alpha \lambda^{-|\alpha|/2} \quad \text{for any } \alpha \in \mathbb{N}.$$

Now, standard integration-by-parts arguments yield the decay estimates

$$|K_\nu^\lambda(x,t)| \lesssim_N \frac{\lambda^{3/2}}{(1 + \lambda|\langle x - t\xi_\nu, \xi_\nu\rangle| + \lambda^{1/2}|\langle x - t\xi_\nu, \xi_\nu^\perp\rangle|)^N} \tag{3.25}$$

for any $N \geq 1$ and $(x,t) \in \mathbb{R}^2 \times [1,2]$. Consequently,

$$\|K_\nu^\lambda(\,\cdot\,,t)\|_{L^1(\mathbb{R}^2)} \lesssim 1 \qquad \text{for each } t \in [1,2]. \tag{3.26}$$

Using the analysis so far, we can prove two easy end-point cases in Theorem 3.3.2:

Case $p=\infty$ *:* Note that the (3.26) combined with Fubini's theorem imply that

$$\|\mathfrak{T}_\nu^\lambda\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} \lesssim 1 \qquad \text{for } p\in[1,\infty].$$

By summing in ν, we also obtain

$$\|\mathfrak{T}^\lambda\|_{L^p(\mathbb{R}^2)\to L^p(\mathbb{R}^2\times[1,2])} \lesssim \lambda^{1/2} \qquad \text{for } p\in[1,\infty], \tag{3.27}$$

which coincides with the required inequality (3.22) when $p=\infty$.

Case $p=2$ *:* In this case, we can improve from (3.27) in view of the finitely overlapping supports of $\{m_\nu^\lambda\}_{\nu\in\mathcal{W}(\lambda)}$. Indeed, by a simple combination of Plancherel's theorem and Fubini's theorem,

$$\|\mathfrak{T}^\lambda\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2\times[1,2])} \lesssim 1. \tag{3.28}$$

In view of (3.27) and (3.28), we may restrict our focus on the estimate

$$\|\mathfrak{T}^\lambda\|_{L^4(\mathbb{R}^2)\to L^4(\mathbb{R}^2\times[1,2])} \lesssim_\epsilon \lambda^\epsilon \qquad \text{for any } \epsilon>0. \tag{3.29}$$

Indeed, by interpolating between these three operator estimates, we can prove Theorem 3.3.2.

Forward, reverse square function estimates and maximal inequalities

To prove (3.29), which asks for a nontrivial improvement from (3.27), we must exploit the oscillations carried by $\mathfrak{T}_\nu^\lambda$, and the argument involves many different elements. The most important among them is a deep result about a related cone square function.

For $0<r<1$ and $\nu\in\mathcal{W}(r^{-1})$, a plank Θ_ν^r centered at $(\xi_\nu,1)$ is defined by

$$\Theta_\nu^r := \Big\{\Xi\in\mathbb{R}^3 : |\langle\Xi, E_j(\nu)\rangle| \lesssim r^{(3-j)/2} \text{ for } j=1,2 \text{ and } |\langle\Xi,E_3(\nu)\rangle|\approx 1\Big\}$$

where

$$E_1(\nu) := \begin{pmatrix}\xi_\nu\\ -1\end{pmatrix}, \quad E_2(\nu) := \begin{pmatrix}\xi_\nu^\perp\\ 0\end{pmatrix}, \quad E_3(\nu) := \begin{pmatrix}\xi_\nu\\ 1\end{pmatrix}.$$

Note that $\{\Theta_\nu^r : \nu\in\mathcal{W}(r^{-1})\}$ is a collection of finitely overlapping convex boxes covering the r-neighbourhood of the truncated piece (at height one) of light cone Λ_1.

The result of Guth–Wang–Zhang [26] gives the sharp (reverse) square function estimates associated to these planks.[5]

Theorem 3.4.1 (Theorem 1.1, [26]). *Let* $0<r<1$. *Consider a collection* $\{f_\nu\in\mathcal{S}(\mathbb{R}^3)\}_{\nu\in\mathcal{W}(r^{-1})}$ *such that* supp $\hat{f}_\nu\subseteq\Theta_\nu^r$. *It follows that for any* $\varepsilon>0$,

[5]Interested reader should also see [37, §1] and [35] for previous results for square function.

there exists $C_\varepsilon > 0$ *such that*

$$\Big\| \sum_{\nu \in \mathcal{W}(r^{-1})} f_\nu \Big\|_{L^4(\mathbb{R}^3)} \le C_\varepsilon r^{-\varepsilon} \Big\| \big(\sum_{\nu \in \mathcal{W}(r^{-1})} |f_\nu|^2 \big)^{1/2} \Big\|_{L^4(\mathbb{R}^3)}.$$

The proof of the square function estimate is highly nontrivial and beyond the scope of this discussion. For this reason, we omit its proof here.

To apply Theorem 3.4.1 for the local smoothing problem, we fix $g \in \mathcal{S}(\mathbb{R}^2)$. Using the support properties of β and χ_ν, we see that the space-time Fourier transform of $\mathfrak{T}^\lambda_\nu(g)$ is supported inside the scaled plank[6] $\lambda \Theta_\nu^{\lambda^{-1}}$. Therefore, after setting $r = \lambda^{-1}$, Theorem 3.4.1 combined with a simple scaling argument gives

$$\big\| \mathfrak{T}^\lambda(g) \big\|_{L^4(\mathbb{R}^3)} \le C_\varepsilon \lambda^\varepsilon \Big\| \big(\sum_{\nu \in \mathcal{W}(r^{-1})} |\mathfrak{T}^\lambda_\nu(g)|^2 \big)^{1/2} \Big\|_{L^4(\mathbb{R}^3)}. \tag{3.30}$$

Now, we claim that

$$\int_{\mathbb{R}^3} \sum_{\nu \in \mathcal{W}(\lambda)} |\mathfrak{T}^\lambda_\nu(g)(x,t)|^2 \cdot f(x,t) \mathrm{d}x \mathrm{d}t \lesssim_\varepsilon \lambda^\varepsilon \|g\|_{L^4(\mathbb{R}^2)} \qquad \text{for any } \varepsilon > 0, \tag{3.31}$$

whenever $f \in L^2(\mathbb{R}^3)$ and $\|f\|_{L^2(\mathbb{R}^3)} \le 1$.

To prove the claim, we first note that the frequency localisation present in the operator can be translated to the function g. Let $\tilde{\tilde{\beta}} \in C_c^\infty(\mathbb{R})$ be defined such that supp $\tilde{\tilde{\beta}} \subseteq [-1/8, 8]$ and $\tilde{\tilde{\beta}} \cdot \tilde{\beta} = \tilde{\beta}$. Similarly, for $\nu \in \mathcal{W}(\lambda)$, define $\tilde{\chi}_\nu \in C^\infty(\mathbb{R}^2)$ in such a manner that it satisfies all the properties[7] of χ_ν and $\tilde{\chi}_\nu \cdot \chi_\nu = \chi_\nu$. Upon defining g^λ_ν to be such that

$$\mathcal{F}_x(g^\lambda_\nu)(\xi) := \mathcal{F}_x(g)(\xi) \tilde{\tilde{\beta}}(\lambda^{-1}|\xi|) \tilde{\chi}_\nu(\xi),$$

it follows immediately from the definitions that $\mathfrak{T}^\lambda_\nu(g)(x,t) = \mathfrak{T}^\lambda_\nu(g^\lambda_\nu)(x,t)$ for almost every $(x,t) \in \mathbb{R}^3$. By the Cauchy–Schwarz inequality and (3.26), we obtain

$$|\mathfrak{T}^\lambda_\nu(g)(x,t)|^2 = |\mathfrak{T}^\lambda_\nu(g^\lambda_\nu)(x,t)|^2 \lesssim (|K^\lambda_\nu(\,\cdot\,,t)| * |g^\lambda_\nu|^2)(x).$$

[6]For a set $E \subseteq \mathbb{R}^d$ and $r \in \mathbb{R}$, recall that rE denotes the scaled set $\{y \in \mathbb{R}^d : r^{-1} y \in E\}$.
[7]with a reasonable modification of item (iii) in its properties

Thus, for $f \in L^2(\mathbb{R}^3)$, we can make the following deductions:

$$\begin{aligned}
&\int_{\mathbb{R}^3} \sum_{\nu \in \mathcal{W}(\lambda)} |\mathfrak{T}_\nu^\lambda(g)(x,t)|^2 \cdot |f(x,t)| \, \mathrm{d}x\mathrm{d}t \\
&\qquad \le \int_{\mathbb{R}^2} \sum_{\nu \in \mathcal{W}(\lambda)} \left[\int_{\mathbb{R}^d} |K_\nu^\lambda(x-y,t)||g_\nu^\lambda(y)|^2 \mathrm{d}y \right] \cdot |f(x,t)| \, \mathrm{d}x\mathrm{d}t \\
&\qquad = \int_{\mathbb{R}^2} \sum_{\nu \in \mathcal{W}(\lambda)} |g_\nu^\lambda(y)|^2 \cdot \left(\int_{\mathbb{R}^2} |K_\nu^\lambda(x-y,t)||f(x,t)| \, \mathrm{d}x\mathrm{d}t \right) \mathrm{d}y. \\
&\qquad \le \int_{\mathbb{R}^2} \left(\sum_{\nu \in \mathcal{W}(\lambda)} |g_\nu^\lambda(y)|^2 \right) \mathcal{N}_{\lambda^{-1},\mathrm{plane}}(f)(y) \, \mathrm{d}y \\
&\qquad \le \Big\| (\sum_{\nu \in \mathcal{W}(\lambda)} |g_\nu^\lambda|^2)^{\frac{1}{2}} \Big\|_{L^4(\mathbb{R}^2)} \cdot \|\mathcal{N}_{\lambda^{-1},\mathrm{plane}}(f)\|_{L^2(\mathbb{R}^2)},
\end{aligned} \tag{3.32}$$

where

$$\mathcal{N}_{\lambda^{-1},\mathrm{plane}}(f)(y) := \sup_{\nu \in \mathcal{W}(\lambda)} \int_{\mathbb{R}^3} |K_\nu^\lambda(x-y,t)||f(x,t)|\mathrm{d}x\mathrm{d}t \qquad \text{for } y \in \mathbb{R}^2.$$

The forward square estimate required to bound the first term in (3.32) is obtained by Cordoba [15].

Theorem 3.4.2 (Cordoba [15]). *For g, g_ν^λ as defined above and for any $\varepsilon > 0$, we have*

$$\|(\sum_{\nu \in \mathcal{W}(\lambda)} |g_\nu^\lambda|^2)^{\frac{1}{2}}\|_{L^4(\mathbb{R}^2)} \lesssim_\varepsilon \lambda^\varepsilon \|g\|_{L^4(\mathbb{R}^2)}.$$

An alternate proof of Theorem 3.4.2 is given by Carbery–Seeger [10, Proposition 4.6]. We omit the proof of the theorem here.

In view of the forward square function estimate, we move to the most important result for this thesis, the L^2 norm bound on $\mathcal{N}_{\lambda^{-1},\mathrm{plane}}$, a Nikodym type maximal operator associated to the circle.

Theorem 3.4.3 (Lemma 1.4, [37]). *For any $\varepsilon > 0$, we have*

$$\|\mathcal{N}_{\lambda^{-1},\mathrm{plane}}\|_{L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^2)} \lesssim_\varepsilon \lambda^\varepsilon.$$

Combining the discussion so far, we complete the proof of the sharp local smoothing theorem.

Proof of Theorem 3.3.2. By combining Theorem 3.4.3, Theorem 3.4.2 and (3.32), we yield the claim (3.31). Combining this inequality with (3.30), we obtain (3.29). Now, using Marcinkiewicz's interpolation theorem [52, §4], we can interpolate between the operator estimates (3.29), (3.28) and (3.27). Thus, we obtain (3.22), concluding the proof of Theorem 3.3.2. □

Brief discussion on the Nikodym maximal estimate

The first step in proving Theorem 3.4.3 is to relate the maximal function with its geometric variant. To this end, we note that (3.25) implies rapid decay of the function K_ν^λ away from the box

$$T_{s_\nu}^\lambda := \Big\{(x,t) \in \mathbb{R}^2 \times [1,2] : |(x,t)\cdot(\xi_\nu,1)| \le \lambda^{-1},\ |(x,t)\cdot(\xi_\nu^\perp,0)| \le \lambda^{-1/2}\Big\},$$

where $s_\nu \in [0,2\pi)$ is chosen such that $\xi_\nu = (\cos s_\nu, \sin s_\nu)$. In view of this, we define the geometric maximal function

$$\mathcal{N}_{\lambda^{-1},\mathrm{plane}}^{\mathrm{geom}}(f)(y) := \sup_{s\in[0,2\pi)} \frac{1}{|T_s^\lambda|}\int_{T_s^\lambda} |f(x-y,t)|\mathrm{d}x\mathrm{d}t,$$

whenever $f \in \mathcal{S}(\mathbb{R}^3)$. In [37], the authors were able to estimate the L^2 operator norm of $\mathcal{N}_{\lambda^{-1},\mathrm{plane}}^{\mathrm{geom}}$ and obtain Theorem 3.4.3 as a consequence.

Theorem 3.4.4 (Nikodym maximal estimate: geometric version)**.** *For any $\varepsilon > 0$, we have*

$$\|\mathcal{N}_{\lambda^{-1},\mathrm{plane}}^{\mathrm{geom}}\|_{L^2(\mathbb{R}^3)\to L^2(\mathbb{R}^2)} \lesssim_\varepsilon \lambda^\varepsilon.$$

Theorem 3.4.4 $\implies$ Theorem 3.4.3. Let us begin by dyadically decomposing the $\mathbb{R}^3$ spaces using dilates of the boxes of the form T_s^λ. For $k \ge 2$, define[8] $W_{k,\nu} := 2^{k+1}T_{s_\nu}^\lambda \setminus 2^k T_{s_\nu}^\lambda$ and set $W_{1,\nu} := T_{s_\nu}^\lambda$. Now, the decay estimate (3.25) implies that

$$|K_\nu^\lambda(x,t)| \lesssim_N \sum_{k\in\mathbb{N}} 2^{-kN}\lambda^{3/2}\chi_{W_{k,\nu}}(x,t) \qquad \text{for } N \ge 1.$$

Temporarily fix N. For $y \in \mathbb{R}^2$, we have

$$\begin{aligned}
\int_{\mathbb{R}^3} &|K_\nu^\lambda(x,t)||f(x-y,t)|\mathrm{d}x\mathrm{d}t \\
&\lesssim \sum_{k\in\mathbb{N}} 2^{-kN}\lambda^{3/2}\int_{W_{k,\nu}} |f(x-y,t)|\mathrm{d}x\mathrm{d}t \\
&\lesssim \sum_{k\in\mathbb{N}} 2^{-kN+3(k+1)}\frac{1}{|2^{k+1}T_{s_\nu}^\lambda|}\int_{2^{k+1}T_{s_\nu}^\lambda} |f(x-y,t)|\mathrm{d}x\mathrm{d}t.
\end{aligned}$$

Taking the supremum in ν, we obtain

$$\mathcal{N}_{\lambda^{-1},\mathrm{plane}}(f)(y) \lesssim \sum_{k\in\mathbb{N}} 2^{-kN+3(k+1)} \sup_{s\in(0,2\pi]}\frac{1}{|2^{k+1}T_s^\lambda|}\int_{2^{k+1}T_s^\lambda} |f(x-y,t)|\mathrm{d}x\mathrm{d}t.$$

[8] Note that the dilates are defined with respect to the centroid of the tube. In particular,

$$2^k T_{s_\nu}^\lambda := \{(x,t)\in\mathbb{R}^2\times[1,2] : |(x,t)\cdot(\xi_\nu,1)| \le 2^k\lambda^{-1},\ |(x,t)\cdot(\xi_\nu^\perp,0)| \le 2^k\lambda^{-1/2}\}$$

However, for each $k \in \mathbb{N}$, a simple change of variables gives

$$\sup_{s\in(0,2\pi]} \frac{1}{|2^{k+1}T_s^\lambda|}\int_{2^{k+1}T_s^\lambda} |f(x-y,t)|\mathrm{d}x\mathrm{d}t = \big(\mathcal{N}^{\mathrm{geom}}_{\lambda^{-1},\mathrm{plane}}(f_{2^{-(k+1)}})\big)_{2^{(k+1)}}(y),$$

where we use the notation[9] h_r to represent the r- dilate of a function $h : \mathbb{R}^m \to \mathbb{R}$, defined by

$$h_r(w_1,\dots,w_m) := h(r^{-1}w_1,\dots,r^{-1}w_m) \qquad \text{for } (w_1,\cdots,w_m)\in\mathbb{R}^m$$

whenever $r > 0$. Using Theorem 3.4.4, we compute that

$$\begin{aligned}\|\big(\mathcal{N}^{\mathrm{geom}}_{\lambda^{-1},\mathrm{plane}}(f_{2^{-(k+1)}})\big)_{2^{(k+1)}}\|_{L^2(\mathbb{R}^2)} &= 2^{(k+1)}\|\big(\mathcal{N}^{\mathrm{geom}}_{\lambda^{-1},\mathrm{plane}}(f_{2^{-(k+1)}})\big)\|_{L^2(\mathbb{R}^2)}.\\ &\lesssim_\varepsilon \lambda^\varepsilon 2^{(k+1)}\|f_{2^{-(k+1)}}\|_{L^2(\mathbb{R}^3)}\\ &= \lambda^\varepsilon 2^{-(k+1)/2}\|f\|_{L^2(\mathbb{R}^3)},\end{aligned}$$

for any $\varepsilon > 0$. Therefore,

$$\begin{aligned}\|\mathcal{N}_{\lambda^{-1},\mathrm{plane}}(f)\|_{L^2(\mathbb{R}^2)} &\le \sum_{k\in\mathbb{N}} 2^{-kN+3(k+1)}\|\big(\mathcal{N}^{\mathrm{geom}}_{\lambda^{-1},\mathrm{plane}}(f_{2^{-(k+1)}})\big)_{2^{(k+1)}}\|_{L^2(\mathbb{R}^2)}\\ &\lesssim_\epsilon \lambda^\varepsilon \sum_{k\in\mathbb{N}} 2^{-kN+5(k+1)/2}\|f\|_{L^2(\mathbb{R}^3)},\end{aligned}$$

for any $\varepsilon > 0$. By setting $N = 10$, we conclude the proof of Theorem 3.4.3. □

As we have hinted before, the purpose of our literature review so far is to introduce this version of the Nikodym-type maximal function. The next chapter considers the study of estimating higher dimensional versions of $\mathcal{N}^{\mathrm{geom}}_{\lambda^{-1},\mathrm{plane}}$. Theorem 3.4.4 then follows as a consequence of the main result (Theorem 4.1.2) over there. Because of this, we omit any discussion on the proof of Theorem 3.4.4 here. Nevertheless, it is worth pointing out that estimating the maximal function $\mathcal{N}^{\mathrm{geom}}_{\lambda^{-1},\mathrm{plane}}$ is comparatively simpler than the circular maximal problem. One reason behind the relative simplicity is that the class of curves associated to $\mathcal{N}^{\mathrm{geom}}_{\lambda^{-1},\mathrm{plane}}$ consists of line segments whose intersection patterns are much easier to study compared to those of curved objects. Furthermore, in contrast with the circular maximal theorem, here we are tasked to estimate the L^2 operator norm of the maximal function, which allows us to efficiently use the machinery of the Fourier transform.

[9] Note that the notation used here differs from that introduced in §1.1.1.

3.5 Nikodym maximal estimates and a local smoothing problem in $\mathbb{R}^d$.

Let $I := [-1, 1]$ and consider a curve $\gamma : I \to \mathbb{R}^d$. We say that γ is non-degenerate if there exists $A > 1$ such that

$$A^{-1}|\xi| \leq \sum_{i=1}^{d} |\langle \gamma^{(i)}(s), \xi \rangle| \leq A|\xi| \qquad \text{for } \xi \in \mathbb{R}^d \text{ and } s \in I.$$

For $s \in I$, let $\{e_1(s), \ldots, e_d(s)\}$ denote the collection of Frenet frame basis vectors, formed by applying Gram–Schmidt process to the set $\{\gamma^{(1)}(s), \ldots, \gamma^{(d)}(s)\}$. The curve γ defines a one-parameter family of directions in $\mathbb{R}^{d+1}$. For $r \in (0, 1)$, we consider an anisotropic tube in $\mathbb{R}^{d+1}$ in the direction of $\begin{pmatrix} \gamma(s) \\ 1 \end{pmatrix}$, defined by

$$T_r(s) := \Big\{(x, t) \in \mathbb{R}^d \times [1, 2] : |\langle x - t\gamma(s), e_j(s) \rangle| \leq r^{\min\{j/2,1\}} \text{ for } 1 \leq j \leq d\Big\}. \tag{3.33}$$

Consider the maximal function

$$\mathcal{N}_r(f)(y) := \sup_{s \in I} \frac{1}{|T_r(s)|} \int_{T_r(s)} |f(x - y, t)| \mathrm{d}x\mathrm{d}t, \qquad y \in \mathbb{R}^d \tag{3.34}$$

for any $f \in L^1_{\mathrm{loc}}(\mathbb{R}^{d+1})$. We will be obtaining L^2 norm estimates for this maximal function.

Theorem 3.5.1. *Let $r \in (0, 1)$. For any $\epsilon > 0$, there exists $C_\epsilon > 1$ such that*

$$\|\mathcal{N}_r\|_{L^2(\mathbb{R}^{d+1}) \to L^2(\mathbb{R}^d)} \leq C_\varepsilon r^{-\varepsilon}$$

The proof of Theorem 3.5.1 is contained in Chapter 3 (See Theorem 4.1.2). The purpose of the present section, however, is to place this result in the study of a local smoothing problem in $\mathbb{R}^d$.

3.5.1 A higher dimensional local smoothing problem

Let $a \in C^\infty(\mathbb{R}^d \setminus \{0\} \times I \times [1, 2])$ be a symbol of order 0. In other words, for every multi-indices $\alpha, \beta, \gamma \in \mathbb{Z}^d$, we have

$$|\partial_\xi^\alpha \partial_s^\beta \partial_t^\gamma a(\xi, s, t)| \lesssim_{\alpha,\beta,\gamma} (1 + |\xi|)^{-|\alpha|} \qquad \text{for } (\xi, s, t) \in \operatorname{supp} a.$$

We are interested in a higher dimensional local smoothing problem that focuses on the Fourier integral operator

$$m[a](D)f(x, t) := (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i\langle x, \xi \rangle} m[a](\xi, t) \hat{f}(\xi) \mathrm{d}\xi \qquad \text{for } (x, t) \in \mathbb{R}^{d+1} \text{ and}$$

$f \in \mathcal{S}(\mathbb{R}^{d+1})$, where the multiplier is defined as

$$m[a](\xi, t) = \int_I e^{-it\langle\gamma(s),\xi\rangle} a(\xi, s, t)\mathrm{d}s \qquad \text{for } (\xi, t) \in \mathbb{R}^{d+1}. \tag{3.35}$$

When $d = 2$ and γ is the unit circle, an asymptotic expression for $m[a]$ is available through stationary phase methods. This fact enabled us to present the proof of the local smoothing problem in the plane neatly. However, when $d \geq 3$, we do not possess such an explicit form to work with. Nevertheless, in view of (3.35), we write

$$m[a](D)f(x, t) = (2\pi)^{-d} \int_{\hat{\mathbb{R}}} \int_{\mathbb{R}} e^{i\langle x - t\gamma(s),\xi\rangle} a(\xi, s, t)\hat{f}(\xi)\mathrm{d}s\mathrm{d}\xi.$$

Recall the definition of β from (3.3) and write

$$a^\lambda(\xi, s, t) := a(\xi, s, t)\beta(\lambda^{-1}|\xi|) \qquad \text{for } (\xi, s, t) \in \operatorname{supp} a,\ \lambda \in 2^{\mathbb{N}}.$$

The sharp local smoothing conjecture in this setup can be formulated as below.

Conjecture 3.5.2 (§1, [32]). *Let $d \geq 3$. For $\lambda \in 2^{\mathbb{N}}$, the operator norm estimate*

$$\|m[a^\lambda](D)\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d\times[1,2])} \lesssim_{p,\sigma} \lambda^{-\sigma} \qquad \textit{hold for } \sigma < \sigma_{\text{crit}}(p),$$

where

$$\sigma_{\text{crit}}(p) = \begin{cases} 1/d & \textit{for } 2 \leq p \leq 2d, \\ 2/p & \textit{for } p > 2d. \end{cases}$$

To the author's best knowledge, the best-known partial result in the way of establishing the conjecture comes from the recent work of Ko–Lee–Oh [32]. Their result obtained sharp local smoothing estimates in the restricted range $p \geq 4d-2$ for any $d \geq 3$.

Through arguments almost identical to the ones in §3.2, the local smoothing conjecture implies L^p estimates for a related maximal function in $\mathbb{R}^d$. To introduce the maximal function, we consider a probability measure σ supported on the non-degenerate curve γ. For $t > 0$, define the dilated measure σ_t by the action

$$\langle\sigma_t, f\rangle = \langle\sigma, f(t\,\cdot)\rangle \qquad \text{for } f \in L^1_{loc}(\mathbb{R}^d).$$

Consider the maximal function

$$\mathcal{M}^\gamma_d f(x) := \sup_{t>0} |\sigma_t * f(x)| \qquad \text{for } x \in \mathbb{R}^d$$

whenever $f \in L^1_{loc}(\mathbb{R}^d)$. The sharp local smoothing conjecture implies sharp L^p estimates for this maximal function.

Conjecture 3.5.3 (§1, [32]). *Let $d \geq 3$. For γ non-degenerate,*

$$\|\mathcal{M}^\gamma_d\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \lesssim_{d,p,\gamma} 1$$

for any $p > d$.

When $d = 3$, the maximal conjecture follows from the independent results of [3] and [31]. When $d \geq 4$, the best partial result comes from [31, Theorem 1.4], where L^p boundedness of $\mathcal{M}_d^\gamma$ is obtained in the restricted range $p > 2(d-1)$.

As the higher dimensional local smoothing conjecture is extremely difficult to prove, we focus on a simpler problem wherein we restrict the support of the symbol to a region where a stronger non-degeneracy condition holds. In particular, we will be only interested in the region where the inequalities

$$A^{-1}|\xi| \leq |\langle \gamma^{(1)}(s), \xi \rangle| + |\langle \gamma^{(2)}(s), \xi \rangle| \leq A|\xi|, \tag{3.36}$$

hold for $(\xi, s) \in \operatorname{supp}_{\xi,s} a$. In this simplified setup, we expect the operator $m[a^\lambda](D)$ to have an L^p operator decay, similar to $\mathfrak{T}^\lambda$, since the non-degeneracy assumptions resemble the planar setup. Having said that, we must also be cautious of the fact that the Fourier Integral Operator $\mathfrak{T}^\lambda$ differs from $m[a^\lambda](D)$ (for $d = 2$) by a symbol of order $1/2$. This difference must be reflected in their decay estimates.

Conjecture 3.5.4. *Let $\gamma \subseteq \mathbb{R}^d$ be a non-degenerate curve and a be a symbol order zero. Assume that the strong non-degeneracy assumptions (3.36) hold in support of a. For $\lambda \in 2^{\mathbb{N}}$, it follows that*

$$\|m[a^\lambda](D)\|_{L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d \times [1,2])} \lesssim_{p,\alpha} \lambda^{-\alpha} \qquad \text{for } \alpha < \alpha_{\text{crit}}(p), \tag{3.37}$$

where

$$\alpha_{\text{crit}}(p) = \begin{cases} 1/2 & \text{for } 2 < p \leq 4, \\ 2/p & \text{for } p > 4. \end{cases}$$

Extending the arguments from §3.4, we sketch a possible proof to this conjecture, conditional on certain square function estimates. As one can imagine, Theorem 3.4.3 will feature in this argument.

3.5.2 Initial reductions

Recall that the specific form of the local smoothing operator $\mathfrak{T}^\lambda$ (in particular, the form of the phase function in (3.19)) made it more or less immediate to identify the 'most singular' part of the operator in the planar case. As we lack a compact expression (such as (3.23)) for the Fourier transform of a measure supported on $\gamma \subseteq \mathbb{R}^d$ when $d \geq 3$, additional deductions are required to achieve the same goal here. By carrying out many step-by-step reductions, we will now show that the most singular part of the operator $m[a^\lambda](D)$ has many characteristics similar to $\mathfrak{T}^\lambda$.

We note that many of the ideas in this subsection are taken from [3].

Strengthening the non-degeneracy condition

To begin with, we reduce the proof to a situation where stronger non-degeneracy conditions hold.

Conjecture 3.5.5. *Let γ be a non-degenerate curve and $a \in S^0$. Suppose there exists $B > 1$ and $s_* \in I$ such that*

$$B^{-1}|\xi| \leq |\langle \gamma^{(2)}(s), \xi\rangle| \leq B|\xi| \qquad \text{for } (\xi, s) \in supp_\xi(a) \times I \tag{3.38}$$

and

$$|\langle \gamma^{(1)}(s_*), \xi\rangle| \leq 10^{-10}B^{-1}|\xi| \qquad \text{for } \xi \in supp_\xi(a). \tag{3.39}$$

For $\lambda \in 2^{\mathbb{N}}$, it follows that

$$\|m[a^\lambda](D)\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d\times[1,2])} \lesssim_p \lambda^{-\alpha} \qquad \text{for } \alpha < \alpha_{\text{crit}}(p). \tag{3.40}$$

The following argument is motivated from [41, §4.3].

Conjecture 3.5.5 $\implies$ Conjecture 3.5.4. Consider the function

$$g(\xi, s) := \langle \gamma^{(1)}(s), \xi/|\xi|\rangle \qquad \text{for } (\xi, s) \in \mathbb{R}^d \setminus \{0\} \times I,$$

and let

$$C_\gamma := \sup_{s,\xi} |(\nabla_\xi g)(s, \xi)|, \qquad D_\gamma := \sup_{s,\xi} |\langle \gamma^{(2)}(s), \xi\rangle|.$$

Fix constants $0 < c_3 < c_2 < c_1 \ll 1$. Let $\{\xi_i\}_{i\in\mathbb{N}}$ denote a $(c_2\lambda)/2$ separated subset of $B(0, 2\lambda) \subseteq \mathbb{R}^d$. Similarly, let $\{s_j\}_{j\in\mathbb{N}}$ denote a $(c_3/2)$ separated subset of $\mathbb{R}$. Suppose $\{\phi_i^\lambda\}_{i\in\mathbb{N}}$ is a smooth partition of unity in $\mathbb{R}^d$ associated to the collection of balls $\{B(\xi_i, c_2\lambda)\}_{i\in\mathbb{N}}$ and let $\{\psi_j\}_{j\in N}$ denote a smooth partition of unity in $\mathbb{R}$ associated to the collection of intervals $\{B(s_j, c_3)\}_{j\in\mathbb{N}}$. Defining $a_{i,j}^\lambda(\xi, s, t) := a^\lambda(\xi, s, t)\phi_i^\lambda(\xi)\psi_j(s)$ for $(\xi, s, t) \in \mathbb{R}^{d+2}$, we can write

$$a^\lambda(\xi, s, t) = \sum_{i,j} a_{i,j}^\lambda(\xi, s, t).$$

Note that only $O((c_2c_3)^{-1})$ many symbols $a_{i,j}^\lambda$ are non-zero, and each of them is of order 0. Thus, it suffices to prove (3.37) after replacing a^λ with $a_{i,j}^\lambda$ for fixed indices $(i, j) \in \mathbb{N}^2$.

Depending on the size of the value g takes at (ξ_i, s_j), two cases arise here:
Case 1 ($|g(\xi_i, s_j)| < c_1$): For $(\xi, s) \in \text{supp}_{\xi,s}\, a_{i,j}^\lambda$, the mean value theorem gives

$$\begin{aligned}
|g(\xi, s) - g(\xi_i, s_j)| &\leq |g(\xi, s) - g(\xi, s_j)| + |g(\xi, s_j) - g(\xi_i, s_j)| \\
&\leq \sup_{t\in[0,1]} |\nabla_\xi g(t\xi_i + (1-t)\xi, s_j)|\lambda c_2 + D_\gamma c_3 \\
&\leq C_\gamma c_2 + D_\gamma c_3.
\end{aligned}$$

Note that in the final step, we used the fact that g is homogeneous of degree zero in ξ variable and $|\xi| \approx \lambda$ on the support of a^λ. Thus, if c_2, c_3 are chosen small enough depending on c_1 and γ, we can estimate

$$|\langle \gamma^{(1)}(s), \xi\rangle| \leq 2c_1|\xi| \qquad \text{for all } (\xi, s) \in \text{supp}_{\xi,s}\, a_{i,j}^\lambda.$$

Now, by choosing c_1 small enough depending on A and combining this inequality with (3.36), we obtain (3.38) for $B := A/2$ and $(\xi, s) \in \mathrm{supp}_{\xi,s}\, a^\lambda_{i,j}$. If we set $s_* = s_j$, (3.39) also becomes clear. Therefore, we can apply Conjecture 3.5.5 in this case, and (3.37) follows from (3.40).

Case 2 ($|g(\xi_i, s_j)| \geq c_1$): By following the argument from case 1, we deduce that

$$|g(\xi, s)| \geq \frac{c_1}{2} \qquad \text{for } (\xi, s) \in \mathrm{supp}_{\xi,s}\, a^\lambda_{i,j}. \tag{3.41}$$

For $f \in \mathcal{S}(\mathbb{R}^d)$, recall that

$$m[a^\lambda_{i,j}](D)f(x,t) := (2\pi)^{-d} \int e^{i\langle x - t\gamma(s), \xi\rangle} a^\lambda_{i,j}(\xi, s, t)\hat{f}(\xi)\mathrm{d}s\mathrm{d}\xi.$$

Note that by (3.41), we have a lower bound on the derivative of the phase function here. If $K^\lambda_{i,j}$ denote the kernel associated to $m[a^\lambda_{i,j}](D)$, a simple integration-by-parts argument gives

$$\left| \int e^{i\langle x - t\gamma(s), \xi\rangle} a^\lambda_{i,j}(\xi, s, t)\mathrm{d}s\mathrm{d}\xi \right| \lesssim_{N,M} \frac{\lambda^{-(N-d)}}{(1+\lambda|x|)^M} \qquad \text{for } x \in \mathbb{R}^d,\ N, M \in \mathbb{N}.$$

The above estimate immediately implies that

$$\|m[a^\lambda_{i,j}](D)\|_{L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d \times [1,2])} \lesssim_N \lambda^{-N},$$

completing the argument for case 2. □

The highlight of the stronger non-degeneracy assumptions in Conjecture 3.5.5 is that by combining (3.38) and (3.39) with the Implicit function theorem we can ensure the existence of a smooth map $s : \mathrm{supp}_\xi\, a^\lambda \to I$ such that

$$\langle \gamma^{(1)} \circ s(\xi), \xi\rangle = 0 \qquad \text{for } \xi \in \mathrm{supp}_\xi\, a^\lambda. \tag{3.42}$$

This map plays a vital role in the reductions to follow. Note that by implicitly differentiating (3.42), we have

$$\partial_{\xi_j} s(\xi) = -\frac{\langle \gamma^{(1)} \circ s(\xi), e_j\rangle}{\langle \gamma^{(2)} \circ s(\xi), \xi\rangle} \qquad \text{for } \xi \in \mathrm{supp}_\xi\, a^\lambda \text{ and } 1 \leq j \leq d, \tag{3.43}$$

where e_j denotes that standard jth basis vector in $\hat{\mathbb{R}}^d$.

Localisation along the curve

Using non-stationary phase arguments, we can restrict the (ξ, s)-support of the symbol to a neighbourhood of the 'degenerate' surface

$$\Sigma^d_\lambda := \{(\xi, s(\xi)) : \xi \in \mathrm{supp}_\xi\, a^\lambda\}.$$

Let $0 < \varepsilon_0 \ll 1$. Recalling the definition (3.3), write

$$\begin{aligned} a^\lambda(\xi, s, t) &= a^\lambda(\xi, s, t)[\eta(c_0^{-1}\lambda^{1/2-\varepsilon_0}(s - s(\xi))) + (1 - \eta(c_0^{-1}\lambda^{1/2-\varepsilon_0}(s - s(\xi)))] \\ &= a^{\lambda,\varepsilon_0}_{\mathrm{main}}(\xi, s, t) + a^{\lambda,\varepsilon_0}_{\mathrm{err}}(\xi, s, t), \end{aligned}$$

where $c_0 = 10^{-2}B^{-1/2}$ is an auxiliary constant. Estimating the operator norm of $m[a^{\lambda,\varepsilon_0}_{\mathrm{err}}](D)$ is an easy matter, as shown by the following lemma.

Lemma 3.5.6. *For $p \in [1, \infty]$, we have*

$$\|m[a^{\lambda,\varepsilon_0}_{\mathrm{err}}](D)\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d\times[1,2])} \lesssim_N \lambda^{-N} \qquad \text{for } N \in \mathbb{N}_0.$$

Proof. Suppose $\mathcal{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}$ denote the kernel associated to $m[a^{\lambda,\varepsilon_0}_{\mathrm{err}}](D)$. In other words,

$$\mathcal{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}(x, t) = (2\pi)^{-d} \int_{\mathbb{R}^d\times I} e^{i\langle x-t\gamma(s),\xi\rangle} a^{\lambda,\varepsilon_0}_{\mathrm{err}}(\xi, s, t)\mathrm{d}\xi\mathrm{d}s \qquad \text{for } (x, t) \in \mathbb{R}^{d+1}. \tag{3.44}$$

We aim for rapid L^1 decay estimates for $\mathcal{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}(\,\cdot\,, t)$, uniformly in t. From the definition, it is clear that $|s - s(\xi)| \geq c_0\lambda^{-(1/2-\varepsilon_0)}$ in the support of the integrand in (3.44). By the mean value theorem and (3.38), we have

$$|\langle\gamma^{(1)}(s), \xi\rangle| \geq B^{-1}\lambda \cdot c_0\lambda^{-(1/2-\varepsilon_0)} = c_0B^{-1}\lambda^{1/2+\varepsilon_0} \qquad \text{for } (\xi, s) \in \mathrm{supp}_{\xi,s}\, a^{\lambda,\varepsilon_0}_{\mathrm{err}}.$$

By Leibniz rule, we deduce that

$$|\partial_s^\beta a^{\lambda,\varepsilon_0}_{\mathrm{err}}(\xi, s, t)| \lesssim_{\beta,c_0} \lambda^{(1/2-\varepsilon_0)\beta} \lesssim_{c_0,B} \lambda^{-2\varepsilon_0\beta}|\langle\gamma^{(1)}(s), \xi\rangle|^\beta,$$

for any $\beta \in \mathbb{N}_0$ and $(\xi, s, t) \in \mathrm{supp}\, a^{\lambda,\varepsilon_0}_{\mathrm{err}}$. Using a standard integration-by-parts argument in the s variable, we obtain

$$\|\mathcal{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}(\,\cdot\,, t)\|_{L^\infty(\mathbb{R}^d)} \lesssim_N \lambda^{-\varepsilon_0 N} \qquad \text{for } N \in \mathbb{N}_0 \text{ and } t \in [1, 2].$$

As a matter of fact, a stronger estimate for the kernel is achievable if we also integrate-by-parts in ξ variable in (3.44). To do so, we record the ξ derivative estimates of the symbol. By elementary computations, we deduce the estimate

$$|\partial_\xi^\alpha s(\xi)| \lesssim_{\alpha,B,\gamma} |\xi|^{-|\alpha|} \qquad \text{for } \xi \in \mathrm{supp}_\xi\, a^\lambda \text{ and } \alpha \in \mathbb{N}_0^d$$

from (3.43) and (3.38). Therefore, by Leibniz rule,

$$|\partial_\xi^\alpha\partial_s^\beta a^{\lambda,\varepsilon_0}_{\mathrm{err}}(\xi, s, t)| \lesssim_{\alpha,\beta,c_0,B} \lambda^{-(1/2+\varepsilon_0)|\alpha|}\lambda^{-2\varepsilon_0\beta}|\langle\gamma^{(1)}(s), \xi\rangle|^\beta$$

for $\alpha \in \mathbb{N}_0^d$ and $\beta \in \mathbb{N}_0$. We can now carry out an integration-by-parts argument in both ξ, s variables in (3.44) (we use the differential operator $\mathcal{L}_{\xi,s} := (\frac{-1}{it\langle\gamma^{(1)}(s),\xi\rangle}\partial_s)\circ \frac{\langle x-t\gamma(s),\nabla_\xi\rangle}{|x-t\gamma(s)|^2}$ to run the integration-by-parts argument). In particular, we obtain

$$\|(1 + |\,\cdot\,|^{10d})\mathcal{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}(\,\cdot\,, t)\|_{L^\infty(\mathbb{R}^d)} \lesssim_N \lambda^{-\varepsilon_0 N} \qquad \text{for } N \in \mathbb{N}_0 \text{ and } t \in [1, 2].$$

The statement of the lemma follows immediately from here. □

Because of Lemma 3.5.6, it suffices to estimate the operator $m[a^{\lambda,\varepsilon_0}_{\mathrm{main}}]$. The next step is to identify the essential support of the spatio-temporal Fourier transform of this operator.

Spatio-temporal localisation

Let $m \in L^\infty(\mathbb{R}^{d+1})$ be a function that takes the frequency variables (ξ, τ) as its arguments. We define the Fourier integral operator $m(D)$ by

$$m(D)f(x,t) := (2\pi)^{-(d+1)} \int_{\mathbb{R}^{d+1}} e^{i(\langle x,\xi\rangle + t\tau)} m(\xi,\tau)\hat{f}(\xi)\mathrm{d}\xi\mathrm{d}\tau \qquad \text{for } f \in \mathcal{S}(\mathbb{R}^d)$$

whenever $(x,t) \in \mathbb{R}^{d+1}$.

Consider the function

$$h(\xi) := \langle \gamma \circ s(\xi), \xi\rangle \qquad \text{for } \xi \in \mathrm{supp}_\xi\, a^\lambda.$$

It is clear that h is a homogeneous function of degree one. Consider the cone

$$\Sigma^d := \{(\xi,\tau) \in \mathbb{R}^{d+1} : \tau + h(\xi) = 0\}.$$

Recall that in the planar problem, the spatio-temporal Fourier transform of $\mathfrak{T}^\lambda$ supports inside a neighborhood of the light cone in $\mathbb{R}^3$. In similar fashion, the lemma below shows that the most degenerate part of $m[a^\lambda](D)f$ has its spatio-temporal Fourier transform supported near Σ^d_λ. Consider the multiplier functions

$$\begin{aligned} m^{\lambda,\varepsilon_0}_{\mathrm{main}}(\xi,\tau) &:= \mathcal{F}_t[m[a^{\lambda,\varepsilon_0}_{\mathrm{main}}](\xi,\cdot)](\tau)\eta(\lambda^{-2\varepsilon_0}(\tau + h(\xi))), \\ m^{\lambda,\varepsilon_0}_{\mathrm{err}} &:= \mathcal{F}_t[m[a^{\lambda,\varepsilon_0}_{\mathrm{main}}]] - m^{\lambda,\varepsilon_0}_{\mathrm{main}}. \end{aligned} \tag{3.45}$$

By the non-stationary phase methods, we can see that the contributions coming from $m^{\lambda,\varepsilon_0}_{\mathrm{err}}$ are negligible.

Lemma 3.5.7. *For $p \in [1,\infty]$, we have*

$$\|m^{\lambda,\varepsilon_0}_{\mathrm{err}}(D)\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^{d+1})} \lesssim_N \lambda^{-N}, \text{ for any } N \in \mathbb{N}.$$

Proof. Let $\mathfrak{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}$ denote the kernel associated to $m^{\lambda,\varepsilon_0}_{\mathrm{err}}(D)$. Write

$$\mathfrak{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}(x,t) = (2\pi)^{-(d+1)} \int_{\mathbb{R}^{d+1}} e^{i(\langle x,\xi\rangle + t\cdot\tau)} m^{\lambda,\varepsilon_0}_{\mathrm{err}}(\xi,\tau)\mathrm{d}\xi\mathrm{d}\tau.$$

We aim to prove rapid L^1 decay estimates for $\mathfrak{K}^{\lambda,\varepsilon_0}_{\mathrm{err}}$, or, equivalently, efficient regularity estimates for the multiplier $m^{\lambda,\varepsilon_0}_{\mathrm{err}}$. Unwinding the definitions, we see that $m^{\lambda,\varepsilon_0}_{\mathrm{err}}(\xi,\tau)$ is the same as

$$(1 - \eta(\lambda^{-2\varepsilon_0}(\tau + h(\xi)))) \int_{I\times\mathbb{R}} e^{-it(\tau + \langle\gamma(s),\xi\rangle)} a^\lambda(\xi,s,t)\eta(c_0^{-1}\lambda^{1/2-\varepsilon_0}(s - s(\xi)))\mathrm{d}s\mathrm{d}t. \tag{3.46}$$

By recalling the support properties of η from (3.3), observe that $|\tau + h(\xi)| \geq \lambda^{2\varepsilon_0}$ in the support of $m^{\lambda,\varepsilon_0}_{\mathrm{err}}$. By Taylor's theorem,

$$\tau + \langle\gamma(s), \xi\rangle = \tau + h(\xi) + (s - s(\xi))^2\langle\gamma^{(2)}(s_0), \xi\rangle \text{ for some } s_0 \in [s, s(\xi)].$$

However, in view of the s-localisation and (3.38),

$$(s - s(\xi))^2|\langle\gamma^{(2)}(s_0), \xi\rangle| \leq 8Bc_0^2\lambda^{-1+2\varepsilon_0}\lambda = 8Bc_0^2\lambda^{2\varepsilon_0}$$

in the support of the integrand of (3.46). As c_0 is chosen sufficiently small, the inequality $|\tau+\langle\gamma(s), \xi\rangle| \gtrsim \lambda^{2\varepsilon_0}$ holds in the same support. By integration-by-parts in the t-variable,

$$|\partial_\tau^\beta\partial_\xi^\alpha(m^{\lambda,\varepsilon_0}_{\mathrm{err}}(\xi, \tau))| \lesssim_N \lambda^{-2\varepsilon_0 N} \qquad \text{for } |\alpha| + |\beta| \leq 10,\ N \in \mathbb{N}$$

and $(\xi, \tau) \in \operatorname{supp} m^{\lambda,\varepsilon_0}_{\mathrm{err}}$. The statement of Lemma 3.5.7 is immediate now. $\square$

By combining Lemma 3.5.6 and Lemma 3.5.7, we deduce that

$$\|m[a^\lambda](D)g\|_{L^p(\mathbb{R}^{d+1})} \lesssim \|m^{\lambda,\varepsilon_0}_{\mathrm{main}}(D)g\|_{L^p(\mathbb{R}^{d+1})} + O_N(\lambda^{-N}\|g\|_{L^p(\mathbb{R}^d)}) \qquad \text{for } N \geq 1$$

whenever $g \in L^p(\mathbb{R}^d)$. Therefore, to prove Conjecture 3.5.5, it suffices to estimate the L^p norm of $m^{\lambda,\varepsilon_0}_{\mathrm{main}}$. The key observation here is that the spatio-temporal frequency support properties of the operator is very similar to that of $\mathfrak{T}^\lambda$ from the $\mathbb{R}^2$ setup. In particular, the co-dimension one surface Σ^d_λ here seems to be the equivalent of the truncated light cone for latter setup. In view of this, our attempt to estimate its operator norm will be along the lines of the method explained (which will be referred to as the M–S–S scheme) in §3.4. Let us also note that our attention will be restricted to obtaining (3.37) solely for the case $p = 4$, which is the critical exponent.

3.5.3 Implementing the M–S–S scheme

The first step is to decompose the neighbourhood of the slow decaying cone Σ^d_λ. The decomposition we have here originates from [3] (See [3, Definition 5.1]).

Plank decomposition and the square functions

Let $\{s_\nu \in I : \nu \in \Omega(\lambda^{-1/2+\varepsilon_0})\}$ be a collection of points in I whose elements are separated by a factor of $c_0\lambda^{-\frac{1}{2}+\varepsilon_0}$, so that the index set $\Omega(\lambda^{-1/2+\varepsilon_0})$ has cardinality $O(\lambda^{\frac{1}{2}-\varepsilon_0})$. Define

$$a^{\lambda,\varepsilon_0}_\nu(\xi, s, t) := a^{\lambda,\varepsilon_0}_{\mathrm{main}}(\xi, s, t)\eta(c_0^{-1}\lambda^{1/2-\varepsilon_0}(s(\xi) - s_\nu))) \tag{3.47}$$

and

$$\begin{aligned} m^{\lambda,\varepsilon_0}_\nu(\xi, \tau) &:= \eta[\lambda^{-2\varepsilon_0}(\tau + h(\xi))]\mathcal{F}_t[m[a^{\lambda,\varepsilon_0}_\nu](\xi, \cdot)](\tau) \\ &= m^{\lambda,\varepsilon_0}_{\mathrm{main}}(\xi, \tau)\eta(c_0^{-1}\lambda^{1/2-\varepsilon_0}(s(\xi) - s_\nu)). \end{aligned} \tag{3.48}$$

To parametrise the planks that cover the slow-decaying cone Σ^d, we introduce the notion of the lift of a curve. For $\gamma : I \to \mathbb{R}^d$, its lift Γ is a curve in $\mathbb{R}^{d+1}$ defined by

$$\Gamma(s) = \begin{pmatrix} \int_0^s \gamma(t)dt \\ s \end{pmatrix} \qquad \text{for } s \in I.$$

Fix $s \in I$ and $r \in (0,1)$. Since γ is assumed to be non-degenerate in $\mathbb{R}^d$, the curve Γ is, by definition, non-degenerate in $\mathbb{R}^{d+1}$. Let

$$V_s := \left[\operatorname{span}(\{\Gamma^{(j)}(s) : 1 \le j \le 3\})\right]^{\perp}.$$

Define a box $\Pi(s,r) \subseteq \mathbb{R}^{d+1}$ to be the collection of all $\Xi \in \mathbb{R}^{d+1}$ satisfying the inequalities

$$|\langle \Gamma^{(1)}(s), \Xi \rangle| \le r^2, \tag{3.49}$$
$$|\langle \Gamma^{(2)}(s), \Xi \rangle| \le r, \tag{3.50}$$
$$1/2 \le |\langle \Gamma^{(3)}(s), \Xi \rangle| \le 1, \tag{3.51}$$
$$|\operatorname{proj}_{V_s}(\Xi)| \le 1, \tag{3.52}$$

where $\operatorname{proj}_{\mathrm{V}} : \mathbb{R}^{d+1} \to V$ is the orthogonal projection onto the subspace V. With this definition, it is not hard to verify that[10]

$$\operatorname{supp} m_\nu^{\lambda,\varepsilon_0} \subseteq C_B \lambda \Pi(s_\nu, \lambda^{-1/2+\varepsilon_0}), \tag{3.53}$$

for a constant C_B that depends only on B. Indeed, (3.38) implies that

$$(2B)^{-1}\lambda \le |\langle \Gamma^{(3)}(s_\nu), \Xi \rangle| \le 2B\lambda, \qquad (\xi,\tau) \in \operatorname{supp}_\xi a^\lambda \times \mathbb{R}. \tag{3.54}$$

On the other hand, for any $(\xi,\tau) \in \operatorname{supp} m_\nu^{\lambda,\varepsilon_0}$, Taylor's theorem gives [11]

$$|\langle \Gamma^{(1)}(s_\nu), \Xi \rangle| = |\tau + \langle \gamma(s_\nu), \xi \rangle| \le |\tau + h(\xi)| + 2B\lambda |s_\nu - s(\xi)|^2 \lesssim \lambda^{2\varepsilon_0} \tag{3.55}$$

where we made use of the s-localisation in (3.48) with τ-localisation in (3.45) and (3.38). Similarly, by combining (3.48), (3.38) and (3.42), we have

$$|\langle \Gamma^{(2)}(s_\nu), \Xi \rangle| = |\langle \gamma^{(1)}(s_\nu), \xi \rangle| \le |\langle \gamma^{(1)} \circ s(\xi), \xi \rangle| + 2B\lambda |s_\nu - s(\xi)| \lesssim \lambda^{1/2+\varepsilon_0} \tag{3.56}$$

for any $(\xi,\tau) \in \operatorname{supp} m_\nu^{\lambda,\varepsilon_0}$. Furthermore, the inequality

$$|\operatorname{proj}_{V_s}(\Xi)| \le |\Xi| \lesssim \lambda \qquad \text{for } \Xi \in \operatorname{supp} m_\nu^{\lambda,\varepsilon_0} \tag{3.57}$$

follows from (3.45) and (3.38). Comparing (3.49), (3.50), (3.51), (3.52) with (3.55), (3.56), (3.54) and (3.57), we finally obtain (3.53).

As in §3.4, we now frame a reverse square function problem to move from studying the whole operator to operators micro-localised to these planks. The

[10]Recall that for a set $E \subseteq \mathbb{R}^d$ and $r \in \mathbb{R}$, rE denotes the scaled set $\{y \in \mathbb{R}^d : r^{-1}y \in E\}$.

[11]provided c_0 is chosen sufficiently small compared to B

reverse square function problem is the hardest part of the argument, and we present its statement as a conjecture.

Conjecture 3.5.8. *Let* $0 < r < 1$*. Consider a collection* $\{f_\nu \in \mathcal{S}(\mathbb{R}^{d+1}) : \nu \in \Omega(r)\}$ *such that* $\operatorname{supp} \hat{f}_\nu \subseteq \Pi(s_\nu, r)$*. It follows that for any* $\epsilon > 0$*, there exists* $C_\epsilon > 0$ *such that*

$$\Big\| \sum_{\nu\in\Omega(r)} f_\nu \Big\|_{L^4(\mathbb{R}^{d+1})} \leq C_\varepsilon r^{-\varepsilon} \Big\| (\sum_{\nu\in\Omega(r)} |f_\nu|^2)^{1/2} \Big\|_{L^4(\mathbb{R}^{d+1})}.$$

For the remainder of the section, we may assume the validity of the reverse square function conjecture. Let $g \in \mathcal{S}(\mathbb{R}^d)$. Recalling (3.53), we apply Conjecture 3.5.8 for $r = \lambda^{-1/2+\varepsilon_0}$ and a scaling argument to conclude that

$$\Big\| m^{\lambda,\varepsilon_0}_{\text{main}}(D)g \Big\|_{L^4(\mathbb{R}^{d+1})} \lesssim_{d,\epsilon} \lambda^{(1/2-\varepsilon_0)\epsilon} \Big\| (\sum_{\nu\in\Omega(\lambda^{-1/2+\varepsilon_0})} |m^{\lambda,\varepsilon_0}_\nu(D)g|^2)^{\frac{1}{2}} \Big\|_{L^4(\mathbb{R}^{d+1})}$$

for any $\epsilon > 0$. By Lemma 3.5.7, the right side can be dominated by

$$\lambda^{(1/2-\varepsilon_0)\epsilon} \Big\| (\sum_{\nu\in\Omega(\lambda^{-1/2+\varepsilon_0})} |m[a^{\lambda,\varepsilon_0}_\nu](D)g|^2)^{\frac{1}{2}} \Big\|_{L^4(\mathbb{R}^{d+1})} + O_N(\lambda^{-N})\|g\|_{L^4(\mathbb{R}^d)} \tag{3.58}$$

for any $N \geq 1$.

As before, the main term in (3.58) is estimated via a duality argument, using a forward square function estimate and Theorem 3.5.1. As the duality argument has already been presented in §3.4, we will only identify the correct formulation of the square function and the maximal problem in this case.

Suppose $\tilde{\eta}, \tilde{\beta} \in \mathcal{S}(\mathbb{R}^1)$ are chosen such that $\operatorname{supp} \tilde{\eta} \subseteq [-2, 2]$, $\operatorname{supp} \tilde{\beta} \subseteq [1/4, 4]$, $\tilde{\eta} \cdot \eta = \eta$ and $\tilde{\beta} \cdot \beta = \beta$. Consider the Fourier multiplier operators $\chi^{\lambda,\varepsilon_0}_\nu(D)$ where

$$\chi^{\lambda,\varepsilon_0}_\nu(\xi) := \tilde{\eta}(\lambda^{1/2-\varepsilon_0}(s(\xi) - s_\nu))\tilde{\beta}(\lambda^{-1}|\xi|).$$

In view of the frequency localisations of $m[a^{\lambda,\varepsilon_0}_\nu]$, we frame the appropriate forward square function problem as below:

Theorem 3.5.9. *For any* $\varepsilon > 0$*, there exists* $C_\varepsilon > 0$ *such that*

$$\Big\| \Big(\sum_{\nu\in\Omega(\lambda^{-1/2+\varepsilon_0})} |\chi^{\lambda,\varepsilon_0}_\nu(D)g|^2 \Big)^{\frac{1}{2}} \Big\|_{L^4(\mathbb{R}^d)} \lesssim_\varepsilon \lambda^\varepsilon \|g\|_{L^4(\mathbb{R}^d)} \qquad \text{for any } g \in \mathcal{S}(\mathbb{R}^d).$$

We may briefly discuss how Theorem 3.5.9 can be proved without going into the details. The argument is an easy generalisation of a related proof presented in [3]. In particular, [3, Proposition 8.10] deals with a similar square function estimate when $d = 3$. The authors prove the required estimate by combining two elements; an iteration scheme originating from Carbery–Seeger [10] and the L^2 boundedness of a related maximal function. It is not hard to see that the iteration scheme can be generalised to higher dimensions from the $d = 3$ case it was presented in. The required maximal estimate, on the other hand, is already presented in great generality in the article; in particular, the maximal function considered in [3, Proposition 10.3] is defined in $\mathbb{R}^d$ for $d \geq 2$ (see the cited reference for further details). Because of these observations, the argument presented in [3] can be modified to provide a proof for Theorem 3.5.9. As the focus of the thesis is in a different direction, we omit the full details here.

Now, we can talk about the Nikodym maximal estimates. As before, to see how we get to the definition (3.34) of the geometric maximal function, we must look at the kernel of the operator $m[a_\nu^{\lambda,\varepsilon_0}](D)$.

Kernel estimates and final arguments

For $s \in I$, set $W_s := \mathrm{span}(\gamma^{(1)}(s), \gamma^{(2)}(s))$. Recall that $\{e_1(s), e_2(s)\}$ form an orthonormal basis for W_s.

Let $K_\nu^{\lambda,\varepsilon_0}$ denote the kernel associated to $m[a_\nu^{\lambda,\varepsilon_0}](D)$. The following lemma describes the essential support of the kernel.

Lemma 3.5.10. *Let $\lambda \in 2^{\mathbb{N}}$, $\varepsilon_0 \in (0,1)$ and $\nu \in \Omega(\lambda^{-1/2+\varepsilon_0})$. For any $N \geq 1$, we have*

$$|K_\nu^{\lambda,\varepsilon_0}(x,t)| \lesssim_N \frac{\lambda^{(d-1)+2\varepsilon_0}\rho(t)}{(1+\sum_{j=1}^2 \lambda^{\frac{j}{2}-(j-2)\varepsilon_0}|\langle x-t\gamma(s_\nu), e_j(s_\nu)\rangle| + \lambda|\mathrm{proj}_{W_{s_\nu}^\perp}(x-t\gamma(s_\nu))|)^N}$$

whenever $(x,t) \in \mathbb{R}^{d+1}$. In particular,

$$\|K_\nu^{\lambda,\varepsilon_0}\|_{L^1(\mathbb{R}^{d+1})} \lesssim_{\varepsilon_0} \lambda^{-1/2+C\varepsilon_0} \qquad \textit{for some constant } C > 1.$$

Proof. Using the definition, we write

$$K_\nu^{\lambda,\varepsilon_0}(x,t) = \int_{\mathbb{R}\times\hat{\mathbb{R}}^d} e^{i\langle x-t\gamma(s),\xi\rangle} a_\nu^{\lambda,\varepsilon_0}(\xi,s,t)\mathrm{d}s\mathrm{d}\xi.$$

For $s \in I$, recall that $\{e_3(s), \cdots, e_d(s)\} \subseteq \mathbb{R}^d$ form an orthonormal basis for $W_s^\perp$. Fix $1 \leq j \leq d$ and consider the differential operator

$$\mathcal{L}_{\nu,j} := [i\langle x - t\gamma(s_\nu), e_j(s_\nu)\rangle]^{-1}\,\partial_{e_j(s_\nu)},$$

so that $\mathcal{L}_{\nu,j}(e^{i\langle x-t\gamma(s_\nu),\xi\rangle}) = e^{i\langle x-t\gamma(s_\nu),\xi\rangle}$. By repeated integration-by-parts,

$$\begin{aligned}
|K_\nu^{\lambda,\varepsilon_0}(x,t)| =& \left|\int e^{i\langle x-t\gamma(s_\nu),\xi\rangle}(\mathcal{L}_{\nu,j}^*)^N\left[e^{it\langle\gamma(s_\nu)-\gamma(s),\xi\rangle}a_\nu^{\lambda,\varepsilon_0}(\xi,s,t)\right]\mathrm{d}\xi\mathrm{d}s\right| \\
\lesssim& |\{\xi \in \hat{\mathbb{R}}^d : |\xi| \approx \lambda, |s(\xi) - s_\nu| \lesssim \lambda^{-\frac{1}{2}+\varepsilon_0}\}| \cdot \\
& \sup_{\xi\in\mathrm{supp}_\xi\ a^\lambda}\left|\int_I(\mathcal{L}_{\nu,j}^*)^N\left[e^{it\langle\gamma(s_\nu)-\gamma(s),\xi\rangle}a_\nu^{\lambda,\varepsilon_0}(\xi,s,t)\right]\mathrm{d}s\right|.
\end{aligned}$$

Using the fact that $\xi \to s(\xi)$ is homogeneous of degree 0, it is not hard to prove the estimate

$$|\{\xi \in \hat{\mathbb{R}}^d : |\xi| \approx \lambda,\ |s(\xi) - s_\nu| \lesssim \lambda^{-\frac{1}{2}+\varepsilon_0}\}| \lesssim \lambda^{(d-1)} \cdot \lambda^{\frac{1}{2}+\varepsilon_0}.$$

Combining this with the s-localisation,

$$|K_\nu^{\lambda,\varepsilon_0}(x,t)| \lesssim \lambda^{(d-1)+\frac{1}{2}+\varepsilon_0-(\frac{1}{2}-\varepsilon_0)} \sup_{(\xi,s,t)\in \text{supp } a^\lambda} \left|(\mathcal{L}_{\nu,j}^*)^N \left[e^{it\langle\gamma(s_\nu)-\gamma(s),\xi\rangle} a_\nu^{\lambda,\varepsilon_0}(\xi,s,t)\right]\right| \tag{3.59}$$

Fix $\xi \in \text{supp}_\xi\, a^\lambda$. Our goal now is to estimate the quantities $\partial_{e_j(s_\nu)}\left[e^{it\langle\gamma(s_\nu)-\gamma(s),\xi\rangle}\right]$ and $\partial_{e_j(s_\nu)}\left[a_\nu^{\lambda,\varepsilon_0}(\xi,s,t)\right]$. By Taylor's theorem,

$$\gamma(s)-\gamma(s_\nu) = \gamma^{(1)}(s_\nu)(s-s_\nu) + \gamma^{(2)}(s_\nu)(s-s_\nu)^2 + \gamma^{(3)}(s^*)(s-s_\nu)^3,$$

for some $s^* \in (s_\nu, s)$. Since $\gamma^{(i)}(s) \in \text{span}\{e_1(s),\cdots,e_i(s)\}$ by construction, we have

$$\langle\gamma^{(i)}(s), e_j(s)\rangle = 0 \qquad \text{whenever } i \le j-1.$$

Therefore,[12]

$$|\langle\gamma(s)-\gamma(s_\nu), e_j(s_\nu)\rangle| \lesssim_\gamma \begin{cases} |s-s_\nu|^j & \text{when } j=1 \text{ and } 2\,, \\ |s-s_\nu|^3 & \text{when } j\ge 3. \end{cases}$$

Consequently, whenever $s \in \text{supp}_s\, a_\nu^{\lambda,\varepsilon_0}$, we deduce that

$$|\partial_{e_j(s_\nu)}\left(e^{it\langle\gamma(s_\nu)-\gamma(s),\xi\rangle}\right)| \lesssim \lambda^{-\min\{j,3\}(1/2+\varepsilon_0)} \quad \text{for } (\xi,s,t)\in \text{supp } a_\nu^{\lambda,\varepsilon_0}. \tag{3.60}$$

To estimate the partial derivatives of the symbol, we compute

$$\partial_{e_j(s_\nu)}(\beta(\lambda^{-1}|\xi|)) = \lambda^{-1}(\partial_{e_j(s_\nu)}\beta)(\lambda^{-1}|\xi|), \tag{3.61}$$

$$\partial_{e_j(s_\nu)}(\eta(\lambda^{1/2-\varepsilon_0}(s(\xi)-s_\nu))) = \lambda^{1/2-\varepsilon_0}\partial_{e_j(s_\nu)}(s(\xi))(\partial_{e_j(s_\nu)}\eta)(\lambda^{1/2-\varepsilon_0}(s(\xi)-s_\nu)) \tag{3.62}$$

By implicitly differentiating (3.42), we obtain

$$\partial_{e_j(s_\nu)}(s(\xi)) = -\frac{\langle\gamma^{(1)}\circ s(\xi), e_j(s_\nu)\rangle}{\langle\gamma^{(2)}\circ s(\xi),\xi\rangle} \qquad \text{for } \xi\in \text{supp}_\xi\, a^\lambda \text{ and } 1\le j\le d.$$

By Taylor's theorem,

$$\gamma^{(1)}\circ s(\xi) = \gamma^{(1)}(s_\nu) + \gamma^{(2)}(s_\nu)(s(\xi)-s_\nu) + \gamma^{(3)}(s^*)(s(\xi)-s_\nu)^2$$

for some $s^* \in (s_\nu, s)$. In view of the s-localisation in (3.47),

$$|\langle\gamma^{(1)}\circ s(\xi), e_j(s_\nu)\rangle| \lesssim \lambda^{-\min\{(j-1),2\}(1/2-\varepsilon_0)} \qquad \text{for } (\xi,s)\in \text{supp}_{\xi,s}\, a_\nu^{\lambda,\varepsilon_0}. \tag{3.63}$$

[12] This is only true when $|s-s_\nu|$ is smaller than 1. However, by decomposing the symbol into many parts at the start, we can always ensure that this is the case.

Combining (3.38), (3.62) and (3.63), we see that

$$\begin{aligned}|\partial_{e_j(s_\nu)}[\eta(\lambda^{1/2-\varepsilon_0}(s(\xi)-s_\nu))]| &\lesssim \lambda^{1/2-\varepsilon_0}\lambda^{-1}\lambda^{-\min\{(j-1),2\}(1/2-\varepsilon_0)} \\ &= \lambda^{-\min\{j/2,3/2\}+\min\{j-2,1\}\varepsilon_0}. \end{aligned} \tag{3.64}$$

Now, combining (3.60), (3.61) and (3.64), we obtain

$$|\partial_{e_j(s_\nu)}(e^{it\langle\gamma(s_\nu)-\gamma(s),\xi\rangle}a_\nu^{\lambda,\varepsilon_0}(\xi,s,t))| \lesssim \begin{cases}\lambda^{-j/2+(j-2)\varepsilon_0} & j=1,2\\ \lambda^{-1} & j\geq 3,\end{cases}$$

for any $(\xi,s,t)\in a_\nu^{\lambda,\varepsilon_0}$. Feeding this estimate back into (3.59), we can conclude the decay estimate of $K_\nu^{\lambda,\varepsilon_0}$ in the Lemma. The estimate on the L^1 norm of $K_\nu^{\lambda,\varepsilon_0}$ follows immediately as a consequence. □

In view of Lemma 3.5.10, the function $K_\nu^{\lambda,\varepsilon_0}$ is essentially supported in the tube $T_r(s_\nu)$ defined by (3.33), for $r=\lambda^{-1}$. In view of this, we see that the maximal function $\mathcal{N}_r$ is the higher dimensional equivalent of $\mathcal{N}^{\text{geom}}_{r,\text{plane}}$.

Now, by arguments very similar to the one presented in §3.4, we see that Theorem 3.5.1, Theorem 3.5.9 and the kernel estimate from Lemma 3.5.10 implies that

$$\Big\|\Big(\sum_{\nu\in\Omega(\lambda^{-1/2+\varepsilon_0})}|m[a_\nu^{\lambda,\varepsilon_0}](D)g|^2\Big)^{\frac12}\Big\|_{L^4(\mathbb{R}^{d+1})} \lesssim_{\varepsilon_0} \lambda^{C\epsilon_0}\|g\|_{L^4(\mathbb{R}^d)}.$$

In view of the discussions until (3.58), we see that this completes the conditional proof of Conjecture 3.5.4, as $\varepsilon_0\in(0,1)$ is chosen arbitrarily.

Chapter 4

L^2 estimates for a Nikodym maximal function associated to space curves

The content of this chapter is based on author's own work in [49].

4.1 Introduction

As mentioned in §3.5, the goal of this chapter is study and estimate a higher dimensional Nikodym maximal function.

Let $I := [-1, 1]$ and $\gamma := I \to \mathbb{R}^d$ be a C^∞ non-degenerate curve. In other words,

$$\det \begin{pmatrix} \gamma^{(1)}(s) & \cdots & \gamma^{(d)}(s) \end{pmatrix} \neq 0 \quad \text{for all } s \in I. \tag{4.1}$$

For each $s \in I$ and $1 \leq i \leq d$, recall the definition of a Frenet frame basis vector $e_i(s)$ from the initial discussions of §3.5. For $\mathbf{r} = (r_1, \dots, r_d) \in (0, 1)^d$, we consider an anisotropic tube in $\mathbb{R}^{d+1}$ in the direction of $\begin{pmatrix} \gamma(s) \\ 1 \end{pmatrix}$, whose axis-lengths are determined by $\mathbf{r}$. Define

$$T_{\mathbf{r}}(s) := \big\{(y, t) \in \mathbb{R}^d \times I : |\langle y - t\gamma(s), e_j(s)\rangle| \leq r_j \text{ for } \ 1 \leq j \leq d\big\}.$$

We define the corresponding averaging and maximal operator as

$$\mathcal{A}_{\mathbf{r}}^\gamma g(x, s) := \frac{1}{|T_{\mathbf{r}}(s)|} \int_{T_{\mathbf{r}}(s)} g(x - y, t)\mathrm{d}y\mathrm{d}t \qquad \text{for } (x, s) \in \mathbb{R}^d \times I \tag{4.2}$$

and

$$\mathcal{N}_{\mathbf{r}}^\gamma g(x) := \sup_{s \in I} |\mathcal{A}_{\mathbf{r}}^\gamma g(x, s)| \qquad \text{for } x \in \mathbb{R}^d,$$

whenever $g \in L^1_{\mathrm{loc}}(\mathbb{R}^{d+1})$.

The main result in this chapter addresses the L^p boundedness problem for $\mathcal{N}_{\mathbf{r}}^\gamma$

under mild conditions on $\mathbf{r}$.

Definition 4.1.1. A d-tuple $\mathbf{r} = (r_1, \ldots, r_d) \in (0,1)^d$ is said to be *admissible* if

$$r_d \leq \cdots \leq r_1 \leq r_2^{1/2} \qquad \text{and} \qquad r_j \leq r_i^{\frac{k-j}{k-i}} r_k^{\frac{j-i}{k-i}} \quad \text{for } 1 \leq i \leq j \leq k \leq d. \tag{4.3}$$

Our main theorem is as follows.

Theorem 4.1.2. *Let $\mathbf{r} = (r_1, \ldots, r_d) \in (0,1)^d$ be admissible. There exists $C_{d,\gamma} > 0$ such that*

$$\|\mathcal{N}_{\mathbf{r}}^{\gamma}\|_{L^2(\mathbb{R}^{d+1}) \to L^2(\mathbb{R}^d)} \leq C_{d,\gamma} (\log r_d^{-1})^{d/2}.$$

By interpolating with the trivial bound at L^∞, we estimate the L^p operator norm for the maximal function as $O((\log r_d^{-1})^{d/p})$ for $2 \leq p \leq \infty$. This is sharp in the sense that the L^p operator norm has polynomial blowup in δ^{-1} for $1 \leq p < 2$. The result is new for $d \geq 4$. The theorem also slightly strengthens the known estimates for $d = 2$ and $d = 3$ (see [37, Lemma 1.4] and [3, Proposition 5.5], respectively) by improving the dependence on r_d^{-1}.

We have already seen that a specific case (see the special cases listed below) of the maximal function arises in the study of local smoothing problems in higher dimensions. It is also natural to consider $\mathcal{N}_{\mathbf{r}}^{\gamma}$ as a variant of the classical Nikodym maximal function considered in [15]. The main difference lies in the dimensional setup of the problem: by the above definition, $\mathcal{N}_{\mathbf{r}}^{\gamma}$ maps functions on $\mathbb{R}^{d+1}$ to functions on $\mathbb{R}^d$, whereas the classical operator considered in [15] is a mapping between functions on the same Euclidean space. A detailed discussion is included in §4.5 where this relation is further explored.

Before proceeding to the main section containing the proof of Theorem 4.1.2, let us discuss some of its special cases:

(i) **Isotropic case:** For $0 < \delta < 1$ and $s \in I$, consider a δ-tube in $\mathbb{R}^{d+1}$ in the direction of $\gamma(s)$, defined as

$$T_\delta(s) := \{(y,t) \in \mathbb{R}^d \times I : |y - t\gamma(s)| \leq \delta\}. \tag{4.4}$$

As before, we introduce the corresponding averaging and maximal operator as

$$\mathcal{A}_\delta^\gamma g(x,s) := \frac{1}{|T_\delta(s)|} \int_{T_\delta(s)} g(x-y,t) \mathrm{d}y \mathrm{d}t \qquad \text{for } (x,s) \in \mathbb{R}^d \times I$$

and

$$\mathcal{N}_\delta^\gamma g(x) := \sup_{s \in I} |\mathcal{A}_\delta^\gamma g(x,s)| \qquad \text{for } x \in \mathbb{R}^d \tag{4.5}$$

whenever $g \in L^1_{\mathrm{loc}}(\mathbb{R}^{d+1})$.

By a simplified version of the argument presented in this chapter, the following result has been proved in [49].

Theorem 4.1.3. *Let $\gamma : I \to \mathbb{R}^d$ be a non-degenerate curve. There exists $C_{d,\gamma} > 0$ such that*

$$\|\mathcal{N}_\delta^\gamma\|_{L^2(\mathbb{R}^{d+1})\to L^2(\mathbb{R}^d)} \le C_{d,\gamma}(\log\delta^{-1})^{d/2} \qquad \text{for all } 0<\delta<1.$$

Theorem 4.1.2 is a stronger version of Theorem 4.1.3. To see this, we first note that $\mathbf{r} = (\delta, \dots, \delta)$ is admissible by (4.3). By applying Theorem 4.1.2 for this special case, Theorem 4.1.3 follows easily.

(ii) **Anisotropic case:** Let $1 \le L \le d$ and $r \in (0,1)$. Suppose

$$r_i = \begin{cases} r^i & \text{when } 1 \le i \le L, \\ r^L & \text{when } L < i \le d. \end{cases}$$

With these choices, it is not hard to check that $\mathbf{r}$ is admissible. But first, observe that by setting $L = 2$, we obtain Theorem 3.5.1 as a consequence of Theorem 4.1.2. Furthermore, if $d = L = 2$, we also reproduce Theorem 3.4.4 from the previous chapter.

In verifying the admissibility of $\mathbf{r}$, the only non-trivial estimate to check is

$$r_j \le r_i^{\frac{k-j}{k-i}} r_k^{\frac{j-i}{k-i}} \qquad \text{for } 1 \le i \le j \le k \le d. \tag{4.6}$$

Its verification is done by cases.

Case a) Let $i, j, k \le L$ or $i, j, k \ge L$. In both cases, (4.6) follows from direct computations.

Case b) Let $i \le L$ and $j, k \ge L$. Here, (4.6) follows from the inequality

$$\frac{i(k-j) + L(j-i)}{k-i} \le \frac{L(k-j) + L(j-i)}{k-i} = L.$$

Case c) Let $i, j \le L$ and $k \ge L$. We can see that (4.6) now follows from the inequality

$$\frac{i(k-j) + L(j-i)}{k-i} \le \frac{i(k-j) + k(j-i)}{k-i} = j,$$

completing the verification of (4.6) in all the cases.

4.2 Initial reductions and Sobolev embedding

Many components of our proof are inspired from the existing literature [37, 3] on the maximal function. However, our method of proof differs from the cited works in two key respects. First, we use a fractional Sobolev embedding argument to dominate the maximal function by a Fourier integral operator (see

Proposition 4.2.1). This allows us to fully access orthogonality in the subsequent decomposition. Secondly, we use an induction scheme, which hides the complexity of the root analysis in [3]. The induction is motivated by [32], where a (more complex) induction argument is used to investigate the local smoothing problem associated to averages along curves in $\mathbb{R}^d$.

4.2.1 Initial reductions

Let $I := [-1,1]$ and $\gamma\colon I \to \mathbb{R}^d$ be a non-degenerate curve, as in §4.1. We begin by replacing the classical averaging operators by Fourier integral operators. Given $a \in L^\infty(\hat{\mathbb{R}}^d \times I \times I)$, consider

$$\mathcal{A}[a,\gamma]g(x,s) := \int_I \int_{\mathbb{R}^d} e^{i\langle x - t\gamma(s), \xi\rangle} a(\xi,s,t)\mathcal{F}_x(g)(\xi,t)\mathrm{d}\xi\mathrm{d}t \qquad \text{for } g \in \mathcal{S}(\mathbb{R}^{d+1}), \tag{4.7}$$

where $\mathcal{F}_x(g)(\xi,t)$ denotes $\mathcal{F}(g(\,\cdot\,,t))(\xi)$, the Fourier transform of g in x only. Define the associated maximal operator

$$\mathcal{N}[a,\gamma]g(x) := \sup_{s\in I} |\mathcal{A}[a,\gamma]g(x,s)|.$$

Choose a function $\psi \in C_c^\infty(\mathbb{R})$ with $\operatorname{supp}\psi \subseteq [-1,1]$ such that its inverse Fourier transform $\check{\psi}$ is non-negative and $\check{\psi}(y) \gtrsim 1$ whenever $|y| \leq 1$. Let $\tilde{\chi}_I$ be a non-negative smooth function that satisfies $\tilde{\chi}_I(x) = 1$ for all $x \in I$ and $\tilde{\chi}_I(x) = 0$ when $x \notin [-2,2]$. Define

$$a_{\mathbf{r}}(\xi,s,t) := \prod_{j=1}^d \psi(\langle \xi, e_j(s)\rangle r_j)\tilde{\chi}_I(s)\tilde{\chi}_I(t). \tag{4.8}$$

Let $K_{\mathbf{r}}$ denote the kernel of the averaging operator $\mathcal{A}_{\mathbf{r}}^\gamma$ defined in (4.2). In particular,

$$K_{\mathbf{r}}(x,s,t) := \frac{1}{|T_{\mathbf{r}}(s)|}\chi_{T_{\mathbf{r}}(s)}(x,t).$$

By integral formula for the inverse Fourier transform and a change of variable,

$$K_{\mathbf{r}}(x,s,t) \lesssim_d \int_{\mathbb{R}^d} e^{i\langle x - t\gamma(s),\xi\rangle} a_{\mathbf{r}}(\xi,s,t)\mathrm{d}\xi\mathrm{d}t.$$

Thus, the pointwise estimate

$$|\mathcal{A}_{\mathbf{r}}^\gamma g(x,s)| \lesssim_d |\mathcal{A}[a_{\mathbf{r}},\gamma]g(x,s)|$$

holds. It is therefore enough to bound the operator $\mathcal{N}[a_{\mathbf{r}},\gamma]$.

We now perform an endpoint Sobolev embedding result to replace the L_s^∞ norm in the maximal function with an L_s^2 norm. Here we write

$$\mathfrak{D}_s\mathcal{A}[a,\gamma] := (1+\sqrt{-\partial_s^2})^{1/2}\mathcal{A}[a,\gamma],$$

where a and γ are as above and $(1+\sqrt{-\partial_s^2})^{1/2}$ is the fractional differential operator in s with multiplier $(1+|\sigma|)^{1/2}$.

Proposition 4.2.1. *For γ a nondegenerate curve, a d-tuple $\mathbf{r} = (r_1, \dots, r_d)$ with $r_{i+1} \le r_i$ and $a_{\mathbf{r}}$ as defined in (4.8), we have*

$$\|\mathcal{N}[a_{\mathbf{r}}, \gamma]g\|_{L^2(\mathbb{R}^d)} \lesssim_{\gamma,d} |\log r_d|^{1/2} \|\mathfrak{D}_s \mathcal{A}[a_{\mathbf{r}}, \gamma]g\|_{L^2(\mathbb{R}^{d+1})} + \|g\|_{L^2(\mathbb{R}^{d+1})}$$

for all $g \in \mathcal{S}(\mathbb{R}^{d+1})$.

Before proceeding to its proof, we make a brief remark about the statement of the proposition. Recall from the classical Sobolev embedding theorem[1] that the embedding $W^{1/2+\epsilon,2}(\mathbb{R}) \subseteq L^\infty(\mathbb{R})$ holds if and only if $\epsilon > 0$. Therefore, Proposition 4.2.1 is an endpoint result in the sense that it essentially embeds a subset of the Sobolev space $W^{1/2,2}$ in L^∞ (both in the s variable). The key observation is that the Fourier transform (in the s variable) of the function we consider in the proposition is essentially localised in the dual variable of s (see (4.9) in the proof below). This is what allows us to achieve an embedding of this form with a permissible loss in the operator norm.

Proof of Proposition 4.2.1. Let $\tilde{\chi} : \mathbb{R} \to [0,1]$ satisfy $\tilde{\chi}(\sigma) = 1$ for all $\sigma \in (-Cr_d^{-2}, Cr_d^{-2})$ and $\tilde{\chi}(\sigma) = 0$ when $\sigma \notin (-2Cr_d^{-2}, 2Cr_d^{-2})$. The constant $C = C(d,\gamma)$ is chosen large enough to satisfy the requirements of the forthcoming argument. Defining

$$\mathcal{A}_{\mathrm{main}}[a_{\mathbf{r}}, \gamma] := \tilde{\chi}(\tfrac{1}{i}\partial_s) \circ \mathcal{A}[a_{\mathbf{r}}, \gamma] \qquad \text{and} \qquad \mathcal{A}_{\mathrm{err}}[a_{\mathbf{r}}, \gamma] := \mathcal{A}[a_{\mathbf{r}}, \gamma] - \mathcal{A}_{\mathrm{main}}[a_{\mathbf{r}}, \gamma],$$

it suffices to prove

$$\|\mathcal{A}_{\mathrm{main}}[a_{\mathbf{r}}, \gamma]g\|_{L^2_x L^\infty_s(\mathbb{R}^d \times I)} \lesssim_d |\log r_d|^{1/2} \|\mathfrak{D}_s \mathcal{A}[a_{\mathbf{r}}, \gamma]g\|_{L^2(\mathbb{R}^{d+1})}, \tag{4.9}$$

$$\|\mathcal{A}_{\mathrm{err}}[a_{\mathbf{r}}, \gamma]\|_{L^2_x L^\infty_s(\mathbb{R}^d \times I)} \lesssim_{\gamma,d} \|g\|_{L^2(\mathbb{R}^{d+1})} \tag{4.10}$$

for all $g \in \mathcal{S}(\mathbb{R}^{d+1})$.

To prove (4.9), fix $g \in \mathcal{S}(\mathbb{R}^{d+1})$ and write

$$\mathcal{A}_{\mathrm{main}}[a_{\mathbf{r}}, \gamma]g(x,s) = \tilde{\chi}_1(\tfrac{1}{i}\partial_s) \circ \mathfrak{D}_s \mathcal{A}[a_{\mathbf{r}}, \gamma]g(x,s) \qquad \text{for } (x,s) \in \mathbb{R}^d \times I,$$

where $\tilde{\chi}_1(\sigma) := (1+|\sigma|)^{-1/2}\tilde{\chi}(\sigma)$. Temporarily fix $x \in \mathbb{R}^d$. The above expression can be written as a convolution product in s variable between $\mathcal{F}_s^{-1}(\tilde{\chi}_1)$ and $\mathfrak{D}_s \mathcal{A}[a_{\mathbf{r}}, \gamma]g(x, \cdot\,)$. Using Young's inequality, Plancherel's theorem and by noting that the L^2 norm $\tilde{\chi}_1$ is $O(|\log r_d|^{1/2})$, we obtain

$$\|\mathcal{A}_{\mathrm{main}}[a_{\mathbf{r}}, \gamma]g(x, \cdot\,)\|_{L^\infty_s(I)} \lesssim |\log r_d|^{1/2} \|\mathfrak{D}_s \mathcal{A}[a_{\mathbf{r}}, \gamma]g(x, \cdot\,)\|_{L^2_s(\mathbb{R})}.$$

Combining Fubini's theorem with the above estimate for each $x \in \mathbb{R}^d$, we obtain (4.9).

[1] which can be found in [24, Theorem 1.3.5].

To prove (4.10), write

$$\begin{aligned}\mathcal{A}_{\mathrm{err}}[a_{\mathbf{r}},\gamma]g &= (1+\sqrt{-\partial_s^2})^{-1}\circ(1+\sqrt{-\partial_s^2})\circ(1-\tilde{\chi})(\tfrac{1}{i}\partial_s)\circ\mathcal{A}[a_{\mathbf{r}},\gamma]g\\ &= \tilde{\chi}_2(\tfrac{1}{i}\partial_s)\circ\tilde{\chi}_3(\tfrac{1}{i}\partial_s)\circ\mathcal{A}[a_{\mathbf{r}},\gamma]g\end{aligned}$$

where

$$\tilde{\chi}_2(\sigma):=(1+|\sigma|)^{-1}(1-\tilde{\chi}(\sigma))^{1/2} \quad \text{and} \quad \tilde{\chi}_3(\sigma):=(1+|\sigma|)(1-\tilde{\chi}(\sigma))^{1/2}$$

for $\sigma\in\mathbb{R}$. Note that $(1+|\sigma|)^{-1}(1-\tilde{\chi}(\sigma))^{1/2}$ has uniformly bounded L^2 norm (in r_d). Thus, an application of Young's convolution inequality gives

$$\|\mathcal{A}_{\mathrm{err}}[a_{\mathbf{r}},\gamma]g(x,\cdot)\|_{L^\infty_s(I)}\lesssim\|\tilde{\chi}_3(\tfrac{1}{i}\partial_s)\circ\mathcal{A}[a_{\mathbf{r}},\gamma]g(x,\cdot)\|_{L^2_s(\mathbb{R})} \qquad \text{for } x\in\mathbb{R}^d.$$

Integrating in x using Fubini's theorem,

$$\|\mathcal{A}_{\mathrm{err}}[a_{\mathbf{r}},\gamma]g\|_{L^2_xL^\infty_s(\mathbb{R}^d\times I)}\lesssim\|\tilde{\chi}_3(\tfrac{1}{i}\partial_s)\circ\mathcal{A}[a_{\mathbf{r}},\gamma]g\|_{L^2(\mathbb{R}^{d+1})}.$$

By Plancherel's theorem, the quantity on the right can be estimated from above by L^2 norm of the function $\mathcal{B}_{\mathrm{err}}[a_{\mathbf{r}},\gamma,\tilde{\chi}_3]g$, where

$$\mathcal{B}_{\mathrm{err}}[a_{\mathbf{r}},\gamma,\tilde{\chi}_3]g(\xi,\sigma):=\int_I b_{\mathbf{r}}(\xi,\sigma,t)\mathcal{F}_x(g)(\xi,t)\mathrm{d}t$$

for

$$b_{\mathbf{r}}(\xi,\sigma,t):=\tilde{\chi}_3(\sigma)\int_I e^{-i(\sigma s+t\langle\gamma(s),\xi\rangle)}a_{\mathbf{r}}(\xi,s,t)\mathrm{d}s. \tag{4.11}$$

By Minkowski's integral inequality and Plancherel's theorem,

$$\|\mathcal{B}_{\mathrm{err}}[a_{\mathbf{r}},\gamma,\tilde{\chi}_3]g\|_{L^2(\mathbb{R}^{d+1})}\lesssim\|b_{\mathbf{r}}\|_{L^\infty_{\xi,t}L^2_\sigma(\mathbb{R}^d\times I\times\mathbb{R})}\|g\|_{L^2(\mathbb{R}^d\times I)}.$$

Thus, the proof of (4.10) boils down to the estimate $\|b_{\mathbf{r}}(\xi,\,\cdot\,,t)\|_{L^2_\sigma(\mathbb{R})}\lesssim 1$ uniformly in $(\xi,t)\in\mathbb{R}^d\times I$. Since $r_d\le r_i$ for all $1\le i\le d$, it follows from (4.8) that

$$|\xi|\lesssim_d r_d^{-1} \qquad \text{for } \xi\in\operatorname{supp}_\xi a_{\mathbf{r}}. \tag{4.12}$$

Thus, by choosing C large (in particular, it suffices to have $C=10d^{1/2}\|\gamma\|_{C^1}$), we have

$$|\sigma+t\langle\gamma'(s),\xi\rangle|\sim|\sigma| \quad \text{whenever } (\xi,s,t)\in\operatorname{supp} a_{\mathbf{r}} \text{ and } \sigma\in\ \operatorname{supp}\tilde{\chi}_3.$$

Furthermore, using (4.8), (4.12) and the support properties of $\tilde{\chi}_3$, one obtains the derivative bounds

$$|\partial_s^\beta a_{\mathbf{r}}(\xi,s,t)|\lesssim_{\beta,\gamma,d} r_d^{-\beta}\lesssim_{\beta,\gamma,d}|\sigma|^{\beta/2} \qquad \text{for } \beta\in\mathbb{N},$$

whenever $(\xi,s,t)\in\operatorname{supp} a_{\mathbf{r}}$ and $\sigma\in\operatorname{supp}\tilde{\chi}_3$. Combining these observations, we apply integration-by-parts to estimate the oscillatory integral in (4.11). In particular,

$$b_{\mathbf{r}}(\xi,\sigma,t) = O_{N,\gamma}((1+|\sigma|)^{-N}) \qquad (\xi,t)\in\mathbb{R}^d\times I \text{ and } N\geq 1.$$

It is evident that the required L^2_σ estimate for $b_{\mathbf{r}}(\xi,\cdot,t)$, uniformly in ξ and t variables follows from this rapid decay. This completes the proof of (4.10). □

Proposition 4.2.1 reduces the analysis to estimating the operator $\mathfrak{D}_s\mathcal{A}[a_{\mathbf{r}},\gamma]$. Now, we radially decompose the frequency space using the smooth Littlewood–Paley functions.

Recall the functions $\eta,\beta\in C_c^\infty(\mathbb{R})$ satisfying (3.3) and (3.4). For $\lambda\in\{0\}\cup 2^{\mathbb{N}}$, introduce the dyadic symbols

$$a_{\mathbf{r}}^\lambda(\xi,s,t) := \begin{cases} a_{\mathbf{r}}(\xi,s,t)\eta(|\xi|) & \text{if } \lambda=0,\\ a_{\mathbf{r}}(\xi,s,t)\beta(|\xi|/\lambda) & \text{if } \lambda\in 2^{\mathbb{N}}.\end{cases} \tag{4.13}$$

Theorem 4.1.2 is a consequence of the following result.

Proposition 4.2.2. *Suppose $\lambda\in\{0\}\cup 2^{\mathbb{N}}$ and let $\mathbf{r}$ be admissible. Then,*

$$\|\mathfrak{D}_s\mathcal{A}[a_{\mathbf{r}}^\lambda,\gamma]\|_{L^2(\mathbb{R}^{d+1})\to L^2(\mathbb{R}^{d+1})} \lesssim_{d,\gamma} (\log(2+\lambda))^{(d-1)/2}.$$

Proposition 4.2.2 $\Longrightarrow$ Theorem 4.1.2 . Let $\tilde\eta,\tilde\beta\in C_c^\infty(\mathbb{R})$ be two non-negative functions such that $\tilde\eta(r)=1$ for $r\in\operatorname{supp}\eta$, $\tilde\beta(r)=1$ for $r\in\operatorname{supp}\beta$ and

$$\tilde\eta(r)+\sum_{\lambda\in 2^{\mathbb{N}}}\tilde\beta(r/\lambda)\lesssim 1 \qquad \text{for all } r\in\mathbb{R}.$$

For $g\in\mathcal{S}(\mathbb{R}^{d+1})$, define

$$g^\lambda := \begin{cases}\tilde\eta\left(|\tfrac{1}{i}\partial_x|\right)g & \text{if } \lambda=0,\\ \tilde\beta\left(|\tfrac{1}{i}\partial_x/\lambda|\right)g & \text{if } \lambda\in 2^{\mathbb{N}}.\end{cases}$$

It is clear from the definitions that $\mathfrak{D}_s\mathcal{A}[a_{\mathbf{r}}^\lambda,\gamma]g=\mathfrak{D}_s\mathcal{A}[a_{\mathbf{r}}^\lambda,\gamma]g^\lambda$. By Plancherel's theorem and the support properties of the $a_{\mathbf{r}}^\lambda$, we have

$$\|\mathfrak{D}_s\mathcal{A}[a_{\mathbf{r}},\gamma]g\|^2_{L^2(\mathbb{R}^{d+1})}\lesssim\sum_{\lambda\in\{0\}\cup 2^{\mathbb{N}}}\|\mathfrak{D}_s\mathcal{A}[a_{\mathbf{r}}^\lambda,\gamma]g^\lambda\|^2_{L^2(\mathbb{R}^{d+1})}.$$

Applying Proposition 4.2.2 for each λ and observing that $a_{\mathbf{r}}^\lambda=0$ when $r_d^{-1}\lesssim_d\lambda$, we obtain

$$\begin{aligned}\|\mathfrak{D}_s\mathcal{A}[a_{\mathbf{r}},\gamma]g\|^2_{L^2(\mathbb{R}^{d+1})} &\lesssim_{d,\gamma}\sum_{\lambda\in\{0\}\cup 2^{\mathbb{N}}}(\log(2+\lambda))^{d-1}\|g^\lambda\|^2_{L^2(\mathbb{R}^{d+1})}\\ &\lesssim_{d,\gamma}(\log r_d^{-1})^{d-1}\|g\|^2_{L^2(\mathbb{R}^{d+1})}.\end{aligned}$$

Combining the above inequality with Proposition 4.2.1, we deduce Theorem 4.1.2. □

The multiplier associated to $\mathfrak{D}_s\mathcal{A}[a^0_{\mathbf{r}}, \gamma]$ is a bounded function and so the $\lambda = 0$ case of Proposition 4.2.2 is immediate. More interesting cases arise when $\lambda \in 2^{\mathbb{N}}$.

4.3 The proof of Proposition 4.2.2

4.3.1 Some basic definitions

Our proof is based on an induction argument. We begin with a series of definitions that describe the general forms of curves and symbols we encounter at several stages of the induction.

Definition 4.3.1. For $1 \le L \le d$, Define $\mathfrak{G}(B, L)$ to be the collection of all curves $\gamma : I \to \mathbb{R}^d$ such that for all $s \in I$, we have

$$\|\gamma\|_{C^{2d}(I)} \le B \qquad \text{and} \qquad \left|\det\left(\gamma^{(1)}(s) \quad \cdots \quad \gamma^{(L)}(s)\right)\right| \ge B^{-1}, \tag{4.14}$$

where the square of the determinant is interpreted as the sum of squares of its $L \times L$ minors.

The driving force of our induction argument is a rescaling method. The form of symbols we encounter during the induction process carries many features of this rescaling. To present these features, we introduce a collection of auxiliary functions. In what follows, $\gamma : I \to \mathbb{R}^d$ is assumed to be smooth and $\mathbf{r} = (r_1, \ldots, r_d) \in (0,1)^d$, $\lambda \in 2^{\mathbb{N}}$ are fixed.

For $1 \le h \le d$, consider a $(d+1-h)$-tuple

$$\boldsymbol{R} := (R_{h+1}, \ldots, R_{d+1}) \in \mathbb{R}^{d+1-h}.$$

For $j, k \in \mathbb{N}$, consider the function

$$N_\gamma^{j,k,\boldsymbol{R},h}(\xi, s) := r_j \left(\prod_{l=h+1}^{d+1} R_l^j\right) |\langle \gamma^{(j+k)}(s), \xi\rangle| \qquad \text{for } (\xi, s) \in \mathbb{R}^d \times I. \tag{4.15}$$

Let $\beta \in \mathbb{N}_0$. Define

$$M_\gamma^{\beta,\boldsymbol{R},h}(\xi, s) := \max_{\substack{1 \le m \le d \\ \boldsymbol{j} \in \mathcal{W}(m) \\ \boldsymbol{k} \in \mathcal{Z}(\beta, m)}} \prod_{i=1}^m N_\gamma^{j_i,k_i,\boldsymbol{R},h}(\xi, s) \qquad \text{for } (\xi, s) \in \mathbb{R}^d \times I, \tag{4.16}$$

where

$$\mathcal{Z}(\beta, m) := \{\boldsymbol{k} = (k_1, \ldots, k_m) \in \mathbb{N}^m : 1 \le k_i \le d \text{ and } \sum_{i=1}^m k_i \le \beta\}, \tag{4.17}$$

$$\mathcal{W}(m) := \{\boldsymbol{j} = (j_1, \ldots, j_m) \in \mathbb{N}^m : 1 \le j_1 < \cdots < j_m \le d\}. \tag{4.18}$$

The following is the motivation behind the definitions (4.15) and (4.16). Recalling (4.13), we can show that

$$|\partial_s^\beta a_{\mathbf{r}}^\lambda(\xi,s,t)| \lesssim_\gamma M_\gamma^{\beta,\boldsymbol{R},d}(\xi,s)+1, \qquad \text{for any } \beta\in\mathbb{N}_0,$$

provided $R_{d+1}=0$ (See the proof of Proposition 4.2.2 in §4.3.2 for a detailed calculation). Similarly, we use the function defined by (4.16) to dictate the growth of the derivatives of the symbol through the induction process. With this in mind, it is natural to expect terms of the form $\langle\gamma^{j+k}(s),\xi\rangle$ in (4.16), considering the form of the symbol (4.8). The coefficients R_i track the factors a symbol gain or lose when it undergoes rescaling. In particular, the term R_i gets incorporated into the symbol while moving from step i to $(i-1)$ of the induction. Typically, we must think of R_i as some power of λ. It is also worth mentioning that the term R_{d+1} has been introduced purely for convenience, and its value (which will be assumed to be independent of λ) plays no significant role in the analysis.

Fix $1\le h\le d$ and $\boldsymbol{R}\in\mathbb{R}^{d+1-h}$. For $h\le j\le k\le d$, let

$$\lambda_{j,k}^{\boldsymbol{R}} := \begin{cases}\lambda & \text{when } j=k,\\ \prod_{i=j+1}^k R_i^i\lambda & \text{when } j<k.\end{cases}$$

For $h\le j\le d$, define the set

$$Z(\boldsymbol{R},j,\lambda) := \left\{ j\le k\le d : \exists\, C_k>0 \text{ such that } r_k\le C_k\prod_{i=k+1}^{d+1} R_i^{-k}(\lambda_{j,k}^{\boldsymbol{R}})^{-1}\right\}. \tag{4.19}$$

Although the definition of $Z(\boldsymbol{R},j,\lambda)$ depends on $\mathbf{r}$, we have not made it explicit in its notation as $\mathbf{r}$ is a fixed tuple throughout the argument. It is also worth emphasising that the constant C_k should not depend on λ or any of the coefficients R_i. Now, define

$$\begin{aligned} J(\boldsymbol{R},j,\lambda) &:= \min\left[(Z(\boldsymbol{R},j,\lambda)\setminus\{j\})\cup\{d+1\}\right], \qquad (4.20)\\ K(\boldsymbol{R},j,\lambda) &:= \max Z(\boldsymbol{R},j,\lambda).\end{aligned}$$

Let us record an easy observation about these quantities here. For $h\le j\le k\le l\le d$, observe that

$$(\lambda_{j,k}^{\boldsymbol{R}})_{k,l}^{\boldsymbol{R}} = \lambda_{j,l}^{\boldsymbol{R}}.$$

Thus, $Z(\boldsymbol{R},k,\lambda_{j,k}^{\boldsymbol{R}})\subseteq Z(\boldsymbol{R},j,\lambda)$, and consequently, $J(\boldsymbol{R},j,\lambda)\le J(\boldsymbol{R},k,\lambda_{j,k}^{\boldsymbol{R}})$.

Due to the nature of our inductive argument, we may restrict our attention to tuples $\boldsymbol{R}$ of a form, which we describe now.

Definition 4.3.2. We say that $\boldsymbol{R}\in\mathbb{R}^{d+1-h}$ is λ-*admissible* if $R_{d+1}=1$ and $R_i\ge 1$ for any $h+1\le i\le d$. Moreover, for $h\le j\le d$, we must have

$$R_i=1 \qquad \text{whenever } j<i<J(\boldsymbol{R},j,\lambda_{h,j}^{\boldsymbol{R}}). \tag{4.21}$$

With these definitions, we are now well-equipped to introduce the class of symbols tailored for our induction argument.

Definition 4.3.3. Let $1 \leq L \leq d$. Suppose $\gamma \in \mathfrak{S}(B, L)$ and $\mathbf{r} \in (0,1)^d$. A symbol $a \in C^{3d}(\mathbb{R}^d \times I \times I)$ is said to be of type (λ, A, L) with respect to γ, if the following hold:

i) There exists a constant $C > 1$, independent of λ, such that

$$\operatorname{supp}_\xi a \subseteq \{\xi \in \mathbb{R}^d : C\lambda \leq |\xi| \leq 2C\lambda\}.$$

ii) There exists a $\lambda-$admissible tuple $\boldsymbol{R} \in \mathbb{R}^{d+1-L}$ such that for any $\beta \in \mathbb{N}_0$ with $\beta \leq 3d$ and $(\xi, s, t) \in \operatorname{supp} a$, we have

$$|\partial_s^\beta a(\xi, s, t)| \lesssim_{A,\beta} M_\gamma^{\beta, \boldsymbol{R}, L}(\xi, s) + 1, \tag{4.22}$$

$$\max_{1 \leq j \leq d} r_j \left(\prod_{i=L+1}^{d+1} R_i^j \right) |\langle \gamma^{(j)}(s), \xi \rangle| \lesssim 1, \tag{4.23}$$

and

$$\max_{\substack{L+1 \leq j \leq d+1 \\ k \in \mathbb{N}}} (\lambda_{L,j}^{\boldsymbol{R}})^{-1} (\prod_{i=L+1}^{j} R_i)^k |\langle \gamma^{(k)}(s), \xi \rangle| \lesssim_A 1. \tag{4.24}$$

iii) The inner product estimates

$$A^{-1}|\xi| \leq \sum_{i=1}^{L} |\langle \gamma^{(i)}(s), \xi \rangle| \leq A|\xi| \qquad \text{for all } (\xi, s) \in \operatorname{supp}_{\xi,s} a. \tag{4.25}$$

Remark 4.3.4. Before proceeding to the next step, let us briefly mention what happens in the isotropic case (where $r_1 = \cdots = r_d$). In this case, (4.23) combined with (4.15) and (4.16) imply the uniform estimate

$$M_\gamma^{\beta, \boldsymbol{R}, L}(\xi, s, t) = O_{\beta,d}(1) \qquad \text{for } (\xi, s, t) \in \operatorname{supp} a,\ \beta \in \mathbb{N}_0 \text{ and } 1 \leq L \leq d.$$

This implies that in the isotropic case, $M_\gamma^{\beta, \boldsymbol{R}, L}$ and the tuple $\boldsymbol{R}$ do not play any significant role in the argument. If the reader is only interested in the proof of the isotropic case, the inequalities (4.22) to (4.24) can be replaced with the uniform estimate

$$|\partial_s^\beta a(\xi, s, t)| = O_\beta(1) \qquad \text{for } (\xi, s, t) \in \operatorname{supp} a,\ \beta \in \mathbb{N}_0$$

in Definition 4.3.3 and proceed to the next step. Furthermore, as we are not required to provide finer estimates for the derivatives of the symbol, the overall argument simplifies to a great extent in this case, and in particular, all the up-coming discussions about estimating $M_\gamma^{\beta, \boldsymbol{R}, L}$ can be avoided. A detailed proof, in this case, can be found in [49].

We record a technical lemma here that can be used to control the derivatives of symbols during later stages of the proof.

Lemma 4.3.5. *Let $\mathbf{r} \in (0,1)^d$ be admissible and $1 \leq L \leq d$. Suppose a is a symbol of type (λ, A, L) with respect to γ. Assume that $L \in Z(\boldsymbol{R}, L, \lambda)$ where $\boldsymbol{R}$ is the λ-admissible tuple associated to a. Then,*

$$Z(\boldsymbol{R}, L, \lambda) = \{L, \dots, d\}. \tag{4.26}$$

Furthermore, for $L \leq j \leq d$ and $k \in \mathbb{N}$, we have

$$|N_\gamma^{j,k,\boldsymbol{R},L}(\xi, s)| \lesssim_{A,B} 1 \qquad \text{for all } (\xi, s) \in \operatorname{supp}_{\xi,s} a. \tag{4.27}$$

Proof. To prove (4.26), assume that it is false, and we will try to reach a contradiction. Simplifying notation, we use λ_j and K to represent $\lambda^{\boldsymbol{R}}_{L,j}$ and $K(\boldsymbol{R}, L, \lambda)$ in this proof. Let $j \in \{L+1, \dots, d\} \setminus Z(\boldsymbol{R}, L, \lambda) \neq \emptyset$. Two cases arise here:

Case 1. Suppose $K < j$. Recall from the definition that $Z(\boldsymbol{R}, K, \lambda_K)$ is the singleton set $\{K\}$. Therefore, by (4.20), we have $J(\boldsymbol{R}, K, \lambda_K) = d+1$. Combining this with (4.21), we see that $R_{K+1} = \cdots = R_{d+1} = 1$, and as a consequence, we see that the values of λ_j and λ_K are the same. Since $r_j \leq r_K$ by (4.3) we therefore deduce that

$$r_j \left(\prod_{i=j+1}^{d+1} R_i^j \right) \leq \left(r_K \prod_{i=K+1}^{d+1} R_i^j \right) \lesssim \lambda_K^{-1} = \lambda_j^{-1}.$$

By (4.19), this implies that $j \in Z(\boldsymbol{R}, L, \lambda)$, reaching a contradiction.

Case 2. Suppose Case 1 does not hold, or, in other words, $L < j < K$. Since $L, K \in Z(\boldsymbol{R}, L, \lambda)$, we can find $j_0, j_1 \in Z(\boldsymbol{R}, L, \lambda)$ such that $L \leq j_0 < j_1 \leq K$ and

$$(j_0, j] \cap Z(\boldsymbol{R}, L, \lambda) = [j, j_1) \cap Z(\boldsymbol{R}, L, \lambda) = \emptyset.$$

Since $J(\boldsymbol{R}, j_0, \lambda_{j_0}) = j_1$, we also have $R_l = 1$ when $j_0 < l < j_1 - 1$ by (4.21). Thus, $\lambda_{j_0} = \lambda_j = \lambda_{j_0 - 1}$. Set $t = (j - j_0)/(j_1 - j_0)$ and recall the estimate

$$r_j \leq r_{j_0}^{1-t} r_{j_1}^t$$

from (4.3). Combining these observations,

$$\begin{aligned}
r_j \left(\prod_{i=j+1}^{d+1} R_i^j \right) &\leq \left(r_{j_0} \prod_{i=j+1}^{d+1} R_i^{j_0} \right)^{1-t} \left(r_{j_1} \prod_{i=j+1}^{d+1} R_i^{j_1} \right)^t \\
&= \left(r_{j_0} \prod_{i=j_0+1}^{d+1} R_i^{j_0} \right)^{1-t} \left(r_{j_1} \prod_{i=j_1+1}^{d+1} R_i^{j_1} \right)^t R_{j_1}^{j_1 t} \\
&\lesssim \lambda_{j_0}^{-(1-t)} \lambda_{j_1}^{-t} R_{j_1}^{j_1 t} = \lambda_{j_0}^{-(1-t)} \lambda_{j_1 - 1}^{-t} \\
&= \lambda_j^{-(1-t)} \lambda_j^{-t} = \lambda_j^{-1}.
\end{aligned}$$

Consequently, $j \in Z(\boldsymbol{R}, L, \lambda)$, reaching a contradiction again. Thus, we have shown that (4.26) holds.

It remains to prove (4.27). For $j \in Z(\boldsymbol{R}, L, \lambda)$ and $k \in \mathbb{N}_0$, it is clear from the definitions (4.15) and (4.19) that

$$N_{\gamma}^{j,k,\boldsymbol{R},L}(\xi,s) \lesssim \Big(\prod_{i=j+1}^{d+1} R_i^{-j}\lambda_j^{-1}\Big)\Big(\prod_{i=L+1}^{d+1} R_i^{j}\Big)|\langle\gamma^{(j+k)}(s),\xi\rangle|$$

for $(\xi,s)\in\mathbb{R}^{d+1}$. When $j=L$, by applying the Cauchy–Schwarz inequality, the quantity on the right is seen to be bounded above by $O_B(1)$ for $(\xi,s)\in\operatorname{supp}_{\xi,s} a$ (recall that $\lambda_L=\lambda$). On the other hand, when $j\geq L+1$, we apply (4.24) so that

$$\begin{aligned}\lambda_j^{-1}\Big(\prod_{i=L+1}^{j} R_i^{-j}\Big)|\langle\gamma^{(j+k)}(s),\xi\rangle| &\lesssim_{A,B} \lambda_j^{-1}\Big(\prod_{i=L+1}^{j} R_i^{-j}\Big)\Big(\prod_{i=L+1}^{j} R_i\Big)^{-(j+k)}\lambda_j\\ &\lesssim \prod_{i=L+1}^{j} R_i^{-k}\leq 1\end{aligned}$$

for $(\xi,s)\in\operatorname{supp}_{\xi,s} a$. Note that the assumption that $R_i\geq 1$ has been used in the final inequality. This completes the proof of (4.27). □

4.3.2 Setting up the induction scheme

Proposition 4.2.2 is a consequence of the following inductive statement.

Proposition 4.3.6. *Let $1\leq L\leq d$ and $\lambda\in 2^{\mathbb{N}}$. Suppose $\gamma\in\mathfrak{G}(B,L)$ and let* $\mathbf{r}$ *be admissible. Let $a\in C^{3d}(\mathbb{R}^d\times I\times I)$ be a symbol of type (λ,A,L) with respect to γ. Then,*

$$\|\mathfrak{D}_s\mathcal{A}[a,\gamma]\|_{L^2(\mathbb{R}^{d+1})\to L^2(\mathbb{R}^{d+1})} \lesssim_{d,L,A,B} (\log\lambda)^{(L-1)/2}. \tag{4.28}$$

Before proceeding to the proof of the proposition, we show that Proposition 4.2.2 can be obtained as a special case of Proposition 4.3.6.

Proof of Proposition 4.2.2. Fix an admissible $\mathbf{r}$. By the final remarks in the previous section, $\lambda=0$ case of Proposition 4.2.2 has been dealt with. Fix $\lambda\in 2^{\mathbb{N}}$.

We claim that the statement of Proposition 4.2.2 corresponds to the special case $L=d$ of Proposition 4.3.6. To see this, we must first verify that $a_{\mathbf{r}}^{\lambda}$ is of type (λ,A,L) with respect to γ for some $A>0$. In view of the frequency localisation of the symbol $a_{\mathbf{r}}^{\lambda}$ and (4.1), it suffices to check the validity of the inequalities (4.22) to (4.24) for $a=a_{\mathbf{r}}^{\lambda}$, $L=d$ and $R_{d+1}=1$. An application of Cauchy–Schwarz inequality directly gives (4.24). We will aim for (4.23) now.

The construction of the Frenet frame along with (4.14) gives

$$\sum_{i=1}^{l}|\langle e_i(s),\xi\rangle| \sim_{B,l} \sum_{i=1}^{l}|\langle\gamma^{(i)}(s),\xi\rangle| \tag{4.29}$$

with constants in the inequalities being uniform for all $(\xi, s) \in \mathbb{R}^d \times I$. Thus,

$$\sum_{i=1}^{d} r_i |\langle \gamma^{(i)}(s), \xi \rangle| \lesssim_{B,d} \sum_{i=1}^{d} \sum_{j=1}^{i} r_i |\langle e_j(s), \xi \rangle| = \sum_{j=1}^{d} \left(\sum_{i=j}^{d} r_i \right) |\langle e_j(s), \xi \rangle|$$
$$\lesssim_d \sum_{j=1}^{d} r_j |\langle e_j(s), \xi \rangle| \tag{4.30}$$

where in the final step, we used the assumption $r_i \leq r_j$ whenever $j \leq i$. The form of the frequency localisation of the symbol (in particular, the localisation arising from (4.8)) implies that the sum on the right in (4.30) is $O_d(1)$ when $(\xi, s) \in \mathrm{supp}_{\xi,s}\, a^\lambda_{\mathbf{r}}$, establishing (4.23). We may proceed to the verification of (4.22) now.

By direct computations that use the basic properties of the Frenet coordinate system, we can estimate the first-order $s-$derivative of the symbol as

$$|\partial_s a^\lambda_{\mathbf{r}}(\xi, s, t)| \lesssim \sum_{i=1}^{d-1} r_i |\langle e_{i+1}(s), \xi \rangle| + \sum_{i=2}^{d} r_i |\langle e_{i-1}(s), \xi \rangle| \quad \text{ for } (\xi, s, t) \in \mathrm{supp}\, a^\lambda_{\mathbf{r}}.$$

More generally, for higher-order derivatives, we can write

$$|\partial_s^\beta a^\lambda_{\mathbf{r}}(\xi, s, t)| \lesssim \sum_{m=1}^{d} \sum_{\boldsymbol{j} \in \mathcal{W}(m)} \sum_{\boldsymbol{k} \in \tilde{\mathcal{Z}}(\beta, \boldsymbol{j}, m)} \prod_{i=1}^{m} r_{j_i} |\langle e_{j_i+k_i}(s), \xi \rangle| \tag{4.31}$$

for $(\xi, s, t) \in \mathrm{supp}\, a^\lambda_{\mathbf{r}}$ and $\beta \in \mathbb{N}_0$. Here, the set $\mathcal{W}(m)$ is as introduced in (4.18) and

$$\tilde{\mathcal{Z}}(\beta, \boldsymbol{j}, m) := \Big\{ \boldsymbol{k} = (k_1, \dots, k_m) \in \mathbb{N}_0^m : k_i \in [1 - j_i, d - j_i],\ \textstyle\sum_{i=1}^m k_i \leq \beta \Big\}.$$

However, by (4.8) and (4.3), we can estimate

$$r_{j_i} |\langle e_{j_i+k_i}(s), \xi \rangle| \lesssim r_{j_i} r^{-1}_{(j_i+k_i)} \leq 1 \qquad \text{for } k_i \leq 0.$$

Therefore, we can simplify (4.31) and write

$$|\partial_s^\beta a^\lambda_{\mathbf{r}}(\xi, s, t)| \lesssim_{\beta,d} \max_{(m,\boldsymbol{j},\boldsymbol{k}) \in \mathcal{V}(\beta)} \prod_{i=1}^{m} r_{j_i} |\langle e_{j_i+k_i}(s), \xi \rangle| + 1 \tag{4.32}$$

for $(\xi, s, t) \in \mathrm{supp}\, a^\lambda_{\mathbf{r}}$, where

$$\mathcal{V}(\beta) := \Big\{ (m, \boldsymbol{j}, \boldsymbol{k}) : 1 \leq m \leq d,\ \boldsymbol{j} \in \mathcal{W}(m) \text{ and }$$
$$\boldsymbol{k} \in \mathcal{Z}(\beta, m) \text{ with } k_i + j_i \leq d \text{ for each } i \Big\}.$$

To prove (4.22), it remains to replace the Frenet vectors $e_{j_i+k_i}$ with $\gamma^{(j_i+k_i)}$ in (4.32). To do this, we use (4.29), (4.23) and the assumption $r_i \leq r_j$ whenever $j \leq i$, so that

$$\begin{aligned} r_j|\langle e_{j+k}(s),\xi\rangle| &\lesssim \max_{1-j\leq k'\leq k} r_j|\langle \gamma^{(j+k')}(s),\xi\rangle| \\ &\lesssim \max_{1\leq k'\leq k} r_j|\langle \gamma^{(j+k')}(s),\xi\rangle| + \max_{1-j\leq k'\leq 0} r_j r_{j+k'}^{-1} \\ &\lesssim \max_{1\leq k'\leq k} r_j|\langle \gamma^{(j+k')}(s),\xi\rangle| + 1. \end{aligned} \tag{4.33}$$

Recalling the definition (4.17), it is easy to see that $(k_1',\dots,k_m') \in \mathcal{Z}(\beta,m)$ when $k_i' \leq k_i$ and $(k_1,\dots,k_m) \in \mathcal{Z}(\beta,m)$. Therefore, we can obtain (4.22) for $a = a_{\mathbf{r}}^{\lambda}$ with $L = d$ by combining (4.33) for each $1 \leq j \leq d-1$ with (4.32).

From the above discussion, we conclude that $a_{\mathbf{r}}^{\lambda}$ is type (λ, A, d) with respect to γ.Thus, statement of Proposition 4.3.6 applies, concluding the proof of Proposition 4.2.2. ☐

Proposition 4.3.6 is proved by inducting on L. Given an arbitrary symbol $a \in C^{3d}(\mathbb{R}^d \times I \times I)$ and a smooth curve γ, we now present a general argument that will be used repeatedly through the induction process to obtain favorable norm bounds for the Fourier integral operator $\mathfrak{D}_s\mathcal{A}[a,\gamma]$. For a Schwartz function g, we aim for the estimate

$$\|\mathfrak{D}_s\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})} \lesssim_d \|g\|_{L^2(\mathbb{R}^{d+1})}. \tag{4.34}$$

By applying Plancherel's theorem and the Cauchy–Schwarz inequality,

$$\begin{aligned} \|\mathfrak{D}_s\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}^2 &= \int_{\mathbb{R}} (1+|\sigma|)|\mathcal{F}_{x,s}(\mathcal{A}[a,\gamma]g)|^2(\sigma,\xi)\mathrm{d}\xi\mathrm{d}\sigma \\ &\lesssim \|\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}\|(1+\sqrt{-\partial_s^2})\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})} \\ &\leq \|\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}^2 \\ &\qquad + \|\sqrt{-\partial_s^2}\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}\|\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}. \end{aligned}$$

Since the Hilbert transform is bounded on L^2,

$$\|\sqrt{-\partial_s^2}\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})} \lesssim \|\partial_s\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}.$$

Thus, to prove (4.34), it suffices to show that there exists $\Lambda > 1$ such that

$$\|\partial_s^{\iota}\mathcal{A}[a,\gamma]\|_{L^2(\mathbb{R}^{d+1})\to L^2(\mathbb{R}^{d+1})} \lesssim_d \Lambda^{(2\iota-1)/2} \quad \text{for } \iota = 0,1.$$

Applying Plancherel's theorem and the Cauchy–Schwarz inequality,

$$\begin{aligned} \|\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}^2 &\sim_d \int_I\int_{\mathbb{R}^d} \mathcal{B}[a]\mathcal{F}_x(g)(\xi,t)\overline{\mathcal{F}_x(g)(\xi,t)}\mathrm{d}\xi\mathrm{d}t \\ &\leq \int_{\mathbb{R}^d} \|\mathcal{B}[a]\mathcal{F}_x(g)(\xi,\cdot)\|_{L^2(\mathbb{R})}\|\mathcal{F}_x(g)(\xi,\cdot)\|_{L^2(\mathbb{R})}\mathrm{d}\xi, \end{aligned}$$

where $\mathcal{B}[a]$ is the operator that integrates (in t' variable) functions against the kernel

$$K[a](\xi, t', t) := \int_I e^{i\langle (t-t')\gamma(s), \xi\rangle} a(\xi, s, t')\overline{a(\xi, s, t)}\mathrm{d}s. \tag{4.35}$$

At this point, note that $\partial_s \mathcal{A}[a, \gamma]g$ can be expressed as $\mathcal{A}[\mathfrak{d}_s a, \gamma]g$, with a symbol

$$\mathfrak{d}_s a(\xi, s, t) := t\langle \gamma'(s), \xi\rangle a(\xi, s, t) + \partial_s a(\xi, s, t) \qquad \text{for } (\xi, s, t) \in \mathbb{R}^{d+2}.$$

Applying Schur's test, we see that (4.34) is a consequence of the estimates

$$\sup_{(\xi,t')\in \operatorname{supp}_\xi a\times I} \|K[\mathfrak{d}_s^\iota a](\xi, t', \cdot)\|_{L^1_t(I)} \lesssim_d \Lambda^{2\iota-1} \qquad \text{for } \iota = 0, 1, \tag{4.36}$$

completing the discussion.

The first application of this reduction is the following lemma.

Lemma 4.3.7 (Base case)**.** *Proposition 4.3.6 holds when $L = 1$.*

Proof of Lemma 4.3.7. Choose a curve γ and a symbol a that satisfies the assumptions of the proposition. In particular, a is of type $(\lambda, A, 1)$ with respect to γ and as a consequence,

$$|\langle \gamma'(s), \xi\rangle| \sim_A \lambda \qquad \text{holds in supp } a. \tag{4.37}$$

Following the previous discussion, we wish to obtain good decay estimates for the function $K[\mathfrak{d}_s^\iota a]$ with $\iota = 0, 1$. The plan is to integrate by parts in (4.35) after obtaining uniform upper bounds for the derivatives of a. To this end, we note an easy observation that $1 \in Z(\boldsymbol{R}, 1, \lambda)$, which follows from (4.25) for $L = 1$ and (4.23). Thus, Lemma 4.3.5 can be applied for $L = 1$. Combining (4.27) for $L = 1$ with (4.16), we deduce that $M_\gamma^{\beta,\boldsymbol{R},1}(\xi, s) = O_{\beta,d}(1)$. By (4.22), we therefore obtain

$$|\partial_s^\beta a(\xi, s, t)| \lesssim_{\beta,d} 1 \qquad \text{for all } (\xi, s, t) \in \text{supp } a. \tag{4.38}$$

Integrating-by-parts in (4.35) using (4.37) and (4.38), we have

$$|K[\mathfrak{d}_s^\iota a](\xi, t', t)| \lesssim_{A,B,d,N} \lambda^{2\iota}(1 + |t - t'|\lambda)^{-N} \qquad \text{for } \iota = 0, 1 \text{ and } N \geq 1.$$

Clearly, these decay estimates imply the required bounds (4.36) with $\Lambda = \lambda$. Consequently, we obtain (4.34), concluding the proof of Lemma 4.3.7. □

Lemma 4.3.7 addresses the base case of Proposition 4.3.6. It remains to establish the inductive step.

Proposition 4.3.8. *Suppose the statement of Proposition 4.3.6 is true for all $L \leq N - 1$. Then it is also true for $L = N$.*

Proposition 4.3.6 and therefore Theorem 4.1.2, follow from Proposition 4.3.8 and Lemma 4.3.7. In the remainder of the section, we present the proof of Proposition 4.3.8, which is broken into steps.

4.3.3 Initial decomposition

Let γ and a be chosen to satisfy the assumptions of the Proposition 4.3.6 with $L = N$. We begin with a natural division of the symbol a. Let $H : \mathbb{R}^{d+1} \to \mathbb{R}$ be defined as the product

$$H(\xi, s) := \prod_{i=1}^{N-1} \eta(A'\lambda^{-1}\langle\gamma^{(i)}(s), \xi\rangle)$$

where A' is large constant which will be chosen depending only on A, B and N. Here η is as defined in (3.3). Note that

$$|\partial_s^\beta H(\xi, s)| \lesssim_{\beta,A,B} 1 \qquad \text{for } (\xi, s) \in \operatorname{supp}_{\xi,s} a \text{ and } \beta \in \mathbb{N} \cup \{0\}. \tag{4.39}$$

The following lemma verifies the type condition for the pair $(a(1 - H), \gamma)$.

Lemma 4.3.9. *The symbol $a(1 - H)$ is of type $(\lambda, A', N - 1)$ with respect to γ.*

Proof. The definition of H combined with the assumption on the type of a give (4.25) for $L = N - 1$, $a = a(1 - H)$ and $A = A'$. Thus, it remains to verify Definition 4.3.3 ii) for $a = a(1 - H)$ and $L = N - 1$.

Let $\boldsymbol{R} = (R_{N+1}, \dots, R_{d+1})$ be the λ-admissible tuple associated to a. Set $R_N := 1$ and write

$$\boldsymbol{R}' := (R_N, R_{N+1}, \dots, R_{d+1}).$$

First, we verify that $\boldsymbol{R}'$ is λ-admissible. Since $\lambda^{\boldsymbol{R}}_{N,j} = \lambda^{\boldsymbol{R}'}_{N-1,j}$ and $\boldsymbol{R}$ is assumed to be λ-admissible, (4.21) clearly holds with $\boldsymbol{R}$ replaced with $\boldsymbol{R}'$ and $h = N - 1$ whenever $N \le j \le d - 1$. When $j = N - 1$, we obtain (4.21) as a consequence of the easy observation

$$J(\boldsymbol{R}', N - 1, \lambda) \le J(\boldsymbol{R}, N, \lambda).$$

Thus, we conclude that $\boldsymbol{R}'$ is λ-admissible.

It now remains to verify the estimates (4.22) to (4.24) for $a = a(1 - H)$, $L = N - 1$ and $\boldsymbol{R} = \boldsymbol{R}'$. The first bound (4.22) is immediate from the type condition on a, (4.39) and the easy observation that $M_\gamma^{\beta,\boldsymbol{R},N} = M_\gamma^{\beta,\boldsymbol{R}',N-1}$. The remaining estimates also follow quite easily since $R_N = 1$. This completes the proof of the lemma. $\square$

Given the above lemma, the induction hypothesis can be applied to deduce the desired estimate (4.28) when $a = a(1 - H)$.

Since (4.25) holds with $L = N$ in supp a by assumption, the inequalities

$$(10A)^{-1}|\xi| \le |\langle\gamma^{(N)}(s), \xi\rangle| \le A|\xi|, \tag{4.40}$$

$$\sum_{i=1}^{N-1} |\langle\gamma^{(i)}(s), \xi\rangle| \le 10^{-10}A^{-1}|\xi| \tag{4.41}$$

also hold for all $(\xi, s) \in \operatorname{supp}_{\xi,s} aH$, provided A' is chosen large enough depending on N and A. Without loss of generality, we can therefore work with the stronger assumptions (4.40) and (4.41) on the support of a. An application of the implicit function theorem now shows that for any $\xi \in \mathrm{supp}_\xi\, a$, there exists $\sigma(\xi) \in I$ with

$$\langle \gamma^{(N-1)} \circ \sigma(\xi), \xi \rangle = 0. \tag{4.42}$$

The strategy now involves a decomposition of the symbol away from the most degenerate regions in $\mathbb{R}^{d+1}$. Set

$$G(\xi, s) := \sum_{i=1}^{N-1} |\varepsilon_0^{-1} \lambda^{-1} \langle \gamma^{(i)} \circ \sigma(\xi), \xi \rangle|^{2/(N-i)} + \varepsilon_0^{-2} |s - \sigma(\xi)|^2,$$

where the constant $\varepsilon_0 = \varepsilon_0(B, d)$ will be chosen small enough to satisfy the forthcoming requirements of the proof.[2] The function G should be interpreted as the function measuring the distance of (ξ, s) from the co-dimension N surface

$$\Gamma := \{(\xi, s) \in \mathbb{R}^d \times I : \langle \gamma^{(i)} \circ \sigma(\xi), \xi \rangle = 0 \text{ for } 1 \leq i \leq N-1 \text{ and } |s - \sigma(\xi)| = 0\}.$$

We now decompose the (ξ, s)-space dyadically away from Γ. Suppose we have $\eta_1, \beta_1 \in C_c^\infty(\mathbb{R})$ chosen such that

$$\mathrm{supp}\; \eta_1 \subseteq \{r \in \mathbb{R} : |r| \leq 4\}, \qquad \mathrm{supp}\; \beta_1 \subseteq \{r \in \mathbb{R} : 1/4 \leq |r| \leq 4\} \tag{4.43}$$

and

$$\eta_1(r) + \sum_{n \in \mathbb{N}} \beta_1(2^{-2n} r) = 1 \qquad \text{for all } r \in \mathbb{R}.$$

Set

$$a^n(\xi, s, t) := a(\xi, s, t) \cdot \begin{cases} \eta_1(\varepsilon_1^2 \lambda^{2/N} G(\xi, s)) & \text{if } n = 0, \\ \beta_1(\varepsilon_1^2 2^{-2n} \lambda^{2/N} G(\xi, s)) & \text{if } n \geq 1. \end{cases} \tag{4.44}$$

where ε_1 will be chosen small enough (depending on ε_0) to satisfy the forthcoming requirements of the proof.[3] Observe that $a = a^0 + \sum_{n \in \mathbb{N}} a^n$ and this automatically induces a similar decomposition for the Fourier integral operator $\mathfrak{D}_s \mathcal{A}[a, \gamma]$. Since

$$|G(\xi, s)| = O_{B,d}(\varepsilon_0^{-2}) \qquad \text{for all } (\xi, s) \in \mathrm{supp}_{\xi, s}\; a, \tag{4.45}$$

the symbols a^n are trivially zero except for $O_{A,B}(\log \lambda)$ many values of n. Thus, by Plancherel's theorem,

$$\begin{aligned} \|\mathfrak{D}_s \mathcal{A}[\sum_{n \geq 0} a^n, \gamma] g\|^2_{L^2(\mathbb{R}^{d+1})} &= \sum_{n \geq 0} \|\mathfrak{D}_s \mathcal{A}[a^n, \gamma] g\|^2_{L^2(\mathbb{R}^{d+1})} \\ &\lesssim_{A,B} |\log \lambda| \max_{n \geq 0} \|\mathfrak{D}_s \mathcal{A}[a^n, \gamma] g\|^2_{L^2(\mathbb{R}^{d+1})}. \end{aligned} \tag{4.46}$$

In light of the above, it remains to bound the fractional operator $\mathfrak{D}_s \mathcal{A}[a^n, \gamma]$ for different values of n. The case of $n = 0$ is dealt with by the following lemma.

[2] As we will note later, $\varepsilon_0 := (\min_{1 \leq N \leq d} c_N)(50(2 + B))^{-1}$ would suffice for our purposes.
[3] $\varepsilon_1 := \varepsilon_0 / (10 d B^{2d+1})$ would be sufficient

Lemma 4.3.10.

$$\|\mathfrak{D}_s\mathcal{A}[a^0,\gamma]\|_{L^2(\mathbb{R}^{d+1})\to L^2(\mathbb{R}^{d+1})} \lesssim_{A,B,d} 1. \tag{4.47}$$

The next lemma addresses the case of all other values of n.

Lemma 4.3.11. *For any $n \geq 1$, we have*

$$\|\mathfrak{D}_s\mathcal{A}[a^n,\gamma]\|_{L^2(\mathbb{R}^{d+1})\to L^2(\mathbb{R}^{d+1})} \lesssim_{A,B,d} (\log\lambda)^{(N-2)/2}. \tag{4.48}$$

Assuming the lemmas for now, we plug (4.47), (4.48) into (4.46) and obtain

$$\|\mathfrak{D}_s\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})} \lesssim_{A,B,d} (\log\lambda)^{(N-1)/2}\|g\|_{L^2(\mathbb{R}^{d+1})}.$$

This concludes the proof of Proposition 4.3.8.

The remaining parts of this section is dedicated to the proofs of Lemma 4.3.10 and Lemma 4.3.11.

4.3.4 Proof of Lemma 4.3.10

To prove Lemma 4.3.10, we do not appeal to the induction hypothesis but directly estimate the operator. There are two key elements in the proof, which are stated below in the form of lemmas. In the first one, uniform bounds on inner products with derivatives of γ are obtained by using the information that a^0 is supported near the degenerate surface Γ.

Lemma 4.3.12. *For any $(\xi,s) \in \operatorname{supp}_{\xi,s} a^0$ and $1 \leq i \leq N$, we have*

$$|\langle\gamma^{(i)}(s),\xi\rangle| \lesssim_{A,B,N} \lambda^{i/N}. \tag{4.49}$$

Proof. By the definition of a^0 from (4.44), it is known that

$$|\langle\gamma^{(i)}\circ\sigma(\xi),\xi\rangle| \lesssim_{A,B,N} \lambda\lambda^{(i-N)/N} \text{ and } |s-\sigma(\xi)| \lesssim_{A,B} \lambda^{-1/N}$$

for $(\xi,s) \in \operatorname{supp}_{\xi,s} a^0$. Using Taylor's theorem,

$$\begin{aligned}|\langle\gamma^{(i)}(s),\xi\rangle| &\leq \sum_{j=i}^{N-1} \big|\langle\gamma^{(j)}\circ\sigma(\xi),\xi\rangle\big|\frac{|s-\sigma(\xi)|^{j-i}}{(j-i)!} + B|\xi|\frac{|s-\sigma(\xi)|^{N-i}}{(N-i)!} \\ &\lesssim_{A,B,N} \lambda^{i/N}\end{aligned}$$

in the support of supp a^0, as required. □

Through the next lemma, we record a property that the admissible tuples satisfy.

Lemma 4.3.13. *For any admissible $\mathbf{r} = (r_1,\dots,r_d)$, we have*

$$r_j \leq r_k^{j/k} \qquad \text{for all } 1 \leq j \leq k \leq d. \tag{4.50}$$

Proof of Lemma 4.3.13. For $2 \le j \le d$, we have $r_1^2 \le r_2$ and $r_2 \le r_1^{\frac{j-2}{j-1}} r_j^{\frac{1}{j-1}}$ as direct applications of (4.3). Combining both inequalities, we deduce that

$$r_1 \le r_j^{1/j} \qquad \text{for } 1 \le j \le d.$$

Combining the above with the second part of (4.3) for $i = 1$, we deduce that

$$r_j \le r_1^{\frac{k-j}{k-1}} r_k^{\frac{j-1}{k-1}} \le r_j^{\frac{k-j}{j(k-1)}} r_k^{\frac{j-1}{k-1}} \qquad \text{for } 1 \le j \le k \le d.$$

Rearranging, we obtain (4.50). □

Proof of Lemma 4.3.10. In view of discussions around (4.34) and (4.36), it suffices to show

$$|K[\mathfrak{d}_s^\iota a^0](\xi, t', t)| \lesssim_B \lambda^{(2\iota-1)/N} \qquad \text{for } \iota = 0, 1 \text{ and } (\xi, t', t) \in \mathbb{R}^d \times I \times I. \quad (4.51)$$

Indeed, (4.51) implies (4.36) with $a = a^0$ and $\Lambda = \lambda^{1/N}$, which in turn gives (4.47).

The estimate (4.51) for $\iota = 0$ is immediate from (4.35) as the $\operatorname{supp}_s a^0(\xi, \cdot, \cdot)$ is contained in an interval of length $\lambda^{-1/N}$ for any fixed $\xi \in \mathbb{R}^d$. By (4.35) again, the case $\iota = 1$ becomes evident once we verify the estimates

$$|\langle \gamma^{(1)}(s), \xi \rangle| + |\partial_s(a^0)(\xi, s, t)| \lesssim_{A,B,d} \lambda^{1/N} \qquad \text{for } (\xi, s, t) \in \operatorname{supp} a^0.$$

The estimate on the first term yields by putting $i = 1$ in (4.49). To estimate the latter term, recall (4.16) for $\beta = 1$ and write

$$M_\gamma^{1,\boldsymbol{R},N} = \max_{1 \le j \le d} N_\gamma^{j,1,\boldsymbol{R},N} = \max_{l \in \{1,2\}} \max_{j \in S_l} N_\gamma^{j,1,\boldsymbol{R},N}$$

where $S_1 := \{1, \ldots, N-1\}$ and $S_2 := \{N, \ldots, d\}$. We claim that

$$\max_{j \in S_1} N_\gamma^{j,1,\boldsymbol{R},h}(\xi, s) = O_{A,B,d}(\lambda^{1/N}) \quad \text{and} \quad \max_{j \in S_2} N_\gamma^{j,1,\boldsymbol{R},h}(\xi, s) = O_{A,B,d}(1) \quad (4.52)$$

for $(\xi, s) \in \operatorname{supp}_{\xi,s} a^0$. Once the claim is verified, we can use (4.22) to conclude that

$$|\partial_s a(\xi, s, t)| \lesssim_{A,B,d} \lambda^{1/N} \qquad \text{for } (\xi, s, t) \in \operatorname{supp} a^0.$$

An application of Leibniz's rule gives a similar estimate where a is replaced with a^0, completing the proof of (4.51) and therefore Lemma 4.3.10.

To prove the claim (4.52), we first assume that $j \in S_1$. Using (4.49), (4.40) and (4.23), we have

$$\begin{aligned}N_\gamma^{j,k,\boldsymbol{R},N}(\xi,s) &= r_j\left(\prod_{l=N+1}^{d+1} R_l^j\right)|\langle\gamma^{(j+1)}(s),\xi\rangle|\\ &\lesssim_{A,B,N} r_j\left(\prod_{l=N+1}^{d+1} R_l^j\right)\lambda^{(j+1)/N}\\ &\lesssim \lambda^{1/N} r_j\left(\prod_{l=N+1}^{d+1} R_l^j\right)\left(r_N^{-1}\prod_{l=N+1}^{d+1} R_l^{-N}\right)^{j/N}\\ &\lesssim \lambda^{1/N} r_j r_N^{-j/N} \lesssim \lambda^{1/N},\end{aligned}$$

where the final inequality follows from (4.50).

To deal with the case $j \in S_2$, we first note that (4.23) combined with (4.40) gives

$$r_N\left(\prod_{i=N+1}^{d+1} R_i^N\right) \lesssim |\langle\gamma^{(N)}(s),\xi\rangle|^{-1} \approx \lambda^{-1} \qquad \text{for } (\xi,s) \in \operatorname{supp}_{\xi,s} a.$$

Combining the above with (4.19), we see that $N \in Z(\boldsymbol{R}, N, \lambda)$. Therefore, Lemma 4.3.5 applies here. By (4.27), we establish the second part of (4.52), completing the argument. □

4.3.5 Further decomposition

In order to prove Lemma 4.3.11 we must introduce a further decomposition of the symbol. Let $\zeta \in C_c^\infty(\mathbb{R})$ be chosen such that $\operatorname{supp}\zeta \subseteq [-1,1]$ and $\sum_{\nu\in\mathbb{Z}}\zeta(\,\cdot\,-\nu) = 1$. For $n \in \mathbb{N}$ and $\nu \in \mathbb{Z}$, consider the symbol

$$a^{n,\nu}(\xi,s,t) := a^n(\xi,s,t)\zeta(2^{-n}\lambda^{1/N}(s-s_{n,\nu})) \tag{4.53}$$

where $s_{n,\nu} := 2^n\lambda^{-1/N}\nu$. Observe that the original symbol is recovered as the sum

$$a = \sum_{n=0}^{C\log(\lambda)} a^n = a^0 + \sum_{n=1}^{C\log(\lambda)}\sum_{\nu\in\mathbb{Z}} a^{n,\nu}, \tag{4.54}$$

where C is a constant that depends only on A, B. The following lemma records a basic property of the localised symbols, which becomes useful later in the proof.

Lemma 4.3.14. *Let $n \geq 1, \nu \in \mathbb{Z}$ and $\rho := 2^n\lambda^{-1/N}$. For $(\xi,s) \in \operatorname{supp}_{\xi,s} a^{n,\nu}$, we have*

$$\sum_{i=1}^{N-1} \rho^{i-N}|\langle\gamma^{(i)}(s),\xi\rangle| \sim_{A,B,d} |\xi| \sim \lambda. \tag{4.55}$$

Proof. The upper bound in (4.55) is easier to prove than the lower bound and follows from a similar argument. Consequently, we will focus only on the lower bound.

Fix $n \geq 1$ and $\nu \in \mathbb{Z}$. Recall from the definitions that

$$\varepsilon_1^{-2}/4 \leq \sum_{i=1}^{N-1} |\varepsilon_0^{-1}\lambda^{-1}\rho^{i-N}\langle\gamma^{(i)} \circ \sigma(\xi), \xi\rangle|^{2/(N-i)} + |\varepsilon_0^{-1}\rho^{-1}(s - \sigma(\xi))|^2 \leq 4\varepsilon_1^{-2} \tag{4.56}$$

for all $(\xi, s) \in \operatorname{supp}_{\xi,s} a^{n,\nu}$. Fixing ξ, we now consider two cases depending on which terms of the above sum dominate.

Case 1. Suppose $(\varepsilon_0\varepsilon_1^{-1}\rho)/4 \leq |s - \sigma(\xi)|$. By Taylor's theorem, there lies $s_* \in I$ between s and $\sigma(\xi)$ such that

$$\langle\gamma^{(N-1)}(s), \xi\rangle - \langle\gamma^{(N-1)} \circ \sigma(\xi), \xi\rangle = \langle\gamma^{(N)}(s_*), \xi\rangle(s - \sigma(\xi)).$$

Combining this with (4.40) and (4.42), we deduce that

$$|\langle\gamma^{(N-1)}(s), \xi\rangle| \gtrsim_A \lambda(\varepsilon_0\varepsilon_1^{-1}\rho).$$

This gives the lower bound in (4.55).

Case 2. Suppose Case 1 fails. Using (4.56), we can find $1 \leq i_0 \leq N-2$ such that

$$c_N\varepsilon_0\lambda(\varepsilon_1^{-1}\rho)^{N-i_0} \leq |\langle\gamma^{(i_0)} \circ \sigma(\xi), \xi\rangle| \leq 2^N\varepsilon_0\lambda(\varepsilon_1^{-1}\rho)^{N-i_0}, \tag{4.57}$$

with $c_N := (4N)^{-N}$, whilst $|s - \sigma(\xi)| \leq \varepsilon_0\varepsilon_1^{-1}\rho$ and

$$|\langle\gamma^{(i)} \circ \sigma(\xi), \xi\rangle| \leq 2^N\varepsilon_0\lambda(\varepsilon_1^{-1}\rho)^{N-i} \qquad \text{for all } i_0 < i \leq N-1.$$

By Taylor's theorem,

$$\begin{aligned} |\langle\gamma^{(i_0)}(s), \xi\rangle - &\langle\gamma^{(i_0)} \circ \sigma(\xi), \xi\rangle| \\ &\leq \sum_{i=i_0+1}^{N-1} 2^N\varepsilon_0^{1+i-i_0}\lambda(\varepsilon_1^{-1}\rho)^{N-i}(\varepsilon_1^{-1}\rho)^{i-i_0} + B\lambda(\varepsilon_0\varepsilon_1^{-1}\rho)^{N-i_0} \\ &\leq (c_N\varepsilon_0/2)\lambda(\varepsilon_1^{-1}\rho)^{N-i_0}, \end{aligned} \tag{4.58}$$

provided the constant ε_0 is chosen small enough depending on B and d (in particular, $\varepsilon_0 := (\min_{1\leq N\leq d} c_N)(50(2+B))^{-1}$ would suffice). Combining (4.57) and (4.58), we deduce that

$$|\langle\gamma^{(i_0)}(s), \xi\rangle| \sim_{\epsilon_0} \lambda(\varepsilon_1^{-1}\rho)^{N-i_0} \qquad \text{for all } s \in \operatorname{supp}_s a^{n,\nu},$$

which implies the lower bound in (4.55). □

In view of (4.54), we restrict our attention to $\mathfrak{D}_s\mathcal{A}[a^{n,\nu}, \gamma]$ for fixed $n \in \mathbb{N}$ and $\nu \in \mathbb{Z}$. Before proceeding to its analysis, we make the following elementary observation about the size of $\rho := 2^n\lambda^{-1/N}$. From the definition (4.43) of β_1, note

that
$$\varepsilon_1^{-2}/4 \le \rho^{-2}G(\xi,s) \quad \text{for} \quad (\xi,s)\in \operatorname{supp}_{\xi,s} a^n.$$
Combining this with (4.45), we deduce that $\rho = O_{B,d}(\varepsilon_1\varepsilon_0^{-1})$. Thus, by choosing ε_1 small enough depending on ε_0, B, d, (in particular, set $\varepsilon_1 := \varepsilon_0/(10dB^{2d+1})$) we can assume that
$$\rho \le B^{-2d}. \tag{4.59}$$

In the following subsections, the norm bounds for the operator $\mathfrak{D}_s\mathcal{A}[a^{n,\nu},\gamma]$ are obtained using the induction hypothesis via a method of rescaling.

4.3.6 Rescaling for the curve

In this subsection, we describe the rescaling map in a generic setting and describe its basic properties, which will play a crucial role in the proof of Lemma 4.3.11.

For $\gamma \in \mathfrak{G}(B,N)$ and $s_\circ \in I$, let
$$V_{s_\circ}^N := \operatorname{span}\{\gamma^{(1)}(s_\circ),\dots,\gamma^{(N)}(s_\circ)\}.$$
Using (4.14), note that $\dim V_{s_\circ}^N = N$. For $0<\rho<1$, define a linear operator $T_{s_\circ,\rho}^N$ such that
$$T_{s_\circ,\rho}^N\left(\gamma^{(i)}(s_\circ)\right) := \rho^i\gamma^{(i)}(s_\circ) \qquad \text{for } 1\le i\le N \tag{4.60}$$
and
$$T_{s_\circ,\rho}^N v = \rho^N v \qquad \text{for } v\in (V_{s_\circ}^N)^\perp.$$
It is clear that $T_{s_\circ,\rho}^N$ is a well-defined map such that
$$\|(T_{s_\circ,\rho}^N)^{-1}\| \lesssim_B \rho^{-N}. \tag{4.61}$$
Supposing $[s_\circ-\rho, s_\circ+\rho]\subseteq I$, we define the rescaled curve
$$\gamma_{s_\circ,\rho}^N(s) := \left(T_{s_\circ,\rho}^N\right)^{-1}(\gamma(s_\circ+\rho s)-\gamma(s_\circ)).$$
For simplicity, we introduce the notations
$$T := T_{s_\circ,\rho}^N, \quad T^* := \left(T_{s_\circ,\rho}^N\right)^{-\top} \quad \text{and} \quad \tilde\gamma := \gamma_{s_\circ,\rho}^N. \tag{4.62}$$
The following lemma verifies nondegeneracy assumptions for the rescaled curve.

Lemma 4.3.15. *For $0<\rho\le B^{-2d}$ and $\gamma\in\mathfrak{G}(B,N)$, the rescaled curve $\tilde\gamma$ lies in $\mathfrak{G}(B_1,N-1)$ where B_1 depends only on B and N.*

Proof of Lemma 4.3.15. We begin by verifying the first part of (4.14) for the curve $\tilde\gamma$. From the definition, we see that $\tilde\gamma^{(i)}(s) = \rho^i T^{-1}(\gamma^{(i)}(s_\circ+\rho s))$ for any

$i \in \mathbb{N}$. Combining this identity with (4.14) and (4.61), we deduce that

$$\|\tilde{\gamma}^{(i)}\|_{L^\infty(I)} = O_B(\rho) \qquad \text{whenever } N+1 \le i \le 2d. \tag{4.63}$$

Let $1 \le i \le N$. By Taylor's theorem, (4.60) and (4.61), we have

$$\begin{aligned}\tilde{\gamma}^{(i)}(s) &= \rho^i \sum_{j=i}^{N} T^{-1}\gamma^{(j)}(s_\circ)\frac{(\rho s)^{j-i}}{(j-i)!} + O_B(\|T^{-1}\|\rho^{N+1}) \\ &= \sum_{j=i}^{N} \gamma^{(j)}(s_\circ)\frac{s^{j-i}}{(j-i)!} + O_B(\rho)\end{aligned} \tag{4.64}$$

Combining (4.64) with (4.14), we obtain uniform size estimates for $\tilde{\gamma}^{(i)}(s)$ when $1 \le i \le N$. Together with (4.63), this implies

$$\|\tilde{\gamma}\|_{C^{2d}(I)} \lesssim_B 1. \tag{4.65}$$

It remains to verify the second part in (4.14) for the curve $\tilde{\gamma}$ and $L = N-1$. In view of (4.65), it suffices to obtain a lower bound for the determinant of the $d \times N$ matrix whose columns vectors are formed by $(\tilde{\gamma}^{(i)}(s))_{1\le i\le N}$. Observe that using the multilinearity of the determinant and elementary column operations, (4.64) gives

$$|\det\begin{pmatrix}\tilde{\gamma}^{(1)}(s) & \cdots & \tilde{\gamma}^{(N)}(s)\end{pmatrix}| = |\det\begin{pmatrix}\gamma^{(1)}(s_\circ) & \cdots & \gamma^{(N)}(s_\circ)\end{pmatrix}| + O_B(\rho).$$

By assumption, ρ is small enough so that the above identity combined with (4.14) gives the estimate

$$|\det\begin{pmatrix}\tilde{\gamma}^{(1)}(s) & \cdots & \tilde{\gamma}^{(N)}(s)\end{pmatrix}| \ge (2B)^{-1}.$$

Now, an application of (4.65) (in particular, $|\tilde{\gamma}^{(N)}(s)| \lesssim_B 1$) completes the proof of (4.14) for $\gamma = \tilde{\gamma}$, $L = N-1$ and B replaced with a new constant B_1. □

The rescaling map $T^N_{s_\circ,\rho}$ can be used to introduce a rescaling for the operators we are interested in. This is done in the next subsection.

4.3.7 Rescaling for the operator

To introduce the operator rescaling, we begin by considering a Schwartz function $u : \mathbb{R} \to \mathbb{R}$. Let $s_\circ \in I$ and $0 < \rho < 1$. Direct computations give

$$[(1+\sqrt{-\partial_s^2})^{1/2}u](s_\circ + \rho s) = \rho^{-1/2}[(\rho + \sqrt{-\partial_s^2})^{1/2}\tilde{u}](s),$$

where $\tilde{u}(s) := u(s_\circ + \rho s)$. Thus,

$$\begin{aligned}\|(1+\sqrt{-\partial_s^2})^{1/2}u\|_{L^2(\mathbb{R})}^2 &\sim \int_{\mathbb{R}} |(\rho+|\sigma|)^{1/2}\mathcal{F}_s(\tilde{u})(\sigma)|^2\mathrm{d}\sigma \\ &\leq \int_{\mathbb{R}} |(1+|\sigma|)^{1/2}\mathcal{F}_s(\tilde{u})(\sigma)|^2\mathrm{d}\sigma \\ &= \|(1+\sqrt{-\partial_s^2})^{1/2}\tilde{u}\|_{L^2(\mathbb{R})}^2. \end{aligned} \tag{4.66}$$

For an arbitrary symbol $a \in C^{3d}(\mathbb{R}^d \times I \times I)$ and $\gamma \in \mathfrak{G}(B, N)$, recall the definition of $\mathcal{A}[a,\gamma]$ from (4.7). Temporarily fixing $x \in \mathbb{R}^d$, set

$$u(s) = \mathcal{A}[a,\gamma]g(x,s) \qquad \text{and} \qquad \tilde{\mathcal{A}}[a,\gamma]g(x,s) := \mathcal{A}[a,\gamma]g(x, s_\circ + \rho s). \tag{4.67}$$

By combining (4.66) for each $x \in \mathbb{R}^d$ with Fubini's theorem,

$$\|\mathfrak{D}_s\mathcal{A}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})} \lesssim \|(1+\sqrt{-\partial_s^2})^{1/2}\tilde{\mathcal{A}}[a,\gamma]g\|_{L^2(\mathbb{R}^{d+1})}. \tag{4.68}$$

We claim that for $(x,s) \in \mathbb{R}^{d+1}$, the identity

$$\tilde{\mathcal{A}}[a,\gamma]g(x,s) = |\det T^*|^{1/2}\mathcal{A}[\tilde{a},\tilde{\gamma}]\tilde{g}(T^{-1}x, s) \tag{4.69}$$

holds with T, $\tilde{\gamma}$ as in (4.62), symbol

$$\tilde{a}(\xi, s, t) := a(T^*\xi, t, s_\circ + \rho s)$$

and input function $\tilde{g}$ defined by

$$\mathcal{F}_x(\tilde{g})(\xi, t) := |\det T^*|^{1/2} e^{it\langle T^{-1}\gamma(s_\circ), \xi\rangle}\mathcal{F}_x(g)(T^*\xi, t).$$

Verifying (4.69) is just a matter of unwinding the definitions. First, we expand $\tilde{\mathcal{A}}[a,\gamma]g(x,s)$ using (4.67) as the oscillatory integral

$$\int_{\mathbb{R}^d \times I} e^{i\langle x - t(\gamma(s_\circ+\rho s)-\gamma(s_\circ)), \xi\rangle} a(\xi, s_\circ+\rho s, t) e^{it\langle\gamma(s_\circ),\xi\rangle}\mathcal{F}_x(g)(\xi,t)\mathrm{d}\xi\mathrm{d}t.$$

Applying change of variables $\xi \to T^*\xi$, the above expression can be written as

$$\begin{aligned} |\det T^*|^{1/2} &\int_{\mathbb{R}^d\times\mathbb{R}} e^{i\langle(T^{-1}x - t\tilde{\gamma}(s), \xi\rangle} a(T^*\xi, s_\circ+\rho s, t)\mathcal{F}_x(\tilde{g})(T^*\xi, t)\mathrm{d}\xi\mathrm{d}t \\ &= |\det T^*|^{1/2}(\mathcal{A}[\tilde{a},\tilde{\gamma}]\tilde{g})(T^{-1}x, s), \end{aligned} \tag{4.70}$$

proving the claim (4.69).

Fix $n \in \mathbb{N}$, $\nu \in \mathbb{Z}$ and recall the definitions of $a^{n,\nu}$ and $s_{n,\nu}$ from §4.3.5. Consider the rescaling map T as defined in §4.3.6 for

$$s_\circ = s_{n,\nu} \qquad \text{and} \qquad \rho = 2^n\lambda^{-1/N}.$$

Furthermore, we consider the operator rescaling as in (4.69) for $a = a^{n,\nu}$. In this setup, we record some of the basic observations about how T^* (as defined in

(4.62)) interacts with $\tilde{a}$.

Lemma 4.3.16. *The rescaling map T^* satisfies the estimate*

$$\rho^{-N}|\xi| \lesssim_{A,B} |T^*\xi| \lesssim_B \rho^{-N}|\xi| \qquad \textit{for all} \quad \xi \in \operatorname{supp}_\xi \tilde{a}. \tag{4.71}$$

Proof. Fix $1 \leq i \leq N$. From the definition of T, we have

$$\langle \gamma^{(i)}(s_{n,\nu}), \xi \rangle = \rho^i \langle T^{-1}\gamma^{(i)}(s_{n,\nu}), \xi \rangle = \rho^i \langle \gamma^{(i)}(s_{n,\nu}), T^*\xi \rangle. \tag{4.72}$$

Fix $\xi \in \operatorname{supp}_\xi \tilde{a}$ so that, by the definition of the rescaled symbol, $T^*\xi \in \operatorname{supp}_\xi a^{n,\nu}$. Using Lemma 4.3.14 when $i \leq N-1$ and the Cauchy–Schwarz inequality (or (4.40)) when $i = N$, we obtain

$$\big|\langle \gamma^{(i)}(s_{n,\nu}), T^*\xi \rangle\big| \lesssim_{A,B} \rho^{N-i}\big|T^*\xi\big| \quad \text{for } 1 \leq i \leq N.$$

Combining this with (4.72), we deduce that

$$|\langle \gamma^{(i)}(s_{n,\nu}), \xi \rangle| \lesssim_{A,B} \rho^N |T^*\xi|. \tag{4.73}$$

On the other hand, if $v \in (V^N_{s_{n,\nu}})^\perp$ is a unit vector, one can argue as in (4.72) to have

$$|\langle v, \xi \rangle| = |\rho^N \langle v, T^*\xi \rangle| \leq |T^*\xi| \tag{4.74}$$

where the fact $\rho < 1$ has been used. Combining (4.73), (4.74) and (4.14), we obtain the lower bound in (4.71). The upper bound follows from (4.61). □

The following lemma now verifies how rescaling improves the type condition of the symbol.

Lemma 4.3.17. *The rescaled symbol $\tilde{a}$ is of type $(\rho^N\lambda, A_1, N-1)$ with respect to $\tilde{\gamma}$, where A_1 depends only on A, B and N.*

Proof of Lemma 4.3.17. By Lemma 4.3.16, it is clear that

$$\operatorname{supp}_\xi \tilde{a} \subseteq \{\xi \in \mathbb{R}^d : |\xi| \sim_{A,B} \rho^N\lambda\}.$$

Thus, the proof of the lemma reduces to verifying Definition 4.3.3 ii) and iii) for $a = \tilde{a}$, $\gamma = \tilde{\gamma}$ and $L = N-1$. For now, we restrict our focus on the latter.

Recall from Lemma 4.3.14 that

$$\sum_{i=1}^{N-1} \rho^{i-N}|\langle \gamma^{(i)}(s), \xi \rangle| \sim_{A,B} \lambda \quad \text{for all } (\xi, s) \in \operatorname{supp}_{\xi,s} a^{n,\nu}. \tag{4.75}$$

However, by unwinding the definition,

$$\langle \tilde{\gamma}^{(i)}(s), \xi \rangle = \rho^i \langle T^{-1}\gamma^{(i)}(s_{n,\nu} + \rho s), \xi \rangle = \rho^i \langle \gamma^{(i)}(s_{n,\nu} + \rho s), T^*\xi \rangle.$$

Thus, by (4.75) and Lemma 4.3.16,

$$\sum_{i=1}^{N-1} |\langle \widetilde{\gamma}^{(i)}(s), \xi\rangle| \sim_{A,B} \rho^N |T^* \xi| \sim |\xi| \quad \text{for all } (\xi, s) \in \operatorname{supp}_{\xi,s} \tilde{a},$$

obtaining the required estimate (4.25) for the rescaled setup for $L = N-1$ and A replaced with a new constant A_1.

It remains to verify Definition 4.3.3 ii) for the rescaled setup for $L = N-1$. Since, a is of type (λ, A, L) with respect to γ by assumption, there exists a λ-admissible tuple $\boldsymbol{R} = (R_{N+1}, \ldots, R_{d+1})$ such that the inequalities (4.22) to (4.24) hold in the support of a for $L = N$. Since a genuine rescaling has happened at this stage, a non-trivial coefficient will be added to $\boldsymbol{R}$. We define

$$R_N := \rho^{-1}.$$

By (4.59), it is known that $R_N \geq B^{2d} > 1$. The modified $(d+1-N)$-tuple is introduced as

$$\tilde{\boldsymbol{R}} := (R_N, R_{N+1}, \ldots, R_{d+1}).$$

It is our aim to establish (4.22) to (4.24) for the rescaled setup for $L = N-1$ and $\boldsymbol{R} = \tilde{\boldsymbol{R}}$. But first, we must check if $\tilde{\boldsymbol{R}}$ is $(\rho^N\lambda)$-admissible. To this end, observe that (4.23) combined with (4.40) gives

$$r_N \left(\prod_{i=N+1}^{d+1} R_i^N \right) \lesssim |\langle \gamma^{(N)}(s), \xi\rangle|^{-1} \approx |\xi|^{-1} \approx \lambda^{-1}$$

for $(\xi, s) \in \operatorname{supp}_{\xi,s} a^{n,\nu}$. Recalling (4.19), we see that $N \in Z(\tilde{\boldsymbol{R}}, N-1, \rho^N\lambda)$ (observe that $(\rho^N\lambda)^{\tilde{\boldsymbol{R}}}_{N-1,N} = \lambda$) and as a consequence,

$$J(\tilde{\boldsymbol{R}}, N-1, \rho^N\lambda) = N. \tag{4.76}$$

Furthermore, it follows from the definitions that

$$(\rho^N\lambda)^{\tilde{\boldsymbol{R}}}_{N-1,j} = \lambda^{\boldsymbol{R}}_{N,j} \qquad \text{for } j \geq N. \tag{4.77}$$

Therefore,

$$J(\tilde{\boldsymbol{R}}, j, (\rho^N\lambda)^{\tilde{\boldsymbol{R}}}_{N-1,j}) = J(\boldsymbol{R}, j, \lambda^{\boldsymbol{R}}_{N,j}) \qquad \text{for } j \geq N. \tag{4.78}$$

Combining (4.76), (4.78) and (4.21), we conclude that $\tilde{\boldsymbol{R}}$ is $(\rho^N\lambda)$-admissible.

We proceed to the proof of (4.22) for this setup. Let $\beta \in \mathbb{N}_0$. From (4.16) and (4.17), it is clear that the function $M_\gamma^{\beta,\boldsymbol{R},N}$ pointwise dominate $M_\gamma^{\alpha,\boldsymbol{R},N}$ whenever $\alpha \in \mathbb{N}_0$ and $\alpha \leq \beta$. Thus, one can use Leibnitz rule and (4.22) for $L = N$ to deduce that

$$\begin{aligned} |\partial_s^\beta \tilde{a}(\xi, s, t)| &= \rho^\beta |\partial_s^\beta a^{n,\nu}(T^*\xi, s_{n,\nu} + \rho s, t)| \\ &\lesssim R_N^{-\beta} M_\gamma^{\beta,\boldsymbol{R},N}(T^*\xi, s_{n,\nu} + \rho s) + 1 \end{aligned} \tag{4.79}$$

for $(\xi, s, t) \in \operatorname{supp} \tilde{a}$. By direct computations, we have

$$|\langle \gamma^{(j+k)}(s_{n,\nu} + \rho s), T^*\xi \rangle| = \rho^{-(j+k)} |\langle \tilde{\gamma}^{(j+k)}(s), \xi \rangle| \tag{4.80}$$

for any $j, k \in \mathbb{N}$ and $(\xi, s) \in \operatorname{supp}_{\xi,s} a^{n,\nu}$. Combining (4.80) with (4.15), we get

$$R_N^{-\beta} \prod_{i=1}^{m} N_\gamma^{j_i,k_i,\boldsymbol{R},N}(T^*\xi, s_{n,\nu} + \rho s) = R_N^{-(\beta - \sum_{i=1}^m k_i)} \prod_{i=1}^{m} N_{\tilde{\gamma}}^{j_i,k_i,\tilde{\boldsymbol{R}},N-1}(\xi, s)$$

for any $m \in \mathbb{N}, \mathbf{k} = (k_1, \dots, k_m) \in \mathcal{Z}(\beta, m)$ and $\mathbf{j} \in \mathcal{W}(m)$. Recall from the definitions that $\sum_{i=1}^m k_i \leq \beta$ and $R_N \geq 1$. Therefore, we deduce that

$$R_N^{-\beta} M_\gamma^{\beta,\boldsymbol{R},N}(T^*\xi, s_{n,\nu} + \rho s) \leq M_{\tilde{\gamma}}^{\beta,\tilde{\boldsymbol{R}},N-1}(\xi, s).$$

Combining this with (4.79), we obtain

$$|\partial_s^\beta \tilde{a}(\xi, s, t)| \lesssim M_{\tilde{\gamma}}^{\beta,\tilde{\boldsymbol{R}},N-1}(\xi, s) + 1$$

for $(\xi, s, t) \in \operatorname{supp} \tilde{a}$, completing the proof of (4.22) for $a = \tilde{a}$, $L = N - 1$ and $\boldsymbol{R} = \tilde{\boldsymbol{R}}$.

Proofs of (4.23) and (4.24) for the rescaled setup are easier and follow almost immediately from the same inequalities for the non-rescaled setup, using (4.80) and (4.77). This completes the proof of Lemma 4.3.17. □

4.3.8 Proof of Lemma 4.3.11

With all the available components, the operator estimate for $\mathfrak{D}_s\mathcal{A}[a^n, \gamma]$ for $n \geq 1$ now follows easily.

Proof of Lemma 4.3.11. Fix $n \geq 1$. Temporarily fix $\nu \in \mathbb{Z}$. In view of (4.68) and (4.69) for $a = a^{n,\nu}$, we have

$$\|\mathfrak{D}_s\mathcal{A}[a^{n,\nu}, \gamma]g\|_{L^2(\mathbb{R}^{d+1})} \lesssim \|\mathfrak{D}_s\mathcal{A}[\tilde{a}, \tilde{\gamma}]\tilde{g}\|_{L^2(\mathbb{R}^{d+1})}. \tag{4.81}$$

Suppose $\tilde{\zeta} \in C_c^\infty(\mathbb{R})$ is chosen such that $\operatorname{supp} \tilde{\zeta} \subseteq [-4, 4]$, $\tilde{\zeta}(r) = 1$ when $|r| \leq 3$ and

$$\sum_{\nu \in \mathbb{Z}} \tilde{\zeta}(\,\cdot\, - \nu) \lesssim 1.$$

In view of the support properties of $a^{n,\nu}$ (in particular (4.44) and (4.53)), we have

$$\tilde{\zeta}(\varepsilon_0^{-1}\varepsilon_1\rho^{-1}(\sigma(T^*\xi) - s_{n,\nu})) = 1 \qquad \text{for } \xi \in \operatorname{supp}_\xi \tilde{a}.$$

Consequently, recalling the integral expression (4.70), it is clear that one can replace $\tilde{g}$ with $\tilde{g}^{n,\nu}$ in (4.81) where

$$\tilde{g}^{n,\nu} := \tilde{\zeta}\left(\varepsilon_0^{-1}\varepsilon_1\rho^{-1}(\sigma \circ T^*(\tfrac{1}{i}\partial_x) - s_{n,\nu})\right)\tilde{g}.$$

Now, Lemma 4.3.15 and Lemma 4.3.17 ensures that the rescaled pair $(\tilde{a}, \tilde{\gamma})$ satisfy the assumptions of Proposition 4.3.6 with $L = N-1$ (note that (4.59) ensures that ρ is of the right size, as assumed in Lemma 4.3.15). Thus, the statement of the proposition applies and we obtain

$$\begin{aligned}\|\mathfrak{D}_s\mathcal{A}[\tilde{a},\tilde{\gamma}]\tilde{g}^{n,\nu}\|_{L^2(\mathbb{R}^{d+1})} &\lesssim_{A,B,d} (\log \rho^N\lambda)^{(N-2)/2}\|\tilde{g}^{n,\nu}\|_{L^2(\mathbb{R}^{d+1})}\\ &\lesssim_{A,B,d} (\log \lambda)^{(N-2)/2}\|\tilde{g}^{n,\nu}\|_{L^2(\mathbb{R}^{d+1})}. \end{aligned} \tag{4.82}$$

Thus, the proof of Lemma 4.3.11 reduces to summing the above estimates in ν without further loss in λ. Using (4.54), Plancherel's theorem and the support properties of symbols $a^{n,\nu}$, we combine (4.82) for different values of ν to deduce that

$$\begin{aligned}\|\mathfrak{D}_s\mathcal{A}[a^n,\gamma]g\|^2_{L^2(\mathbb{R}^{d+1})} &\lesssim_d \sum_{\nu\in\mathbb{Z}}\|\mathfrak{D}_s\mathcal{A}[a^{n,\nu},\gamma]g\|^2_{L^2(\mathbb{R}^{d+1})}\\ &\lesssim_{A,B,d} (\log\lambda)^{N-2}\sum_{\nu\in\mathbb{Z}}\|\tilde{g}^{n,\nu}\|^2_{L^2(\mathbb{R}^{d+1})}.\end{aligned}$$

After a change of variable, it is evident that $\|\tilde{g}^{n,\nu}\|_{L^2(\mathbb{R}^{d+1})} = \|g^{n,\nu}\|_{L^2(\mathbb{R}^{d+1})}$, where

$$g^{n,\nu} := \tilde{\zeta}\left(\varepsilon_0^{-1}\varepsilon_1\rho^{-1}(\sigma(\tfrac{1}{i}\partial_x) - s_{n,\nu})\right) g.$$

Thus, by another application of Plancherel's theorem,

$$\begin{aligned}\|\mathfrak{D}_s\mathcal{A}[a^n,\gamma]g\|^2_{L^2(\mathbb{R}^{d+1})} &\lesssim_{A,B,d} (\log\lambda)^{N-2}\sum_{\nu\in\mathbb{Z}}\|g^{n,\nu}\|^2_{L^2(\mathbb{R}^{d+1})}\\ &\lesssim_{A,B,d} (\log\lambda)^{N-2}\|g\|^2_{L^2(\mathbb{R}^{d+1})}\end{aligned}$$

concluding the proof. □

4.4 Sharpness of Theorem 4.1.2 and Theorem 4.1.3

By acting the maximal operator on standard test functions, here we discuss the sharpness of Theorem 4.1.2 in the range of p and the sharpness of Theorem 4.1.3 in the dependence of the operator norm on $\log\delta^{-1}$.

4.4.1 Sharpness of the range of p in Theorem 4.1.2

Fix $p \in [1,\infty)$, a curve $\gamma \in \mathbb{R}^d$ and assume that given any $\epsilon > 0$, we have

$$\|\mathcal{N}^\gamma_{\mathbf{r}}\|_{L^p(\mathbb{R}^{d+1})\to L^p(\mathbb{R}^d)} \lesssim_\epsilon r_d^{-\epsilon} \qquad \text{for all } \mathbf{r} \in (0,1)^d. \tag{4.83}$$

Temporarily fix ϵ and $\mathbf{r}$. Let $\mathbf{h} = (0,\dots,0,1) \in \mathbb{R}^{d+1}$ and define $g_{\mathbf{r}} := \chi_{\mathbf{h}+B(0,r_d)}$. It is easy to see that the r_d-neighbourhood of the curve $-\gamma$ is a subset of the super-level set

$$\{x \in \mathbb{R}^d : |\mathcal{N}^\gamma_{\mathbf{r}} g_{\mathbf{r}}(x)| \gtrsim r_d\}.$$

Applying Chebyshev's inequality and using (4.83), we have

$$r_d r_d^{(d-1)/p} \lesssim_\epsilon r_d^{(d+1)/p-\epsilon}.$$

Letting $r_d \to 0$, we see that $p \geq 2-\epsilon$. Letting $\epsilon \to 0$, we conclude that $p \geq 2$. Thus, L^p operator norm of $\mathcal{N}_{\mathbf{r}}^\gamma$ has polynomial blowup in r_d^{-1} for $p \in [1,2)$.

4.4.2 Sharpness of the operator norm in Theorem 4.1.3

Fix $\delta \in (0,1)$. Consider the vectors $\boldsymbol{w} := (x,0)$, $\boldsymbol{z} := (y,0)$ in $\mathbb{R}^{d+1}$. It follows from the definition that

$$\boldsymbol{w} + T_\delta(r) \cap \boldsymbol{z} + T_\delta(s) \neq \emptyset$$

if and only if there exists a $t \in [-1,1]$ such that

$$(x-y) + t(\gamma(r) - \gamma(s)) = O(\delta). \tag{4.84}$$

Assuming $|\gamma(s)| \sim 1$ for all $s \in [-1,1]$, it is also not hard to see that

$$\mathrm{Vol}_{\mathbb{R}^{d+1}}(\boldsymbol{w} + T_{10\delta}(r) \cap \boldsymbol{z} + T_\delta(s)) \sim \frac{\delta^{d+1}}{\delta + |\gamma(r) - \gamma(s)|} \tag{4.85}$$

whenever (4.84) holds.

Fixing $(x,r) \in \mathbb{R}^d \times I$, set $f_\delta := \chi_{\boldsymbol{w}+T_{10\delta}(r)}$ and note that $\|f_\delta\|_{L^2(\mathbb{R}^{d+1})} \sim \delta^{d/2}$. Fix $0 \leq k \leq \lfloor \log(\delta^{-1}) \rfloor$, define

$$A_k := \{y \in \mathbb{R}^d : |\mathcal{N}_\delta^\gamma f_\delta(y)| \sim 2^{-k}\}.$$

We claim that

$$|A_k| \gtrsim 2^{2k}\delta^d.$$

Indeed, in view of (4.85), A_k contains all points $y \in \mathbb{R}^d$ for which there exist $s, r \in [-1,1]$ such that (4.84) holds and $|\gamma(s) - \gamma(r)| \sim 2^k\delta$. The latter condition ensures that the admissible directions $\gamma(s)$ belong to a portion of the curve which is contained inside a ball of radius $\sim 2^k\delta$. Moreover, for a fixed direction $\gamma(s)$, any $y \in \mathbb{R}^d$ that lies in the δ-neighbourhood of the tube $x + \{t(\gamma(r) - \gamma(s)) : t \in [-1,1]\}$ satisfies (4.84). Therefore, A_k contains the δ-neighbourhood of a two-dimensional cone in $\mathbb{R}^d$ of diameter $\sim 2^k\delta$, justifying our claim. Thus,

$$(\log \delta^{-1})\delta^d \lesssim \sum_{k=0}^{\lfloor \log(\delta^{-1}) \rfloor} 2^{-2k}|A_k| \leq \|\mathcal{N}_\delta^\gamma\|^2_{L^2(\mathbb{R}^{d+1}) \to L^2(\mathbb{R}^d)} \|f_\delta\|^2_{L^2(\mathbb{R}^{d+1})}.$$

Consequently, we see that

$$\|\mathcal{N}_\delta^\gamma\|_{L^2(\mathbb{R}^{d+1}) \to L^2(\mathbb{R}^d)} \gtrsim (\log \delta^{-1})^{1/2}.$$

In view of the above, we may conjecture that $(\log \delta^{-1})^{1/2}$ is the sharp L^2 operator norm of $\mathcal{N}_\delta^\gamma$. In other words, it is possible that Theorem 4.1.3 gives only a partial result in this direction.

4.5 A discussion on Geometric methods

In this section, we discuss the scope and challenges of crafting a geometric proof for Theorem 4.1.2. We begin by considering a simpler setup where an easy geometric argument is sufficient to obtain sharp operator estimates for the maximal function.

4.5.1 Case of the classical Nikodym maximal function in the plane

Fix $0 < \delta < 1$. For $v \in \mathbb{S}^{d-1}$, consider the δ-tube

$$T_\delta(v) := \{y \in \mathbb{R}^d : |y \cdot v| \leq 1 \text{ and } |\mathrm{proj_{v^\perp} y}| \leq \delta\}.$$

By the *classical* Nikodym maximal function, we refer to the maximal function defined as

$$\mathcal{N}_\delta^{\mathrm{clas}} g(x) := \sup_{v \in \mathbb{S}^{d-1}} \frac{1}{|T_\delta(v)|} \int_{T_\delta(v)} g(x-y)\mathrm{d}y \qquad \text{for } x \in \mathbb{R}^d$$

for $g \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Following Córdoba's well known argument [14] for the Kakeya maximal theorem, we present here a geometric argument, which obtains sharp L^2 estimates for $\mathcal{N}_\delta^{\mathrm{clas}}$ when $d = 2$. Although the author was unable to find a reference to this argument, it is understood that the argument is widely known in the field. The result is as follows.

Proposition 4.5.1. *There exists $C > 0$ such that for any $0 < \delta < 1$, we have*

$$\|\mathcal{N}_\delta^{\mathrm{clas}}\|_{L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)} \leq C(\log \delta^{-1})^{\frac{1}{2}}. \tag{4.86}$$

Proof. Fix $0 < \delta < 1$. The argument can be split into several steps.
Localisation: Note that $\mathcal{N}_\delta^{\mathrm{clas}}$ is a local operator. In other words, if Q and $Q^* \subseteq \mathbb{R}^2$ denote the cubes centered at the origin with side-lengths 1 and 3 respectively, we have

$$\mathcal{N}_\delta^{\mathrm{clas}} \chi_Q(x) = 0 \qquad \text{for } x \notin Q^*.$$

In view of the above, we claim that the proof of (4.86) follows from the operator norm estimate

$$\|\mathcal{N}_\delta^{\mathrm{clas}} g\|_{L^2(Q)} \lesssim (\log \delta^{-1})^{\frac{1}{2}} \|g\|_{L^2(\mathbb{R}^2)}, \qquad \text{for } g \in \mathcal{S}(\mathbb{R}^2) \tag{4.87}$$

where the implicit constant is independent of δ. To see this, we fix $g \in \mathcal{S}(\mathbb{R}^2)$ and write

$$\|\mathcal{N}_\delta^{\mathrm{clas}} g\|^2_{L^2(\mathbb{R}^2)} = \sum_{m \in 2\mathbb{Z}^2} \|\mathcal{N}_\delta^{\mathrm{clas}} g\|^2_{L^2(m+Q)}.$$

Now, translation invariance of $\mathcal{N}_\delta^{\text{clas}}$ combined with its local behaviour allows one to extend (4.87) to the bound

$$\|\mathcal{N}_\delta^{\text{clas}} g\|_{L^2(m+Q)}^2 \lesssim (\log \delta^{-1}) \|g\|_{L^2(m+Q^*)}^2 \qquad \text{uniformly in } m \in \mathbb{R}^2.$$

Thus,

$$\|\mathcal{N}_\delta^{\text{clas}} g\|_{L^2(\mathbb{R}^2)}^2 \lesssim (\log \delta^{-1}) \sum_{m \in 2\mathbb{Z}^2} \|g\|_{L^2(m+Q^*)}^2 \lesssim (\log \delta^{-1}) \|g\|_{L^2(\mathbb{R}^2)}^2,$$

obtaining the bound (4.86).

Discretisation and Dualisation. It is easy to see that the maximal function is locally constant at scale δ. Motivated by this observation, we discretize the setup at the same scale. Suppose $\{B_i \subset Q : 1 \le i \lesssim \delta^{-2}\}$ be a maximal cover of Q consisting of balls of radius δ. For each ball B_i, let x_i be chosen such that

$$\sup_{x \in B_i} |\mathcal{N}_\delta^{\text{clas}} g(x)| \le 2 |\mathcal{N}_\delta^{\text{clas}} g(x_i)|.$$

For each x_i, we can find a direction $v_i \in \mathbb{S}^1$ such that

$$|\mathcal{N}_\delta^{\text{clas}} g(x_i)| = \frac{1}{|T_\delta(v_i)|} \int_{T_\delta(v_i)} |g(x_i - y)| \mathrm{d}y.$$

Set $T_\delta^i := x + T_\delta(v_i)$. Combining the above estimates and the duality principle,

$$\begin{aligned}
\|\mathcal{N}_\delta^{\text{clas}} g\|_{L^2(Q)} &\lesssim \left(\sum_{1 \le i \lesssim \delta^{-2}} |\mathcal{N}_\delta^{\text{clas}} g(x_i)|^2 \delta^2 \right)^{1/2} \\
&\lesssim \left(\sum_{1 \le i \lesssim \delta^{-2}} \left(\int \chi_{T_\delta^i}(y) |g(y)| \mathrm{d}y \right)^2 \right)^{1/2} \\
&= \sup_{a_i \in \mathbb{C} : \|a_i\|_{\ell^2} \le 1} \left[\int_{\mathbb{R}^2} \sum_{1 \le i \lesssim \delta^{-2}} a_i \chi_{T_\delta^i}(y) \cdot |g(y)| \mathrm{d}y \right] \\
&\le \| \sum_{1 \le i \lesssim \delta^{-2}} a_i \chi_{T_\delta^i} \|_{L^2(\mathbb{R}^2)} \|g\|_{L^2(\mathbb{R}^2)}.
\end{aligned}$$

Therefore, the proof of (4.86) reduces to the estimate

$$\|\sum_{T \in \mathcal{T}} a_T \chi_T\|_{L^2(\mathbb{R}^2)} \lesssim [\log \delta]^{\frac{1}{2}} \qquad \text{for } (a_T)_{T \in \mathcal{T}} \in \mathbb{C}^{(\#\mathcal{T})} \text{ with } \|a_T\|_{\ell^2} \le 1, \tag{4.88}$$

where $\mathcal{T}$ is a collection of δ-tubes whose centers form a δ-net inside Q.

Counting incidences with an angle parameter. Fix a collection $\mathcal{T}$ with the properties mentioned above. For two δ-tubes $T_1, T_2 \in \mathcal{T}$, let $\theta(T_1, T_2)$ denote the angle made by the long directions of the T_1 and T_2. It is easy to see that for $k \ge 0$, we have

$$\theta(T_1, T_2) \approx 2^k\delta \implies |T_1 \cap T_2| \lesssim 2^{-k}\delta. \tag{4.89}$$

Since the angle of separation determines the volume of intersection between two tubes, it is reasonable to make at attempt at controlling the number of incidence pairs of tubes with a specific angular separation.

For a fixed $T_1 \in \mathcal{T}$, let us denote by $\mathcal{T}_k^{T_1}$ the collection of all $T_2 \in \mathcal{T}$ such that $T_1 \cap T_2 \neq \emptyset$ and

$$\theta(T, T_1) \approx 2^k\delta \text{ when } k \geq 1 \qquad \text{or} \qquad \theta(T, T_1) \lesssim \delta \text{ when } k = 0.$$

It is not hard to see that the center of any tube in $\mathcal{T}_k^{T_1}$ lies inside a rectangle of dimensions $1 \times 2^k\delta$. Since the centers of any two tubes in the collection are δ-seperated, we easily obtain the bound

$$\#\mathcal{T}_k^{T} \lesssim 2^k\delta^{-1} \qquad \text{for any } T \in \mathcal{T}. \tag{4.90}$$

We have all the necessary ingredients to prove the required L^2 estimate. Fix a complex sequence $(a_T)_{T\in\mathcal{T}}$. Squaring and expanding the L^2 term in (4.88),

$$\begin{aligned} \|\sum_{T\in\mathcal{T}} a_T\chi_T\|^2_{L^2(\mathbb{R}^2)} &= \sum_{T_1,T_2\in\mathcal{T}} a_{T_1}\bar{a}_{T_2}|T_1\cap T_2| \\ &= \sum_{k=1}^{C\log\delta^{-1}} \sum_{T_1\in\mathcal{T}} a_{T_1} \sum_{T_2\in\mathcal{T}_k^{T_1}} \bar{a}_{T_2}|T_1\cap T_2|, \end{aligned}$$

where C is an absolute constant. Applying the Cauchy–Schwarz inequality in the T_1 sum,

$$\|\sum_{T\in\mathcal{T}} a_T\chi_T\|^2_{L^2(\mathbb{R}^2)} \lesssim \sum_{k=1}^{C\log\delta^{-1}} \left(\sum_{T_1\in\mathcal{T}} |a_{T_1}|^2\right)^{1/2} \left(\sum_{T_1\in\mathcal{T}} (\sum_{T_2\in\mathcal{T}_k^{T_1}} \bar{a}_{T_2}|T_1\cap T_2|)^2\right)^{1/2}. \tag{4.91}$$

However, by another application of the Cauchy–Schwarz inequality and noting the ℓ^2-normalisation of $(a_T)_{T\in\mathcal{T}}$, we have

$$\begin{aligned} \sum_{T_1\in\mathcal{T}} (\sum_{T_2\in\mathcal{T}_k^{T_1}} \bar{a}_{T_2})^2 &\leq \sum_{T_1\in\mathcal{T}} \sum_{T_2\in\mathcal{T}_k^{T_1}} (\#\mathcal{T}_k^{T_1})|a_{T_2}|^2 \\ &\leq (\max_{T\in\mathcal{T}} \#\mathcal{T}_k^{T}) \sum_{T_2\in\mathcal{T}} \sum_{T_1\in\mathcal{T}_k^{T_2}} |a_{T_2}|^2 \\ &\leq (\max_{T\in\mathcal{T}} \#\mathcal{T}_k^{T})^2 \sum_{T_2\in\mathcal{T}} |a_{T_2}|^2 \leq (\max_{T\in\mathcal{T}} \#\mathcal{T}_k^{T})^2. \end{aligned}$$

Combining the above with (4.91) and recalling (4.89) and (4.90), we deduce that

$$\begin{aligned}\|\sum_{T\in\mathcal{T}} a_T\chi_T\|^2_{L^2(\mathbb{R}^2)} &\lesssim \sum_{k=1}^{C\log\delta^{-1}} (\max_{T\in\mathcal{T}} \#\mathcal{T}_k^T)2^{-k}\delta \\ &\lesssim \log\delta^{-1},\end{aligned}$$

completing the proof of (4.88) and thereby (4.86). □

In the following subsection, we attempt to establish L^2 estimates for the maximal function considered in Theorem 4.1.3 by extending the argument just described.

4.5.2 Runnning the Córdoba-type argument for $\mathcal{N}_\delta^\gamma$.

Recall the definition of the maximal function $\mathcal{N}_\delta^\gamma$ from (4.5). For simplicity, we investigate $\mathcal{N}_\delta^\gamma$ only when $d = 2$ and γ is the unit circle in the plane. In this simpler setup, we discuss the possibility of carrying out the argument similar to the one described in the previous subsection (the argument will be referred to as the Córdoba-type argument henceforth) and the necessary modification our problem seeks for.

Fix $0 < \delta < 1$. Let $x \in \mathbb{R}^2$ and $w \in \mathbb{S}^1$. In view of (4.4), we are interested in the tubes $T_w(x)$ defined as the δ-neighbourhood of the line segment

$$L_w(x) := \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} + t \begin{pmatrix} w \\ 1 \end{pmatrix} : t \in [-1, 1] \right\}.$$

Let $Q \subseteq \mathbb{R}^2$ denote the unit square centered at the origin. Suppose $\mathcal{T}$ denotes a collection δ-tubes of the form $T_w(x)$ for $w \in \mathbb{S}^1$ and $x \in \mathbb{R}^2$ such that their 'centers' x form a δ-net inside Q. By following the steps of *localisation, discretisation* and *dualisation* as in the proof of Proposition 4.5.1, we see that the estimate of Theorem 4.1.3,

$$\|\mathcal{N}_\delta^\gamma\|_{L^2(\mathbb{R}^{2+1})\to L^2(\mathbb{R}^2)} \lesssim (\log\delta^{-1}).$$

for centered unit circle γ can be obtained as a consequence of the estimate

$$\|\sum_{T\in\mathcal{T}} a_T\chi_T\|_{L^2(\mathbb{R}^3)} \lesssim \delta(\log\delta^{-1}), \tag{4.92}$$

where $\mathcal{T}$ is a collection of δ-tubes with the above-mentioned properties and $(a_T)_{T\in\mathcal{T}}$ is ℓ^2-normalised. Compared to (4.88), we note that there is an additional δ term in the required dual estimate. This additional factor arises due to the fact that the geometric objects of interest in this setup have a higher co-dimension to what was the case for the classical Nikodym maximal function.

Further simplifying the setup, we restrict our attention to the case when $a_T = [\#\mathcal{T}]^{-\frac{1}{2}}$ for each $T \in \mathcal{T}$. For this special case, we re-write our goal (4.92) as

$$\|\sum_{T\in\mathcal{T}} \chi_T\|_{L^2(\mathbb{R}^3)} \lesssim \delta(\log\delta^{-1})[\#\mathcal{T}]^{\frac{1}{2}}. \tag{4.93}$$

Recall that the Córdoba-type argument relied on two key components. One, a volume estimate for the region of intersection between two δ-tubes. Two, an effective control over the number of tubes incident to a typical tube under a constaint on the volume of intersection between the two (In the proof of Propostion 4.5.1, this translated to a constraint on the angular separation). We will argue along similar lines.

For $T \in \mathcal{T}$, let $l_T := \mathrm{proj}_{\mathbb{R}^2}(T)$ and define $\Theta(T_1, T_2)$ denote the angle between long directions of l_{T_1} and l_{T_2} whenever $T_1, T_2 \in \mathcal{T}$. For a fixed $T_1 \in \mathcal{T}$, define $\mathcal{T}_k^{T_1}$ to be the collection of all tubes $T \in \mathcal{T}$ such that $T_1 \cap T_2 \neq \emptyset$ and

$$\Theta(T, T_1) \approx 2^k\delta \quad \text{when } k \geq 1 \qquad \text{and} \qquad \Theta(T, T_1) \lesssim \delta \quad \text{when } k = 0.$$

In the following lemma, we record a volume estimate on the intersecting region between two tubes and also an estimate on the number of incident neighbours to a tube.

Lemma 4.5.2 (Intersection lemma)**.** *For $T_1, T_2 \in \mathcal{T}$, we have*

$$|T_1 \cap T_2| \lesssim \frac{\delta^3}{\delta + \Theta(T_1, T_2)}. \tag{4.94}$$

Furthermore, for any $T \in \mathcal{T}$ and $k \in \mathbb{N}$, we have

$$\#\mathcal{T}_k^T \lesssim 2^{2k}. \tag{4.95}$$

Proof. The proof of the volume bound (4.94) is immediate once we project the tubes onto the $\mathbb{R}^2$ plane. Thus, we may restrict our attention to (4.95).

For $\alpha \in [0, 1)$, let $w(\alpha) := e^{2\pi i\alpha}$. To simplify notations, we use the notation T_α to represent $T_{w(\alpha)}$ for the rest of the discussion. From the definition, it follows that $T_\alpha(x) \cap T_\beta(y) \neq \emptyset$ whenever there exists $t \in [-1, 1]$ with

$$(x - y) + t(w(\alpha) - w(\beta)) = O(\delta).$$

Fix $x \in \mathbb{R}^2, \alpha \in [0, 1)$. If $T_\beta(y) \in \mathcal{T}_k^{T_\alpha}$, we have

$$2^k\delta \leq |\theta(T_\alpha(x), T_\beta(y))| = |w(\alpha) - w(\beta)| \leq 2^{k+1}\delta.$$

Thus, without losing generality, we may assume that β lies in a sub-interval (β_k, β_{k+1}) of $[0, 1]$ with length $2^k\delta$. By multiple applications of law of cosines, we can see that[4] the angle between $w(\beta_k) - w(\alpha)$ and $w(\beta_{k+1}) - w(\alpha)$ is at most $O(2^k\delta)$. Therefore, whenever $T_\beta(y) \in \mathcal{T}_k^{T_\alpha}$, we sees that y lies inside the δ neighbourhood of a sector of angular width $O(2^k\delta)$, centered around x, at a distance of at most approximately $2^k\delta$ from x. This region is contained inside a rectangle of side-lengths $O((2^k\delta)^2 + \delta)$ and $O(2^k\delta)$, there its area is bounded above by $O((2^k\delta)^3 + 2^k\delta^2)$. Computing its δ-entropy gives us the desired estimate (4.95). □

[4] Here one has to additionally assume that $|2^k\delta| \leq \frac{\pi}{4}$ for all k, but this is fine as we can decompose the maximal function into finitely many maximal functions for which it holds.

We feed the estimates (4.94) and (4.95) into the Córdoba-argument to see what comes out of it. Expanding the L^2 sum,

$$\begin{aligned}
\|\sum_{T\in\mathcal{T}}\chi_T\|^2_{L^2(\mathbb{R}^3)} &= \sum_{k=1}^{|\log\delta|}\sum_{T_1\in\mathcal{T}}\sum_{T_2\in\mathcal{T}_k^{T_1}}|T_1\cap T_2| \\
&\lesssim \sum_{k=1}^{|\log\delta|}[\#\mathcal{T}](\max_{T\in\mathcal{T}}\#\mathcal{T}_k^T)2^{-k}\delta^2 \\
&\lesssim \delta^2[\#\mathcal{T}]\sum_{k=1}^{|\log\delta|}2^k \\
&\lesssim \delta[\#\mathcal{T}]. \qquad (4.96)
\end{aligned}$$

Clearly, the bound (4.96) is much weaker than what is required by (4.93) and therefore fails to reprove this specific case of the maximal theorem.

The above calculation clearly demonstrates difficulties of executing a simple Córdoba-type argument for the maximal problem that we are interested in. It also conveys the reason why we have presented a proof using Fourier analytic tools rather than geometric argument, although the latter seems to be the most natural method to study a geometric maximal function like $\mathcal{N}_{\mathbf{r}}^{\gamma}$. Nevertheless, in view of the results on the Wolff's circular maximal function [33, 55] or Bourgain's circular maximal function [44, 45], it is reasonable to expect that a geometric argument, if exists, has to carefully investigate the geometry of large collections $\mathcal{T}$ where most tubes intersects with every other tube in the family. This is well beyond the scope of this thesis and we end the discussion here.

Bibliography

[1] Theresa C. Anderson, Laura Cladek, Malabika Pramanik, and Andreas Seeger. Spherical means on the Heisenberg group: stability of a maximal function estimate. *J. Anal. Math.*, 145(1):1–28, 2021.

[2] Sayan Bagchi, Sourav Hait, Luz Roncal, and Sundaram Thangavelu. On the maximal function associated to the spherical means on the Heisenberg group. *New York J. Math.*, 27:631–675, 2021.

[3] David Beltran, Shaoming Guo, Jonathan Hickman, and Andreas Seeger. Sharp L^p bounds for the helical maximal function. *arXiv e-prints*, page arXiv:2102.08272, February 2021.

[4] David Beltran, Jonathan Hickman, and Christopher D. Sogge. Sharp local smoothing estimates for fourier integral operators. In Paolo Ciatti and Alessio Martini, editors, *Geometric Aspects of Harmonic Analysis*, pages 29–105, Cham, 2021. Springer International Publishing.

[5] Nicolas Bourbaki. *Lie groups and Lie algebras. Chapters 1–3.* Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 1989. Translated from the French, Reprint of the 1975 edition.

[6] Jean Bourgain. Averages in the plane over convex curves and maximal operators. *J. Analyse Math.*, 47:69–85, 1986.

[7] Jean Bourgain and Ciprian Demeter. The proof of the l^2 decoupling conjecture. *Ann. of Math. (2)*, 182(1):351–389, 2015.

[8] Calixto P. Calderón. Lacunary spherical means. *Illinois J. Math.*, 23(3):476–484, 1979.

[9] Anthony Carbery. A weighted inequality for the maximal bochner-riesz operator on $\mathbb{R}^2$. *Transactions of the American Mathematical Society*, 287(2):673–680, 1985.

[10] Anthony Carbery and Andreas Seeger. Weighted inequalities for Bochner-Riesz means in the plane. *Q. J. Math.*, 51(2):155–167, 2000.

[11] Michael Christ. Hilbert transforms along curves. I. Nilpotent groups. *Ann. of Math. (2)*, 122(3):575–596, 1985.

[12] Michael Christ. Weak type $(1, 1)$ bounds for rough operators. *Ann. of Math. (2)*, 128(1):19–42, 1988.

[13] Michael Christ. Lecture notes in harmonic analysis and differential equations on nilpotent groups, *Cortona summer school in mathematics*, 1992.

[14] Antonio Córdoba. The Kakeya maximal function and the spherical summation multipliers. *Amer. J. Math.*, 99(1):1–22, 1977.

[15] Antonio Córdoba. Geometric Fourier analysis. *Ann. Inst. Fourier (Grenoble)*, 32(3):vii, 215–226, 1982.

[16] Javier Duoandikoetxea and José L. Rubio de Francia. Maximal and singular integral operators via Fourier transform estimates. *Invent. Math.*, 84(3):541–561, 1986.

[17] Veronique Fischer and Michael Ruzhansky. *Quantization on nilpotent Lie groups*, volume 314 of *Progress in Mathematics*. Birkhäuser/Springer, [Cham], 2016.

[18] G. B. Folland and Elias M. Stein. *Hardy spaces on homogeneous groups*, volume 28 of *Mathematical Notes*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982.

[19] Robert Fraser and Kyle Hambrook. Explicit Salem sets in $\mathbb{R}^n$. *Adv. Math.*, 416:Paper No. 108901, 23, 2023.

[20] Pritam Ganguly and Sundaram Thangavelu. On the lacunary spherical maximal function on the Heisenberg group. *J. Funct. Anal.*, 280(3):Paper No. 108832, 32, 2021.

[21] D. J. H. Garling. *Inequalities: a journey into linear analysis*. Cambridge University Press, Cambridge, 2007.

[22] Gustavo Garrigos, Wilhelm Schlag, and Andreas Seeger. Improvements in wolff's inequality for decompositions of cone multipliers. *preprint*, 2009.

[23] Gustavo Garrigos and Andreas Seeger. On plate decompositions of cone multipliers. *Proceedings of the Edinburgh Mathematical Society*, 52, 07 2007.

[24] Loukas Grafakos. *Modern Fourier analysis*, volume 250 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[25] Larry Guth. Decoupling estimates in fourier analysis, 2022.

[26] Larry Guth, Hong Wang, and Ruixiang Zhang. A sharp square function estimate for the cone in $\mathbb{R}^3$. *Ann. of Math. (2)*, 192(2):551–581, 2020.

[27] Ya Ryong Heo. An endpoint estimate for some maximal operators associated to submanifolds of low codimension. *Pacific J. Math.*, 201(2):323–338, 2001.

[28] L. Hörmander. *The Analysis of Linear Partial Differential Operators IV: Fourier Integral Operators*. Classics in Mathematics. Springer Berlin Heidelberg, 2009.

[29] Alex Iosevich. Maximal operators associated to families of flat curves in the plane. *Duke Mathematical Journal*, 76(2):633 – 644, 1994.

[30] A. W. Knapp and Elias M. Stein. Intertwining operators for semisimple groups. *Ann. of Math. (2)*, 93:489–578, 1971.

[31] Hyerim Ko, Sanghyuk Lee, and Sewook Oh. Maximal estimates for averages over space curves. *arXiv e-prints*, page arXiv:2102.07175, February 2021.

[32] Hyerim Ko, Sanghyuk Lee, and Sewook Oh. Sharp smoothing properties of averages over curves. *arXiv e-prints*, page arXiv:2105.01628, May 2021.

[33] Lawrence Kolasa and Thomas Wolff. On some variants of the Kakeya problem. *Pacific J. Math.*, 190(1):111–154, 1999.

[34] Thomas William Körner. Hausdorff and Fourier dimension. *Studia Math.*, 206(1):37–50, 2011.

[35] Sanghyuk Lee and Ana Vargas. On the cone multiplier in $\mathbb{R}^3$. *J. Funct. Anal.*, 263(4):925–940, 2012.

[36] Akihiko Miyachi. On some estimates for the wave equation in L^p and H^p. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 27(2):331–354, 1980.

[37] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge. Wave front sets, local smoothing and Bourgain's circular maximal theorem. *Ann. of Math. (2)*, 136(1):207–218, 1992.

[38] Detlef Müller and Andreas Seeger. Singular spherical maximal operators on a class of two step nilpotent Lie groups. *Israel J. Math.*, 141:315–340, 2004.

[39] Daniel M. Oberlin. An endpoint estimate for some maximal operators. *Rev. Mat. Iberoamericana*, 12(3):641–652, 1996.

[40] Juan C. Peral. L^p estimates for the wave equation. *J. Functional Analysis*, 36(1):114–145, 1980.

[41] Malabika Pramanik and Andreas Seeger. L^p regularity of averages over curves and bounds for associated maximal operators. *Amer. J. Math.*, 129(1):61–103, 2007.

[42] Fulvio Ricci and Elias M. Stein. Harmonic analysis on nilpotent groups and singular integrals. II. Singular kernels supported on submanifolds. *J. Funct. Anal.*, 78(1):56–84, 1988.

[43] Joris Roos, Andreas Seeger, and Rajula Srivastava. Lebesgue space estimates for spherical maximal functions on Heisenberg groups. *Int. Math. Res. Not. IMRN*, 2022(24):19222–19257, 2022.

[44] Wilhelm Schlag. A generalization of Bourgain's circular maximal theorem. *J. Amer. Math. Soc.*, 10(1):103–122, 1997.

[45] Wilhelm Schlag. $L^p \to L^q$ estimates for the circular maximal function. In *Topics in analysis and its applications*, pages 1–59. World Sci. Publ., River Edge, NJ, 2000.

[46] Silvia Secco. L^p-improving properties of measures supported on curves on the Heisenberg group. II. *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8)*, 5(2):527–543, 2002.

[47] Andreas Seeger, Christopher D. Sogge, and Elias M. Stein. Regularity properties of Fourier integral operators. *Ann. of Math. (2)*, 134(2):231–251, 1991.

[48] Andreas Seeger and James Wright. Problems on averages and lacunary maximal functions. *Banach Center Publications*, 95(1):235–250, 2011.

[49] Aswin Govindan Sheri. l^2 estimates for a nikodym maximal function associated to space curves, 2022.

[50] Aswin Govindan Sheri, Jonathan Hickman, and James Wright. Lacunary maximal functions on homogeneous groups, 2023.

[51] Christopher D. Sogge. Propagation of singularities and maximal functions in the plane. *Invent. Math.*, 104(2):349–376, 1991.

[52] Elias M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[53] Elias M. Stein. Maximal functions. I. Spherical means. *Proc. Nat. Acad. Sci. U.S.A.*, 73(7):2174–2175, 1976.

[54] 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.

[55] Thomas Wolff. A Kakeya-type problem for circles. *Amer. J. Math.*, 119(5):985–1026, 1997.

[56] Thomas Wolff. Local smoothing type estimates on L^p for large p. *Geom. Funct. Anal.*, 10(5):1237–1288, 2000.