Estimates for some rough operators with modulation symmetries

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Lars Becker
aus
Mainz

Bonn 2025

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

Gutachter/Betreuer: Prof. Dr. Christoph Thiele
Gutachter: Prof. Dr. Herbert Koch
Gutachter: Prof. Dr. Andreas Seeger

Tag der Promotion: 05.05.2025
Erscheinungsjahr: 2025

List of publications in this thesis

This is a cumulative thesis consisting of an introduction as well as the following three articles.

- L. Becker. *A degree one Carleson operator along the paraboloid.* preprint, (2023). arXiv:2312.01134
- L. Becker, P. Durcik and F. Y.-H. Lin. *On trilinear singular Brascamp-Lieb integrals.* preprint, (2024). arXiv:2411.00141
- L. Becker. *Sharp Fourier extension for functions with localized support on the circle.* In: *Revista Matemática Iberoamericana.* (2024). DOI 10.4171/RMI/1532

Summary

This thesis contains three articles on inequalities for operators in Euclidean harmonic analysis.

Chapter 1 consists of the article 'A degree one Carleson operator along the paraboloid' [5]. It is concerned with a problem historically motivated by the proof of Carleson's theorem, stating that the Fourier series of a square-integrable function f converges pointwise almost everywhere to f. Carleson's theorem is essentially equivalent to an estimate for the so-called maximally modulated Hilbert transform. We study very rough generalizations, namely maximally modulated singular integrals along certain submanifolds of $\mathbb{R}^d$. Our main result are new L^p estimates for such operators. The proof combines an overarching strategy due to Fefferman with several in this context new ingredients, most notably so-called sparse bounds due to Oberlin and a new square function estimate.

Chapter 2 contains the article 'On trilinear singular Brascamp-Lieb integrals' [8]. It deals with a classification problem for singular Brascamp-Lieb forms and several related problems. Classical examples of such forms are paraproducts, and more singular representatives arise in connection with elliptic partial differential equations on Lipschitz domains. In this more singular context, the theory draws heavily from methods introduced first in the context of the maximal modulation operators relevant to the first article. However, these methods do not always apply, different methods are needed depending on the form in question. We solve the implied classification problem for trilinear singular Brascamp-Lieb forms, working out the relevant features of the form for different methods to apply. Then we use this new insight to prove new estimates and some abstract transference principles.

Chapter 3 consists of the article 'Sharp Fourier extension for functions with localized support on the circle' [7]. This article is about the Tomas-Stein restriction inequality for the circle, one of the starting points of the area of Fourier restriction theory. Among its many applications are, much in the spirit of the motivation of our first article, some optimal convergence results for Bochner-Riesz sums of Fourier series in two dimensions. We are interested in the folklore conjecture that the optimal constant in the Tomas-Stein inequality is attained by constant functions. Our main result is that the conjectured sharp inequality certainly holds for functions supported in a small arc on the circle.

The three articles are preceded by an introduction in which we give the historical motivation for the problems considered in this thesis, elaborate on their connection, and give a more detailed overview of our results.

Acknowledgment

First and foremost, I am deeply grateful to my advisor Christoph Thiele. He taught me a great deal of mathematics, and we had many helpful discussions about the problems in this thesis and harmonic analysis at large. I also thank him for involving me in collaborations that do not appear in this thesis, through which I got in touch with many other stimulating topics. Finally, I would like to thank him for his mentorship during my studies in Bonn.

This thesis and many other projects would not exist without my collaborators.

I thank Alexander Volberg for sharing some of his countless open problems and interesting ideas with me and for introducing me to Joseph Slote and Haonan Zhang. He and Eugenia Malinnikova made it possible for me to visit Stanford and I am extremely thankful for that opportunity. I also want to express my gratitude to Josef Greilhuber for the fun discussions we had at Stanford.

I am also very grateful to Paata Ivanishvili for hosting me in Irvine. The time I spent there was very stimulating and I am still thinking about the questions he introduced me to.

Finally, I am very thankful to Polona Durcik, who taught me about the twisted paraproduct while we wrote the second article in this thesis.

I express my gratitude to all members of the Harmonic Analysis and PDE group in Bonn. In particular, I would like to thank Dimitri Cobb for our many interesting discussions. I am also especially thankful to my academic siblings Valentina Ciccone and Gevorg Mnatsakanyan, and also to Shao Liu and Lorenzo Pompili for being truly great colleagues.

I thank Fred Lin for being the best office mate imaginable and a great friend. Without him, my time as a PhD student would have been a lot less fun.

Finally, I want to express my gratitude to my friends and family, who continuously supported me throughout the writing of this thesis.

Contents

Introduction

In this introduction, we will lay out the motivation for the results in the later chapters of this thesis and explain their relation to each other. The results in the first article contribute to a theory that grew out of questions about the convergence of Fourier series.

0.1 Fourier series and maximal modulation operators

Fourier series

Given an integrable function $f : [0,1] \to \mathbb{C}$, its Fourier coefficients are defined by

$$\hat{f}(n) = \int_0^1 e^{-2\pi i n x} f(x) \, \mathrm{d}x,$$

and its Fourier series is the, a priori formal, sum

$$\sum_{n \in \mathbb{Z}} \hat{f}(n) e^{2\pi i n x}.$$

Fourier series are an incredibly useful tool in both theoretical mathematics and applied disciplines, due to the fact that the Fourier series of many functions f converges, in some appropriate way, to f. A simple rigorous statement to this effect is that the Fourier series of a function $f \in L^2([0,1])$ converges to f with respect to the L^2 norm. This follows from abstract Hilbert space theory, because the functions $x \mapsto e^{2\pi i n x}$ form an orthonormal basis of $L^2([0,1])$.

On the other hand the, from a naive standpoint, more natural question of convergence of the partial Fourier sums

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

as $N \to \infty$ at fixed points x is much harder. It is easy to see that it has a negative answer if convergence is required at all points. For example, let $f \in L^2([0,1])$ with

$$\hat{f}(n) = \begin{cases} 0 & \text{if } n \leq 0 \\ \frac{\epsilon_n}{n} & \text{if } n \geq 1 \end{cases}$$

for some signs $\epsilon_n \in \{-1, 1\}$. Since the harmonic series diverges, the signs ϵ_n can easily be chosen so that

$$\limsup_{N \to \infty} S_N f(0) = \limsup_{N \to \infty} \sum_{n=1}^{N} \frac{\epsilon_n}{n} = +\infty \tag{0.1.1}$$

and

$$\liminf_{N\to\infty} S_N f(0) = \liminf_{N\to\infty} \sum_{n=1}^{N} \frac{\epsilon_n}{n} = -\infty. \tag{0.1.2}$$

Taking linear combinations of shifts of this f, one can construct functions with Fourier series that satisfy (0.1.1) and (0.1.2) in any finite set of points and, with some more care, in any countable set. Katznelson [70] gave an elementary construction showing that even for every null set with respect to Lebesgue measure there exists a continuous function f with Fourier series diverging in all points in that set.

This still leaves open the possibility that the Fourier series converges almost everywhere when $f \in L^2([0,1])$. Luzin [86] formulated this as a conjecture in 1915, and it was answered positively in a celebrated article by Carleson [21] only in 1966.

Theorem 0.1.1 (Carleson [21])**.** *Let $f \in L^2([0,1])$. Then for almost every $x \in [0,1]$*

$$\lim_{N\to\infty} S_N f(x) = f(x).$$

The Hilbert transform

The theory of Fourier series in one dimension is intimately connected to the Hilbert transform, and it plays an important role in the proof of Theorem 0.1.1. To emphasize more clearly the similarity to our later results, we describe the connection in the setting of the Fourier transform on the real line $\mathbb{R}$ rather than for Fourier series on $[0,1]$. In this setting, the Fourier transform of a function $f : \mathbb{R} \to \mathbb{C}$ is defined as

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-ix\xi} f(x) \,\mathrm{d}x$$

and Carleson's theorem takes the following form.

Theorem 0.1.2 (Carleson [21])**.** *Let $f \in L^2(\mathbb{R})$. Then for almost every $x \in \mathbb{R}$*

$$\lim_{N\to\infty} \int_{-N}^{N} \hat{f}(\xi) e^{ix\xi} \,\mathrm{d}\xi = f(x).$$

There is no serious difference between Theorem 0.1.1 and Theorem 0.1.2, they follow from each other by straightforward limiting arguments.

The Hilbert transform is the Fourier multiplier operator $H : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ defined by

$$\widehat{Hf}(\xi) = -i \operatorname{sgn}(\xi) \hat{f}(\xi). \tag{0.1.3}$$

It spans, together with the identity operator, the space of simultaneously dilation- and translation-invariant operators on $L^2(\mathbb{R})$. This space also contains Fourier truncation operators that appear implicitly in Theorem 0.1.2. Explicitly, we can write

$$\int_0^{\infty} \hat{f}(\xi) e^{ix\xi} \,\mathrm{d}\xi = \frac{1}{2}(1 + iH) f(x).$$

The modulation operator $M_N f(x) = e^{-iNx} f(x)$ translates the Fourier transform of f by N:

$$\widehat{M_N f}(\xi) = \hat{f}(\xi + N).$$

This allows expressing also the Fourier truncation at N in terms of the Hilbert transform

$$\int_N^\infty \hat{f}(\xi)e^{ix\xi}\,\mathrm{d}\xi = \frac{1}{2}M_{-N}(1+iH)M_N f(x). \tag{0.1.4}$$

Now, all truncated Fourier integrals in Theorem 0.1.2 are differences of expressions as on the left-hand side of (0.1.4). Thus, Carleson's theorem can be rephrased as a statement about the operators on the right-hand side of (0.1.4), that is, the identity operator and $M_{-N}HM_N$.

If $\hat{f}$ is integrable, then for every x the integral

$$\int_{-\infty}^\infty \hat{f}(\xi)e^{ix\xi}\,\mathrm{d}\xi$$

converges absolutely, so the conclusion of Theorem 0.1.2 holds for f. Such functions are dense in $L^2(\mathbb{R})$. To extend the conclusion of Theorem 0.1.2 to all of $L^2(\mathbb{R})$, one then only needs uniform in N upper bounds for the partial Fourier integrals in terms of the L^2 norm of f, or by equation (0.1.4), for the maximally modulated Hilbert transform $\sup_N |HM_N f|$.

These bounds were the main ingredient in Carleson's proof of Theorem 0.1.2. A slightly stronger version, due to Hunt, is the following.

Theorem 0.1.3 (Carleson [21], Hunt [68])**.** *There exists a constant $C > 0$ such that for every function $f \in L^2(\mathbb{R})$,*

$$\Big\| \sup_{N\in\mathbb{R}} |HM_N f| \Big\|_{L^2(\mathbb{R})} \le C\|f\|_{L^2(\mathbb{R})}.$$

Other proofs of Carleson's theorem, also via slightly weaker bounds for the same operator, were later obtained by Fefferman [51] and Lacey and Thiele [77]. Theorem 0.1.3 continues to hold on L^p for all $p > 1$, see [68], and in fact for functions in the Orlicz space $L\log L\log^3 L$, see [2, 82, 41].

Maximal modulation operators

The Hilbert transform H is an example of a singular integral operator. It can be expressed as

$$Hf(x) = \frac{1}{\pi}\int f(x-y)\frac{1}{y}\,\mathrm{d}y\,.$$

Some care is needed to make sense of this integral, because the kernel $\frac{1}{y}$ is not locally integrable. If f is a Schwartz function, then the integral is well-defined as a principal value integral, or alternatively via the formula (0.1.3).

Theorem 0.1.3 can be generalized by replacing H by certain other singular integral operators. One commonly considered class are Calderón-Zygmund operators. They are defined by convolution with tempered distributions K which agree with a function away from 0 and satisfies

$$|\hat{K}(\xi)| \le 1, \qquad \xi \in \mathbb{R}^d \tag{0.1.5}$$

and

$$|\partial^\alpha K(x)| \le |x|^{-d-|\alpha|}, \qquad x \in \mathbb{R}^d \setminus \{0\} \tag{0.1.6}$$

for all $|\alpha| \le m$, where $m \ge 1$. Such K are called m-Calderón-Zygmund kernels.

The assumption (0.1.5) implies that Calderón-Zygmund operators are bounded on L^2, and the assumptions (0.1.6) on the kernel imply that Calderón-Zygmund operators are bounded on $L^p(\mathbb{R}^d)$ for $1 < p < \infty$. The Hilbert transform is an example of a Calderón-Zygmund operator on $\mathbb{R}$. In higher dimensions, examples arise naturally as $(-d)$-homogeneous versions of differential operators.

A generalization of Theorem 0.1.3 to Calderón-Zygmund operators was given by Sjölin.

Theorem 0.1.4 (Sjölin [107])**.** *For every $d \ge 1$ there exist $m \ge 1$ and $C > 0$ such that the following holds. Let K be an m-Calderón-Zygmund kernel and $Tf = K * f$ the corresponding Calderón-Zygmund operator. Then for every $f \in L^2(\mathbb{R}^d)$*

$$\Big\| \sup_{N \in \mathbb{R}^d} |T M_N f| \Big\|_{L^2(\mathbb{R}^d)} \le C \|f\|_{L^2(\mathbb{R}^d)}.$$

However, we note that for $d \ge 2$ Theorem 0.1.4 no longer has any implications for the convergence of Fourier series. Further far-reaching generalizations of Carleson's theorem for Calderón-Zygmund operators have been obtained in [83, 84, 116, 9].

Singular integrals along submanifolds

The first article of this thesis extends the Carleson-Sjölin bound for maximally modulated singular integrals beyond Calderón-Zygmund operators to the rougher class of singular integral operators along submanifolds.

We will fix the manifold to be the paraboloid

$$\mathbb{P} = \{(x, |x|^2) \ : \ x \in \mathbb{R}^d\} \subset \mathbb{R}^{d+1}$$

with $d \ge 2$. This is mainly for concreteness; relevant is mainly the positive curvature, that the dimension of $\mathbb{P}$ is at least two and that its codimension is one. Let K be an m-Calderón-Zygmund kernel on $\mathbb{R}^d$ as defined above in (0.1.6), satisfying in addition the cancellation condition

$$\int_{r<|x|<R} K(x) \, \mathrm{d}x = 0 \tag{0.1.7}$$

for all $r < R$. We consider the operator T defined on Schwartz functions f by

$$Tf(x', x_{d+1}) = \int_{\mathbb{R}^d} f(x' - y, x_{d+1} - |y|^2) K(y) \, \mathrm{d}y, \tag{0.1.8}$$

where we write $x = (x', x_{d+1}) \in \mathbb{R}^{d+1}$. Thus, T is given by convolution with the tempered distribution

$$W(x) = \delta(x_{d+1} - |x'|^2) K(x')$$

supported on the paraboloid $\mathbb{P}$. The operator T is a singular integral operator that shares many properties with Calderón-Zygmund operators.

The cancellation condition (0.1.7) together with the curvature of the paraboloid imply that $\widehat{W}$ is bounded, thus T defines a bounded operator on $L^2(\mathbb{R}^{d+1})$. Calderón-Zygmund theory can be adapted to prove that T is also bounded on $L^p(\mathbb{R}^{d+1})$ when $1 < p < \infty$ and that various maximal operators associated to T are bounded in the same range, see

[26, 110]. However, it is, for example, an open problem whether the maximal averaging operator associated to T maps $L^1(\mathbb{R}^{d+1})$ into $L^{1,\infty}(\mathbb{R}^{d+1})$, see [30], hinting at the additional difficulties in dealing with operators along submanifolds.

Let $V \subset \mathbb{R}^{d+1}$ be a linear subspace. We will consider the partial maximal modulation operators of T, modulated with frequencies only in V, defined by

$$T_V f(x) = \sup_{N \in V} |T M_N f|(x). \tag{0.1.9}$$

Motivated by Carleson's and Sjölin's results, one is led to ask the following question.

Question 1. *Does there exist a constant $C > 0$ such that for all $f \in L^2(\mathbb{R}^{d+1})$*

$$\|T_{\mathbb{R}^{d+1}} f\|_{L^2(\mathbb{R}^{d+1})} \le C \|f\|_{L^2(\mathbb{R}^{d+1})} \,?$$

This question was explicitly asked, for $d = 1$, for the first time in [104], motivated by a bound for a related maximal operator with polynomial phases due to Pierce and Yung [101]. Various other results for related operators have appeared since then, see [1, 6, 10, 66, 103, 102].

There are two key obstacles towards Question 1. Firstly, the convolution kernel W of the operator $T_{\mathbb{R}^{d+1}}$, being supported on a submanifold, is very rough. The existing proofs of Carleson's theorem rely heavily on the kernel being locally constant at the correct scales. This is used, for example, in integration-by-parts arguments. These arguments fail for W and have to be circumvented.

Secondly, the operator $T_{\mathbb{R}^{d+1}}$ has an exceptionally large group of symmetries. We now go into detail on this second point.

Symmetries

Every maximal modulation operator as in Question 1 or Theorem 0.1.4 is clearly invariant under the respective modulations M_N. This is a key feature for proving estimates for such operators, because it prevents techniques such as Littlewood-Paley decompositions, which rely on the existence of a distinguished frequency 0, from working. The modulation symmetry implies that the absolute frequency at which a function oscillates has no meaning to the symmetric operator. Instead, one needs to work with finer frequency localizations into boxes, and exploit relative oscillation of the different localized pieces. The resulting methods are called time-frequency analysis.

An exceptional feature of Question 1 is an additional symmetry which similarly implies that time-frequency analysis as used in the proofs of Carleson's theorem does not directly apply to it. Namely, the operator $T_{\mathbb{R}^{d+1}}$ is invariant under the quadratic modulations Q_A defined by

$$Q_A f(x', x_{d+1}) = e^{iA(|x'|^2 + x_{d+1})} f(x', x_{d+1}), \quad A \in \mathbb{R}. \tag{0.1.10}$$

Indeed,

$$\begin{aligned} &M_N Q_A f(x' - y, x_{d+1} - |y|^2) \\ &= e^{iA(|x'+y|^2 + x_{d+1} - |y|^2)} M_N f(x' - y, x_{d+1} - |y|^2) \\ &= e^{iA(3|x'|^2 + x_{d+1})} M_{N - 2Ax'} f(x' - y, x_{d+1} - |y|^2)\,, \end{aligned}$$

which implies that

$$T_{\mathbb{R}^{d+1}} Q_A f = T_{\mathbb{R}^{d+1}} f.$$

In fact, it shows that the same is already true for the operator $T_{\mathbb{R}^d \times \{0\}}$ with only 'horizontal' modulations.

Unfortunately, we are still not able to deal with such larger groups of symmetries. In this thesis, we will sidestep the issues caused by them by only proving estimates for T_V when V is a strict subspace of $\mathbb{R}^d \times \{0\}$. In that case there are still modulation symmetries under M_N for $N \in V$, however there are no exceptional symmetries under Q_A.

A more precise explanation of the obstruction caused by the exceptional symmetry under the quadratic modulations Q_A defined in (0.1.10) in our proof is as follows. In proving bounds for maximal modulation operators, one starts by fixing for each point x a frequency $N(x)$, so that the supremum in (0.1.9) is attained at $N = N(x)$ up to a factor two, say. This is also done in this thesis. The function N is called the linearizing function. The above computation shows that the operator with linearizing function N satisfies exactly the same estimates as the operator with linearizing function $N_A(x) = N(x) + Ax'$, for every $A \in \mathbb{R}$. Moreover, the same is true for all auxiliary operators constructed in our proof. However, some estimates in the proof are clearly false in the limit $A \to \infty$. Specifically, consulting our article and the technical definitions therein, it is easy to check that the density of a set of tiles, which depends implicitly on the function N_A, will tend to zero as $A \to \infty$. Thus upper bounds by the density will fail in the presence of the additional symmetry, however the proof heavily relies on such bounds.

0.1.1 Our results

Maximal modulations of singular integrals along paraboloids

Sidestepping the issue of additional modulation symmetries, there remain the obstacles that come from the roughness of W. We are able to address them in the first article of this thesis. Our main theorem is as follows.

Theorem 0.1.5 (Becker [5])**.** *Let $d \geq 2$ and let $m > \frac{d}{2}$. Suppose that $V = \{0\}^d \times \mathbb{R}$ or V is a proper subspace of $\mathbb{R}^d \times \{0\}$. Then for all p with*

$$\frac{d^2 + 4d + 2}{(d+1)^2} < p < 2(d+1)$$

there exists $C > 0$ such that for all m-Calderón-Zygmund kernels K and all Schwartz functions f, we have

$$\|T_V f\|_{L^p(\mathbb{R}^{d+1})} \leq C \|f\|_{L^p(\mathbb{R}^{d+1})}.$$

Our proof of Theorem 0.1.5 is based on a modification of Fefferman's proof of Carleson's theorem in [51]. In addition to various technical complications, the key new ingredients needed to address the roughness of W are sparse bounds and a square function estimate, which we will present now.

Sparse bounds

The first new tool we use are certain localized estimates for singular integrals and maximal averages along the paraboloid, which follow from refinements of so-called sparse bounds.

A collection $\mathcal{S}$ of subcubes of $\mathbb{R}^d$ is called sparse if for every $Q \in \mathcal{S}$ there exists a measurable subset $E(Q) \subset Q$ such that

$$2|E(Q)| \geq |Q|$$

and such that the sets $E(Q)$ are pairwise disjoint. For a cube Q and an exponent $1 \leq p < \infty$, we denote the p-average of a function f over Q by

$$\langle f \rangle_{Q,p} = \Big(\frac{1}{|Q|} \int_Q |f|^p \, \mathrm{d}x \Big)^{1/p}.$$

Let X be some operator mapping measurable functions to measurable functions. The term 'sparse bound' refers to upper bounds for X of the form

$$\Big| \int (Xf) g \, \mathrm{d}x \Big| \leq C \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{3Q,p} \langle g \rangle_{Q,q}, \tag{0.1.11}$$

where $C > 0$, $1 \leq p, q < \infty$ and $\mathcal{S}$ is a sparse collection which might depend on f and g.

Sparse bounds were introduced in 2013 by Lerner [80, 81], who used them to give a new proof of the A_2 theorem, stating that the operator norm of any Calderón-Zygmund operator on the weighted space $L^2(w)$ grows at most linearly in the so-called A_2 characteristic of the weight w. His technique proved to be very successful, and since then sparse bounds have been established for many operators in harmonic analysis, often leading to new or simplified proofs of effective weighted bounds. We refer to [11] for an overview of the literature. For the singular integral operators along submanifolds of Question 1, sparse bounds were essentially established by Lacey [74] and Oberlin [97], see also the work of Cladek and Ou [31].

In this thesis we will use the sparse bounds to obtain good estimates for the operators $\mathbf{1}_E X$ for certain sets E. The sets E will be, in some appropriate sense, small, and we need estimates that quantitatively capture this smallness. To be more precise, we will be given a partition of $\mathbb{R}^{d+1}$ into dyadic cubes, and E will satisfy a thinness condition

$$|E \cap Q| \leq \delta |Q| \tag{0.1.12}$$

for all cubes Q in the partition. The operators X we will estimate are likewise 'adapted' to the same partition, meaning that there is no contribution of scales smaller than the scales of the cubes in the partition. We will then establish a refined version of the sparse bounds of Oberlin [97], which adds to his result the information that in this situation the sparse collection $\mathcal{S}$ can be chosen to only contain cubes refined by the given partition, thus satisfying (0.1.12). After that, Hölder's inequality in the L^q averages in (0.1.11) leads to the improved bound $C\delta^\epsilon$ for the operator norm of the localized operator $\mathbf{1}_E X$ for some positive ϵ, which suffices to prove the L^2 estimates in Theorem 0.1.5.

In the classical setting of Carleson's theorem, such localized estimates with decay in δ are significantly easier to prove. This is because in that setting, the relevant Xf are essentially constant on each cube of the given partition, which directly gives an estimate (0.1.11) with optimal $p = q = 1$ and a disjoint collection of cubes $\mathcal{S}$, rather than just a sparse one.

A square function

We introduce another new ingredient to deal with the rough nature of W, a square function inspired by the article [110] by Stein and Wainger.

Stein and Wainger proved bounds on L^2 for maximal averages along the parabola

$$\sup_{k\in\mathbb{Z}} 2^{-k}\int_0^{2^k} f(x-t,y-t^2)\,\mathrm{d}t =: \sup_{k\in\mathbb{Z}} f*\mu_k.$$

Their proof goes as follows. Replacing μ_k by a function $\varphi_k(x,y) = 2^{-3k}\varphi(2^{-k}x, 2^{-2k}y)$ with rapidly decaying Fourier transform and $\int \varphi_k = \mu_k(\mathbb{R}^2)$ essentially results in the parabolic Hardy-Littlewood maximal function, so this operator satisfies the desired L^2 bounds. It only remains to control the difference, which can be dominated by a square function:

$$\sup_{k\in\mathbb{Z}} |f*(\mu_k-\varphi_k)| \le \Big(\sum_{k\in\mathbb{Z}} |f*(\mu_k-\varphi_k)|^2\Big)^{1/2}.$$

By Plancherel's theorem, the L^2 norm of the square function is at most

$$\Big(\int |\hat{f}(\xi)|^2 \sum_{k\in\mathbb{Z}} |\hat{\mu}_k(\xi)-\hat{\varphi}_k(\xi)|^2\,\mathrm{d}\xi\Big)^{1/2}.$$

Finally, it follows from standard estimates for the decay of the Fourier transform of $\mu_k-\varphi_k$, coming from the curvature of the parabola and the assumption $\int \varphi_k = \mu_k(\mathbb{R}^2)$, that the sum over k is bounded. This completes the proof.

A key step of our argument uses an estimate for a similar maximal function in higher dimensions and with an additional supremum over modulations. Thus, μ_k will now be the measure defined by

$$\int f\,\mathrm{d}\mu_k = |B(0,2^k)|^{-1}\int_{B(0,2^k)} f(x_1,x_2,x_1^2+x_2^2)\,\mathrm{d}x.$$

Like in Stein and Wainger's argument, the low frequency contributions to μ_k can be dealt with using known arguments, here from the proof of Carleson's theorem in [51]. There remain certain high frequency truncations μ_k^δ of μ_k, which for the sake of this discussion can be thought of as being truncated in frequency so that

$$|\hat{\mu}_k^\delta(\xi)| \le \delta,$$

see (1.5.3) for the correct definition. A square function argument as above shows that the maximal function $\sup_{k\in\mathbb{Z}} f*\mu_k^\delta$ satisfies an improved bound on L^2, its norm is at most $C\delta$. In our proof we need, however, estimates for the larger maximal function

$$\sup_{k\in\mathbb{Z}} \sup_{N\in 2^{-k}\mathbb{Z}\times\{0\}^2} |f*M_N\mu_k^\delta| \tag{0.1.13}$$

involving also a supremum over modulations. With the precise definition of μ_k^δ at hand it is not hard to see that the maximal function (0.1.13) is controlled by the positive maximal average along the paraboloid. Estimating in that way, however, loses the information about the Fourier support of μ_k^δ and therefore the δ decay. To obtain estimates with good dependence on δ we instead use a more complex square function argument, keeping track more carefully of the essential Fourier support of the measures μ_k^δ the Fourier transforms of the measures $M_N\mu_k^\delta$.

0.2 Singular Brascamp-Lieb forms

The modulation invariant operators, discussed in the previous section are closely related to modulation invariant bilinear operators such as the bilinear Hilbert transform

$$\mathrm{BHT}_{\alpha,\beta}(f_1, f_2)(x) = \int f_1(x - \alpha t) f_2(x - \beta t) \frac{1}{t} \, \mathrm{d}t. \tag{0.2.1}$$

Besides the similar symmetries of these operators, all known proofs of their boundedness also draw from the same set of techniques. A more direct manifestation of this connection, observed by Kovač, Thiele and Zorin-Kranich in [72, Appendix B3], is that certain estimates for the so-called triangular Hilbert transform

$$\mathrm{THT}(f_1, f_2)(x, y) = \int f_1(x - t, y) f_2(x, y - t) \frac{1}{t} \, \mathrm{d}t \tag{0.2.2}$$

would imply Carleson's Theorem 0.1.3 directly.

The second article in this thesis deals with a joint generalization of the forms (0.2.1) and (0.2.2), so-called trilinear singular Brascamp-Lieb forms

$$\Lambda(f_1, f_2, f_3) = \int_{\mathbb{R}^d} f_1(\Pi_1(x)) f_2(\Pi_2(x)) f_3(\Pi_3(x)) K(\Pi_0(x)) \, \mathrm{d}x. \tag{0.2.3}$$

Here, $f_i : \mathbb{R}^{d_i} \to \mathbb{C}$ are measurable functions, $\Pi_i : \mathbb{R}^d \to \mathbb{R}^{d_i}$ are linear maps, and K is a tempered distribution with

$$|\partial^\alpha \widehat{K}(\xi)| \leq |\xi|^{-|\alpha|}, \qquad |\alpha| \leq m,$$

for some $m \geq 0$. We will call such distributions m-Calderón-Zygmund kernels, noting that the definition differs slightly from the one given in Section 0.1. The notion of singular Brascamp-Lieb form was introduced in [47, 48]. The name is inspired by the non-singular variant without K, for which a fairly complete theory is presented in [13].

The goal of our article is to give criteria on the maps Π_i and on exponents p_1, p_2 and p_3 that determine whether the form Λ satisfies bounds

$$|\Lambda(f_1, f_2, f_3)| \leq C \|f_1\|_{L^{p_1}(\mathbb{R}^{d_1})} \|f_2\|_{L^{p_2}(\mathbb{R}^{d_2})} \|f_3\|_{L^{p_3}(\mathbb{R}^{d_3})} \tag{0.2.4}$$

for some constant $C > 0$ and all functions f_i, or not. We call forms Λ satisfying (0.2.4) $\mathbf{p}$-bounded, where $\mathbf{p} = (p_1, p_2, p_3)$. Proving bounds (0.2.4) is in general a difficult open problem. Our main contribution is towards the implied classification problem: We work out the relevant characteristics of Λ for existing methods to apply.

Before turning to the classification, we briefly give two applications that motivate the study of singular Brascamp-Lieb forms.

Calderón's conjecture and superposition arguments

The bilinear Hilbert transform (0.2.1) was introduced by Calderón in an attempt to prove bounds for the Calderón commutator [18], which comes up naturally in the theory of elliptic partial differential equations on domains with Lipschitz boundary; see, for example, the discussion in [91].

Using Fourier inversion, the bilinear Hilbert transform can expressed - up to a constant factor - as

$$\mathrm{BHT}_{\alpha,\beta}(f_1, f_2)(x) = \iint \hat{f}_1(\xi)\hat{f}_2(\eta)\,\mathrm{sgn}(\alpha\xi + \beta\eta)e^{ix(\xi+\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta.$$

On the other hand, the Calderón commutator has the form

$$B_m(f_1, f_2)(x) = \iint \hat{f}_1(\xi)\hat{f}_2(\eta)m(\xi, \eta)e^{ix(\xi+\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta$$

for a function m satisfying $m(\xi, \eta) = m(\lambda\xi, \lambda\eta)$ for all $\xi, \eta \in \mathbb{R}$, $\lambda > 0$. This function m is determined by its values on the circle $\xi^2 + \eta^2 = 1$. Suppose that its restriction to that circle is odd and of bounded variation. Then there exists a finite measure μ such that

$$m(\xi, \eta) = \int_0^\pi \mathrm{sgn}(\cos(\theta)\xi + \sin(\theta)\eta)\,\mathrm{d}\mu(\theta).$$

Hence, writing $\mathrm{BHT}_\theta = \mathrm{BHT}_{\cos\theta,\sin\theta}$:

$$B_m = \int_0^\pi \mathrm{BHT}_\theta\ \mathrm{d}\mu(\theta).$$

Calderón conjectured that the bilinear operators BHT_θ for $\theta \in [0, \pi]$ are uniformly bounded from $L^2 \times L^2 \to L^1$, which by this argument would imply bounds for all operators B_m, of linear growth in the variation of m on the circle $\xi^2 + \eta^2 = 1$.

In this argument the bilinear Hilbert transforms play the role of elementary building blocks, convex combinations of which give rise to a large class of operators occurring in applications. We will give a generalization of such superposition arguments in Section 0.2.1.

Calderón's conjecture was only resolved in 1997 by Lacey and Thiele [76, 78]. They proved a weaker form of the conjecture, with no control on the θ-dependence of the bound. Uniform bounds for $\theta \in [0, \pi]$ were shown by Thiele [112] into $L^{1,\infty}$ and finally by Grafakos and Li [62] in the precise form needed by Calderón.

Multilinear Ergodic Averages

Another motivation for the study of singular Brascamp-Lieb forms are certain quantitative pointwise convergence results in ergodic theory. We illustrate this connection using the example of the triangular Hilbert transform (0.2.2). No bounds for the triangular Hilbert transform are known. The goal of this section is to motivate its study from applications that are likewise open problems.

Let $(X, \mathcal{A}, \mu)$ be a probability space. Two commuting and invertible measure preserving transformations S, T induce a $\mathbb{Z}^2$ action on X. Our object of interest are the bilinear ergodic averages

$$A_N(f_1, f_2)(x) = \frac{1}{N}\sum_{n=1}^{N} f_1(T^n x)f_2(S^n x).$$

Somewhat similarly as in the setting of Fourier series, $L^2(X, \mu)$ convergence of these averages is known, see [34], but pointwise convergence is open.

Question 2. *Let $f_1, f_2 \in L^2(X)$. Is it true that the sequence $A_N(f_1, f_2)(x)$ converges for μ-almost every x as $N \to \infty$?*

One approach to proving pointwise convergence, introduced in this setting by Bourgain [17], is through estimates for the r-variation of the sequence $A_N(f,g)$. The r-variation of a sequence $(a_n)_{n\in\mathbb{N}}$ is defined as

$$\|a\|_{V^r} = \sup_J \sup_{n_0<\cdots<n_J} \Big(\sum_{j=1}^{J} |a_{n_{j-1}} - a_{n_j}|^r\Big)^{1/r}.$$

Then a positive answer to Question 2 would follow, for example, from an estimate

$$\|\|A_N(f_1, f_2)(x)\|_{V^r(N)}\|_{L^1(X)} \leq C\|f_1\|_{L^2(X)}\|f_2\|_{L^2(X)}. \tag{0.2.5}$$

for some finite r. The Calderón transference principle [19] allows to lift such estimates to the acting group $\mathbb{Z}^2$, to which one can transfer from $\mathbb{R}^2$. In this way (0.2.5) would follow from the estimate

$$\Big\|\Big\|\frac{1}{T}\int_0^T f_1(x-t,y)f_2(x,y-t)\,\mathrm{d}t\Big\|_{V^r(T)}\Big\|_{L^1(\mathbb{R}^2)} \leq C\|f_1\|_{L^2(\mathbb{R}^2)}\|f_2\|_{L^2(\mathbb{R}^2)}. \tag{0.2.6}$$

The expression on the left-hand side bears an obvious resemblance to the triangular Hilbert transform (0.2.2) and motivates its study.

We note that by the same argument there are n-linear ergodic averages associated to all n-linear singular integral operators. For classical, linear, ergodic averages the argument here was executed by Birkhoff [15] and in the variational form by Bourgain [17]. For the bilinear Hilbert transform and the corresponding bilinear ergodic averages it was done in the papers of Lacey [75], Demeter [37] and Do-Oberlin-Palsson [42].

The two-dimensional situation

We return to the topic of our article, the classification of singular Brascamp-Lieb forms.

An important motivation for us is the article [39] of Demeter and Thiele, where they considered the two dimensional variants of the bilinear Hilbert transform

$$\mathrm{BHT}^2_{A,B}(f_1, f_2)(x) = \int_{\mathbb{R}^2} f_1(x+Ay)f_2(x+By)K(y)\,\mathrm{d}y \tag{0.2.7}$$

for linear maps $A, B : \mathbb{R}^2 \to \mathbb{R}^2$. After introducing a dualizing function they are special cases of the notion of singular Brascamp-Lieb form (0.2.3). The latter is more general, allowing, for example, functions with different-dimensional arguments. We will however show, as a byproduct of our classification, that the higher-dimensional bilinear Hilbert transforms are among a small number of interesting cases in which estimates are not either trivially false or elementary.

Certain linear transformations in the integral and functions in (0.2.7) preserve the form of the operator but change A, B. The application of such transformations has no effect on the boundedness properties of the operator. The same applies to singular Brascamp-Lieb forms (0.2.3). It is therefore natural to classify pairs (A, B) modulo the implied equivalence relation, and Demeter and Thiele obtained such a classification in [39] in the two dimensional case (0.2.7).

Their classification identifies a generic, nondegenerate case in which the one-dimensional methods apply. The new feature in two dimensions are several degenerate cases. For some of them Demeter and Thiele were able to adapt the one dimensional methods with a more careful analysis, which they named 'one and a half-dimensional time-frequency analysis'. They however left open one very degenerate case, the 'twisted paraproduct'

$$\int f_1(x-s,y)f_2(x,y-t)K(s,t)\,\mathrm{d}s\,\mathrm{d}t\,.$$

Boundedness in this case was later shown by Kovač [71].

Another case with two dimensional functions, which is not mentioned in the classification of Demeter and Thiele however included in our notion of singular Brascamp-Lieb form, is the triangular Hilbert transform (0.2.2). In this case estimates are still open.

Given the understanding of the two-dimensional situation, it is natural to ask the following slightly imprecise questions.

Question 3. *Are there additional, qualitatively different phenomena in higher dimensions? That is, are the nondegenerate methods of Lacey and Thiele, the degenerate fractional-dimensional time-frequency analysis of Demeter and Thiele and the twisted techniques of Kovac enough to bound all forms other than the triangular Hilbert transform? Given a specific form, how can one determine whether they apply or not?*

0.2.1 Our results

We will show that the short answer to these questions is no, there are no fundamentally new phenomena in higher dimensions. We now make this precise.

Quiver representations

A quiver is a directed graph. A quiver representation assigns to each vertex of the quiver a vector space, and to each arrow a linear map between the corresponding spaces. Quiver representations are studied in the representation theory of finite dimensional algebras, they are the modules of certain algebras associated to the quiver. Because of this, we will from now on call them modules.

The algebraic data $(\Pi_i)_{i=0,\dots,3}$ associated to a singular Brascamp-Lieb form, together with the implicit vector spaces, constitutes a module of the quiver in Figure 1. There is a natural notion of isomorphism of modules. This notion corresponds exactly to the equivalence relation discussed above: Two modules are isomorphic if and only if the corresponding singular Brascamp-Lieb forms can be transformed into one another by linear changes of variables. Thus, the problem of classifying singular Brascamp-Lieb forms is equivalent to the problem of classifying modules of the quiver in Figure 1 up to isomorphism.

This is a classical problem in representation theory. Indeed, up to taking adjoints this is the so-called four subspace problem of classifying configurations of four finite dimensional subspaces of a finite dimensional vector space. It was solved by Gelfand and Ponomarev [58] over algebraically closed fields, and Nazarova [94, 95] over $\mathbb{R}$. We note that with certain mild assumptions on the singular Brascamp-Lieb form, which are satisfied when it

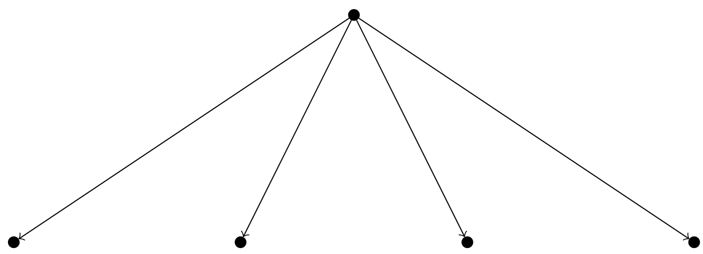

Figure 1: Dual of the four subspace quiver

is (p_1, p_2, p_3)-bounded in the Hölder range $1/p_1+1/p_2+1/p_3 = 1$, the classification problem reduces to the simpler Kronecker normal form [73].

Classification results

The solution of the four subspace problem in [58] is in terms of a list of indecomposable modules, each module can be expressed in a unique way as a direct sum of indecomposables.

We prove the following projection theorem, which implies that the direct summands of the associated module indicate the difficulty of proving estimates for a singular Brascamp-Lieb form.

For technical reasons, we distinguish the data $\mathbf{H} = (\Pi_i)_{i=0,\dots,3}$ of singular Brascamp-Lieb forms and the associated modules $\mathbf{M}$. The module corresponding to the datum $\mathbf{H}$ is denoted $\mathbf{M_H}$ and conversely.

Theorem 0.2.1 (Becker-Durcik-Lin [8])**.** *Let* $\mathbf{M}, \mathbf{M}'$ *be two modules and let* $\mathbf{p} < \infty$*. Let* $\mathbf{H}$ *and* $\mathbf{H} \oplus \mathbf{H}'$ *be data with* $\mathbf{M_H} \cong \mathbf{M}$ *and* $\mathbf{M_{H \oplus H'}} \cong \mathbf{M} \oplus \mathbf{M}'$*. Suppose that for each* l*-Calderón-Zygmund kernel* K *we have*

$$|\Lambda_{\mathbf{H}\oplus\mathbf{H}'}(K, f_1, f_2, f_3)| \le C\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3} .$$

Then there exists a constant C' *such that for each* $2l$*-Calderón-Zygmund kernel* K *we have*

$$|\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3)| \le C'\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3} .$$

Using necessary boundedness conditions for singular Brascamp-Lieb forms that follow from the non-singular results in [13], Theorem 0.2.1 allows us to exclude many indecomposable modules as direct summands of bounded singular Brascamp-Lieb forms. This leads to the following result.

Theorem 0.2.2 (Becker-Durcik-Lin [8])**.** *Let* $1 \le \mathbf{p} < \infty$ *and let* $\mathbf{H}$ *be a* $\mathbf{p}$*-bounded singular Brascamp-Lieb datum with* $H_1, H_2, H_3 \neq \{0\}$*. Then one of the following holds, with the notation from Appendix 2.8.*

i) (Bilinear Hölder-type) There exists an assignment $\{i, j, k\} = \{1, 2, 3\}$ *such that* $\frac{1}{p_j} = \frac{1}{p_k} = 1 - \frac{1}{p_i}$ *and* $n_1, n_2, n_3, n_4 \ge 0$ *such that*

$$\mathbf{M_H} \cong (\mathbf{P}^{(j)})^{\oplus n_1} \oplus (\mathbf{K}^{(j)})^{\oplus n_2} \oplus (\mathbf{P}^{(k)})^{\oplus n_3} \oplus (\mathbf{K}^{(k)})^{\oplus n_4} .$$

ii) (Young-type) We have $\mathbf{p} = (p_1, p_2, p_3)$ *with* $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 2$*. If* $p_1, p_2, p_3 \neq 1$ *then there exist* $n_1, n_2 \geq 0$ *such that*

$$\mathbf{M_H} \cong \mathbf{Y}^{\oplus n_1} \oplus \mathbf{Z}^{\oplus n_2} \,.$$

If there is some $i \in \{1,2,3\}$ *with* $p_i = 1$*, then there exist* $n_1, n_2, n_3, n_4 \geq 0$ *such that*

$$\mathbf{M_H} \cong \mathbf{Y}^{\oplus n_1} \oplus \mathbf{Z}^{\oplus n_2} \oplus (\mathbf{P}^{(i)})^{\oplus n_3} \oplus (\mathbf{K}^{(i)})^{\oplus n_4} \,.$$

iii) (Loomis-Whitney-type) We have $\mathbf{p} = (2,2,2)$ *and there exist* $n_1, n_2 \geq 0$ *and a list of modules* $\mathbf{M}_1, \dots, \mathbf{M}_k$ *from Table 2.4 with*

$$\mathbf{M_H} \cong \mathbf{L}^{\oplus n_1} \oplus \mathbf{B}^{\oplus n_2} \oplus \mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_k \,.$$

iv) (Hölder-type) We have $\mathbf{p} = (p_1, p_2, p_3)$ *with* $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$*. In this case, there exists a finite list of modules* $\mathbf{M}_1, \dots, \mathbf{M}_k$ *from Table 2.2 such that*

$$\mathbf{M_H} \cong \mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_k \,.$$

The first two cases are trivial, being essentially built by combining a product and a linear singular integral operator, or a convolution and a linear singular integral operator. In these cases the necessary conditions are already sufficient; bounds follow from linear singular integral bounds and Hölder's or Young's inequality, respectively. The third case is more interesting, but all forms in this case can still be bounded by elementary methods, by combining Plancherel's theorem and the Loomis-Whitney inequality. These observations yield the following theorem.

Theorem 0.2.3 (Becker-Durcik-Lin [8])**.** *Let* $\mathbf{M_H}$ *and* $\mathbf{p}$ *be as in case i), ii) or iii) of Theorem 2.1.15. Then* $\mathbf{H}$ *is* $\mathbf{p}$*-bounded.*

This leaves the Hölder exponent case (iv) as the most interesting one. In that case there are the four families $\mathbf{N}_n, \mathbf{C}_n, \mathbf{T}_n, \mathbf{J}_n^{(i)}$ of indecomposable modules, see Table 2.2. The first two correspond to nondegenerate cases, and boundedness of all singular Brascamp-Lieb forms associated to direct sums of them is by now well understood, see for example [56]. The family $\mathbf{T}_n$ includes the triangular Hilbert transform $\mathbf{T}_1$ and higher dimensional versions of it. Bounds for forms in this family, and by Theorem 0.2.1 for any forms containing direct summands from this family, are therefore likely outside of reach of current methods.

Taking direct sums of modules $\mathbf{J}_1^{(i)}$ yields forms with 'twisted' behavior. We prove new bounds for all such sums as a special case of the following result.

Theorem 0.2.4 (Becker-Durcik-Lin [8])**.** *Let* $n \geq 1$ *and let* $\mathbf{M} = (\mathbf{J}_1^{(1)} \oplus \mathbf{J}_1^{(2)} \oplus \mathbf{J}_1^{(3)} \oplus \mathbf{C}_1)^{\oplus n}$*. Let* $2 < \mathbf{p} < \infty$ *and let* $\mathbf{H}$ *be a singular Brascamp-Lieb datum associated with* $\mathbf{M}$*. Then there exists* l *and* $C > 0$ *such that for all* l*-Calderón-Zygmund kernels* K*, we have*

$$|\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3)| \leq C \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_2} \,.$$

A single $\mathbf{J}_n^{(j)}$ leads to fractional time-frequency analysis, whereas multiple modules $\mathbf{J}_n^{(j)}$ with $n \geq 2$, or direct sums including also $\mathbf{N}_n$ for some n or $\mathbf{C}_n$ with $n \geq 2$ introduces both twisted and modulation invariant features. Currently, there are no bounds for forms of the latter type in the literature, but there is no fundamental obstruction to proving them. Together with Theorems 0.2.2 and 0.2.3 this provides our answer to Question 3.

Method of rotations

We complement the projection Theorem 0.2.1 with a general superposition result, showing that every Caldéron-Zygmund kernel on $\mathbb{R}^d$ can be expressed as a superposition of Caldéron-Zygmund kernels on $d-1$-dimensional subspaces of $\mathbb{R}^d$.

At its heart is the following proposition about decompositions of mean zero functions on spheres. Fix $d \geq 3$ and denote

$$S^{d-1} = \{x \in \mathbb{R}^d \ : \ |x| = 1\}.$$

Let σ be the normalized $(d-1)$-dimensional Hausdorff probability measure on S^{d-1}. Further, if $\nu \in S^{d-1}$, let σ_ν be the normalized $(d-2)$-dimensional Hausdorff probability measure on the great circle

$$(\operatorname{span} \nu)^\perp \cap S^{d-1}\,.$$

Finally, denote by $\mathcal{M}_d$ the manifold

$$\{(x,y) \in S^{d-1} \times S^{d-1} \ : \ x \cdot y = 0\}$$

and by $H_0^s(S^{d-1})$ the mean zero functions in the Sobolev space $H^s(S^{d-1})$.

Proposition 0.2.5 (Becker-Durcik-Lin [8])**.** *Let $d \geq 3$ and $s > 1/2$. There exists a constant $C > 0$ such that the following holds. Let $\Omega \in H_0^s(S^{d-1})$. Then there exists a function $\Gamma : \mathcal{M}_d \to \mathbb{C}$ such that*

- *for all $\nu \in S^{d-1}$*

$$\int_{(\operatorname{span}\nu)^\perp \cap S^{d-1}} \Gamma(\nu, \theta)\, d\sigma_\nu(\theta) = 0\,.$$

  and

$$\|\Gamma(\nu, \cdot)\|_{H_0^{s-1/2}((\operatorname{span}\nu)^\perp \cap S^{d-1})} \leq C \|\Omega\|_{H_0^s(S^{d-1})}\,.$$

- *as measures, we have*

$$\Omega(\theta)\sigma(\theta) = \int_{S^{d-1}} \Gamma(\nu, \theta)\sigma_\nu(\theta)\, d\sigma(\nu)\,.$$

Moreover, Γ can be chosen so that the mapping $\Omega \mapsto \Gamma$ is continuous from $C^k(S^{d-1})$ into $C^k(\mathcal{M}_d)$, for every k.

The proposition states that every mean zero function Ω on S^{d-1} can be expressed as a superposition of functions $\Gamma(\nu, \cdot)$ that are supported on the great circles $\nu^\perp \cap S^{d-1}$ and have mean zero, and that one has control of the $H^{s-1/2}$ norm of these functions.

A corollary of Proposition 0.2.5 are superposition arguments for singular Brascamp-Lieb forms in a similar spirit as for the Calderón commutator. However, one should expect the estimates for the forms with kernels supported on subspaces to be significantly harder to prove than the estimates for the original form. So this method is unlikely to give new bounds.

0.3 Fourier restriction inequalities

We turn to the last article of this thesis, which is about a sharp version of the Tomas-Stein Fourier restriction inequality.

The classical Fourier restriction problem asks about the possibility of defining the restriction of the Fourier transform of a function to a subset of $\mathbb{R}^d$. By Plancherel's theorem, the Fourier transform is a bijective isometry $L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. Thus, it is not possible to meaningfully restrict the Fourier transform of an $L^2(\mathbb{R}^d)$ function to a measure zero subset of $\mathbb{R}^d$. On the other extreme, the Fourier transform of a function in $L^1(\mathbb{R}^d)$ is always continuous, so it is well-defined pointwise. In between, the Hausdorff-Young inequality implies that the Fourier transform of an $L^p(\mathbb{R}^d)$ function is in $L^{p'}(\mathbb{R}^d)$. However, the Fourier transform is no longer surjective, so it might still make sense to restrict it to some sets of measure zero.

Stein observed [50, Page 28] that this is indeed sometimes possible. His result was later improved by Tomas and then made optimal by Stein [113]. Let σ be the arc length measure on the unit circle

$$S^1 = \{(x, y) \ : \ x^2 + y^2 = 1\} \subset \mathbb{R}^2.$$

Theorem 0.3.1 (Tomas, Stein [113]). *There exists a constant $C > 0$ such that for all functions $f \in L^1(\mathbb{R}^2) \cap L^{6/5}(\mathbb{R}^2)$*

$$\Big(\int |\hat{f}(\xi)|^2 \, \mathrm{d}\sigma(\xi) \Big)^{1/2} \le C \|f\|_{L^{6/5}(\mathbb{R}^2)}. \tag{0.3.1}$$

Since $L^1(\mathbb{R}^2) \cap L^{6/5}(\mathbb{R}^2)$ is dense in $L^{6/5}(\mathbb{R}^2)$, it follows that there exists a unique operator

$$R : L^{6/5}(\mathbb{R}^2) \to L^2(\sigma)$$

which agrees with the restriction of the Fourier transform for $L^1(\mathbb{R}^2)$ functions.

The adjoint of the Fourier restriction operator R is the Fourier extension operator

$$E : L^2(\sigma) \to L^6(\mathbb{R}^2), \qquad f \mapsto \widehat{f\sigma}.$$

By duality, the Tomas-Stein inequality (0.3.1) is then equivalent to the adjoint bound for the extension operator, which reads

$$\|\widehat{f\sigma}\|_{L^6(\mathbb{R}^2)} \le C \|f\|_{L^2(\sigma)}. \tag{0.3.2}$$

To illustrate its relevance beyond restricting Fourier transforms of $L^{6/5}$ functions, we want to mention two classical applications of the Tomas-Stein inequality.

Bochner-Riesz summation of Fourier series

For the first application we return to the question of convergence of Fourier series, now of functions $f : [0, 1]^2 \to \mathbb{C}$. Since there is no natural ordering of $\mathbb{Z}^2$, one has to choose the order in which to sum to Fourier series. One natural choice is to sum in order of increasing magnitude of the frequency, leading to the following question.

Question 4. *Let $f \in L^p([0,1]^2)$. Is it true that the circular Fourier sums*

$$\sum_{k_1^2+k_2^2 \le R^2} \hat{f}(k_1,k_2) e^{2\pi i(k_1x_1+k_2x_2)} \tag{0.3.3}$$

converge to f in $L^p([0,1]^2)$ as $R \to \infty$?

When $p = 2$ the answer is yes, this is a simple consequence of Plancherel's theorem. Surprisingly, however, the answer is no as soon as $p < 2$, as shown by Fefferman in his celebrated paper on the ball multiplier [52]. Thus, one is led to introduce a weight of smoothness $\lambda > 0$

$$\Big(1 - \frac{|k|^2}{R^2}\Big)_+^\lambda = \begin{cases} \Big(1 - \frac{|k|^2}{R^2}\Big)^\lambda & \text{if } |k| \le R, \\ 0 & \text{if } |k| \ge R, \end{cases}$$

in the sums (0.3.3), to make them better behaved.

Question 5. *Let $1 \le p \le \infty$. For which $\lambda > 0$ is it true that for all $f \in L^p([0,1]^2)$, the Bochner-Riesz sums*

$$\sum_{k \in \mathbb{Z}^2} \Big(1 - \frac{|k|^2}{R^2}\Big)_+^\lambda \hat{f}(k_1,k_2) e^{2\pi i(k_1x_1+k_2x_2)} \tag{0.3.4}$$

converge to f in $L^p([0,1]^2)$ as $R \to \infty$?

In the two dimensional setting considered here, this question was resolved for all $1 \le p \le \infty$ by Carleson and Sjölin [22]. If one is only interested in the range $p \le 6/5$, however, their result already follows as a simple application of the Tomas-Stein restriction inequality (0.3.1), see [108, Page 422]. The argument exploits localization of pieces of the Bochner-Riesz sum to pass from $L^{6/5}$ to L^2, and then expresses the radial Fourier multipliers inherent in the Bochner-Riesz sum as superpositions of the Fourier restriction operators onto concentric circles.

Strichartz estimates for the Schrödinger equation

Fourier restriction inequalities find another important application in the theory of dispersive partial differential equations. The Tomas-Stein inequality is not specific to the arc length measure on the circle, but holds for many measures supported on manifolds, the relevant property being the nonvanishing curvature of the manifold. In particular, a similar theorem applies to the parabola. In that setting inequality (0.3.2) gives information about solutions of the Schrödinger equation in $1+1$ dimension

$$\begin{aligned} i\partial_t u + \Delta u &= 0, \\ u(0,x) &= f(x). \end{aligned}$$

Taking the Fourier transform of the Schrödinger equation yields

$$(\eta - \xi^2)\hat{u}(\eta, \xi) = 0.$$

Thus distributional solutions of the Schrödinger equation are supported on the parabola $\xi^2 = \eta$. Taking a Fourier transform only in x and solving an ODE also yields

$$\mathcal{F}_x[u](t,\xi) = e^{it|\xi|^2} \hat{f}(\xi),$$

giving the more precise formula

$$\hat{u}(\eta, \xi) = \delta(\eta - |\xi|^2)\hat{f}(\xi).$$

The version of the Tomas-Stein extension estimate for the paraboloid can then be stated either as an L^6 estimate for the solution of the Schrödinger equation

$$\|u(t,x)\|_{L^6(\mathbb{R}^2)} \leq C\|\hat{f}\|_{L^2(\mathbb{R})} = C\|f\|_{L^2(\mathbb{R})}, \tag{0.3.5}$$

or, in a form more closely resembling (0.3.2), as

$$\|\widehat{g\mu}\|_{L^6(\mathbb{R}^2)} \leq C\|g\|_{L^2(\mu)} \tag{0.3.6}$$

with $\mu = \delta(\eta - \xi^2)\mathrm{d}\eta\mathrm{d}\xi$ and $g(\xi, \eta) = \hat{f}(\xi)$.

Sharp constants

Having introduced Fourier restriction inequalities in general, we now turn to the specific problem considered in this thesis. We are interested in the exact optimal constant C in the inequality (0.3.2).

This problem is motivated by work of Foschi [54], who found the optimal constant in the Tomas-Stein extension inequality from the two dimensional sphere

$$S^2 = \{x \in \mathbb{R}^3 \ : \ |x| = 1\}.$$

Let σ_2 denote the surface measure on S^2.

Theorem 0.3.2 (Foschi [54])**.** *If*

$$\|f\|_{L^2(\sigma_2)} \leq \|1\|_{L^2(\sigma_2)}$$

then

$$\|\widehat{f\sigma_2}\|_{L^4(\mathbb{R}^3)} \leq \|\widehat{\sigma_2}\|_{L^4(\mathbb{R}^3)}.$$

In other words, constant functions are maximizers for the Fourier extension inequality for S^2. Foschi also found the sharp constants in inequality (0.3.5) for the parabola and the cone in two and three dimensions, in that case the maximizers are Gaussians [55]. The results for the parabola were also independently obtained by Hundertmark and Zharnitsky [67]. There are numerous other results on sharp constants in extension inequalities in the literature, for higher-dimensional spheres, other manifolds, some giving full characterizations of maximizers, some just showing existence or regularity of maximizers, some showing nonexistence of maximizers. We refer to [96] for a survey of the literature.

For the Tomas-Stein inequality for the circle, Theorem 0.3.1, a characterization of the extremizers and the sharp constant are, however, still open. It is known that maximizers exist and are smooth, this was shown by Shao [105, 106]. It is also known by work of Carneiro, Foschi, Oliveira e Silva and Thiele [23] that constant functions are local maximizers. In light also of Theorem 0.3.2 it is then natural to make the following conjecture.

Conjecture 0.3.3. *Constant functions maximize the Tomas-Stein extension inequality for S^1. That is, if*

$$\|f\|_{L^2(\sigma)} \leq \|1\|_{L^2(\sigma)}$$

then

$$\|\widehat{f\sigma}\|_{L^6(\mathbb{R}^2)} \leq \|\widehat{\sigma}\|_{L^6(\mathbb{R}^2)}.$$

In their paper [23], Carneiro, Foschi, Oliveira e Silva and Thiele proposed a program to prove Conjecture 0.3.3 along the lines of Foschi's proof of Theorem 0.3.2. They managed to reduce Conjecture 0.3.3 to a conjecture about positive semidefiniteness of a certain quadratic form Q on a subspace of $L^2(\mathbb{T}^3)$, where $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$. A function f on $\mathbb{T}^3$ is called antipodal if it is π-periodic in each argument.

Conjecture 0.3.4. *Let*

$$\mathrm{d}\Sigma = \delta(\sum_{j=1}^{6} e^{i\theta_j}) \prod_{j=1}^{6} \mathrm{d}\theta_j$$

and

$$Q(f) = \iint_{\mathbb{T}^6} (|e^{i\theta_1} + e^{i\theta_2} + e^{i\theta_3}|^2 - 1)(|f(\theta_1,\theta_2,\theta_3)|^2 - f(\theta_1,\theta_2,\theta_3)\overline{f(\theta_4,\theta_5,\theta_6)})\,\mathrm{d}\Sigma.$$

Then $Q(f) \geq 0$ for all antipodal $f \in L^2(\mathbb{T}^3)$.

0.3.1 Our results

The final article of this thesis establishes a partial result towards Conjecture 0.3.4. It draws inspiration for a guiding strategy from the two papers [3, 99] of Barker, Oliveira e Silva, Thiele and Zorin-Kranich. These two papers contain rigorous numerical computations of the eigenvalues and eigenfunctions of the quadratic form Q when restricted to the finite dimensional space of trigonometric polynomials of degree at most 120. After various reductions, these numerical results paint a clear picture: There are a small number of eigenfunctions with very small eigenvalues, and they can be identified to be very close to the eigenfunctions of the quadratic form

$$M(f) = \int_{\mathbb{T}^3} m(\theta)|f(\theta)|^2 \prod_{j=1}^{3} \mathrm{d}\theta_j,$$

where m is an explicit nonnegative function. In addition to this, there is a large space of functions without any noticeable structure, on which the quadratic form seems to be safely positive, and a few very large eigenvalues. For a more detailed discussion, we refer to [3].

The eigenfunctions of M are, of course, easy to understand. They concentrate around the level sets of m, which can be computed in the asymptotic range of eigenvalues close to zero. This analysis shows that they correspond to functions in the original problem that concentrate near two antipodal points. Motivated by this, we study in our article the quadratic form Q on such functions. We obtain, after various reductions, an asymptotic expansion with M as the main term. Then we prove effective bounds to obtain positivity of the quadratic form for functions under an explicit small support assumption.

The precise statement is as follows. Let V be the space of all antipodal functions in $L^2(\mathbb{T}^3, \mathbb{R})$. Let C_ε be the cylinder of radius ε centered at the line $\mathbb{R}(1,1,1)$, and define

$$V_\varepsilon := \left\{ f \in V \ : \ \operatorname{supp} f \subset C_\varepsilon / (2\pi\mathbb{Z})^3 \right\}.$$

Theorem 0.3.5 (Becker [7])**.** *Let $\varepsilon = 1/20$. Then for all $f \in V_\varepsilon$ it holds that $Q(f) \geq 0$.*

As a consequence, we obtain the following partial result towards Conjecture 0.3.3.

Corollary 0.3.6 (Becker [7])**.** *Let $\varepsilon' = \sqrt{3/8}\varepsilon \approx 0.031$. If*

$$\|g\|_{L^2(\sigma)} \leq \|1\|_{L^2(\sigma)}$$

and $g(e^{i\theta})$ is supported in $(-\varepsilon', \varepsilon') + \pi\mathbb{Z}$, then

$$\|\widehat{g\sigma}\|_{L^6(\mathbb{R}^2)} \leq \|\widehat{\sigma}\|_{L^6(\mathbb{R}^2)}.$$

Chapter 1

A degree one Carleson operator along the paraboloid

This chapter consists of the article [5].

1.1 Introduction

This paper advances the program of Pierce and Yung [101] of studying maximally modulated singular Radon transforms along paraboloids. While their work focuses on certain polynomial modulations without linear terms and uses TT^* methods, our result is the first instance in this program for degree one polynomials featuring symmetries that mandate the use of time-frequency analysis.

Our main result is as follows. An m-Calderón-Zygmund kernel on $\mathbb{R}^d$ is a function $K : \mathbb{R}^d \setminus \{0\} \to \mathbb{C}$ satisfying the estimates

$$|\partial^\alpha K(x)| \le |x|^{-d-|\alpha|}, \quad |\alpha| \le m, \tag{1.1.1}$$

and the cancellation property

$$\int_{B(0,R)\setminus B(0,r)} K(x)\, \mathrm{d}x = 0, \quad 0 < r < R. \tag{1.1.2}$$

Let $V \subset \mathbb{R}^{d+1}$ be a linear subspace. We consider the maximally modulated singular integral along the paraboloid defined a priori on Schwartz functions f on $\mathbb{R}^{d+1}$ by

$$T_V f(x) = \sup_{N \in V} \sup_{r<R} \left| \int_{r<|y|<R} f(x' - y, x_{d+1} - |y|^2) e^{iN\cdot(y,|y|^2)} K(y)\, \mathrm{d}y \right|, \tag{1.1.3}$$

where $x = (x', x_{d+1}) \in \mathbb{R}^d \times \mathbb{R} = \mathbb{R}^{d+1}$, and K is a Calderón-Zygmund kernel.

Theorem 1.1.1. *Let $d \ge 2$ and let $m > \frac{d}{2}$. Suppose that $V = \{0\}^d \times \mathbb{R}$ or V is a proper subspace of $\mathbb{R}^d \times \{0\}$. Then for all p with*

$$\frac{d^2+4d+2}{(d+1)^2} < p < 2(d+1) \tag{1.1.4}$$

there exists $C > 0$ such that for all m-Calderón-Zygmund kernels K and all Schwartz functions f, we have with T_V as defined in (1.1.3)

$$\|T_V f\|_{L^p(\mathbb{R}^{d+1})} \le C\|f\|_{L^p(\mathbb{R}^{d+1})} .$$

Note that the singular integral along the paraboloid in (1.1.3) is given by convolution with the tempered distribution $k(z) = \delta(z_{d+1} - |z|^2)K(z)$ on $\mathbb{R}^{d+1}$. In terms of the argument z, the modulation argument $N \cdot (y, |y|^2) = N \cdot z$ in (1.1.3) is a degree one polynomial. This, and its consequences for the method of proof below, is why we call T_V a degree one operator.

1.1.1 Motivation

Our interest in the operator (1.1.3) stems from the following result of Pierce and Yung [101], see also [1]. They prove L^p bounds for $p \in (1, \infty)$ for maximally *polynomially* modulated singular integral operators along the paraboloid. More precisely, they consider the operator

$$f \mapsto \sup_P \left| \int_{\mathbb{R}^d} f(x' - y, x_{d+1} - |y|^2) e^{iP(y)} K(y) \, \mathrm{d}y \right| , \tag{1.1.5}$$

where $d \ge 2$ and P ranges over a certain set of polynomials of fixed degree without linear terms, and without a monomial $c|y|^2$. Note that this excludes exactly the monomials that are present in (1.1.3). Very recently, Beltran, Guo and Hickman [10] gave a version of the Pierce-Yung theorem with $d = 1$, and P ranging over $\{cy^3 \, : \, c \in \mathbb{R}\}$.

The study of maximally modulated singular integrals such as (1.1.3), (1.1.5) has a long history, starting with Carleson's [21] proof of pointwise almost everywhere convergence of Fourier series of L^2 functions. His proof relies crucially on an L^2 to $L^{2,\infty}$ estimate for the maximally modulated Hilbert transform

$$f \mapsto \sup_{N \in \mathbb{R}} \left| \int f(x - y) e^{iNy} \frac{1}{y} \, \mathrm{d}y \right| . \tag{1.1.6}$$

Carleson's theorem was subsequently extended to L^p for $p \in (1, \infty)$ by Hunt [68], and to singular integrals in higher dimension by Sjölin [107]. Other essentially different proofs of Carleson's theorem were later given by Fefferman [51] and by Lacey and Thiele [77]. Endpoint questions were considered in [2, 82, 41].

Pierce and Yung's theorem is motivated by a variation of this theme due to Stein and Wainger [109]. They investigated maximally polynomially modulated singular integrals

$$f \mapsto \sup_P \left| \int f(x - y) e^{iP(y)} K(y) \, \mathrm{d}y \right| , \tag{1.1.7}$$

where P ranges over the set of all polynomials of fixed degree *without linear terms.* They prove L^p bounds for $p \in (1, \infty)$. The harder extension to the operator (1.1.7) with P ranging over *all* polynomials of fixed degree was accomplished by Lie [83, 84]. Zorin-Kranich subsequently gave a version of Lie's proof with very weak regularity assumptions in [116].

1.1.2 Modulation symmetries

We want to discuss the relevance of excluding linear terms in the polynomials in (1.1.5) and (1.1.7). The methods used by Stein-Wainger to bound the operator (1.1.7), without linear terms in the modulations, are fundamentally different from the methods employed by Carleson and Sjölin for the same operator with only linear terms. Stein and Wainger use a TT^* argument exploiting almost orthogonality of contributions of different scales, and decay when the polynomial modulation is large [109, Theorem 1]. Sjölin and Lie on the other hand use time frequency analysis. This difference in methods is dictated by the symmetries of the operator. Carleson's operator (1.1.6), and Sjölin's higher dimensional variant, are invariant under the modulations

$$f \mapsto [x \mapsto e^{iNx} f(x)]\,, \quad N \in \mathbb{R}^d\,. \tag{1.1.8}$$

On the other hand, a quick computation shows that the operator (1.1.7) has no symmetries under the transformations (1.1.8), as long as no linear terms are present in the polynomials P. To illustrate how this symmetry affects the proof, the reader is invited to use the modulation symmetry under (1.1.8) to show that [109, Theorem 1] would not be true for degree one polynomials.

Pierce and Yung sidestep the issue of modulations symmetries via their restrictions on the linear and quadratic terms of the polynomials. This allows for a TT^* argument in the same spirit as in [109], but using more sophisticated oscillatory integral estimates. In contrast, our operators T_V are invariant under linear modulations (1.1.8) with $N \in V$, so our proof will use time frequency analysis.

1.1.3 Previous results

Pierce and Yung's bound for (1.1.5) sparked interest in the corresponding operators with linear modulations.

Roos [103] proved a version of Sjölin's theorem for anisotropic Calderón-Zygmund operators with scaling symmetry preserving the paraboloid. Roos's result assumes enough regularity of the Calderón-Zygmund operator, and his bounds blow up if one approximates a singular integral along a paraboloid by such operators. In [6], the author improved the control of this blowup.

Guo, Roos, Pierce and Yung proved in [64] a number of interesting related results for Hilbert transforms on planar curves. They weaken the maximal operator by first taking an L^p norm in one of the variables, before taking the supremum in the modulation parameter. Most relevant to us is the following special case of their Theorem 1.2:

$$\int \sup_{N \in \mathbb{R}} \int \left| \int f(x_1 - y, x_2 - y^2) e^{iNy^2} \frac{\mathrm{d}y}{y} \right|^p \mathrm{d}x_1 \, \mathrm{d}x_2 \le C \|f\|_p^p\,, \tag{1.1.9}$$

where $p \in (1, \infty)$. They also noted that for $p = 2$, (1.1.9) holds with the order of $\mathrm{d}x_1\,\mathrm{d}x_2$ reversed, and also with modulations e^{iNy}, by a combination of Plancherel and Lie's polynomial Carleson theorem.

Finally, Ramos [102] proved bounds for maximal modulations of the one dimensional Fourier multipliers obtained by restricting the multiplier of the Hilbert transform along the parabola to a line, uniformly in the choice of line.

Our Theorem 1.1.1 is the first L^p estimate for maximal linear modulations of a singular integral operator along a submanifold, with - differently from (1.1.9) - the supremum fully inside the L^p norm. We cannot directly compare our result to [64], because they consider singular integrals along planar curves, while Theorem 1.1.1 assumes that the dimension of the paraboloid is at least two. However, Theorem 1.1.1 implies a stronger version of the hypothetical generalization of (1.1.9) to paraboloids of dimension ≥ 2 in the range of p given by (1.1.4). We stress that our proof does not apply to the one dimensional parabola, so we do not recover (1.1.9) itself. However, our Theorem 1.1.1 does recover, via projection and limiting arguments, Sjölin's theorem [107] in the full range of exponents $p \in (1, \infty)$. In particular, it implies the Carleson-Hunt theorem.

1.1.4 Overview of the proof of Theorem 1.1.1

Our proof is an adaptation of Fefferman's [51] proof of Carleson's theorem to the setting of singular integrals along paraboloids. In some parts we also follow the presentation in [116]. We establish variants of Fefferman's key time-frequency analysis estimates in this more singular setting.

The proof starts with a discretization of the operator, in Section 1.2, and a combinatorial decomposition into so called forest operators and antichain operators in Section 1.4 (see Sections 1.2 and 1.3 for the definitions of these operators). These two steps are straightforward variation of the corresponding steps in Fefferman's proof. However, differently from Fefferman, we test the operator with the indicator function of a set F and use a modified decomposition adapted to F. This allows us to directly prove weak type L^2 bounds.

After the decomposition of the operator, the argument proceeds by proving bounds for antichain operators in Section 1.5, for so called tree operators in Section 1.6 and finally for forest operators in Section 1.7, see Propositions 1.3.2, 1.3.3 and 1.3.4. Each of these three steps needs new ingredients for singular integrals supported on submanifolds.

Antichains: A square function argument

To control the antichain operators in Section 1.5, we decompose the kernel k with singular support on the paraboloid into a smoothened kernel, which is no longer supported just on a submanifold, and a remainder. The versions of the operator with smoothened kernel are estimated using the argument of Fefferman (see Lemma 1 and Lemma 2 in [51]). The contribution of the remainder is estimated by a square function. This square function is closely related to a classical (unmodulated) square function used in [110] to prove boundedness of the maximal average along the paraboloid. In our square function, there is an additional summation over modulations compared to the one in [110]. To bound this larger square function, we rely on the precise decay rate in each direction of the Fourier transform of the kernel supported on the paraboloid, coming from stationary phase. In addition, we use that the Fourier transform of the remainder vanishes near a certain subspace.

Trees: Sparse bounds

In Section 1.6 we estimate the tree operators. These are pieces of the operator which can be modeled by truncated singular integral operators. Fefferman's argument for this step (see Lemma 3 in [51]) relies on the fact that the convolution of a function with a truncated singular integral kernel is essentially constant at the scale of the lower truncation parameter. This makes the integral of said convolution over a set which is thin at this scale small. In contrast, truncated singular integrals along paraboloids are not essentially constant at the lower truncation scale. Our new ingredient to solve this issue is a Sobolev smoothing estimate for truncated singular integrals along paraboloids. On a technical level, we use certain sparse bounds for singular Radon transforms, see Lemma 1.6.2, due to Oberlin [97], see also [74]. The arguments in Section 1.6 are formulated for a general class of singular Radon transforms, similar to the setting of [44].

Forests: Square functions and oscillatory integrals

In Section 1.7 we follow Fefferman's argument (Lemma 4 and Lemma 5 in [51]) to prove almost orthogonality estimates between tree operators, and combine the estimates for tree operators to an estimate for forest operators. We prove that tree operators essentially only act on frequencies close to a central frequency associated to the tree, leading to almost orthogonality for trees with sufficiently separated central frequencies. This requires more work than in [51], because of the singular support of the kernels on the paraboloid. However, it is within the scope of the square function arguments in [110], [44], used there to bound maximal averages and maximally truncated singular integrals along paraboloids. Fefferman's argument requires certain upper bounds to be local, in the sense that they only depend on the values of the functions involved on certain sets associated to the trees. To ensure this locality we further use estimates for certain oscillatory integrals along paraboloids, see Section 1.7.2. They replace easier partial integration arguments in [51].

Finally, we deduce L^p bounds in Section 1.8, using interpolation and a localization argument as in [116].

1.1.5 Limitations and further questions

The limitations of our methods are still dictated by the symmetries of the operator. The restrictions on V in Theorem 1.1.1 ensure that the operator T_V has only linear modulation symmetries. If we had $\mathbb{R}^d \times \{0\} \subset V$, then T_V would be invariant under the transformations

$$f \mapsto [(x', x_{d+1}) \mapsto e^{iN(|x'|^2 + x_{d+1})} f(x', x_{d+1})] , \quad N \in \mathbb{R} . \tag{1.1.10}$$

Time frequency analysis in its current form seems to be unable to handle such additional symmetries. For example, with our setup Proposition 1.3.2 fails in the presence of the symmetry (1.1.10), because the definition (1.3.2) of density used therein is not invariant under (1.1.10).

Note that for the modulation subspace V to be compatible with the anisotropic dilation symmetry of the paraboloid, it can only be the vertical subspace $\{0\}^d \times \mathbb{R}$, a subspace of $\mathbb{R}^d \times \{0\}$, or a sum of two such spaces. Then the only remaining case (up to rotation around the vertical axis) that could conceivably be within reach of current methods is $V = \mathbb{R}^{d-1} \times \{0\} \times \mathbb{R}$. Our argument does not handle this V, the particular point of failure is the square function argument in Subsection 1.5.3. However, we know of no fundamental obstruction. It is tempting to ask whether the results of this paper can be extended to that case.

The restriction to paraboloids of dimension two or higher in Theorem 1.1.1 is needed in the square function argument in Subsection (1.5.3). For the parabola this square function argument fails, because the Fourier transform of a measure on the parabola has too little decay.

The range (1.1.4) of p in Theorem 1.1.1, which one might conjecture should really be $(1,\infty)$, is a consequence of restrictions on exponents in the sparse bounds in Lemma 1.6.2. These restrictions are related to the L^p-improving range for averages along paraboloids. It is an interesting question whether the range of p can be improved by other methods.

Acknowledgement. The author thanks Christoph Thiele and Rajula Srivastava for various helpful discussions about this paper. The author was supported by the Collaborative Research Center 1060 funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and the Hausdorff Center for Mathematics, funded by the DFG under Germany's Excellence Strategy - GZ 2047/1, ProjectID 390685813.

1.2 Reduction to a discretized operator

We may assume that $V = \{0\}^d \times \mathbb{R}$ or $V = \mathbb{R}^{d-1} \times \{0\}^2$. Define anisotropic dilations $\delta_s(x', x_{d+1}) := (2^s x', 2^{2s} x_{d+1})$. Define the dyadic cubes of scale 0 to be

$$\mathbf{D}_0 := \{k + [0,1)^{d+1} \,:\, k \in \mathbb{Z}^{d+1}\}\,,$$

and the dyadic cubes of scale s to be $\mathbf{D}_s := \delta_s(\mathbf{D}_0)$. The collection of all dyadic cubes is denoted $\mathbf{D} = \cup_{s\in\mathbb{Z}} \mathbf{D}_s$. Given a dyadic cube $I \in \mathbf{D}_s$, we denote by $s(I) = s$ its scale.

We define dyadic frequency cubes of scale 0 as

$$\mathbf{\Omega}_0 := \{k + \{0\}^d \times [0,1) \,:\, k \in \{0\}^d \times \mathbb{Z}\}$$

if $V = \{0\}^d \times \mathbb{R}$ and as

$$\mathbf{\Omega}_0 := \{k + [0,1)^{d-1} \times \{0\}^2 \,:\, k \in \mathbb{Z}^{d-1} \times \{0\}^2\}\,,$$

if $V = \mathbb{R}^{d-1} \times \{0\}^2$. The dyadic frequency cubes of scale s are $\mathbf{\Omega}_s := \delta_{-s}(\mathbf{\Omega}_0)$ and the collection of all dyadic frequency cubes is defined by $\mathbf{\Omega} := \bigcup_{s\in\mathbb{Z}} \mathbf{\Omega}_s$. A tile is a pair $p = (I, \omega)$, where $I \in \mathbf{D}_s$ and $\omega \in \mathbf{\Omega}_s$ for some $s = s(p)$ called the scale of p. The collection of all tiles is denoted by

$$\mathbf{P} := \bigcup_{s\in\mathbb{Z}} \{(I,\omega) \,:\, I \in \mathbf{D}_s,\, \omega \in \mathbf{\Omega}_s\}\,.$$

We decompose the kernel K into pieces localized in dyadic annuli. Fix a smooth function η supported in $[1/8, 1/3]$ such that

$$\sum_{s\in\mathbb{Z}} \eta(2^{-s}t) = 1\,, \quad t \in (0,\infty)\,,$$

and define $K_s(x) = K(x)\eta(2^{-s}|x|)$. Then (1.1.1) and (1.1.2) imply that there exists a constant $C = C(\eta)$ with

$$|\partial^\alpha K_s(x)| \le C 2^{-s(d+|\alpha|)}\,, \quad |\alpha| \le m\,,\, s \in \mathbb{Z}\,, \tag{1.2.1}$$

and that

$$\int K_s(x)\,\mathrm{d}x = 0\,, \quad s \in \mathbb{Z}\,. \tag{1.2.2}$$

It suffices to estimate the operator

$$\sup_{N\in V}\sup_{\underline{s}<\overline{s}} \left| \sum_{s=\underline{s}}^{\overline{s}} \int f(x'-y, x_{d+1}-|y|^2) e^{iN\cdot(y,|y|^2)} K_s(y)\,\mathrm{d}y \right|\,, \tag{1.2.3}$$

since the difference to T_V is controlled by the maximal average along the paraboloid, by (1.1.1). Since the sum of integrals in (1.2.3) is continuous in N, we may restrict the supremum in N to a countable dense subset of V. By monotone convergence, we may further restrict both suprema to finite subsets of V and $\mathbb{Z}\times\mathbb{Z}$. Choosing maximizers for each x, it suffices to bound the operator

$$\sum_{s=\underline{s}(x)}^{\overline{s}(x)} \int f(x'-y, x_{d+1}-|y|^2) e^{-iN(x)\cdot(y,|y|^2)} K_s(y)\,\mathrm{d}y \tag{1.2.4}$$

uniformly over all measurable functions $N : \mathbb{R}^{d+1} \to V$, $\underline{s} : \mathbb{R}^{d+1} \to \mathbb{Z}$ and $\overline{s} : \mathbb{R}^{d+1} \to \mathbb{Z}$ with finite range. We fix such functions $N, \underline{s}, \overline{s}$.

For a tile $p \in \mathbf{P}$, we define

$$E(p) = \{x \in I(p) \,:\, N(x) \in \omega(p),\, \underline{s}(x) \le s(p) \le \overline{s}(x)\}$$

and the operator associated to the tile p

$$T_p f(x) = \mathbf{1}_{E(p)} \int f(x'-y, x_{d+1}-|y|^2) e^{-iN(x)\cdot(y,|y|^2)} K_{s(p)}(y)\,\mathrm{d}y\,.$$

For a subset $\mathbf{C} \subset \mathbf{P}$ we write $T_{\mathbf{C}} = \sum_{p\in\mathbf{C}} T_p$. The operator (1.2.4) is then simply $T_{\mathbf{P}}$.

Finally, we remove some tiles with unfavourable properties, using an argument due to Fefferman [51]. A tile is called admissible if $3\omega(p) \subset \hat{\hat{\omega}}(p)$, where $\hat{\omega}$ denotes the unique frequency cube of scale $s(\omega)-1$ containing ω. Then it suffices to show the estimate $\|T_{\mathbf{P}_{\mathrm{ad}}} f\|_{L^p} \le C\|f\|_{L^p}$, where $\mathbf{P}_{\mathrm{ad}}$ is the set of admissible tiles, as follows from an averaging argument analogous to the one in Section 5 of [51]. From now on by tile we always mean an admissible tile.

1.3 Outline of the proof of weak type L^2-bounds for (1.2.4)

By duality and the reductions in the previous section, it suffices to show that there exists $C > 0$ such that for each compact set $F \subset \mathbb{R}^{d+1}$, there exists a set $\tilde{F} \subset F$ with $|\tilde{F}| \le \frac{1}{2}|F|$ and

$$\|\mathbf{1}_{F\setminus\tilde{F}} T_{\mathbf{P}_{\mathrm{ad}}}\|_{2\to 2} \le C\,. \tag{1.3.1}$$

We fix F. The operators $\mathbf{1}_F T_p$ do not change upon replacing $E(p)$ by $E(p)\cap F$, so we will assume from now on that $E(p) \subset F$ for all tiles p. Note that since F is compact and $\underline{s}, \overline{s}$ and N have finite range, only finitely many operators T_p are not zero.

We now make some definitions to state a decomposition for the operator on the left hand side of (1.3.1). Denote by $\dim_h V$ the homogeneous dimension of V with respect to the dilations δ_s, that is

$$\dim_h(\mathbb{R}^{d-1} \times \{0\}^2) = d-1\,, \qquad \dim_h(\{0\}^d \times \mathbb{R}) = 2\,,$$

and define the convex cylinder associated to a collection $\mathbf{C}$ of tiles by

$$C(\mathbf{C}) = \{p \in \mathbf{P}_{\mathrm{ad}} \,:\, \exists p', p'' \in \mathbf{C}, I(p') \subset I(p) \subset I(p'')\}\,.$$

Then the density of a collection $\mathbf{C}$ of tiles is defined as

$$\mathrm{dens}(\mathbf{C}) = \sup_{p \in \mathbf{C}} \sup_{\substack{\lambda \geq 1 \\ \text{odd}}} \sup_{\substack{p' \in C(\mathbf{C}) \,:\, I(p) \subset I(p') \\ \lambda\omega(p') \subset \lambda\omega(p)}} \lambda^{-\dim_h V} \frac{|E(\lambda, p')|}{|I(p')|}\,, \tag{1.3.2}$$

where we use the notation $\lambda\omega = c(\omega) + \delta_{\log_2 \lambda}(\omega - c(\omega))$ where $c(\omega)$ is the center of ω and

$$E(\lambda, p') = \{x \in I(p) \,:\, N(x) \in \lambda\omega(p)\}\,.$$

We define a partial order on the set of tiles by

$$(I, \omega) \leq (I', \omega') \iff I \subset I' \text{ and } \omega' \subset \omega\,.$$

An antichain is a set of tiles that are pairwise not comparable with respect to this order. A set of tiles $\mathbf{C}$ is called convex, if $p, p' \in \mathbf{C}$ and $p \leq p'' \leq p'$ implies $p'' \in \mathbf{C}$. A tree is a convex collection $\mathbf{T}$ of tiles together with an upper bound $\mathrm{top}(\mathbf{T})$, i.e. a tile m such that for all $p \in \mathbf{T}$ we have $p \leq m$. We denote $\omega(\mathbf{T}) = \omega(\mathrm{top}(\mathbf{T}))$ and $I(\mathbf{T}) = I(\mathrm{top}(\mathbf{T}))$. A tree is called normal if $3I(p) \subset I(\mathbf{T})$ for all $p \in \mathbf{T}$.

A pair of trees $\mathbf{T}_1, \mathbf{T}_2$ is called Δ-separated, if for $\{i, j\} = \{1, 2\}$ and each tile $p_i \in \mathbf{T}_i$ with $I(p_i) \subset I(\mathbf{T}_j)$ we have $\Delta\omega(p_i) \cap \omega(\mathbf{T}_j) = \emptyset$. An n-forest is a collection of pairwise 2^{10dn}-separated, normal trees of density at most 2^{-n} that satisfy the overlap estimate

$$\sum_{\mathbf{T} \in \mathbf{F}} \mathbf{1}_{I(\mathbf{T})} \leq 2^n \log(n+2)\,. \tag{1.3.3}$$

Proposition 1.3.1. *There exists a constant $C = C(d)$ and an exceptional set $\tilde{F}$ with $|\tilde{F}| \leq |F|/2$, such that the set*

$$\mathbf{P}_{F \setminus \tilde{F}} := \{p \in \mathbf{P}_{\mathrm{ad}} \,:\, I(p) \cap (F \setminus \tilde{F}) \neq \emptyset\}$$

can be decomposed as a disjoint union

$$\mathbf{P}_{F \setminus \tilde{F}} = \bigcup_{n \geq 0} \left(\bigcup_{l=1}^{C(n+1)^2} \mathbf{F}_{n,l} \cup \bigcup_{l=1}^{C(n+1)^3} \mathbf{A}_{n,l} \right)\,,$$

where each $\mathbf{F}_{n,l}$ is an n-forest and each $\mathbf{A}_{n,l}$ is an antichain of density at most 2^{-n}.

Estimate (1.3.1), and therefore weak type L^2 boundedness of the operator defined in (1.1.3) then follows from the following estimates for antichain and forest operators.

Proposition 1.3.2. *There exists $\varepsilon = \varepsilon(d)$ and $C > 0$ such that the following holds. Let $\mathbf{A}$ be an antichain of density δ. Then*

$$\|T_{\mathbf{A}}\|_{2\to 2} \le C\delta^{\varepsilon}\,.$$

Proposition 1.3.3. *For each $\varepsilon < \frac{1}{2} - \frac{1}{2(d+1)}$ there exist $C > 0$ such that the following holds. Let $\mathbf{F}$ be an n-forest. Then*

$$\|T_{\mathbf{F}}\|_{2\to 2} \le C2^{-\varepsilon n}\,.$$

We will independently prove Proposition 1.3.1 in Section 1.4 and Proposition 1.3.2 in Section 1.5. Proposition 1.3.3 is proven using an estimate for trees and an almost orthogonality argument. The almost orthogonality argument and the deduction of Proposition 1.3.3 are carried out in Section 1.7. The estimate for single trees is proven in Section 1.6, and is a mild generalization of the following.

Proposition 1.3.4. *For each $\varepsilon < \frac{1}{2} - \frac{1}{2(d+1)}$ there exist $C > 0$ such that the following holds. Let $\mathbf{T}$ be a tree of density δ. Then*

$$\|T_{\mathbf{T}}\|_{2\to 2} \le C\delta^{\varepsilon}\,.$$

1.4 Tile organization: Proof of Proposition 1.3.1

Let $k \ge 0$ and let $\mathbf{D}_k(F)$ be the set of maximal dyadic cubes Q with $|Q \cap F|/|Q| \ge 2^{-k-1}$. Let $\tilde{\mathbf{P}}_{\le k}$ be the set of tiles $p \in \mathbf{P}_{\mathrm{ad}}$ such that $I(p)$ is contained in some $Q \in \mathbf{D}_k(F)$ and let $\tilde{\mathbf{P}}_k := \tilde{\mathbf{P}}_{\le k} \setminus \tilde{\mathbf{P}}_{\le k-1}$. Then we have $\mathbf{1}_F T_{\mathbf{P}_{\mathrm{ad}}} = \sum_{k\ge 0} \mathbf{1}_F T_{\tilde{\mathbf{P}}_k}$, and each $p \in \tilde{\mathbf{P}}_k$ satisfies that $|I(p) \cap F|/|I(p)| < 2^{-k}$. We define

$$\overline{E}(p) := \{x \in I(p) \cap F \,:\, N(x) \in \omega(p)\}\,,$$

and define $\tilde{\mathbf{M}}_{n,k}$ to be the set of maximal tiles p in $\tilde{\mathbf{P}}_k$ such that $|\overline{E}(p)|/|I(p)| \ge 2^{-n-1}$.

Lemma 1.4.1. *The exceptional set*

$$\tilde{F}_1 := F \cap \bigcup_{k\ge 0} \bigcup_{Q\in \mathbf{D}_k(F)} \bigcup_{n\ge k} \{x \in Q \,:\, \sum_{\substack{p\in\tilde{\mathbf{M}}_{n,k}\\ I(p)\subset Q}} \mathbf{1}_{I(p)} \ge 1000\cdot 2^n \log(n+2)\}$$

satisfies $|\tilde{F}_1| \le |F|/4$.

Proof. We have for each $J \in \mathbf{D}$

$$\sum_{\substack{p\in\tilde{\mathbf{M}}_{n,k}\\ I(p)\subset J}} |I(p)| \le 2^{n+1} \sum_{\substack{p\in\tilde{\mathbf{M}}_{n,k}\\ I(p)\subset J}} |\overline{E}(p)| \le 2^{n+1}|J|\,,$$

since tiles in $\tilde{\mathbf{M}}_{n,k}$ are pairwise not comparable and hence the sets $\overline{E}(p)$, $p \in \tilde{\mathbf{M}}_{n,k}$ are pairwise disjoint. We estimate for each $Q \in D_k(F)$ and each $n \ge k$ using the John-Nirenberg inequality:

$$|\{x \in Q \,:\, \sum_{\substack{p\in\tilde{\mathbf{M}}_{n,k}\\ I(p)\subset Q}} \mathbf{1}_{I(p)} \ge 1000\cdot 2^n \log(n+2)\}|$$
$$\le c_1 \exp(-c_2 \frac{1000\cdot 2^n \log(n+2)}{2^{n+1}})|Q| \le (n+2)^{-100}|Q|\,. \quad (1.4.1)$$

For the last inequality we have used that in this version of the John-Nirenberg inequality, one can choose $c_1 = e^2$ and $c_2 = (2e)^{-1}$.

Note that the set on the left hand side of (1.4.1) is a disjoint union of cubes $I(p)$ each of which satisfies $|I(p) \cap F|/|I(p)| \le 2^{-k}$, hence

$$|F \cap \{x \in Q : \sum_{\substack{p \in \tilde{\mathbf{M}}_{n,k} \\ I(p) \subset Q}} \mathbf{1}_{I(p)} \ge 1000 \cdot 2^n \log(n+2)\}|$$

$$\le 2^{-k}(n+2)^{-100}|Q| \le 2(n+2)^{-100}|Q \cap F| .$$

Summing up, we obtain

$$\sum_{k \ge 0} \sum_{Q \in \mathbf{D}_k(F)} \sum_{n \ge k} |F \cap \{x \in Q : \sum_{\substack{p \in \tilde{\mathbf{M}}_{n,k} \\ I(p) \subset Q}} \mathbf{1}_{I(p)} \ge 1000 \cdot 2^n \log(n+2)\}|$$

$$\le 2 \sum_{k \ge 0} \sum_{n \ge k} (n+2)^{-100}|F| \le \frac{1}{4}|F| .$$

This completes the proof. □

After removing the exceptional set $\tilde{F}_1$, only the tiles in

$$\mathbf{P}_k := \{p \in \tilde{\mathbf{P}}_k : I(p) \not\subset \tilde{F}_1\}$$

contribute, i.e. we have $\mathbf{1}_{\mathbb{R}^{d+1} \setminus \tilde{F}_1} T_{\tilde{\mathbf{P}}_k} = \mathbf{1}_{\mathbb{R}^{d+1} \setminus \tilde{F}_1} T_{\mathbf{P}_k}$. Thus it suffices to decompose the sets $\mathbf{P}_k$ into forests and antichains.

We define the k-density of a tile $p \in \mathbf{P}_k$ to be

$$\mathrm{dens}_k(p) := \sup_{\substack{\lambda \ge 1 \\ \text{odd}}} \sup_{\substack{p' \in \mathbf{P}_k : I(p) \subset I(p') \\ \lambda\omega(p') \subset \lambda\omega(p)}} \lambda^{-\dim_h V} \frac{|E(\lambda, p')|}{|I(p')|} .$$

Then we split each $\mathbf{P}_k$ into sets $\mathbf{H}_{n,k} := \{p \in \mathbf{P}_k : 2^{-n-1} < \mathrm{dens}_k(p) \le 2^{-n}\}$, and decompose each of them separately into forests and antichains. Note that $\mathrm{dens}(\mathbf{C}) \le \sup_{p \in \mathbf{C}} \mathrm{dens}_k(p)$ for each $\mathbf{C} \subset \mathbf{P}_k$, so all subsets forests and antichains obtained in the decomposition of $\mathbf{H}_{n,k}$ have density at most 2^{-n}.

Lemma 1.4.2. *For each $k \ge 0$ and $n \ge k$, there exists an exceptional set $\tilde{F}_{n,k}$ such that $|\tilde{F}_{n,k}| \le 2^{-k-n-2}|F|$, and such that the set of tiles $p \in \mathbf{H}_{n,k}$ with $I(p) \not\subset \tilde{F}_{n,k}$ can be decomposed as a disjoint union of $O(n+1)$ many n-forests and $O_d((n+1)^2)$ many antichains.*

Proof. We first note that $\mathbf{H}_{n,k}$ is convex: Since dens is a decreasing function on tiles with respect to $\le$, $\mathbf{H}_{n,k}$ is the difference of two down-sets, and as such convex.

Next, we prune the top $n+2$ layers off $\mathbf{H}_{n,k}$: Let $\mathbf{H}_{n,k}^+$ be the set of tiles $p \in \mathbf{H}_{n,k}$ for which there exists no chain $p < p_1 < \cdots < p_{n+2}$ with all $p_i \in \mathbf{H}_{n,k}$, where we use $p < p'$ to say that $p \le p'$ and $p \ne p'$. Clearly, $\mathbf{H}_{n,k}^+$ is the union of at most $n+2$ antichains, so it suffices to decompose $\mathbf{H}_{n,k}^0 := \mathbf{H}_{n,k} \setminus \mathbf{H}_{n,k}^+$, and this set is still convex.

For $n \geq k \geq 0$, let $\mathbf{M}_{n,k}$ to be the set of maximal tiles p in $\mathbf{P}_k$ that satisfy $|\overline{E}(p)|/|I(p)| \geq 2^{-n-1}$. By Lemma 1.4.1 and the definition of $\mathbf{P}_k$, we then have the overlap estimate

$$\sum_{p \in \mathbf{M}_{n,k}} \mathbf{1}_{I(p)} \leq 1000 \cdot 2^n \log(n+2)\,. \tag{1.4.2}$$

We claim that for each $p \in \mathbf{H}^0_{n,k}$, there exists $m \in \mathbf{M}_{n,k}$ with $p \leq m$. Since $p \in \mathbf{H}^0_{n,k}$, there exists a chain $p < p_1 < \cdots < p_{n+2}$ with $p_{n+2} \in \mathbf{H}_{n,k}$. Since all tiles are admissible, we have $3^{\lfloor (n+2)/2 \rfloor}\omega(p_{n+2}) \subset \omega(p)$. Since $p_{n+2} \in \mathbf{H}_{n,k}$, there exists an odd $\lambda \geq 1$ and a tile $p' \in \mathbf{P}_k$ with $I(p_{n+2}) \subset I(p')$ and $\lambda\omega(p') \subset \lambda\omega(p_{n+2})$ such that

$$\frac{|E(\lambda, p')|}{|I(p')|} \geq \lambda^{\dim_h V} 2^{-n-1}\,. \tag{1.4.3}$$

Since λ is odd, the set $\lambda\omega(p')$ is the disjoint union of $\lambda^{\dim_h V}$ cubes $\omega(p'')$ of tiles p'' with $I(p'') = I(p')$, thus there exists one such $p'' \in \mathbf{P}_k$ with $|\overline{E}(p'')|/|I(p'')| \geq 2^{-n-1}$. By definition of $\mathbf{M}_{n,k}$, there is a tile $m \in \mathbf{M}_{n,k}$ with $p'' \leq m$. Equation (1.4.3) implies that $\lambda \leq 2^{(n+1)/\dim_h V}$, so that

$$\omega(m) \subset \omega(p'') \subset 2^{(n+1)/\dim_h V}\omega(p_{n+2}) \subset \omega(p)\,.$$

Combining all of this we obtain that $p \leq m$, so the claim holds.

For a tile $p \in \mathbf{H}^0_{n,k}$, let $\mathbf{B}(p)$ be the set of tiles $m \in \mathbf{M}_{n,k}$ with $p \leq m$. Decompose $\mathbf{H}^0_{n,k}$ into $2n+10$ collections

$$\mathbf{C}_j := \{p \in \mathbf{H}_{n,k} \,:\, 2^{j-1} \leq |\mathbf{B}(p)| < 2^j\}\,, \qquad j = 1, \ldots, 2n+10\,.$$

The collections $\mathbf{C}_j$ are convex, since $p \leq p' \leq p''$ implies $B(p'') \subset B(p') \subset B(p)$, and by the overlap estimate (1.4.2) and the claim, their union is $\mathbf{H}^0_{n,k}$.

Let $\mathbf{U}_j$ be the set of maximal tiles in $\mathbf{C}_j$, clearly $\mathbf{U}_j$ also has overlap bounded by $100 \cdot 2^n \log(n+2)$. For each $m \in \mathbf{U}_j$ let

$$\tilde{\mathbf{T}}(m) := \{p \in \mathbf{C}_j \,:\, p \leq m\}\,.$$

Then the sets $\tilde{\mathbf{T}}(m)$ are disjoint: If $p \in \tilde{\mathbf{T}}(m) \cap \tilde{\mathbf{T}}(m')$ for $m \neq m'$, then $\mathbf{B}(m) \cup \mathbf{B}(m') \subset \mathbf{B}(p)$. But the sets $\mathbf{B}(m)$ and $\mathbf{B}(m')$ are disjoint: Else there would be m'' with $p \leq m, m' \leq m''$, which implies that m, m' are comparable. But they are both maximal in $\mathbf{C}_j$, so they cannot be comparable. Hence $2^{j+1} > |\mathbf{B}(p)| \geq |\mathbf{B}(m)| + |\mathbf{B}(m')| \geq 2^j + 2^j$, a contradiction. In particular, tiles $p \in \tilde{\mathbf{T}}(m)$, $p' \in \tilde{\mathbf{T}}(m')$ for $m \neq m'$ are not comparable.

The sets $\tilde{\mathbf{T}}(m)$ are of course also convex, so they are trees with top m. To obtain the separation property, we prune the bottom $20dn$ layers of the trees: We define $\mathbf{T}^-(m)$ to be the set of tiles $p \in \tilde{\mathbf{T}}(m)$ for which there exists no chain $p_{20dn} < \cdots < p_1 < p$. Clearly, $\mathbf{T}^-(m)$ is the union of at most $20dn$ antichains. As tiles in different $\tilde{\mathbf{T}}(m)$ are never comparable, $\cup_{m \in \mathbf{U}_j} \mathbf{T}^-(m)$ is still a union of at most $20dn$ antichains. Let $\tilde{\mathbf{T}}^0(m) := \tilde{\mathbf{T}}(m) \setminus \mathbf{T}^-(m)$, this is still a convex tree with top m. If $p \in \tilde{\mathbf{T}}^0(m)$, then there exists a chain $p_{20dn} < \cdots < p_1 < p$ in $\tilde{\mathbf{T}}(m)$. If $m' \neq m$ is such that $I(p) \subset I(m') = I(\tilde{\mathbf{T}}(m'))$, then by the last paragraph we must have $\omega(p_{20dn}) \cap \omega(m') = \emptyset$. Since all tiles are admissible, it follows that $3^{10dn}\omega(p) \cap \omega(m') = \emptyset$. Hence the trees $\tilde{\mathbf{T}}^0(m)$ are 3^{10dn}-separated.

Finally, we make the trees normal. For this purpose, let $r = 100 + d + 6n$. We first prune the top r layers $\tilde{\mathbf{T}}^{0+}(m)$ off each $\tilde{\mathbf{T}}^0(m)$ similarly as in the last paragraph. This produces another $r = O_d(n+1)$ antichains. Then we define the exceptional set

$$\tilde{F}_{n,k} = \bigcup_j \bigcup_{m \in \mathbf{U}_j} (I(m) \setminus (1 - 2^{-r})I(m)) \,,$$

where $aQ = c(Q) + \delta_{\log_2 a}(Q - c(Q))$ is the anisotropic dilate of Q by a factor a about its center $c(Q)$. We have, using that $1 \le j \le 2n + 10$, the Bernoulli inequality and (1.4.2)

$$\begin{aligned} |\tilde{F}_{n,k}| &\le (2n+10) \cdot \sum_{m \in \mathbf{U}_j} (d+2) 2^{-r} |I(m)| \\ &\le (2n+10) \cdot 1000 \cdot 2^n \log(n+2) \cdot (d+2) \cdot 2^{-r} \cdot \sum_{Q \in \mathbf{D}_k(F)} |Q| \\ &\le 2^{-3n-3} \sum_{Q \in \mathbf{D}_k(F)} |Q| \,. \end{aligned}$$

Using the definition of $\mathbf{D}_k(F)$ and that $n \ge k$ we estimate this by

$$\le 2^{-3n-2} 2^k |F| \le 2^{-n-k-2} |F| \,.$$

We finally define

$$\mathbf{T}(m) := \{p \in \tilde{\mathbf{T}}^0(m) \setminus \tilde{\mathbf{T}}^{0+}(m) \, : \, I(p) \not\subset \tilde{F}_{n,k}\} \,.$$

Then $\mathbf{T}(m), m \in \mathbf{U}_j$ is still a collection of 2^{10dn}-separated trees, and they are now normal: If $p \in \mathbf{T}(m)$ then $s(p) \le s(m) - r$ and $I(p) \subset (1 - 2^{-r})I(m)$, and therefore $3I(p) \subset I(m)$. By the overlap estimate (1.4.2), the collection $\{\mathbf{T}(m) \, : \, m \in \mathbf{U}_j\}$ is the union of at most 1000 n-forests. Thus $\mathbf{C}_j$ can be decomposed into 1000 n-forests and $O_d(n+1)$ antichains, which completes the proof. □

Recall that all antichains in Lemma 1.4.2 have density at most 2^{-n}, since they are contained in $\mathbf{H}_{n,k}$. Taking into account all $k \ge 0$, Lemma 1.4.2 then yields a total of $O_d((n+1)^3)$ antichains of density at most 2^{-n}, a total of $O((n+1)^2)$ many n-forests, and an exceptional set $\tilde{F}_2 = \bigcup_{n,k} \tilde{F}_{n,k}$ with $|\tilde{F}_2| \le |F|/4$. Combining this with the estimate for the measure of $\tilde{F}_1$ from Lemma 1.4.1, we obtain Proposition 1.3.1 with $\tilde{F} := \tilde{F}_1 \cup \tilde{F}_2$.

1.5 Antichains: Proof of Proposition 1.3.2

1.5.1 Decomposition of the kernel

We fix an antichain $\mathbf{A}$ of density δ. We will decompose $T_{\mathbf{A}}$, based on a decomposition of the kernel of the singular Radon transform. Let μ_s be the measure defined by

$$\int f \, \mathrm{d}\mu_s = \int_{\mathbb{R}^d} f(y, |y|^2) K_s(y) \, \mathrm{d}y \,.$$

Then the operator T_p can be expressed as

$$T_p f(x) = \mathbf{1}_{E(p)}(x) \int f(x-y) e^{iN(x) \cdot y} \mathrm{d}\mu_{s(p)}(y) \,. \tag{1.5.1}$$

We now decompose the measures μ_s into a smooth part and a high-frequency part, choosing different decompositions depending on V. For this, we let $\varepsilon_0 > 0$ be a small enough positive number such that the exponent of δ in Lemma 1.5.3 below is positive.

In the case $V = \mathbb{R}^{d-1} \times \{0\}^2$, we let φ^1 be a smooth bump function supported on $B(0, 100^{-1}) \subset \mathbb{R}$ with integral 1, and let

$$\varphi^1_{s,\varepsilon_0}(x) = \delta^{-2\varepsilon_0} 2^{-2s} \varphi^1(2^{-2s}\delta^{-2\varepsilon_0} x_{d+1})\delta(x') \,. \tag{1.5.2}$$

We define the low frequency part of μ_s by

$$\mu^l_s(x) := \mu_s * \varphi^1_{s,\varepsilon_0}(x) = \delta^{-2\varepsilon_0} 2^{-2s} K_s(x')\varphi^1(2^{-2s}\delta^{-2\varepsilon_0}(x_{d+1} - |x'|^2)) \tag{1.5.3}$$

and the high frequency part by $\mu^h_s = \mu_s - \mu^l_s$. Note that $\varphi^1_{s,\varepsilon_0}$ is a measure supported on the line $x' = 0$, so μ^l_s is essentially frequency localized near the hyperplane $\xi_{d+1} = 0$.

In the case $V = \{0\}^d \times \mathbb{R}$, we let φ^d be a smooth bump function supported on $B(0, 100^{-1}) \subset \mathbb{R}^d$ with integral 1, and let

$$\varphi^d_{s,\varepsilon_0}(x) = \delta^{-d\varepsilon_0} 2^{-ds} \varphi^d(2^{-s}\delta^{-\varepsilon_0} x')\delta(x_{d+1}) \,. \tag{1.5.4}$$

We define the low frequency part of μ_s by

$$\begin{aligned} \mu^l_s(x) &:= \mu_s * \varphi^d_{s,\varepsilon_0}(x) \\ &= \delta^{-d\varepsilon_0} 2^{-ds} \int \delta(|y'|^2 - x_{d+1}) K_s(y') \varphi^d(2^{-s}\delta^{-\varepsilon_0}(x' - y')) \,\mathrm{d}y' \end{aligned} \tag{1.5.5}$$

and the high frequency part by $\mu^h_s = \mu_s - \mu^l_s$. In this case, φ^d_s is a measure supported on the hyperplane $x_{d+1} = 0$, and μ^l_s is essentially frequency localized near the line $\xi' = 0$.

In both cases the low frequency part μ^l_s is a function, and one can check using (1.5.3) or (1.5.5) and (1.2.1) that

$$|\mu^l_s(x)| \lesssim \delta^{-d\varepsilon_0} |I(p)|^{-1} \tag{1.5.6}$$

and

$$|\partial_i \mu^l_{s(p)}(x)| \lesssim \begin{cases} \delta^{-(d+2)\varepsilon_0} 2^{-s} |I(p)|^{-1} & \text{if } i = 1, \dots, d, \\ \delta^{-(d+2)\varepsilon_0} 2^{-2s} |I(p)|^{-1} & \text{if } i = d+1. \end{cases} \tag{1.5.7}$$

We define operators T^l_p, T^h_p by replacing $\mu_{s(p)}$ with $\mu^l_{s(p)}, \mu^h_{s(p)}$ in (1.5.1), and we define $T^l_{\mathbf{A}}, T^h_{\mathbf{A}}$ as the sum of T^l_p, T^h_p over all $p \in \mathbf{A}$. Then we have $T_{\mathbf{A}} = T^l_{\mathbf{A}} + T^h_{\mathbf{A}}$. We prove estimates for $T^l_{\mathbf{A}}$ and $T^h_{\mathbf{A}}$ separately in the next two subsections.

1.5.2 The smooth part

Here we estimate the smooth part $T^l_{\mathbf{A}}$. The argument is the same for $V = \mathbb{R}^{d-1} \times \{0\}^2$ and $V = \{0\}^d \times \mathbb{R}$. While μ^l_s is defined differently in these two cases, we will only use the properties (1.5.6) and (1.5.7) of μ^l_s, which hold in both cases.

We define the separation $\Delta(p, p')$ of a pair of tiles p, p' to be 1 if $\omega(p) \cap \omega(p') \neq \emptyset$, and else we define it as

$$\Delta(p, p') := \sup\{\lambda > 0 \,:\, \lambda\omega(p) \cap \omega(p') = \emptyset \text{ and } \omega(p) \cap \lambda\omega(p') = \emptyset\} \,.$$

Then we have the following almost orthogonality estimate.

Lemma 1.5.1. *There exists a constant $C > 0$ such that for all tiles p, p' with $s(p) \geq s(p')$, we have*

$$|\langle T_p^{l*} g, T_{p'}^{l*} g\rangle| \leq C\delta^{-(2d+2)\varepsilon_0} \frac{\Delta(p,p')^{-1}}{|I(p)|} \int_{E(p)} |g| \int_{E(p')} |g| \,. \tag{1.5.8}$$

Proof. We abbreviate $\Delta = \Delta(p, p')$. We have, by Fubini and the definition of T_p^l

$$|\langle T_p^{l*} g, T_{p'}^{l*} g\rangle| \leq \int\int |\mathbf{1}_{E(p)} g(x_1)||\mathbf{1}_{E(p')} g(x_2)| \\ \left|\int e^{i(N(x_1)-N(x_2))\cdot y} \mu_{s(p)}^l(x_1 - y)\overline{\mu_{s(p')}^l(x_2 - y)} \,\mathrm{d}y\right| \mathrm{d}x_1 \,\mathrm{d}x_2 \,. \tag{1.5.9}$$

The inner integral in (1.5.9) is bounded by $C\delta^{-2d\varepsilon_0}|I(p)|^{-1}$, by (1.5.6). This implies (1.5.8) if $\Delta \leq 3$, so we will assume from now on that $\Delta > 3$. Fix $x_1 \in E(p), x_2 \in E(p')$ and let

$$\Phi(y) = \mu_{s(p)}^l(x_1 - y)\overline{\mu_{s(p')}^l(x_2 - y)} \,.$$

By (1.5.6) and (1.5.7), we have the bounds

$$|\partial_i \Phi(y)| \lesssim \delta^{-(2d+2)\varepsilon_0} \frac{1}{|I(p)||I(p')|} 2^{-s(p')} \,, \quad i = 1, \dots, d$$

and

$$|\partial_i \Phi(y)| \lesssim \delta^{-(2d+2)\varepsilon_0} \frac{1}{|I(p)||I(p')|} 2^{-2s(p')} \,, \quad i = d+1 \,.$$

Since $\Delta = \Delta(p, p') > 3$, we have by definition

$$\max_{i=1,\dots,d} 2^{s(p)} |N_i(x_1) - N_i(x_2)| \geq (\Delta - 1)/2 \geq \Delta/3$$

or

$$2^{2s(p)} |N_{d+1}(x_1) - N_{d+1}(x_2)| \geq (\Delta^2 - 1)/2 \geq \Delta/3 \,.$$

Integrating by parts in the corresponding direction, we find that

$$\left|\int e^{i(N(x_1)-N(x_2))\cdot y} \mu_{s(p)}^l(x_1 - y)\overline{\mu_{s(p')}^l(x_2 - y)} \,\mathrm{d}y\right| \\ \lesssim \Delta^{-1}\delta^{-(2d+2)\varepsilon_0}|I(p)|^{-1} \,,$$

which together with (1.5.9) gives the claimed estimate (1.5.8). □

Lemma 1.5.2. *Let $1 \leq q \leq \infty$. There exists $C > 0$ such that for every antichain $\mathbf{A}$ of density δ and every tile p*

$$\Bigg\| \sum_{p' \in \mathbf{A}, s(p') \leq s(p)} \Delta(p, p')^{-1} \mathbf{1}_{E(p')} \Bigg\|_q \leq C\delta^{\frac{\kappa}{q}} \Bigg| \bigcup_{p' \in \mathbf{A}, s(p') \leq s(p)} I(p') \Bigg|^{1/q} \,,$$

where $\kappa = (1 + \dim_h V)^{-1}$.

Proof. The estimate holds for $q = \infty$ with $C = 1$, since $\Delta(p, p') \geq 1$ and the sets $E(p')$, $p' \in \mathbf{A}$ are disjoint. By Hölder's inequality it therefore only remains to show the estimate for $q = 1$:

$$\sum_{p' \in \mathbf{A}, s(p') \leq s(p)} \Delta(p, p')^{-1} |E(p')| \lesssim \delta^{\kappa} \left| \bigcup_{p' \in \mathbf{A}, s(p') \leq s(p)} I(p') \right| .$$

The contribution of tiles with $\Delta(p, p') \geq \delta^{-\kappa}$ is clearly bounded by the right hand side, since $E(p') \subset I(p')$ and the sets $E(p')$ are disjoint. So it remains only to estimate the contribution of the tiles

$$\mathbf{A}' := \{p' \in \mathbf{A} \,:\, s(p') \leq s(p)\,,\ \Delta(p, p') < \delta^{-\kappa}\} .$$

Let $\mathbf{L}$ be the set of maximal dyadic cubes L for which there exists $p' \in \mathbf{A}'$ with $L \subsetneq I(p')$, and there exists no $p' \in \mathbf{A}'$ with $I(p') \subset L$. The set $\mathbf{L}$ is a partition of $\cup_{p' \in \mathbf{A}'} I(p')$, so it will be enough to show for each $L \in \mathbf{L}$ the estimate

$$\left| L \cap \bigcup_{p' \in \mathbf{A}'} E(p') \right| \leq \delta^{1 - \kappa \dim_h V} |L| . \tag{1.5.10}$$

Fix $L \in \mathbf{L}$. There exists a tile $p' \in \mathbf{A}'$ with $I(p') \subset \hat{L}$. If $I(p') = \hat{L}$, define $p_L = p'$, and else let p_L be the unique tile with $I(p_L) = \hat{L}$ and $\omega(p) \subset \omega(p_L)$. If λ is the smallest odd number such that $\lambda \geq 5\delta^{-\kappa}$, then we have in both cases that $\lambda\omega(p') \supset \lambda\omega(p_L)$. For each $p'' \in \mathbf{A}'$ with $L \cap I(p'') \neq \emptyset$, it holds that $L \subsetneq I(p'')$ and $\omega(p'') \subset \lambda\omega(p_L)$. It follows that for all such p'', we have that $L \cap E(p'') \subset E(\lambda, p_L)$. Thus

$$\begin{aligned} |L \cap \bigcup_{p'' \in \mathbf{A}'} E(p'')| &\leq |E(5\delta^{-\kappa} p_L)| \\ &\leq \lambda^{-\dim_h V} \operatorname{dens}(\mathbf{A}) |L| \lesssim \delta^{1 - \kappa \dim_h V} |L| , \end{aligned}$$

giving (1.5.10) and hence the lemma. □

Using the last two lemmas, we can prove our main estimate for $T^l_{\mathbf{A}}$.

Lemma 1.5.3. *For all $1 \leq q < 2$, there exists a constant $C > 0$ such that*

$$\|T^l_{\mathbf{A}}\|_{2 \to 2} \leq C \delta^{\frac{\kappa}{2q'} - (d+1)\varepsilon_0} ,$$

where $\kappa = 1/(1 + \dim_h V)$.

Proof. Define for each $p \in \mathbf{A}$ the set

$$\mathbf{A}(p) = \{p' \in \mathbf{A} \,:\, s(p') \leq s(p),\, 3I(p) \cap 3I(p') \neq \emptyset\} .$$

Since $T^{l*}_p g$ is always supported in $3I(p)$, we have that

$$|\langle T^{l*}_{\mathbf{A}} g, T^{l*}_{\mathbf{A}} g \rangle| \leq 2 \sum_{p \in \mathbf{A}} \sum_{p' \in \mathbf{A}(p)} |\langle T^{l*}_p g, T^{l*}_{p'} g \rangle| .$$

By Lemma 1.5.1, this is bounded up to constant factor by

$$\delta^{-(2d+2)\varepsilon_0} \sum_{p \in \mathbf{A}} \int \mathbf{1}_{E(p)} |g| \frac{1}{|I(p)|} \int |g| \sum_{p' \in \mathbf{A}(p)} \Delta(p,p')^{-1} \mathbf{1}_{E(p')} \,.$$

Cubes $I(p')$ associated to $p' \in \mathbf{A}(p)$ are contained in $5I(p)$, so we may apply Hölder's inequality in the inner integral with $q < 2$ to bound this by a constant times

$$\delta^{-(2d+2)\varepsilon_0} \sum_{p \in \mathbf{A}} \int \mathbf{1}_{E(p)} |g| M^q |g| \frac{\| \sum_{p' \in \mathbf{A}(p)} \Delta(p,p')^{-1} \mathbf{1}_{E(p')} \|_{L^{q'}}}{|I(p)|^{1/q'}} \,,$$

where $M^q g = M(g^q)^{1/q}$ is the q-maximal function. Using Lemma 1.5.2 to estimate the $L^{q'}$-norm, we estimate this by a constant times

$$\delta^{\frac{\kappa}{q'} - (2d+2)\varepsilon_0} \sum_{p \in \mathbf{A}} \int \mathbf{1}_{E(p)} |g| M^q |g|$$
$$\leq \delta^{\frac{\kappa}{q'} - (2d+2)\varepsilon_0} \int |g| M^q |g| \lesssim \delta^{\frac{\kappa}{q'} - (2d+2)\varepsilon_0} \|g\|_2^2 \,.$$

Here we used disjointness of the sets $E(p)$ and L^2-boundedness of M^q for $q < 2$. This completes the proof. $\square$

1.5.3 The high frequency part

Here we estimate the high frequency part $T^h_{\mathbf{A}}$.

We start by discretizing modulation frequencies. For this, we let $\varepsilon_1 > 0$ be a small positive number, much smaller than ε_0, it will be chosen at the end of this section. We fix finite subsets $M(\omega) \subset \omega$ of each dyadic frequency cube ω, such that the following holds with $\rho = \delta^{\varepsilon_1}$:

(i) $|M(\omega)| \lesssim \rho^{-\dim_h V}$,

(ii) For each $N \in \omega$, there exists $N' \in M(\omega)$ with

$$d_{\mathrm{par}}(N, N') \leq \rho 2^{-s(\omega)} \,, \tag{1.5.11}$$

where d_{par} denotes the parabolic distance

$$d_{\mathrm{par}}(N, N') = \max\{ \max_{1 \leq i \leq d} |N_i - N'_i|, |N_{d+1} - N'_{d+1}|^{1/2} \} \,.$$

(iii) If $\omega = a + b\omega'$ then $M(\omega) = a + bM(\omega')$.

For each tile p, we partition the set $E(p)$ into $E(p^c)$, $c \in M(\omega(p))$, such that for all $x \in E(p^c)$

$$d_{\mathrm{par}}(N(x), c) \leq \delta^{\varepsilon_1} 2^{-s(p)} \,. \tag{1.5.12}$$

We define for a tile p and $c \in M(\omega(p))$ the operator

$$T^{h,c}_p f(x) := \mathbf{1}_{E(p^c)} \int f(x-y) e^{ic \cdot y} \, \mathrm{d}\mu^h_{s(p)}(y) \,,$$

and we define also

$$T^{h,d}_{\mathbf{A}} = \sum_{p\in\mathbf{A}} \sum_{c\in M(\omega(p))} T^{h,c}_p \,.$$

By (1.5.12), we have for $x \in E(p^c)$ and all y in the support of $\mu^h_{s(p)}$ that

$$|c \cdot y - N(x) \cdot y| \le \delta^{\varepsilon_1}\,,$$

hence

$$|T_p f(x) - \sum_{c\in M(\omega(p))} T^{h,c}_p f(x)| \le \delta^{\varepsilon_1} \mathbf{1}_{E(p)}(x) \int |f(x-y)|\,\mathrm{d}|\mu^h_{s(p)}|(y)\,.$$

Summing this over all $p \in \mathbf{A}$ and using that the corresponding sets $E(p)$ are disjoint, it follows that $|T^h_{\mathbf{A}} - T^{h,d}_{\mathbf{A}}|$ is dominated by δ^{ε_1} times a mollified maximal average along the parabola. Thus

$$\|T^h_{\mathbf{A}} - T^{h,d}_{\mathbf{A}}\|_{2\to 2} \lesssim \delta^{\varepsilon_1}\,. \tag{1.5.13}$$

It now remains to estimate the discretized operator $T^{h,d}_{\mathbf{A}}$. By disjointness of the sets $E(p^c)$ and the third property of the sets $M(\omega)$, we have

$$T^{h,d}_{\mathbf{A}} f(x) \le Sf(x)\,, \tag{1.5.14}$$

where the square function Sf is defined depending on V as follows: In the case $V = \mathbb{R}^{d-1} \times \{0\}^2$ it is defined by

$$(Sf(x))^2 = \sum_{s\in\mathbb{Z}} \sum_{N\in\mathbb{Z}^{d-1}\times\{0\}^2} \sum_{c\in M(\omega_0)} |f * (e^{i\delta_s(c+N)\cdot y} \mu^h_s(y))(x)|^2 \tag{1.5.15}$$

with $\omega_0 = [0,1)^{d-1} \times \{0\}^2$, and in the case $V = \{0\}^d \times \mathbb{R}$ it is defined by

$$(Sf(x))^2 = \sum_{s\in\mathbb{Z}} \sum_{N\in\{0\}^d\times\mathbb{Z}} \sum_{c\in M(\omega_0)} |f * (e^{i\delta_s(c+N)\cdot y} \mu^h_s(y))(x)|^2 \tag{1.5.16}$$

with $\omega_0 = \{0\}^d \times [0,1)$.

Lemma 1.5.4. *There exist $C > 0$ such that the following holds. Let S be defined by* (1.5.15) *or by* (1.5.16)*. Then we have*

$$\|S\|_{2\to 2} \le C\delta^{\frac{2}{d+2}\varepsilon_0 - \frac{\dim_h V}{2}\varepsilon_1}\,.$$

Proof. We first treat the case (1.5.15). We have

$$\hat{\mu}_0(\xi) = \int_{1/8<|y|<1/3} e^{i(\xi'\cdot y + \xi_{d+1}|y|^2)} K_0(y)\,\mathrm{d}y\,.$$

Thus we obtain from the method of stationary phase (see e.g. [65], Lemma 7.7.3) and (1.2.1):

$$|\hat{\mu}_0(\xi)|^2 \lesssim \min\{1, |\xi_{d+1}|^{-d}\}\,. \tag{1.5.17}$$

Combining this with (1.5.2) yields

$$|\hat{\mu}^h_0(\xi)|^2 = |(1 - \hat{\varphi}^1(\delta^{2\varepsilon_0}\xi_{d+1}))\hat{\mu}_0(\xi)|^2 \lesssim \min\{|\delta^{2\varepsilon_0}\xi_{d+1}|^2, |\xi_{d+1}|^{-d}\}\,. \tag{1.5.18}$$

Summing up (1.5.18) gives

$$\sum_{\substack{N\in\mathbb{Z}^{d-1}\times\{0\}^2\\ |\xi'+c'+N'|\le|\xi_{d+1}|}} |\hat{\mu}_0^h(\xi+c+N)|^2 \lesssim |\xi_{d+1}|^{d-1}\min\{|\delta^{2\varepsilon_0}\xi_{d+1}|^2, |\xi_{d+1}|^{-d}\}\,. \tag{1.5.19}$$

If $|\xi'|>|\xi_{d+1}|$ then we have

$$|\nabla(\xi'\cdot y+\xi_{d+1}|y|^2)| = |\xi'+\xi_{d+1}y| > \frac{1}{2}|\xi'|\,.$$

Integrating by parts (see e.g. [65], Theorem 7.7.1), we obtain

$$|\hat{\mu}_0(\xi)|^2 \lesssim \min\{1, |\xi'|^{-d-1}\}\,, \tag{1.5.20}$$

and hence

$$|\hat{\mu}_0^h(\xi)|^2 \lesssim \min\{|\delta^{2\varepsilon_0}\xi_{d+1}|\min\{1,|\xi'|^{-d-1}\}, |\xi'|^{-d-1}\}\,. \tag{1.5.21}$$

Summing up estimate (1.5.21) gives

$$\sum_{\substack{N\in\mathbb{Z}^{d-1}\times\{0\}^2\\ |\xi'+c'+N'|>|\xi_{d+1}|}} |\hat{\mu}_0^h(\xi+c+N)|^2 \lesssim \min\{|\delta^{2\varepsilon_0}\xi_{d+1}|^2, |\xi_{d+1}|^{-1}\}\,. \tag{1.5.22}$$

Combining (1.5.19) and (1.5.22) and using $\min\{a,b\}\le\min\{(a^3b^{d-1})^{1/(d+2)},b\}$ yields

$$\sum_{N\in\mathbb{Z}^{d-1}\times\{0\}^2} |\hat{\mu}_0^h(\xi+c+N)|^2 \lesssim \min\{\delta^{2\varepsilon_0/d}|\xi_{d+1}|^2,\ |\xi_{d+1}|^{-1}\}\,.$$

Using dilation invariance of all assumptions on K this implies

$$\sum_{N\in\mathbb{Z}^{d-1}\times\{0\}^2} |\hat{\mu}_s^h(\xi+\delta_{-s}(c+N))|^2 \lesssim \min\{\delta^{\frac{12}{d+2}\varepsilon_0}|2^{2s}\xi_{d+1}|^2,\ |2^{2s}\xi_{d+1}|^{-1}\}\,.$$

Summing in s, and also summing in the $\lesssim \delta^{-\varepsilon_1\dim_h V}$ choices of c, we obtain

$$\sum_{s\in\mathbb{Z}}\ \sum_{N\in\mathbb{Z}^{d-1}\times\{0\}^2}\ \sum_{c\in M(\omega_0)} |\hat{\mu}_s^h(\xi+\delta_{-s}(c+N))|^2 \lesssim \delta^{\frac{4}{d+2}\varepsilon_0-\varepsilon_1\dim_h V}\,.$$

This completes the proof in the case (1.5.15) by Plancherel.

Now we deal with the vertical modulation case (1.5.16). The stationary phase estimate (1.5.17) combined with (1.5.4) yields

$$|\hat{\mu}_0^h(\xi)|^2 \lesssim \min\{1,|\xi_{d+1}|^{-d}\}|\delta^{\varepsilon_0}\xi'|^2\,,$$

and hence

$$\sum_{N\in\{0\}^d\times\mathbb{Z}} |\hat{\mu}_0^h(\xi+c+N)|^2 \lesssim |\delta^{\varepsilon_0}\xi'|^2\,.$$

Combining the estimates (1.5.20) in the range $|\xi'|>|(\xi+c+N)_{d+1}|$ and (1.5.17) in the range $|\xi'|\le|(\xi+c+N)_{d+1}|$ gives

$$\sum_{N\in\{0\}^d\times\mathbb{Z}} |\hat{\mu}_0^h(\xi+c+N)|^2 \lesssim |\xi'|^{1-d}\,.$$

Therefore we have

$$\sum_{N \in \{0\}^d \times \mathbb{Z}} |\hat{\mu}_0^h(\xi + c + N)|^2 \lesssim \min\{\delta^{2\varepsilon_0}|\xi'|^2, |\xi'|^{1-d}\}\,.$$

As before, this implies the square function estimate (with better dependence on δ) using dilation invariance of all assumptions to sum in s, the bound on the number of summands to sum in c, and then Plancherel. □

Combining the estimate for $T^l_{\mathbf{A}}$ in Lemma 1.5.3, for $T^h_{\mathbf{A}} - T^{h,d}_{\mathbf{A}}$ in (1.5.13) and for $T^{h,d}_{\mathbf{A}}$ in (1.5.14) and Lemma 1.5.4, and choosing ε_0 and then ε_1 sufficiently small, we obtain Proposition 1.3.2.

1.6 Trees: Proof of Proposition 1.3.4

Here we prove a mild generalization of Proposition 1.3.4. We will estimate what we call generalized tree operators, with general kernels satisfying conditions described below. We need the estimates in Section 1.7 both for the singular Radon transform and for a version of it with a smoothened kernel.

1.6.1 General setup

Let in the following μ_s, λ_s, $s \in \mathbb{Z}$ be any finite measures satisfying the following conditions:

i) μ_s, λ_s are supported in $\delta_s(B(0, 1/2))$,

ii) $|\mu_s| \le \lambda_s$,

iii) $\mu_s(\mathbb{R}^{d+1}) = 0$,

iv) we have
$$\max\{|\widehat{D_{-s}\lambda_s}(\xi))|, |\widehat{D_{-s}\mu_s}(\xi)|\} \le |\xi|^{-1}\,, \tag{1.6.1}$$
where $D_{-s}\mu_s := 2^{-(d+2)s}\mu_s \circ \delta_{-s}$,

v) and there exist $\gamma > 0$ and $p, q < 2$ such that convolution with $D_{-s}\lambda_s$ and $D_{-s}\mu_s$ is bounded from L^p to $L^{q'+\gamma}$ with norm at most 1 for each s.

Relevant for the proof of Theorem 1.1.1 are the measures defined by

$$\int f \,\mathrm{d}\mu_s = \int_{\mathbb{R}^d} f(y, |y|^2) K_s(y)\,\mathrm{d}y\,, \tag{1.6.2}$$

$$\int f \,\mathrm{d}\lambda_s = \int_{\mathbb{R}^d} f(y, |y|^2) 2^{-ds}\eta(2^{-s}y)\,\mathrm{d}y\,, \tag{1.6.3}$$

and smoothened versions thereof. They clearly satisfy conditions i) to iii), condition iv) follows from standard stationary phase estimates (e.g. [65], Lemma 7.7.3) and for condition v) one can take any (p, q) such that $(\frac{1}{p}, \frac{1}{q})$ is in the interior of the convex hull of $(1, 0), (0, 1), (\frac{d+1}{d+2}, \frac{d+1}{d+2})$ (see [85]).

1.6.2 Generalized tree operators

With the choice of μ_s from (1.6.2), the tile operators defined in Section 1.2 can be written as

$$T_p f(x) = \mathbf{1}_{E(p)}(x) \int f(x-y) e^{iN(x)\cdot y} \mathrm{d}\mu_s(y) \,,$$

and tree operators can be written as

$$T_{\mathbf{T}} f(x) = \sum_{p \in \mathbf{T}} T_p f(x) = \sum_{s \in \sigma(x)} \int f(x-y) e^{iN(x)\cdot y} \mathrm{d}\mu_s(y) \,,$$

where $\sigma(x) := \{s(p) \,:\, p \in \mathbf{T}, x \in E(p)\}$. Since $\mathbf{T}$ is a subset of the set of all admissible tiles, the set $\sigma(x)$ is always contained in

$$J = J(\mathbf{T}) := \{s \in \mathbb{Z} \,:\, \exists \omega \text{ admissible}, s(\omega) = s,\, \omega(\mathbf{T}) \subset \omega\} \,,$$

and convexity of the tree $\mathbf{T}$ implies that there are $\underline{\sigma}(x) \le \overline{\sigma}(x)$ with

$$\sigma(x) = J \cap [\underline{\sigma}(x), \overline{\sigma}(x)] \,.$$

Motivated by this, we define a generalized tree to be a pair $(\mathbf{T}, \sigma')$, where $\mathbf{T}$ is a tree and σ' is a function associating to each x a set

$$\sigma'(x) = J \cap [\underline{\sigma}'(x), \overline{\sigma}'(x)] \subset \sigma(x) \,.$$

The generalized tree operator associated to $(\mathbf{T}, \sigma')$ is defined by

$$T_{\mathbf{T},\sigma'} f(x) = \sum_{s \in \sigma'(x)} \int f(x-y) e^{iN(x)\cdot y} \mathrm{d}\mu_s(y) \,.$$

The above discussion shows that this includes in particular the tree operators defined in Section 1.3, by choosing μ_s as in (1.6.2) and $\sigma'(x) = \sigma(x)$ for each x.

1.6.3 Single tree estimate

We now dominate the generalized tree operators $T_{\mathbf{T},\sigma'}$ by a sum of two simpler operators $T_{*\mathbf{T},\sigma'}$ and $M_{\mathbf{T}}$, resembling a maximally truncated singular integral and a maximal average along the paraboloid.

Define the unmodulated operator associated to $(\mathbf{T}, \sigma')$ by

$$T_{*\mathbf{T},\sigma'} f(x) := \sum_{s \in \sigma'(x)} f * \mu_s(x) = \sum_{s=\underline{\sigma}'(x)}^{\overline{\sigma}'(x)} f * (\mathbf{1}_J(s)\mu_s)(x) \,.$$

This is a maximally truncated singular integral along the paraboloid, and it follows from standard square function arguments as in [110] that $T_{*\mathbf{T}}$ is bounded on L^2.

Define also the maximal average $M_{\mathbf{T}}$ associated to $\mathbf{T}$:

$$M_{\mathbf{T}} f(x) := \sup_{s \in \sigma(x)} |f| * \lambda_s(x) \,.$$

The operator $M_{\mathbf{T}}$ is bounded above pointwise by the maximal average associated to $(\lambda_s)_{s\in\mathbb{Z}}$, which is bounded on L^2 under our assumptions.

We have the following pointwise estimate for $T_{\mathbf{T},\sigma'}$ in terms of $T_{*\mathbf{T},\sigma'}$ and $M_{\mathbf{T}}$.

Lemma 1.6.1. *There exists $C > 0$ such that for each generalized tree $(\mathbf{T}, \sigma')$*

$$|T_{\mathbf{T},\sigma'} f(x)| \le |T_{*\mathbf{T},\sigma'}(e^{-iN(\mathbf{T})(y)} f(y))(x)| + CM_{\mathbf{T}} f(x) \,,$$

where $N(\mathbf{T})$ is the smallest element, in lexicographic order, of $\omega(\mathbf{T})$.

Proof. Suppose that $p \in \mathbf{T}$ with $s(p) = \max \sigma(x) - k$. Then there exists some $p' \in \mathbf{T}$ with $s(p') = s(p) + k$ and $x \in E(p')$. Since $N(x) \in \omega(p')$ and $N(\mathbf{T}) \in \omega(p')$, it follows that

$$|(N(\mathbf{T}) - N(x)) \cdot y| \le d2^{-k}$$

for all y in the support of $\mu_{s(p)}$. Hence we have

$$\begin{aligned}
&\left| T_p f(x) - e^{iN(\mathbf{T})\cdot x} \mathbf{1}_{E(p)}(x) \int f(x-y) e^{-iN(\mathbf{T})\cdot(x-y)} \mathrm{d}\mu_{s(p)}(y) \right| \\
&\qquad = \left| \mathbf{1}_{E(p)}(x) \int f(x-y)(e^{iN(x)\cdot y} - e^{iN(\mathbf{T})\cdot y}) \mathrm{d}\mu_{s(p)}(y) \right| \\
&\qquad\qquad \lesssim 2^{-k} \mathbf{1}_{E(p)}(x) |f| * \lambda_s(x) \,,
\end{aligned}$$

using that $|\mu_s| \le \lambda_s$ and $|e^{iN(x)\cdot y} - e^{iN(\mathbf{T})\cdot y}| \lesssim 2^{-k}$ on the support of $\mu_{s(p)}$. Summing over all tiles $p \in \mathbf{T}$, we obtain:

$$\begin{aligned}
&|T_{\mathbf{T},\sigma'} f(x) - e^{N(\mathbf{T})\cdot x} T_{*\mathbf{T},\sigma'}(e^{-N(\mathbf{T})\cdot y} f(y))(x)| \\
&\qquad \le \sum_{p \in \mathbf{T}} \left| T_p f(x) - e^{iN(\mathbf{T})\cdot x} \mathbf{1}_{E(p)}(x) \int f(x-y) e^{-iN(\mathbf{T})\cdot(x-y)} \mathrm{d}\mu_{s(p)}(y) \right| \\
&\qquad\qquad \lesssim \sum_{k \ge 0} 2^{-k} \sum_{\substack{p \in \mathbf{T} \\ s(p) = \max \sigma(x) - k}} \mathbf{1}_{E(p)}(x) |f| * \lambda_s(x) \,.
\end{aligned}$$

By disjointness of the sets $E(p)$ for tiles p of a fixed scale, the inner sum is bounded by $M_{\mathbf{T}}$. This completes the proof. □

1.6.4 Low density trees and sparse bounds

Lemma 1.6.1 implies that the operators $T_{\mathbf{T},\sigma'}$ are bounded on L^2 uniformly over all generalized trees $(\mathbf{T}, \sigma')$. We will now improve this estimate for trees of small density δ, using sparse bounds for the operators $T_{*\mathbf{T},\sigma'}$ and $M_{\mathbf{T}}$. These sparse bounds are variants of bounds for prototypical singular Radon transforms that were shown by Oberlin [97]. Our setting is slightly more general than in [97], and we need a more precise estimate, but it still follows from Oberlin's proof with only minor modifications.

A collection $\mathcal{S}$ of dyadic cubes is called sparse if for each $Q \in \mathcal{S}$ there exists a subset $U(Q) \subset Q$ with $|U(Q)| \ge |Q|/2$, such that the sets $U(Q)$ are pairwise disjoint. For every cube Q and $p \ge 1$, we denote by

$$\langle f \rangle_{Q,p} := \left(\frac{1}{|Q|} \int_Q |f|^p \,\mathrm{d}x \right)^{1/p}$$

the p-average of a function f over Q. Finally, given a dyadic cube Q and a tree $\mathbf{T}$, we define

$$E(Q) = E(\mathbf{T}, Q) := \bigcup_{\substack{p \in \mathbf{T} \\ I(p) \subset 3Q}} E(p)\,.$$

Lemma 1.6.2 ([97])**.** *Suppose that (p, q) are as in condition v) in Subsection 1.6.1. Then there exists a constant $C > 0$ such that for every generalized tree $(\mathbf{T}, \sigma')$ and all f, g there exists a sparse collection of cubes $\mathcal{S} \subset \mathbf{D}$ such that*

$$\left| \int T_{*\mathbf{T},\sigma'} f(x) g(x)\,\mathrm{d}x \right| \leq C \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{Q,p} \langle \mathbf{1}_{E(Q)} g \rangle_{3Q,q}\,. \tag{1.6.4}$$

The same statement holds with $T_{*\mathbf{T},\sigma'}$ replaced by $M_{\mathbf{T}}$.

Proof. This follows from the proof of Theorems 1.3 and 1.4 in [97] (note that our q is his q'), with the following modifications. Firstly, Oberlin considers operators

$$\sum_{s \in \mathbb{Z}} \varepsilon_s(x) f * \mu_s\,,$$

with $\varepsilon_s(x) \in [-1, 1]$, where μ_s is an isotropic dilate of a fixed measure μ_0. In our setting the μ_s are *anisotropic* dilates of $D_{-s}\mu_s$, but Oberlin's proof goes through in the anisotropic setting as well. Furthermore, in our setting the measures $D_{-s}\mu_s$ are not identical, they depend on s. However, they satisfy all assumptions of Oberlin's theorems uniformly in s, by i) - v) above, and the proof still goes through with this assumption.

It remains to explain why we can insert $\mathbf{1}_{E(Q)}$ in the q-average over $3Q$ in (1.6.4). We explain it for $T_{*\mathbf{T},\sigma'}$, the argument for $M_{\mathbf{T}}$ is very similar. Oberlin constructs the sparse collection $\mathcal{S}$ by an iterative argument starting from a large cube Q_0, such that f, g are supported in Q_0, $3Q_0$ respectively. The expression on the left hand side of (1.6.4) does not change if f is restricted to $I(\mathbf{T})$ and g to $3I(\mathbf{T})$, so we can choose $Q_0 = I(\mathbf{T})$. Oberlin then defines operators T_Q, which in our notation are

$$T_Q f(x) = \sum_{\substack{p \in \mathbf{T} \\ s(p) \leq s(Q)}} \mathbf{1}_{E(p)}(x) \mathbf{1}_{s(p) \in \sigma'(x)} \cdot \mu_{s(p)} * (\mathbf{1}_Q f)(x)\,, \tag{1.6.5}$$

so that $T_{*\mathbf{T},\sigma'} = T_{Q_0}$. Note that in the sum in (1.6.5) only tiles p with $Q \cap 3I(p) \neq \emptyset$ contribute. But if $s(p) \leq s(Q)$ and $Q \cap 3I(p) \neq \emptyset$, then $I(p) \subset 3Q$. Thus, $T_Q = \mathbf{1}_{E(Q)} T_Q$, and hence

$$\langle T_Q f, g \rangle = \langle T_Q f, \mathbf{1}_{E(Q)} g \rangle\,. \tag{1.6.6}$$

Our goal is to estimate

$$|\langle T_{Q_0} f, g \rangle| = |\langle T_{Q_0} f, \mathbf{1}_{E(Q_0)} g \rangle|\,.$$

Oberlin shows that for every dyadic cube Q and every f there exists a collection $Q_1(Q)$ of dyadic cubes such that

$$|\langle T_Q f, g \rangle| \leq |Q| \langle f \rangle_{Q,p} \langle g \rangle_{3Q,q} + \sum_{\substack{Q' \in Q_1(Q) \\ Q' \subset Q}} \langle T_{Q'} f, g \rangle\,. \tag{1.6.7}$$

This follows from his equations (3.2) and (3.3) and the claims below them. The sparse collection $\mathcal{S}$ is then constructed by starting with Q_0 and iteratively adding for all $Q \in \mathcal{S}$ the cubes $Q_1(Q)$ to $\mathcal{S}$. Combining (1.6.6) and (1.6.7) one obtains (1.6.4) for this $\mathcal{S}$, and Oberlin shows that $\mathcal{S}$ is sparse. This completes the proof. Note that the only change in our argument compared to Oberlin's is that we use (1.6.6) to insert $\mathbf{1}_{E(Q)}$ into the q-averages. $\square$

Corollary 1.6.3. *Let q be as in v) in Subsection 1.6.1. Then for each $\varepsilon < \frac{1}{q} - \frac{1}{2}$ there exists $C > 0$ such that for each generalized tree $(\mathbf{T}, \sigma')$ with $\mathbf{T}$ of density δ, we have*

$$\|T_{*\mathbf{T},\sigma'}\|_{2\to 2} \le C\delta^{\varepsilon}\,.$$

The same statement holds with $T_{*\mathbf{T},\sigma'}$ replaced by $M_{\mathbf{T}}$.

Proof. Let $\mathbf{L} = \mathbf{L}(\mathbf{T})$ be the collection of maximal dyadic cubes L such that there exists some $p \in \mathbf{T}$ with $L \subset I(p)$ but there exists no $p \in \mathbf{T}$ with $I(p) \subset L$. This is a partition of $I(\mathbf{T})$. Define

$$E(L) = L \cap \bigcup_{p' \in \mathbf{T}} E(p')\,.$$

We claim that for each $L \in \mathbf{L}$ we have

$$|E(L)| \le \delta |L|\,. \tag{1.6.8}$$

To prove this, fix $L \in \mathbf{L}$. By definition of $\mathbf{L}$ there exists $p \in \mathbf{T}$ with $I(p) \subset \hat{L}$. Define p_L to be p if $I(p) = \hat{L}$, and else let p_L be the unique tile with $I(p_L) = \hat{L}$ and $\omega(\mathbf{T}) \subset \omega(p_L)$. If $p' \in \mathbf{T}$ with $I(p') \cap L \neq \emptyset$ then we have $p_L \le p'$. Thus

$$|E(L)| \le |E(p_L)| \le \mathrm{dens}(\mathbf{T})|L| \le \delta |L|\,,$$

giving the claim (1.6.8). Since $\mathbf{L}$ forms a partition of $I(\mathbf{T})$, we have for all $Q \in \mathbf{D}$

$$|E(Q)| \le \sum_{L \subset 3Q} |E(L)| \le \delta \sum_{L \subset 3Q} |L| \le \delta 3^{d+2} |Q|\,.$$

Thus, for every $\tilde{q} > q$, we have by Hölder's inequality

$$\langle \mathbf{1}_{E(Q)} g\rangle_{3Q,q} \le (3^{d+2}\delta)^{1/q - 1/\tilde{q}} \langle g \rangle_{3Q,\tilde{q}}\,. \tag{1.6.9}$$

Now pick (p,q) as in v), and pick $\tilde{q}$ with $q < \tilde{q} < 2$. We obtain from Lemma 1.6.2 and (1.6.9)

$$\begin{aligned}\left|\int T_{*\mathbf{T},\sigma'} f(x) g(x)\,\mathrm{d}x\right| &\lesssim \delta^{1/q-1/\tilde{q}} \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{Q,p} \langle g \rangle_{3Q,\tilde{q}} \\ &\le 2\delta^{1/q-1/\tilde{q}} \sum_{Q \in \mathcal{S}} |U(Q)| \langle f \rangle_{Q,p} \langle g \rangle_{3Q,\tilde{q}} \\ &\le 2\delta^{1/q-1/\tilde{q}} \sum_{Q \in \mathcal{S}} \int_{U(Q)} M_p f M_{\tilde{q}} g \\ &\lesssim \delta^{1/q-1/\tilde{q}} \|f\|_2 \|g\|_2\,.\end{aligned}$$

Here we used boundedness of the p- and $\tilde{q}$- maximal function on L^2 for $p, \tilde{q} < 2$. This yields the desired result for $T_{*\mathbf{T},\sigma'}$.

For $M_{\mathbf{T}}$ the proof is exactly the same. $\square$

Combining Corollary 1.6.3 with Lemma 1.6.1 we obtain an estimate for a single tree with decay in the density parameter.

Corollary 1.6.4. *For each* $\varepsilon < \frac{1}{q} - \frac{1}{2}$ *there exist* $C > 0$ *such for each generalized tree* $(\mathbf{T}, \sigma')$ *with* $\mathbf{T}$ *of density* δ*, we have*

$$\|T_{\mathbf{T},\sigma'}\|_{2\to 2} \le C\delta^{\varepsilon} .$$

Recall from Subsection 1.6.1 that for the measures μ_s defined by (1.6.2), condition v) holds for all $p, q < 2$ such that $(1/p, 1/q)$ is in the interior of the convex hull of $(1,0), (0,1), (\frac{d+1}{d+2}, \frac{d+1}{d+2})$. In this case the condition in Corollary 1.6.4 becomes $\varepsilon < \frac{1}{2} - \frac{1}{2(d+1)}$.

1.7 Forests: Proof of Proposition 1.3.3

To prove Proposition 1.3.3, it remains to combine the bounds for the operators $T_{\mathbf{T}}$ from Proposition 1.3.4 for all trees in a forest $\mathbf{F}$. For this, we will show almost orthogonality estimates for tree operators associated to separated trees.

1.7.1 Basic orthogonality estimates

As a first step, we show that the adjoint generalized tree operator $T^*_{\mathbf{T},\sigma'}$ is frequency localized near $N_{\mathbf{T}}$. We use the definitions from Subsection 1.6.2.

We let Ψ denote the set of Schwartz functions ψ such that

(i) ψ is supported in $B(0,1)$,

(ii) ψ has integral 1,

(iii) $|\partial^\alpha \psi(x)| \le L$ for all $|\alpha| \le 10d$,

where L is chosen sufficiently large such that for each ξ with $|\xi| \ge 1$, there exists $\psi \in \Psi$ with $\hat{\psi}(\xi) = 0$.

For $\psi \in \Psi$, we define the frequency projection $\Pi_{R,N} = \Pi_{R,N,\psi}$ by

$$\widehat{\Pi_{R,N} f}(\xi) = (1 - \hat{\psi}(R^{-1}(\xi - N)))\hat{f}(\xi) .$$

Lemma 1.7.1. *For each* $\kappa < \frac{1}{2(d+1)}$ *there exist* C *such that the following holds. For each generalized tree* $(\mathbf{T}, \sigma')$ *such that each tile in* $\mathbf{T}$ *has scale at least* 0*, all* $R \ge 1$ *and all* $\psi \in \Psi$

$$\|T_{\mathbf{T},\sigma'} \Pi_{R,N(\mathbf{T})}\|_{2\to 2} \le CR^{-\kappa} .$$

Proof. We start by separating the $\log_2(R)$ largest scales:

$$T_{\mathbf{T},\sigma'} f(x) = \sum_{\substack{s\in\sigma'(x)\\ s\ge\overline{\sigma}'(x)-\log_2(R)}} \int f(x-y) e^{iN(x)\cdot y} \mathrm{d}\mu_s(y) \tag{1.7.1}$$

$$+ \sum_{\substack{s\in\sigma'(x)\\ s<\overline{\sigma}'(x)-\log_2(R)}} \int f(x-y) e^{iN(x)\cdot y} \mathrm{d}\mu_s(y) . \tag{1.7.2}$$

The small scale contribution (1.7.2) is close to the $N(\mathbf{T})$-modulated operator:

$$\left|\sum_{\substack{s\in\sigma'(x)\\ s<\overline{\sigma}'(x)-\log_2(R)}} \int f(x-y)(e^{-iN(x)\cdot y}-e^{-iN(\mathbf{T})\cdot y})\,\mathrm{d}\mu_s(y)\right|$$
$$\lesssim \sum_{\substack{s\in\sigma'(x)\\ s<\overline{\sigma}'(x)-\log_2(R)}} 2^{s-\overline{\sigma}'(x)}\int |f(x-y)|\,\mathrm{d}|\mu_s|(y) \lesssim R^{-1}M^\lambda f\,, \quad (1.7.3)$$

where M^λ denotes the maximal average associated to λ. Here we have used that if $s\in\sigma'(x)$ with $s=\overline{\sigma}'(x)-k$, then there exists a frequency cube of scale $s+k$ containing $N(\mathbf{T})$ and $N(x)$, which implies that for all y in the support of μ_s we have $|(N(x)-N(\mathbf{T}))\cdot y|\lesssim 2^{-k}$. So to bound (1.7.2), it suffices to show the following estimate for the $N(\mathbf{T})$-modulated operator:

$$\left\|\sum_{\substack{s\in\sigma'(x)\\ s<\overline{\sigma}'(x)-\log_2(R)}} \int e^{iN(\mathbf{T})\cdot(x-y)}(\Pi_{R,N(\mathbf{T})}f)(x-y)\,\mathrm{d}\mu_s(y)\right\|_2 \lesssim R^{-1/2}\|f\|_2\,.$$

Replacing f by $e^{iN(\mathbf{T})\cdot x}f(x)$ and taking differences it suffices to show

$$\left\|\sup_{0\le\underline{\sigma}}\left|\sum_{\underline{\sigma}\le s}\mathbf{1}_J(s)\int(\Pi_{R,0}f)(x-y)\,\mathrm{d}\mu_s(y)\right|\right\|_2 \lesssim R^{-1/2}\|f\|_2\,.$$

This follows from a standard square function argument, using that by (1.6.1) we have for all $\underline{\sigma}\ge 0$

$$|1-\hat{\psi}(R^{-1}\xi)|\sum_{\underline{\sigma}\le s}|\hat{\mu}_s(\xi)|\lesssim R^{-1/2}2^{-\underline{\sigma}/2}\,.$$

Now we treat the large scales (1.7.1), using that only logarithmically many scales contribute at each point. We discretize modulation frequencies. Let $\gamma>0$, to be chosen later. For each dyadic frequency cube ω, we fix finite subsets $M(\omega)$ satisfying conditions (i), (ii), (iii) in Section 1.5.3, with $\rho=R^{-\gamma}$. For each $s\in\mathbb{Z}$, let ω_s be the unique frequency cube of scale s containing $N(\mathbf{T})$. For each x and each $s\in\sigma'(x)$, we pick a frequency $c(x,s)\in M(\omega_s)$ such that $d_{\mathrm{par}}(N(x),c(x,s))\le R^{-\gamma}2^{-s}$. By a similar computation as in (1.7.3)

$$\left|\sum_{\substack{s\in\sigma'(x)\\ s\ge\overline{\sigma}'(x)-\log_2(R)}} \int f(x-y)(e^{-iN(x)\cdot y}-e^{ic(x,s)\cdot y})\,\mathrm{d}\mu_s(y)\right|$$
$$\le \log_2(R)R^{-\gamma}M^\lambda f\,.$$

Thus we may replace $N(x)$ by $c(x,s)$ in (1.7.1). We bound the resulting sum by the number of nonzero summands, which is at most $\log_2(R)$ at each point x, times the maximal summand. The maximal summand is controlled by the estimate

$$\left\| \sup_{s\geq 0} \sup_{c\in M(s)} \left| \int (\Pi_{R,N} f)(x-y) e^{ic\cdot y} \,\mathrm{d}\mu_s(y) \right| \right\|_2 \lesssim R^{\gamma \dim_h V} R^{-1/2} \|f\|_2 \,.$$

This follows from $|(1-\hat{\psi}(R^{-1}(\xi - N + c)))||\hat{\mu}_s(\xi)| \lesssim R^{-1/2} 2^{-s/2}$ and a standard square function argument, using that there are at most $R^{\gamma \dim_h V}$ choices of c for each s. Optimizing γ, we obtain the lemma. □

An immediate corollary are almost orthogonality estimate for separated trees with the same top cube.

Corollary 1.7.2. *For each $\kappa < \frac{1}{2d+3}$ there exists $C > 0$ such that the following holds. Let $\mathbf{T}_1, \mathbf{T}_2$ be a pair of Δ-separated, normal trees with $I(\mathbf{T}_1) = I(\mathbf{T}_2) =: I$. Let $(\mathbf{T}_1, \sigma_1), (\mathbf{T}_2, \sigma_2)$ be generalized tree operators, with possibly different μ_s. Then*

$$\left| \int T^*_{\mathbf{T}_1,\sigma_1} g_1 \overline{T^*_{\mathbf{T}_2,\sigma_2} g_2} \right| \leq C \Delta^{-\kappa} \|g_1\|_{L^2(I)} \|g_2\|_{L^2(I)} \,. \tag{1.7.4}$$

Proof. We drop the σ_i from the notation and write $T_{\mathbf{T}_i,\sigma_i} = T_{\mathbf{T}_i}$. By scaling, we may assume that the minimal scale of a tile in $\mathbf{T}_1 \cup \mathbf{T}_2$ is 0. Furthermore we may assume that $\Delta > 3$, otherwise (1.7.4) follows from L^2 boundedness of generalized tree operators. Then it holds by Δ-separation that $|N(\mathbf{T}_1) - N(\mathbf{T}_2)| \geq \Delta/3$.

We let $\kappa' = \kappa/(1-\kappa) < 1/(2(d+1))$ and $\gamma = 1/(1+\kappa')$, define $\Pi_i = \Pi_{\Delta^\gamma, N(\mathbf{T}_i)}$ for some function $\psi \in \Psi$, and split up

$$\begin{aligned}\langle T^*_{\mathbf{T}_1} g_1, T^*_{\mathbf{T}_2} g_2 \rangle = {} & \langle \Pi_1 T^*_{\mathbf{T}_1} g_1, \Pi_2 T^*_{\mathbf{T}_2} g_2 \rangle + \langle (1-\Pi_1) T^*_{\mathbf{T}_1} g_1, \Pi_2 T^*_{\mathbf{T}_2} g_2 \rangle \\ & + \langle \Pi_1 T^*_{\mathbf{T}_1} g_1, (1-\Pi_2) T^*_{\mathbf{T}_2} g_2 \rangle + \langle (1-\Pi_1) T^*_{\mathbf{T}_1} g_1, (1-\Pi_2) T^*_{\mathbf{T}_2} g_2 \rangle \,.\end{aligned}$$

The first three terms are bounded by $\Delta^{-\gamma\kappa'} \|g_1\|_{L^2(I)} \|g_2\|_{L^2(I)}$ by Lemma 1.7.1, and the last term is bounded by

$$\|(1-\Pi_1)(1-\Pi_2)\|_{2\to 2} \|g_1\|_{L^2(I)} \|g_2\|_{L^2(I)} \lesssim \Delta^{\gamma - 1} \|g_1\|_{L^2(I)} \|g_2\|_{L^2(I)} \,,$$

because $\psi \in \Psi$ and $|N(\mathbf{T}_1) - N(\mathbf{T}_2)| \geq \Delta/3$. This gives the desired estimate since $\gamma\kappa' = \gamma - 1 = \kappa$. □

1.7.2 Auxiliary estimates for oscillatory integrals

In this subsection we prepare the proof of a version of Corollary 1.7.2 for a general pair of separated trees by showing some estimates for oscillatory integrals along paraboloids.

From now on, μ_s and λ_s are fixed to be the measures defined in (1.6.2) and (1.6.3). We define

$$\chi_{s,\varepsilon} = \varepsilon^{-d-1} \mathbf{1}_{B(0,\varepsilon)} * (\lambda_{s-1} + \lambda_s + \lambda_{s+1}) \tag{1.7.5}$$

and $\chi_{s,\varepsilon,\delta} = \delta^{-1} \chi_{s,\varepsilon} \mathbf{1}_{|x_d| < \delta 2^s}$. We also define the associated maximal convolution operators

$$M^{\chi,1} f(x) := \sup_{s\in\mathbb{Z}} \sup_{\varepsilon > 0} |f| * \tilde{\chi}_{s,\varepsilon} \,, \tag{1.7.6}$$

$$M^{\chi,2}f(x) := \sup_{s\in\mathbb{Z}}\sup_{\varepsilon>0}\sup_{\delta>0}|f| * \tilde{\chi}_{s,\varepsilon,\delta}\,, \tag{1.7.7}$$

where $\tilde{\chi}(x) = \chi(-x)$. Both $M^{\chi,1}$ and $M^{\chi,2}$ are bounded on L^2, because they are dominated by the composition of the Hardy-Littlewood maximal function and maximal averages along the parabolas in the direction of the coordinate axes $\{te_i + t^2 e_{d+1} \, : \, t \in \mathbb{R}\}$ for $i = 1, \dots, d$.

Lemma 1.7.3. *There exists $C > 0$ such that the following holds. For all $a, b > 0$ with $b \geq 4a$, all $N \in V$, all $\psi \in \Psi$ and all $s \geq -1$ we have*

$$|(e^{iN\cdot y}\mu_s(y)) * \psi_a(x)| \lesssim b\chi_{s,a,b}(x) + \frac{1}{ab|N|}\chi_{s,a}(x)\,. \tag{1.7.8}$$

Furthermore, if $\varphi \geq 0$ has integral 1 *and $\mu_s^l = \mu_s * \varphi$, then*

$$|(e^{iN\cdot y}\mu_s^l(y)) * \psi_a(x)| \lesssim b\chi_{s,a,b} * \varphi(x) + \frac{1}{ab|N|}\chi_{s,a} * \varphi(x)\,. \tag{1.7.9}$$

Here we write $\psi_t(x) = t^{-d-1}\psi(t^{-1}x)$.

Proof. The estimate (1.7.9) follows from (1.7.8) and the triangle inequality. For $|x_d| < b2^s$ estimate (1.7.8) also follows directly from the triangle inequality. For $|x_d| \geq b2^s$ we distinguish cases depending on V.

If $V = \mathbb{R}^{d-1} \times \{0\}^2$ then $N_d = N_{d+1} = 0$, and without loss of generality we have that $|N_{d-1}| \geq |N|/d$. We let $\tilde{y} = (y_1, \dots, y_{d-2})$ and put

$$y = (\tilde{y}, \sqrt{h^2 - |\tilde{y}|^2}\cos(\theta), \sqrt{h^2 - |\tilde{y}|^2}\sin(\theta))\,.$$

We abbreviate $r = \sqrt{h^2 - |\tilde{y}|^2}$ and change variables to obtain

$$\begin{aligned}\int \psi_a(x-z)e^{iN\cdot z}\,\mathrm{d}\mu_s(z) &= \int \psi_a(x' - y, x_{d+1} - |y|^2)e^{iN(x)\cdot(y,|y|^2)}K_s(y)\,\mathrm{d}y \\ &= \iiint \Gamma_{h,\tilde{y}}(\theta)e^{i\phi_{h,\tilde{y}}(\theta)}\,\mathrm{d}\theta\, h\,\mathrm{d}h\,\mathrm{d}\tilde{y}\,, \end{aligned} \tag{1.7.10}$$

where we put $\phi_{h,\tilde{y}}(\theta) = N_{d-1}r\cos(\theta)$ and

$$\begin{aligned}\Gamma_{h,\tilde{y}}(\theta) = \psi_a(\tilde{x} - \tilde{y}, x_{d-1} - r\cos(\theta), x_d - r\sin(\theta), x_{d+1} - h) \\ K_s(\tilde{y}, r\cos(\theta), r\sin(\theta))e^{i\tilde{N}\cdot\tilde{y}}\,.\end{aligned}$$

Using (1.2.1) and the assumption $s \geq -1$, we find that

$$\left|\Gamma'_{h,\tilde{y}}(\theta)\right| \lesssim a^{-d-2}2^{-ds}r\,. \tag{1.7.11}$$

Since $|x_d| > b2^s$ and since $|x_d - r\sin(\theta)| \leq a$ on the support of $\Gamma_{h,\tilde{y}}$, we have additionally that $|\sin(\theta)| \geq b - 2^{-s}a \geq b/2$. Thus

$$\left|\phi'_{h,\tilde{y}}(\theta)\right| \geq \frac{|N|br}{2d} \qquad \text{and} \qquad \left|\phi''_{h,\tilde{y}}(\theta)\right| \leq \frac{|N|r}{d}\,. \tag{1.7.12}$$

Integrating by parts in (1.7.10), we obtain with (1.7.11) and (1.7.12)

$$
\begin{aligned}
(1.7.10) \le & \iiint \frac{|\Gamma'_{h,\tilde{y}}(\theta)|}{|\phi'_{h,\tilde{y}}(\theta)|} + \frac{|\Gamma_{h,\tilde{y}}(\theta)|}{|\phi'_{h,\tilde{y}}(\theta)|^2}|\phi''_{h,\tilde{y}}(\theta)| \,\mathrm{d}\theta h \,\mathrm{d}h \,\mathrm{d}\tilde{y} \\
\lesssim & \iiint (\mathbf{1}_{B(0,1)})_a(\tilde{x}-\tilde{y}, x_{d-1} - r\cos(\theta), x_d - r\sin(\theta), x_{d+1} - h) \\
& \qquad 2^{-ds}\big(\frac{1}{ab|N|} + \frac{1}{b^2|N|}\big)\mathbf{1}_{2^{s-3}<|y|<2^s} \,\mathrm{d}\theta \, h \,\mathrm{d}h \,\mathrm{d}\tilde{y} \\
\lesssim & \int_{2^{s-3}<|y|<2^s} 2^{-ds}\frac{1}{ab|N|}(\mathbf{1}_{B(0,1)})_a(x'-y, x_{d+1} - |y|^2) \,\mathrm{d}y \\
\lesssim & \frac{1}{ab|N|}\chi_{s,a}\,.
\end{aligned}
$$

Now we assume that $V = \{0\}^d \times \mathbb{R}$. Then we have $N' = 0$ and hence

$$
\begin{aligned}
\int \psi_a(x-z)e^{iN\cdot z}\,\mathrm{d}\mu_s(z) &= \int \psi_a(x'-y, x_{d+1} - |y|^2)e^{iN_{d+1}|y|^2}K_s(y)\,\mathrm{d}y \\
&= \int \Gamma(t)e^{iN_{d+1}t}\,\mathrm{d}t\,, \quad (1.7.13)
\end{aligned}
$$

where this time we have set

$$
\Gamma(t) = \int \psi_a(x'-y, x_{d+1}-t)K_s(y)\delta(t-|y|^2)\,\mathrm{d}y\,.
$$

Using (1.2.1) and that $s \ge -1$, one finds that $|\Gamma'(t)| \lesssim a^{-d-2}2^{-ds}$. Integrating by parts in (1.7.13) gives the desired estimate (1.7.8) after a similar, but simpler computation as in the case $V = \mathbb{R}^{d-1}\times\{0\}^2$. □

Lemma 1.7.4. *There exists $C > 0$ such that the following holds. Let $s \in \mathbb{Z}$, $a > 0$ and $b > 2^{-s}$. Let $\psi, \varphi \in \Psi$ with $\hat{\psi}(-aN) = 0$. Define*

$$
\varphi_{b,s}(x) = (2^s b)^{-d-2}\varphi(2^{-s}b^{-1}x', 2^{-2s}b^{-2}x_{d+1})
$$

and $\mu_s^l = \mu_s * \varphi_{b,s}$. Then we have that

$$
|\psi_a * (e^{iN\cdot y}\mu_s^l(y))| \le Ca2^{-s}b^{-(d+3)}2^{-(d+2)s}\mathbf{1}_{[-2^s,2^s]^d\times[-2^{2s},2^{2s}]}(y)\,. \quad (1.7.14)
$$

Proof. We have that

$$
\begin{aligned}
|\psi_a * e^{iN\cdot y}\mu_s^l(y)| &\lesssim \int |\hat{\psi}(a\xi - aN)||\hat{\varphi}_{b,s}(\xi)|\,\mathrm{d}\xi \\
&\lesssim a\int |\xi||\hat{\varphi}_{b,s}(\xi)|\,\mathrm{d}\xi \\
&= a2^{-(d+3)s}b^{-(d+3)}\int |(\xi', 2^{-s}b^{-1}\xi_{d+1})||\hat{\varphi}(\xi)|\,\mathrm{d}\xi\,.
\end{aligned}
$$

The lemma now follows since $2^{-s}b^{-1} < 1$ and $\||\xi|\hat{\varphi}(\xi)\|_1 \lesssim 1$ for $\psi \in \Psi$, and since the left hand side of (1.7.14) is clearly supported in $[-2^s, 2^s]^d \times [-2^{2s}, 2^{2s}]$. □

1.7.3 Main almost orthogonality estimate for separated trees

Lemma 1.7.5. *There exists $C > 0$ such that the following holds. Let $\mathbf{T}_1, \mathbf{T}_2$ be a pair of Δ-separated, normal trees. Then*

$$\left| \int T^*_{\mathbf{T}_1} g_1 \overline{T^*_{\mathbf{T}_2} g_2} \right| \le C \Delta^{-\frac{1}{10d}} \|W_{\mathbf{T}_1} g_1\|_{L^2(I(\mathbf{T}_1)\cap I(\mathbf{T}_2))} \|W_{\mathbf{T}_2} g_2\|_{L^2(I(\mathbf{T}_1)\cap I(\mathbf{T}_2))}$$

for certain operators $W_{\mathbf{T}_i}$ depending only on $\mathbf{T}_i$ with

$$\|W_{\mathbf{T}_i}\|_{2\to 2} \le 1 .$$

Proof. Since the trees are normal, the left hand side vanishes if $I(\mathbf{T}_1) \cap I(\mathbf{T}_2) = \emptyset$. Thus we may assume without loss of generality that $I := I(\mathbf{T}_2) \subset I(\mathbf{T}_1)$. Note also that we can always assume that Δ is sufficiently large, by adding $|T^*_{\mathbf{T}_i}|$ to $W_{\mathbf{T}_i}$. Finally, we assume that the minimal scale of a tile in $\mathbf{T}_2$ is 0, by scaling.

We fix $\gamma := \frac{2(d+1)}{(2d+1)(2d+7)} > \frac{2}{10d} > 0$ and $\varphi \in \Psi$ and define $\mu_s^l := \mu_s * \varphi_{\Delta^{-\gamma},s}$, where

$$\varphi_{\Delta^{-\gamma},s}(x) = 2^{-(d+2)s} \Delta^{\gamma(d+2)} \varphi(2^{-s}\Delta^{\gamma} x', 2^{-2s}\Delta^{2\gamma} x_{d+1}) .$$

Then we decompose $T_{\mathbf{T}_1}$ into a large scales part, a smoothened small scales part and an error term:

$$\begin{aligned} T_{\mathbf{T}_1} f = & \sum_{\substack{s\in\sigma(x)\\ s>\overline{\sigma}(x)-\log_2(\Delta)}} \int f(x-y) e^{iN(x)\cdot y} \,\mathrm{d}\mu_s(y) \\ & + \sum_{\substack{s\in\sigma(x)\\ s\le\overline{\sigma}(x)-\log_2(\Delta)}} \int f(x-y) e^{iN(\mathbf{T}_1)\cdot y} \,\mathrm{d}\mu_s^l(y) \\ & + \mathcal{E}_{\mathbf{T}_1}(f) \\ =: & \; T^1_{\mathbf{T}_1} f + T^2_{\mathbf{T}_1} f + \mathcal{E}_{\mathbf{T}_1}(f) . \end{aligned}$$

Notice that μ_s, λ_s are supported in $\delta_s(B(0,1/3))$, so for Δ sufficiently large the measures μ_s^l and λ_s^l also satisfy conditions i) to v). Thus both $T^1_{\mathbf{T}_1}$ and $T^2_{\mathbf{T}_1}$ are generalized tree operators.

We first show that the error term $\mathcal{E}_{\mathbf{T}_1}$ is small. We have

$$\begin{aligned}|\mathcal{E}_{\mathbf{T}_1}(f)| \le & \left| \sum_{\substack{s\in\sigma(x)\\ s\le\overline{\sigma}(x)-\log_2(\Delta)}} \int f(x-y)(e^{iN(x)\cdot y} - e^{iN(\mathbf{T}_1)\cdot y})\,\mathrm{d}\mu_s(y)\right| \\ & + \left| \sum_{\substack{s\in\sigma(x)\\ s\le\overline{\sigma}(x)-\log_2(\Delta)}} \int f(x-y)e^{iN(\mathbf{T}_1)\cdot y}\,\mathrm{d}(\mu_s-\mu_s^l)(y)\right| \\ \lesssim & \sum_{\substack{s\in\sigma(x)\\ s\le\overline{\sigma}(x)-\log_2(\Delta)}} 2^{s-\overline{\sigma}(x)} \int |f(x-y)|\,\mathrm{d}|\mu_s|(y) \\ & + \sup_{\underline{s}<\overline{s}} \left| \sum_{\substack{\underline{s}<s\le\overline{s}\\ s\in J}} (\mu_s-\mu_s^l) * (f(y)e^{iN(\mathbf{T}_1)\cdot y})\right| .\end{aligned}$$

Here $J = J(\mathbf{T}_1)$ is as in Subsection 1.6.2. The first term is bounded by $\Delta^{-1}M^\lambda f(x)$. The second term is the maximally truncated singular integral associated to the single scale operators $\mathbf{1}_J(s)(\mu_s-\mu_s^l)$, applied to $e^{iN(\mathbf{T}_1)\cdot y}f(y)$. Since

$$|\hat{\mu}_s(\xi) - \hat{\mu}_s^l(\xi)| \le \min\{|\xi|^{-1}, |\Delta^{-\gamma}\xi|\}\,,$$

this operator has norm $\lesssim \Delta^{-\gamma/2}$ on L^2. Thus we have

$$\|\mathcal{E}_{\mathbf{T}_1}(f)\|_2 \lesssim \Delta^{-\gamma/2}\|f\|_2 \le \Delta^{-\frac{1}{10d}}\|f\|_2\,.$$

It now remains to bound $\langle T^{1*}_{\mathbf{T}_1}g_1, T^*_{\mathbf{T}_2}g_2\rangle + \langle T^{2*}_{\mathbf{T}_1}g_1, T^*_{\mathbf{T}_2}g_2\rangle$. We decompose $\mathbf{T}_1 = \mathbf{T}_1' \cup \mathbf{T}_1'' \cup \mathbf{T}_1'''$, where

$$\mathbf{T}_1' = \{p \in \mathbf{T}_1 \,:\, s(p) < -1\,, 3I(p) \subset I\}\,,$$
$$\mathbf{T}_1'' = \{p \in \mathbf{T}_1 \,:\, s(p) \ge -1,\, 3I(p)\cap I \ne \emptyset\}$$

and $\mathbf{T}_1''' = \mathbf{T}_1 \setminus (\mathbf{T}_1' \cup \mathbf{T}_1'')$. Then $\mathbf{T}_1'$ is a normal tree with top I. For all $p \in \mathbf{T}_1''', p' \in \mathbf{T}_2$ we have $2I(p) \cap (2+\frac{1}{4})I(p') = \emptyset$, this follows from normality of $\mathbf{T}_2$ and the fact that all tiles in $\mathbf{T}_1'''$ with $3I(p)\cap I \ne \emptyset$ have scale at most $s(p')-2$. Thus $\langle T^*_{\mathbf{T}_1'''}g_1, T^*_{\mathbf{T}_2}g_2\rangle = 0$.

According to the decomposition of $\mathbf{T}_1$ we split $T^1_{\mathbf{T}_1} = T^1_{\mathbf{T}_1'} + T^1_{\mathbf{T}_1''} + T^1_{\mathbf{T}_1'''}$, and similarly for $T^2_{\mathbf{T}_1}$. We still have $\langle T^{1*}_{\mathbf{T}_1'''}g_1, T^*_{\mathbf{T}_2}g_2\rangle = 0$, thus

$$\langle T^{1*}_{\mathbf{T}_1}g_1, T^*_{\mathbf{T}_2}g_2\rangle = \langle T^{1*}_{\mathbf{T}_1'}g_1, T^*_{\mathbf{T}_2}g_2\rangle + \langle T^{1*}_{\mathbf{T}_1''}g_1, T^*_{\mathbf{T}_2}g_2\rangle\,. \tag{1.7.15}$$

For the first summand in (1.7.15) we obtain with Corollary 1.7.2

$$|\langle T^{1*}_{\mathbf{T}_1'}g_1, T^*_{\mathbf{T}_2}g_2\rangle| \lesssim \Delta^{-\frac{1}{10d}}\|g_1\|_{L^2(I)}\|g_2\|_{L^2(I)}\,. \tag{1.7.16}$$

For the second summand in (1.7.15) we put $\nu = \frac{d+1}{2d+1}$ and choose $\Pi_2 = \Pi_{\Delta^\nu, N(\mathbf{T}_2)}$ for some function $\psi \in \Psi$. If Δ is sufficiently large then $\Pi_2 T^*_{p'} g_2$ is supported in $(2+\frac{1}{4})I(p')$ for each tile $p' \in \mathbf{T}_2$, thus we also have $\langle T^{1*}_{\mathbf{T}_1''} g_1, \Pi_2 T^*_{\mathbf{T}_2} g_2\rangle = 0$. Using this, we expand

$$\langle T^{1*}_{\mathbf{T}_1''} g_1, T^*_{\mathbf{T}_2} g_2\rangle = \langle g_1, T^{1}_{\mathbf{T}_1''}(1-\Pi_2) T^*_{\mathbf{T}_2} g_2\rangle + \langle T^{1*}_{\mathbf{T}_1} g_1 - T^{1*}_{\mathbf{T}_1'} g_1, \Pi_2 T^*_{\mathbf{T}_2} g_2\rangle\,. \tag{1.7.17}$$

The second term in (1.7.17) is by Cauchy-Schwarz, Lemma 1.7.1 and L^2-boundedness of $T^{1*}_{\mathbf{T}_1'} = T^{1*}_{\mathbf{T}_1'} \mathbf{1}_{I(\mathbf{T}_2)}$ bounded by

$$\Delta^{-\frac{1}{10d}} \| |T^{1*}_{\mathbf{T}_1} g_1| + |g_1| \|_{L^2(I)} \|g_2\|_{L^2(I)}\,. \tag{1.7.18}$$

For the first term in (1.7.17) we may assume that $N(\mathbf{T}_2) = 0$. Then we have

$$T^1_{\mathbf{T}_1''}(1-\Pi_2) f(x) = \sum_{\substack{s \in \sigma'(x) \\ s > \overline{\sigma}(x) - \log_2(\Delta)}} (e^{iN(x)\cdot y} \mu_s(y)) * \psi_{\Delta^{-\nu}} * f(x)\,,$$

where $\sigma'(x) = \{s(p) : x \in E(p), p \in \mathbf{T}_1''\}$. Using the estimate for the convolution kernels proven in Lemma 1.7.3 with $a = \Delta^{-\nu}$, $b = \Delta^{-1/2}$, we obtain with the notation defined at (1.7.5)

$$\begin{aligned} &|(e^{iN(x)\cdot y} \mu_s(y)) * \psi_{\Delta^{-\nu}} * f(x)| \\ &\qquad\qquad \lesssim (\Delta^{-1/2} \chi_{s,\Delta^{-\nu},\Delta^{-1/2}} + \Delta^{\nu - 1/2} \chi_{s,\Delta^{-\nu}}) * |f|(x)\,. \end{aligned}$$

Passing the convolutions to the other side in the inner product, we find that the first term in (1.7.17) is bounded by

$$\begin{aligned} &\log(\Delta) \Delta^{\nu-1/2} \langle M^{1,\chi} g_1 + M^{2,\chi} g_1, |T^*_{\mathbf{T}_2} g_2|\rangle \\ &\qquad\qquad \lesssim \Delta^{-\frac{1}{10d}} \|M^{1,\chi} g_1 + M^{2,\chi} g_1\|_{L^2(I)} \|g_2\|_{L^2(I)} \end{aligned}$$

where $M^{1,\chi}$, $M^{2,\chi}$ are defined in (1.7.6) and (1.7.7). Since $M^{1,\chi}$ and $M^{2,\chi}$ are bounded on L^2, this gives the desired estimate for $T^1_{\mathbf{T}_1}$.

Now we turn to $\langle T^{2*}_{\mathbf{T}_1} g_1, T^*_{\mathbf{T}_2} g_2\rangle$. As above, we have

$$\langle T^{2*}_{\mathbf{T}_1} g_1, T^*_{\mathbf{T}_2} g_2\rangle = \langle T^{2*}_{\mathbf{T}_1'} g_1, T^*_{\mathbf{T}_2} g_2\rangle + \langle T^{2*}_{\mathbf{T}_1''} g_1, T^*_{\mathbf{T}_2} g_2\rangle\,. \tag{1.7.19}$$

For the first term in (1.7.19) we have again from Corollary 1.7.2 that

$$|\langle T^{2*}_{\mathbf{T}_1'} g_1, T^*_{\mathbf{T}_2} g_2\rangle| \lesssim \Delta^{-\frac{1}{10d}} \|g_1\|_{L^2(I)} \|g_2\|_{L^2(I)}\,. \tag{1.7.20}$$

For the second term in (1.7.19) we split as before

$$\langle T^{2*}_{\mathbf{T}_1''} g_1, T^*_{\mathbf{T}_2} g_2\rangle = \langle g_1, T^{2}_{\mathbf{T}_1''}(1-\Pi_2) T^*_{\mathbf{T}_2} g_2\rangle + \langle T^{2*}_{\mathbf{T}_1} g_1 - T^{2*}_{\mathbf{T}_1'} g_1, \Pi_2 T^*_{\mathbf{T}_2} g_2\rangle\,. \tag{1.7.21}$$

The second term in (1.7.21) is by Cauchy-Schwarz and Lemma 1.7.1 bounded by

$$\Delta^{-\frac{1}{10d}} \| |T^{2*}_{\mathbf{T}_1} g_1| + |g_1| \|_{L^2(I)} \|g_2\|_{L^2(I)}\,. \tag{1.7.22}$$

For the first term in (1.7.21), we assume again that $N(\mathbf{T}_2) = 0$. Then

$$T^2_{\mathbf{T}_1''}(1-\Pi_2)f(x) = \sum_{\substack{s\in\sigma'(x)\\ s\leq\overline{\sigma}(x)-\log_2(\Delta)}} \psi_{\Delta^{-\nu}} * (e^{iN(\mathbf{T}_1)\cdot y}\mu_s^l(y)) * f(x)\,.$$

By our assumptions on Ψ, we may then further assume that we have $\hat{\psi}(\Delta^{-\nu}N(\mathbf{T}_1)) = 0$, since for Δ sufficiently large $|\Delta^{-\nu}N(\mathbf{T}_1)| \gtrsim \Delta^{1-\nu} \geq 1$. By Lemma 1.7.4 with $a = \Delta^{-\nu}$, $b = \Delta^{-\gamma}$ we have

$$\sum_{s>\gamma\log_2(\Delta)} |\psi_{\Delta^{-\nu}} * (e^{iN(\mathbf{T}_1)\cdot y}\mu_s^l(y)) * f(x)| \lesssim \Delta^{(d+3)\gamma-\nu}\Phi * |f|(x)\,,$$

where $\Phi(y) = \sum_{s\geq 0} 2^{-s(d+3)}\mathbf{1}_{[-2^s,2^s]^d\times[-2^{2s},2^{2s}]}(y)$. By Lemma 1.7.3 with $a = \Delta^{-\nu}$, $b = \Delta^{-1/2}$ we have for $-1 \leq s \leq \gamma\log(\Delta)$ that

$$\begin{aligned}|\psi_{\Delta^{-\nu}} * (e^{iN(\mathbf{T}_1\cdot y}\mu_s^l(y)) * f(x)|&\\ &\lesssim (\Delta^{-1/2}\chi_{s,\Delta^{-\gamma},\Delta^{-1/2}} + \Delta^{\nu-1/2}\chi_{s,\Delta^{-\gamma}}) * M|f|(x)\,.\end{aligned}$$

Similarly as for $T^1_{\mathbf{T}_1}$, we obtain from these estimates that the first term in (1.7.21) is bounded by

$$\Delta^{-\frac{1}{10d}}\|Mg_1 + M^{1,\chi}g_1 + M^{2,\chi}g_1\|_{L^2(I)}\|g_2\|_{L^2(I)}\,.$$

This completes the proof. $\square$

1.7.4 Completing the argument for forests

Here we use Lemma 1.7.5 to complete the proof of Proposition 1.3.3. We follow the presentation in [116].

A row is a union of normal trees with pairwise disjoint top cubes.

Lemma 1.7.6. *There exists $C > 0$ such that the following holds. Let $\mathbf{R}_1, \mathbf{R}_2$ be two rows such that the trees in $\mathbf{R}_1$ are Δ-separated from the trees in $\mathbf{R}_2$. Then*

$$\left|\int T^*_{\mathbf{R}_1}g_1\overline{T^*_{\mathbf{R}_2}g_2}\right| \leq C\Delta^{-\frac{1}{10d}}\|g_1\|_2\|g_2\|_2\,.$$

Proof. We have by Lemma 1.7.5, with $W_{\mathbf{T}_i}$ as defined there:

$$\begin{aligned}\left|\int T^*_{\mathbf{R}_1}g_1\overline{T^*_{\mathbf{R}_2}g_2}\right| &\leq \sum_{\mathbf{T}_1\in\mathbf{R}_1}\sum_{\mathbf{T}_2\in\mathbf{R}_2}\left|\int T^*_{\mathbf{T}_1}\mathbf{1}_{I(\mathbf{T}_1)}g_1\overline{T^*_{\mathbf{T}_2}\mathbf{1}_{I(\mathbf{T}_2)}g_2}\right|\\ &\leq C\Delta^{-\frac{1}{10d}}\sum_{\mathbf{T}_1\in\mathbf{R}_1}\sum_{\mathbf{T}_2\in\mathbf{R}_2}\prod_{i=1,2}\|W_{\mathbf{T}_i}\mathbf{1}_{I(\mathbf{T}_i)}g_i\|_{L^2(I(\mathbf{T}_1)\cap I(\mathbf{T}_2))}\,.\end{aligned}$$

Using Cauchy-Schwarz, disjointness of the cubes $I(\mathbf{T}_i)$ for $\mathbf{T}_i \in \mathbf{R}_i$ and $\|W_{\mathbf{T}_i}\|_{2\to 2} \lesssim 1$, we estimate:

$$
\begin{aligned}
&\le C\Delta^{-\frac{1}{10d}} \prod_{i=1,2} \left(\sum_{\mathbf{T}_1 \in \mathbf{R}_2} \sum_{\mathbf{T}_2 \in \mathbf{R}_2} \|W_{\mathbf{T_i}} \mathbf{1}_{I(\mathbf{T}_i)} g_i\|^2_{L^2(I(\mathbf{T}_1)\cap I(\mathbf{T}_2))} \right)^{1/2} \\
&\le C\Delta^{-\frac{1}{10d}} \prod_{i=1,2} \left(\sum_{\mathbf{T}_i \in \mathbf{R}_i} \|W_{\mathbf{T_i}} \mathbf{1}_{I(\mathbf{T}_i)} g_i\|^2_{L^2} \right)^{1/2} \\
&\le C\Delta^{-\frac{1}{10d}} \|g_1\|_2 \|g_2\|_2 \,. \qquad \square
\end{aligned}
$$

Proof of Proposition 1.3.3. Let $\mathbf{F}$ be an n-forest. By the overlap estimate (1.3.3), we can decompose $\mathbf{F}$ into $2^n \log(n+2)$ rows $\mathbf{R}_i$. Since $\mathbf{F}$ is an n-forest, the trees in different rows are 2^{10dn} separated and have density at most 2^{-n}. Note also that separation of the rows implies that the sets $E(\mathbf{R}_i) := \cup_{p\in\mathbf{R}_i} E(p)$ are pairwise disjoint.

By orthogonality and Corollary 1.6.4, we have that

$$
\|T_{\mathbf{R}_i}\|_{2\to 2} \lesssim_\varepsilon 2^{-\epsilon n}
$$

for each i and each $\varepsilon < \frac{1}{2} - \frac{1}{2(d+1)}$. Using this and Lemma 1.7.6 yields

$$
\begin{aligned}
\|T^*_{\mathbf{F}} g\|_2^2 \le \sum_{i,j} &\left| \int T^*_{\mathbf{R}_i} \mathbf{1}_{E(\mathbf{R}_i)} g T^*_{\mathbf{R}_j} \mathbf{1}_{E(\mathbf{R}_j)} g \right| \\
&\lesssim (2^{-2\varepsilon n} + 2^n \log(n+2) 2^{-\frac{1}{10d}10dn}) \sum_i \|\mathbf{1}_{E(\mathbf{R}_i)} g\|_2^2 \\
&\lesssim 2^{-2\varepsilon n} \|g\|_2^2 \,. \qquad (1.7.23)
\end{aligned}
$$

This completes the proof. $\square$

1.8 L^p-bounds: Proof of Theorem 1.1.1

Theorem 1.1.1 follows from interpolation, Proposition 1.3.1 and the following L^p-estimates for antichain and forest operators replacing Proposition 1.3.2 and Proposition 1.3.3.

Proposition 1.8.1. *For each $1 < p < \infty$ there exists $\varepsilon = \varepsilon(d,p) > 0$ and $C > 0$ such that for every antichain $\mathbf{A}$ of density δ*

$$
\|T_{\mathbf{A}}\|_{p\to p} \le C\delta^{-\varepsilon} \,.
$$

Proof. This follows from interpolation between Proposition 1.3.2 and the trivial bound $\|T_{\mathbf{A}}\|_{p\to p} \le C(p)$, which holds since $T_{\mathbf{A}}$ is dominated pointwise by the maximal average along the paraboloid. $\square$

Proposition 1.8.2. *Let p be such that*

$$
\frac{d^2+4d+2}{(d+1)^2} < p < 2(d+1) \,. \qquad (1.8.1)
$$

There exists $C > 0$ such that the following holds. Let for each $n \geq 0$ and $1 \leq l \leq (n+1)^2$ an n-forest $\mathbf{F}_{n,l}$ of normal trees be given. Then

$$\Big\| \sum_{n \geq 0} \sum_{l=1}^{(n+1)^2} T_{\mathbf{F}_{n,l}} \Big\|_{p \to p} \leq C \,. \tag{1.8.2}$$

Proof. Suppose first that $p < 2$. We will show that for all $0 \leq \alpha < d^2/(2(d^2+4d+2))$ there exists $\varepsilon > 0$ such that for each $G \subset \mathbb{R}^{d+1}$ with $0 < |G| < \infty$ it holds with $\tilde{G} = \{x : M\mathbf{1}_G > t\}$ that

$$\|\mathbf{1}_{\mathbb{R}^{d+1} \setminus \tilde{G}} T_{\mathbf{F}_{n,l}} \mathbf{1}_G\|_{2 \to 2} \lesssim t^\alpha 2^{-\varepsilon n} \,. \tag{1.8.3}$$

Using Bateman's extrapolation argument from [4] (see also [116], Appendix B), (1.8.3) implies that for $p < 2$ satisfying (1.8.1)

$$\|T_{\mathbf{F}_{n,l}}\|_{L^{p,1} \to L^{p,\infty}} \lesssim 2^{-\varepsilon n} \,.$$

Then (1.8.2) follows by interpolation and summation in n.

To prove (1.8.3), recall that by Lemma 1.6.2 for each tree $\mathbf{T}$, all $p > 1 + \frac{1}{d+1}$ and all f, g, there exists a sparse collection $\mathcal{S}$ with

$$|\langle \mathbf{1}_{\mathbb{R}^{d+1} \setminus \tilde{G}} T_{*\mathbf{T}} \mathbf{1}_G f, g \rangle| \lesssim \sum_{Q \in \mathcal{S}} |Q| \langle \mathbf{1}_G f \rangle_{p,Q} \langle \mathbf{1}_{(\mathbb{R}^{d+1} \setminus \tilde{G}) \cap E(Q)} g \rangle_{p,3Q} \,. \tag{1.8.4}$$

If $3Q \cap (\mathbb{R}^{d+1} \setminus \tilde{G}) \neq \emptyset$, then $|Q \cap G| \lesssim t|Q|$. Using this and Hölder's inequality, we obtain with a similar argument as in the proof of Corollary 1.6.3 that for each $\varepsilon < \frac{1}{2} - \frac{1}{d+2}$ and each tree $\mathbf{T}$ of density at most 2^{-n}

$$\|\mathbf{1}_{\mathbb{R}^{d+1} \setminus \tilde{G}} T_{*\mathbf{T}} \mathbf{1}_G\|_{2 \to 2} \lesssim_\varepsilon t^\varepsilon 2^{-\varepsilon n} \,.$$

By the same argument, the same holds for $M_{\mathbf{T}}$ and hence for $T_{\mathbf{T}}$. By orthogonality, the same estimate holds for row operators $T_{\mathbf{R}}$. Finally, each forest $\mathbf{F}_{n,l}$ can be decomposed into at most $2^n \log(n+2)$ rows, so

$$\|\mathbf{1}_{\mathbb{R}^{d+1} \setminus \tilde{G}} T_{\mathbf{F}_{n,l}} \mathbf{1}_G\|_{2 \to 2} \leq 2^{n/2 - \varepsilon n} \log(n+2)^{1/2} t^\varepsilon \,.$$

We obtain (1.8.3) by taking a geometric average of this estimate and (1.7.23).

For the case $p > 2$ we use that for all $0 \leq \alpha < \frac{1}{2} - \frac{1}{2(d+1)}$, there exists $C > 0$ such that for each $G \subset \mathbb{R}^{d+1}$ with $0 < |G| < \infty$

$$\|\mathbf{1}_G \sum_{n \geq 0} \sum_{l=1}^{(n+1)^2} T_{\mathbf{F}_{n,l}} \mathbf{1}_{\mathbb{R}^{d+1} \setminus \tilde{G}}\|_{2 \to 2} \leq C t^\alpha \,. \tag{1.8.5}$$

Indeed, if we replace the sets $E(p)$ by $E(p) \setminus \tilde{G}$, then each tile has density $\lesssim t$, so (1.8.5) follows immediatly from Proposition 1.3.3. Using Bateman's extrapolation argument for the adjoint operator and interpolation, we obtain (1.8.2) for all $p > 2$ satisfying (1.8.1). $\square$

Chapter 2

On trilinear singular Brascamp-Lieb integrals

This chapter consists of a joint article [8] with Polona Durcik and Fred Yu-Hsiang Lin.

2.1 Introduction

This article continues the investigation of generalizations of the bilinear Hilbert transform

$$\mathrm{BHT}(f_1, f_2)(x) = \int_{\mathbb{R}} f_1(x+t) f_2(x+\alpha t) \frac{1}{t} dt\,, \qquad \alpha \neq 0, 1\,, \tag{2.1.1}$$

where the integral is understood as the principal value. Lacey and Thiele in their breakthrough works [76, 78] proved boundedness of BHT from $L^{p_1} \times L^{p_2}$ into L^{p_0}, provided that $\frac{1}{p_0} = \frac{1}{p_1} + \frac{1}{p_2} = 1$ and $p_0 > \frac{2}{3}$. This partially resolved a conjecture of Calderón [69].

It is then very natural to ask about higher dimensional versions of (2.1.1), namely the operators

$$\mathrm{BHT}_d(f_1, f_2)(x) = \int_{\mathbb{R}^d} f_1(x + A_1 t) f_2(x + A_2 t) K(t) dt\,, \tag{2.1.2}$$

where $A_1, A_2 : \mathbb{R}^d \to \mathbb{R}^d$ are linear maps and K is a Calderón-Zygmund kernel on $\mathbb{R}^d$, defined below in (2.1.3). Demeter and Thiele [39] studied the two dimensional case $d = 2$ of (2.1.2). The class of such operators is richer than in the one dimensional case, in that various levels of degeneracies occur depending on A_1 and A_2. Demeter and Thiele found four qualitatively different cases, and prove boundedness for three of them using different tools. The final case was later resolved by Kovač [71], using again different techniques.

In the present paper we extend this classification to the d-dimensional case and in fact to more general singular Brascamp-Lieb forms, in Theorem 2.1.15. We require some definitions, which are set up in Sections 2.1.1 to 2.1.3. We use our classification to fully characterize boundedness at exponents p_1, p_2, p_3 that do *not* satisfy the Hölder relation $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$, in Theorem 2.1.16. In Section 2.1.4 we further give three conditional bounds, Theorem 2.1.17, Theorem 2.1.18 and Theorem 2.1.19. They indicate how the difficulty of algebraically related cases in the classification compares. We put our classification into context and discuss which cases are covered by the existing literature in Section 2.1.5. Finally, we give new bounds for a large class of cases with Hölder exponents in Theorem 2.1.23.

2.1.1 Singular Brascamp-Lieb forms

By duality, bounds for the bilinear operators (2.1.2) are equivalent to bounds for the trilinear forms

$$\int_{\mathbb{R}^{2d}} f_1(x + A_1 t) f_2(x + A_2 t) f_3(x) K(t) dt\, dx\,.$$

Motivated by this, we make the following general definitions.

Definition 2.1.1. An l-Calderón-Zygmund kernel is a tempered distribution K on a Hilbert space H, such that K agrees with a function away from 0 and such that for any choice of orthonormal basis, the corresponding partial derivatives of the Fourier transform $\widehat{K}$ satisfy for all $\xi \neq 0$

$$|\partial^\alpha \widehat{K}(\xi)| \le |\xi|^{-|\alpha|}\,, \quad |\alpha| \le l\,. \tag{2.1.3}$$

Here the Fourier transform of a Schwartz function is defined by

$$\widehat{f}(\xi) = \int e^{-2\pi i \xi \cdot x} f(x)\, dx\,,$$

and this definition is extended to tempered distributions by density.

Definition 2.1.2. We define a (trilinear) *singular Brascamp-Lieb datum* to be a tuple $\mathbf{H} = (H; H_0, H_1, H_2, H_3; \Pi_0, \Pi_1, \Pi_2, \Pi_3)$ of five finite dimensional Hilbert spaces H, H_i and of four surjective linear maps $\Pi_i : H \to H_i$.

Definition 2.1.3. Given a singular Brascamp-Lieb datum $\mathbf{H}$ and a Calderón-Zygumnd kernel K on H_0, the associated *singular Brascamp-Lieb form* $\Lambda_{\mathbf{H}}$ is the trilinear form defined a priori on Schwartz functions $f_i \in \mathcal{S}(H_i)$ by

$$\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3) = \int_H f_1(\Pi_1 x) f_2(\Pi_2 x) f_3(\Pi_3 x) K(\Pi_0 x)\, dx\,. \tag{2.1.4}$$

Our goal is to study Lebesgue space estimates

$$|\Lambda(K, f_1, f_2, f_3)| \le C(l) \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3} \tag{2.1.5}$$

for singular Brascamp-Lieb forms and exponents $\mathbf{p} = (p_1, p_2, p_3)$. This motivates the following definition.

Definition 2.1.4. We say that a form $\Lambda_{\mathbf{H}}$ and the datum $\mathbf{H}$ are $\mathbf{p}$-*bounded* if there exists l such that (2.1.5) holds for all f_1, f_2, f_3 and all l-Calderón-Zygmund kernels K. We say that it is is of *Hölder type* if it is $\mathbf{p}$-bounded for some $1 < p_1, p_2, p_3 < \infty$ with $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$. We will abbreviate $a < p_1, p_2, p_3 < b$ by $a < \mathbf{p} < b$.

The methods used in previous literature to prove or disprove bounds (2.1.5) vary substantially depending on $\mathbf{H}$. This shows in the very different methods used in [39] and [71], and also in the analysis in [39] for different $\mathbf{H}$. The following notion of equivalence is relevant for deciding boundedness of a singular Brascamp-Lieb form, as expressed in Lemma 2.1.6.

Definition 2.1.5. We call two singular Brascamp-Lieb data $\mathbf{H}$, $\mathbf{H}'$ equivalent if there exist invertible linear maps

$$\varphi : H \to H' , \quad \varphi_i : H_i \to H_i', \quad i = 0, 1, 2, 3 , \tag{2.1.6}$$

such that

$$\Pi_i' \circ \varphi = \varphi_i \circ \Pi_i , \quad i = 0, 1, 2, 3 . \tag{2.1.7}$$

Lemma 2.1.6. *Suppose that* $\mathbf{H}$ *and* $\mathbf{H}'$ *are equivalent singular Brascamp-Lieb data. Then for all* $\mathbf{p}$*, the form* $\Lambda_{\mathbf{H}}$ *is* $\mathbf{p}$*-bounded if and only* $\Lambda_{\mathbf{H}'}$ *is* $\mathbf{p}$*-bounded.*

Lemma 2.1.6 is a direct consequence of changes of variables in the functions and the integral defining the singular Brascamp-Lieb form.

Our goal is to classify $\mathbf{p}$-bounded singular Brascamp-Lieb forms up to equivalence. Note that the notions of $\mathbf{p}$-boundedness and equivalence of data are insensitive to the Hilbert space structures on the spaces in $\mathbf{H}, \mathbf{H}'$. Hence, only the underlying vector spaces and linear maps will be relevant for our classification. However, to make sense of (2.1.3) and (2.1.5), we need Lebesgue measures on the spaces H, H_i, and a norm on H_0^*. The H, H_i are defined to be Hilbert spaces to simplify the exposition, because Hilbert spaces canonically have this additional structure. (The same choice is made in [13], for similar reasons.)

Remark 2.1.7. To study quantitative estimates, that is, the size of the constant C in (2.1.5), one needs a finer equivalence relation than the one given by (2.1.6), (2.1.7). Namely one should assume that φ_0 is a scalar multiple of an orthogonal transformation, $\varphi_0 \in \mathbb{R} \cdot O(H_0, H_0')$. This is because only scalar multiples of isometries preserve all quantitative assumptions on the Calderón-Zygmund kernels. Equivalence classes modulo this finer equivalence relation are parametrized by equivalence classes according to Definition 2.1.5 together with an element of $Gl(H_0')/(\mathbb{R} \cdot O(H_0, H_0'))$. The latter can be parametrized by nonsingular lower triangular matrices with a 1 in the upper left corner.

2.1.2 The four subspace problem

The classification of Brascamp-Lieb data up to equivalence is equivalent to the so-called four subspace problem, which we now describe.

Definition 2.1.8. A *module* is a tuple $\mathbf{M} = (M; M_0, M_1, M_2, M_3)$ of a finite dimensional vector space M and four subspaces $M_i \subseteq M$, $i = 0, 1, 2, 3$.

Structures $\mathbf{M}$ are also called representations (of the four subspace quiver). We call them modules, because they are modules over the path algebra associated with that quiver. The interested reader is refered to [40] for a short survey on quiver representations.

Definition 2.1.9. Two modules $\mathbf{M}$ and $\mathbf{M}'$ are *isomorphic* if there exists an invertible linear map $\psi : M \to M'$ such that

$$\psi(M_i) = M_i' , \quad i = 0, 1, 2, 3 .$$

If $\mathbf{M}$ is isomorphic to $\mathbf{M}'$, we write $\mathbf{M} \cong \mathbf{M}'$.

The four subspace problem asks for a classification of all modules up to isomorphism. It was solved by Gelfand and Ponomarev [58] for algebraically closed fields. In the case of general fields (we are interested in $\mathbb{R}$), the solution was given by Nazarova [94, 95]. See also [87] for an elementary proof. The solution consists of a list of indecomposable modules, such that each module is isomorphic to a unique (up to permutation) finite direct sum of indecomposables.

Definition 2.1.10. The *direct sum* of two modules $\mathbf{M} = (M; M_0, M_1, M_2, M_3)$ and $\mathbf{M}' = (M'; M'_0, M'_1, M'_2, M'_3)$ is defined to be the module

$$\mathbf{M} \oplus \mathbf{M}' = (M \oplus M'; M_0 \oplus M'_0, M_1 \oplus M'_1, M_2 \oplus M'_2, M_3 \oplus M'_3)\,.$$

Theorem 2.1.11 (Gelfand, Ponomarev [58]; Nazarova [94, 95])**.** *Let* $\mathbf{M}$ *be a module. Then there exists a finite sequence of modules* $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_k$*, from the list in Table 2.1 (possibly after permuting the subspaces), such that*

$$\mathbf{M} \cong \mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_k\,.$$

For every such representation

$$\mathbf{M} \cong \mathbf{M}'_1 \oplus \cdots \oplus \mathbf{M}'_k\,,$$

there exists a permutation $\pi : \{1, \ldots, k\} \to \{1, \ldots, k\}$ *such that* $\mathbf{M}_i = \mathbf{M}'_{\pi(i)}$, $i = 1, \ldots, k$.

Remark 2.1.12. The indecomposable modules $\mathbf{M}_j$ are listed in Table 2.1 only up to permutation of the subspaces. We give the additional information which permutations give rise to isomorphic modules for each case in Lemma 2.8.1 in Appendix 2.8. This information will be used in the proof of Theorem 2.1.15.

We alert the reader that the same classification problem for more than four subspace, relevant to more-than-three linear forms, is significantly harder. More precisely, the four subspace problem is the last in this sequence which is *tame*. For tame classification problems, there exist for every dimension tuple of the involved subspaces only finitely many one parameter families of indecomposable modules, plus possibly finitely many additional indecomposable modules (at least if the underlying field is algebraically closed), see [35]. If a classification problem is not tame, then it is *wild*. One can show that every wild classification problem is at least as hard as the classification of finite dimensional modules up to isomorphism over *any* finitely generated algebra. For both of these facts, and further references, see [43]. In that sense wild classification problems are substantially more difficult.

Of course, we are not interested in all modules, but just in those corresponding to $\mathbf{p}$-bounded forms with $\mathbf{p} < \infty$. This imposes some restrictions, see Lemma 2.2.1 and Lemma 2.2.2. However, even with these additional restrictions, even if we additionally assume the Hölder condition $1/p_1 + 1/p_2 + 1/p_3 = 1$, the classification problem remains wild for four- and higher linear forms, see Remark 2.2.3.

2.1.3 Classification of singular Brascamp-Lieb forms

Taking adjoint gives rise to a natural correspondence between the underlying vector spaces of singular Brascamp-Lieb data and modules.

Definition 2.1.13. Let $\mathbf{H}$ be a singular Brascamp-Lieb datum. The associated module $\mathbf{M_H}$ is defined to be

$$\mathbf{M_H} = (H^*; \Pi_0^* H_0^*, \Pi_1^* H_1^*, \Pi_2^* H_2^*, \Pi_3^* H_3^*)\,.$$

Here * denotes adjoints and dual spaces. Conversely, if $\mathbf{M}$ is a module, then we associate to it a singular Brascamp-Lieb datum

$$\mathbf{H_M} = (M^*; M_0^*, M_1^*, M_2^*, M_3^*; \iota_0^*, \iota_1^*, \iota_2^*, \iota_3^*)\,.$$

Here ι_j denotes the inclusion map $\iota_j : M_j \to M$, and we equip the finite dimensional vector spaces M^*, M_j^* with any Hilbert space structure.

If we define morphisms of Brascamp-Lieb data to be tuples of (not necessarily invertible) linear maps φ, φ_i satisfying (2.1.6) and (2.1.7), then the maps $\mathbf{H} \mapsto \mathbf{M_H}$ and $\mathbf{M} \mapsto \mathbf{H_M}$ become mutually inverse dualities of categories. As a consequence of this fact and Theorem 2.1.11 we immediately obtain a classification of all singular Brascamp-Lieb data.

Theorem 2.1.14. *Singular Brascamp-Lieb data $\mathbf{H}$ and $\mathbf{H}'$ are equivalent if and only if $\mathbf{M_H}$ and $\mathbf{M_{H'}}$ are isomorphic. For each module $\mathbf{M}$, there exists a singular Brascamp-Lieb datum $\mathbf{H}$ with $\mathbf{M_H} \cong \mathbf{M}$. As a consequence, for every singular Brascamp-Lieb datum $\mathbf{H}$, there exists a finite list of modules $\mathbf{M}_1, \dots, \mathbf{M}_k$ from Table 2.1 such that*

$$\mathbf{M_H} \cong \mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_k\,. \tag{2.1.8}$$

This list is unique, up to permutation. Conversely, for every finite list $\mathbf{M}_1, \dots, \mathbf{M}_k$ there exists a unique up to equivalence singular Brascamp-Lieb datum $\mathbf{H}$ such that (2.1.8) *holds.*

Most of the singular Brascamp-Lieb forms as in (2.1.8) are not $\mathbf{p}$-bounded for any $\mathbf{p} < \infty$. We exclude the case where some $p_i = \infty$ to avoid certain cases where the maps Π_i are not surjective on the kernel of Π_0, which would complicate our analysis while offering little additional insight. We have the following classification of $\mathbf{p}$-bounded forms with $\mathbf{p} < \infty$, which will be proved in Section 2.2.

Theorem 2.1.15. *Let $1 \le \mathbf{p} < \infty$ and let $\mathbf{H}$ be a $\mathbf{p}$-bounded singular Brascamp-Lieb datum with $H_1, H_2, H_3 \neq \{0\}$. Then one of the following holds, with the notation from Appendix 2.8.*

i) (Bilinear Hölder-type) There exists an assignment $\{i,j,k\} = \{1,2,3\}$ such that $\frac{1}{p_j} = \frac{1}{p_k} = 1 - \frac{1}{p_i}$ and $n_1, n_2, n_3, n_4 \ge 0$ such that

$$\mathbf{M_H} \cong (\mathbf{P}^{(j)})^{\oplus n_1} \oplus (\mathbf{K}^{(j)})^{\oplus n_2} \oplus (\mathbf{P}^{(k)})^{\oplus n_3} \oplus (\mathbf{K}^{(k)})^{\oplus n_4}\,. \tag{2.1.9}$$

ii) (Young-type) We have $\mathbf{p} = (p_1, p_2, p_3)$ with $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 2$. If $p_1, p_2, p_3 \neq 1$ then there exist $n_1, n_2 \ge 0$ such that

$$\mathbf{M_H} \cong \mathbf{Y}^{\oplus n_1} \oplus \mathbf{Z}^{\oplus n_2}\,. \tag{2.1.10}$$

If there is some $i \in \{1,2,3\}$ with $p_i = 1$, then there exist $n_1, n_2, n_3, n_4 \ge 0$ such that

$$\mathbf{M_H} \cong \mathbf{Y}^{\oplus n_1} \oplus \mathbf{Z}^{\oplus n_2} \oplus (\mathbf{P}^{(i)})^{\oplus n_3} \oplus (\mathbf{K}^{(i)})^{\oplus n_4}\,. \tag{2.1.11}$$

iii) (Loomis-Whitney-type) We have $\mathbf{p} = (2, 2, 2)$ *and there exist* $n_1, n_2 \geq 0$ *and a list of modules* $\mathbf{M}_1, \dots, \mathbf{M}_k$ *from Table 2.4 with*

$$\mathbf{M}_{\mathbf{H}} \cong \mathbf{L}^{\oplus n_1} \oplus \mathbf{B}^{\oplus n_2} \oplus \mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_k \,. \tag{2.1.12}$$

iv) (Hölder-type) We have $\mathbf{p} = (p_1, p_2, p_3)$ *with* $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$*. In this case, there exists a finite list of modules* $\mathbf{M}_1, \dots, \mathbf{M}_k$ *from Table 2.2 such that*

$$\mathbf{M}_{\mathbf{H}} \cong \mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_k \,.$$

The proof of Theorem 2.1.15 uses necessary conditions for boundedness of nonsingular Brascamp-Lieb forms from [13]. They can be applied to singular Brascamp-Lieb forms because the Dirac δ distribution is a Calderón-Zygmund kernel, and singular Brascamp-Lieb forms with kernel δ simplify to nonsingular Brascamp-Lieb forms. A similar argument previously appeared in [48].

The singular Brascamp-Lieb forms corresponding to cases i), ii) of Theorem 2.1.15 are easily seen to be bounded by Hölder's inequality, Young's convolution inequality, and classical linear singular integral theory. The forms corresponding to case iii) are also bounded, by an elementary argument using Plancherel and the Loomis-Whitney inequality. This is summarized by the following theorem, which we prove in Section 2.3.

Theorem 2.1.16. *Let* $\mathbf{M}_{\mathbf{H}}$ *and* $\mathbf{p}$ *be as in case i), ii) or iii) of Theorem 2.1.15. Then* $\mathbf{H}$ *is* $\mathbf{p}$*-bounded.*

As Theorem 2.1.16 shows, forms of Hölder type are the most interesting ones. Showing boundedness for them is in general open, and contains some difficult problems. We collect results from the literature, proving bounds in some cases, in Section 2.1.5.

2.1.4 Projection results and method of rotations

The difficulty of estimating singular Brascamp-Lieb forms increases when taking direct sums of the corresponding modules, in the following precise sense.

Theorem 2.1.17. *Let* $\mathbf{M}, \mathbf{M}'$ *be two modules and let* $\mathbf{p} < \infty$*. Let* $\mathbf{H}$ *and* $\mathbf{H} \oplus \mathbf{H}'$ *be data with* $\mathbf{M}_{\mathbf{H}} \cong \mathbf{M}$ *and* $\mathbf{M}_{\mathbf{H} \oplus \mathbf{H}'} \cong \mathbf{M} \oplus \mathbf{M}'$*. Suppose that for each* l*-Calderón-Zygmund kernel* K *we have*

$$|\Lambda_{\mathbf{H} \oplus \mathbf{H}'}(K, f_1, f_2, f_3)| \leq C \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3} \,.$$

Then there exists a constant C' *such that for each* $2l$*-Calderón-Zygmund kernel* K *we have*

$$|\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3)| \leq C' \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3} \,.$$

Also, in the case $\mathbf{J}_n^{(2)}$ of the classification, the difficulty increases with the parameter n.

Theorem 2.1.18. *Let* $\mathbf{M}$ *be a module and let* $\mathbf{p} < \infty$*. Let* $\mathbf{H}_n$ *and* $\mathbf{H}_{n-1}$ *be data with* $\mathbf{M}_{\mathbf{H}_n} \cong \mathbf{M} \oplus \mathbf{J}_n^{(2)}$ *and* $\mathbf{M}_{\mathbf{H}_{n-1}} \cong \mathbf{M} \oplus \mathbf{J}_{n-1}^{(2)}$*. Suppose that there exists* C *such that for each* l*-Calderón-Zygmund kernel* K *we have*

$$|\Lambda_{\mathbf{H}_n}(K, f_1, f_2, f_3)| \leq C \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3} \,.$$

Then there exists a constant C' *such that for each* $2l$*-Calderón-Zygmund kernel* K *we have*

$$|\Lambda_{\mathbf{H}_{n-1}}(K, f_1, f_2, f_3)| \leq C' \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3} \,.$$

The same is true for $\mathbf{J}_n^{(1)}$ and $\mathbf{J}_n^{(3)}$, because they are isomorphic to modules that can be obtained from $\mathbf{J}_n^{(2)}$ by permuting the subspaces, see Lemma 2.8.1. We do not state similar theorems for $\mathbf{C}_n$ or $\mathbf{N}_n$, because Theorem 2.6.1 already gives unconditional bounds in these cases. For $\mathbf{T}_n$, we do not expect an analogue of Theorem 2.1.18 to be true. The reason is that the associated forms become more singular as n gets smaller, at least judging only by the number of arguments of the kernel compared to the functions.

In a different direction, it is possible to express forms with a kernel taking d arguments as superpositions of certain forms with kernels taking $d-1$ arguments. Thus bounds for the former are at most as hard as integrable bounds for the latter. A classical instance of this idea is the method of rotations, introduced by Calderón and Zygmund in [20], in which one expresses an odd Calderón-Zygmund kernel as a superposition of Hilbert transforms. Using this, one can deduce bounds for odd kernel Calderón-Zygmund operators in higher dimensions from the boundedness of the Hilbert transform. We prove a stronger version of this fact. Namely, *every* Calderón-Zygmund kernel in dimension 3 or higher can be expressed as a superposition of 2-dimensional Calderón-Zygmund kernels on 2-dimensional subspaces.

This yields the following theorem for singular Brascamp-Lieb forms, which is proved in Section 2.5. We denote by $\mathrm{Gr}_d(V)$ the Grassmann-manifold of d-dimensional subspaces of some vector space V. A Calderón-Zygmund kernel on a d-dimensional Hilbert space is called homogeneous if for all $x \neq 0$ it holds $K(tx) = t^{-d}K(x)$.

Theorem 2.1.19. *Let $\mathbf{H}$ be a singular Brascamp-Lieb datum and suppose that $d = \dim H_0 \geq 3$. Let $l \geq d+1$. There exists $C' > 0$ such that the following holds. For each $\theta \in \mathrm{Gr}_{d-1}(H_0)$, consider the datum*

$$\mathbf{H}(\theta) = \mathbf{H} \cap \Pi_0^{-1}(\theta) = (\Pi_0^{-1}(\theta), \theta, H_1, H_2, H_3; \Pi_0, \Pi_1, \Pi_2, \Pi_3)\,.$$

Here we abuse notation and denote the restriction of Π_i to $\Pi_0^{-1}(\theta)$ still by Π_i. Suppose that for all $\theta \in \mathrm{Gr}_{d-1}(M_0)$ there exists $C(\theta)$ such that for all homogeneous $l - \lceil \frac{d+2}{2} \rceil$-Calderón-Zygmund kernels K on θ, we have

$$|\Lambda_{\mathbf{H}(\theta)}(K, f_1, f_2, f_3)| \leq C(\theta) \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3}\,.$$

Then for all homogeneous l-Calderón-Zygmund kernels K on H_0, we have

$$|\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3)| \leq C' \int C(\theta)\, d\theta \cdot \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3}\,,$$

where the integration is over $\mathrm{Gr}_{d-1}(H_0)$ with respect to the unique rotation invariant probability measure. If $d = 2$, then the same result is true under the additional assumption that K is odd.

Remark 2.1.20. The loss of derivatives is a technical consequence of the fact that we assume Mikhlin bounds (2.1.3) on the Fourier transform of the kernel. If the smoothness assumptions on the kernels are formulated on the spatial side, then there is no loss of derivatives.

There is a crucial difficulty in applying Theorem 2.1.19: It assumes quantitative, integrable estimates for the norms of $\Lambda_{\mathbf{H}(\theta)}$. Such estimates are notoriously hard to prove. See

[115], [56] for the strongest currently known results in that direction, which still only apply to the case $\mathbf{N}_1$. Recall also that to study quantitative bounds, one should use the finer equivalence relation with φ_0 a scalar multiple of an orthogonal transformation, as described in Remark 2.1.7.

2.1.5 Positive boundedness results in the literature

It is tempting to conjecture that the conditions in Theorem 2.1.15 are already sufficient. By Theorem 2.1.16, this would follow from the following conjecture.

Conjecture 2.1.21. *All singular Brascamp-Lieb data $\mathbf{H}_{\mathbf{M}}$ with $\mathbf{M}$ as in case iv) of Theorem 2.1.15 are $\mathbf{p}$-bounded for all $1 < \mathbf{p} < \infty$ with $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$.*

We now give a list of known boundedness results for forms of Hölder type. With the exception of Theorem 2.1.23, these results are not new, however some of them have not been stated in this form anywhere in the literature. In what follows, we will always fix a basis and identify finite dimensional Hilbert spaces with $\mathbb{R}^n$, for some n.

Note first that the module $\mathbf{C}_0$ corresponds simply to Hölder's inequality in three functions. If a datum $\mathbf{H}_{\mathbf{M}}$ is $\mathbf{p}$-bounded then so is $\mathbf{H}_{\mathbf{M}\oplus\mathbf{C}_0}$, by Fubini and Hölder's inequality. Keeping this in mind, we can ignore $\mathbf{C}_0$ in the following discussion.

Coifman and Meyer

The first result on multilinear singular integral operators, due to Coifman and Meyer [33, 32], treats the case $\mathbf{M} = \mathbf{C}_1^{\oplus n}$ for $n \geq 1$. The singular Brascamp-Lieb form corresponding to this module is

$$\Lambda(K, f_1, f_2, f_3) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f_1(x) f_2(x+y) f_3(x+z) K(y,z)\, dy\, dz\, dx\,.$$

These forms are in a sense the least singular among all singular Brascamp-Lieb forms of Hölder type, because the kernel K has the maximum possible number of arguments compared to the functions.

Time-frequency analysis

Lacey and Thiele [76, 78] proved bounds for the Bilinear Hilbert transform

$$\Lambda(f_1, f_2, f_3) = \int_{\mathbb{R}} \int_{\mathbb{R}} f_1(x) f_2(x+t) f_3(x+\alpha t) \frac{1}{t}\, dt\, dx\,,$$

where $\alpha \neq 0, 1$. This corresponds to the module $\mathbf{N}_1$, with $X = \alpha$. Their methods were subsequently extended to treat also the cases $\mathbf{N}_n$ for $n \geq 2$, and direct sums thereof. For $n = 2$, this was done in [39]. For larger n, proofs can be found in [104, 56].

The techniques introduced by Lacey and Thiele apply to a certain class of multilinear Fourier multiplier operators more general than (2.1.4). This was first observed in [59, 92], where it is shown that it suffices if the multipliers satisfy symbol estimates away from some subspace, which in particular holds if they satisfy symbol estimates away from some smaller subspace. Using this observation, one can deduce also bounds for forms $\Lambda_{\mathbf{H}}$ with $\mathbf{M}_{\mathbf{H}}$ including summands $\mathbf{C}_n$. The following theorem summarizes this.

Theorem 2.1.22. *Suppose that* $\mathbf{M}$ *is a direct sum of modules* $\mathbf{N}_{n_i}$ *and* $\mathbf{C}_{m_i}$*, for some finite sequences* $n_i, m_i \in \mathbb{N}_{\geq 1}$*. For each* $2 < \mathbf{p} < \infty$ *and each singular Brascamp-Lieb datum* $\mathbf{H}$ *associated with* $\mathbf{M}$*, there exists* $C > 0$ *and* l *such that for each* l*-Calderón-Zygmund kernel* K

$$|\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3)| \leq C\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}.$$

One and a half dimensional time-frequency analysis

The conditions of Theorem 2.1.22 are generically satisfied. 'Degenerate' cases were first studied by Demeter and Thiele in [39], for functions of two arguments. There are, up to permutation of the functions and equivalence, only three degenerate cases when all functions and kernels have two arguments: $\mathbf{J}_2^{(i)}$, $\mathbf{N}_1 \oplus \mathbf{J}_1^{(i)}$ and $\mathbf{J}_1^{(i)} \oplus \mathbf{J}_1^{(j)}$ for $i \neq j$. This can be read off of Table 2.4, noting that changing i and j only amounts to permuting the functions. Demeter and Thiele develop a 'one and a half-dimensional' time frequency analysis, to prove bounds for the cases $\mathbf{J}_2^{(i)}$ and $\mathbf{N}_1 \oplus \mathbf{J}_1^{(i)}$. Similarly as discussed before Theorem 2.1.22, their proof implies also boundedness of the less singular forms Λ corresponding to $\mathbf{C}_1 \oplus \mathbf{J}_1^{(i)}$. Indeed, by performing a discretization of such Λ as in [39], one arrives at a model form that still specializes the form (3) in [39, Section 3.1.1], and is consequently bounded by the argument given there.

Demeter and Thiele further observe that $\mathbf{p}$-bounds for the form $\Lambda_{\mathbf{H}}$, with $\mathbf{M}_{\mathbf{H}} = \mathbf{J}_2^{(i)}$, imply Carleson's theorem [21] on pointwise convergence of Fourier series of L^{p_1} functions. By Theorems 2.1.17 and 2.1.18, the same is then true whenever $\mathbf{M}_{\mathbf{H}}$ has a direct summand $\mathbf{J}_n^{(i)}$ for any $n \geq 2$.

Twisted techniques

Demeter and Thiele left open the last case in their classification, $\mathbf{J}_1^{(i)} \oplus \mathbf{J}_1^{(j)}$ for $i \neq j$. They called this case the 'twisted-paraproduct'. It was later shown to be bounded by Kovač [71], using very different techniques. Variations of Kovač's techniques can by applied to many other multilinear singular Brascamp-Lieb forms with so-called cubical structure, see [45]. We expect that bounding forms associated with modules containing a direct summand other than $\mathbf{J}_1^{(i)}$ and $\mathbf{C}_1$ requires extensions of the time-frequency analysis methods described above, perhaps in combination with twisted techniques. For $\mathbf{N}_n$ this is suggested by the fact that all known proofs use such techniques, while for $\mathbf{J}_n^{(i)}$ the implication for Carleson's theorem offers some justification. Thus, the only remaining trilinear singular Brascamp-Lieb forms that should be attackable using twisted techniques are the ones associated with $(\mathbf{J}_1^{(1)} \oplus \mathbf{J}_1^{(2)} \oplus \mathbf{J}_1^{(3)} \oplus \mathbf{C}_1)^{\oplus n}$, $n \geq 1$.

The following Theorem, which we will prove in Section 2.7, shows that they are indeed always $\mathbf{p}$-bounded.

Theorem 2.1.23. *Let* $n \geq 1$ *and let* $\mathbf{M} = (\mathbf{J}_1^{(1)} \oplus \mathbf{J}_1^{(2)} \oplus \mathbf{J}_1^{(3)} \oplus \mathbf{C}_1)^{\oplus n}$*. Let* $2 < \mathbf{p} < \infty$ *and let* $\mathbf{H}$ *be a singular Brascamp-Lieb datum associated with* $\mathbf{M}$*. Then there exists* l *and* $C > 0$ *such that for all* l*-Calderón-Zygmund kernels* K*, we have*

$$|\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3)| \leq C\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_2}.$$

Note that Theorem 2.1.23 recovers via Theorem 2.1.17 boundedness of the twisted paraproduct, as well as of certain higher dimensional versions. It further gives a new proof of boundedness of the form associated with $\mathbf{J}_1^{(j)} \oplus \mathbf{C}_1$, different from the one implicit in [39].

By a cone decomposition, the proof of Theorem 2.1.23 reduces to two essentially different cases. The first case can be treated using bounds for the standard maximal and square functions, in analogy with the Coifman-Meyer multipliers. In this case we have, in fact, boundedness in a larger range $1 < \mathbf{p} < \infty$. The second case is bounded using twisted techniques, tailored to the specific structure of the form. The arguments rely on intertwined applications of the Cauchy-Schwarz inequality, integration-by-parts identity, and positivity arguments. These arguments are applied in a localized setting, which in turn gives the claimed range of boundedness. It is an open problem to further lower the range of exponents in Theorem 2.1.23. For the twisted paraproduct [71], fiber-wise Calderón-Zygmund decomposition [14] can be used to extend the range of some exponents. However, similar arguments do not seem to directly apply in our setting.

The triangular Hilbert transform

The final family of the classification, $\mathbf{T}_n$, contains and generalizes the so-called triangular Hilbert transform

$$\Lambda(f_1, f_2, f_3) = \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f_1(x_1, x_2) f_2(x_1 + t, x_2) f_3(x_1, x_2 + t) \frac{1}{t} \, dt \, dx_1 \, dx_2 \, .$$

Proving any $\mathbf{p}$-bounds for this form is a hard open problem. We refer to [72] for some discussion and a partial result. By Theorem 2.1.17, bounding any form associated with a module $\mathbf{M}$ with a direct summand $\mathbf{T}_n$ is at least as hard as bounding $\mathbf{T}_n$, and therefore also open. This applies in particular, but not exclusively, to all forms where $\dim H_0 < \dim H_1$ (note that we have $\dim H_1 = \dim H_2 = \dim H_3$ for every $\mathbf{H}$ of Hölder type). Indeed, all such forms must contain an indecomposable direct summand satisfying the same inequality, and the only such summands are of type $\mathbf{T}_n$. We note that this invalidates a certain claim in the paper [38], in the case of trilinear forms. More precisely, in [38] boundedness of certain multilinear singular Brascamp-Lieb forms is shown, which do have a dimension deficit as above, and it is claimed that the assumptions placed on the forms are generically satisfied. We believe that this genericity claim is false for trilinear forms.

Further questions

At the time of writing, the above list of cases where $\mathbf{p}$-bounds are known is complete, to the best of our knowledge. This gives rise to a number of open questions. We consider it an interesting question whether time frequency techniques and twisted techniques can be combined, natural test cases are $\mathbf{N}_1 \oplus \mathbf{J}_1^{(1)} \oplus \mathbf{J}_1^{(2)}$ or $\mathbf{N}_1 \oplus \mathbf{J}_1^{(1)} \oplus \mathbf{J}_1^{(2)} \oplus \mathbf{J}_1^{(3)}$. More difficult seems to be the question of bounds for the forms associated with $\mathbf{T}_n$, $n \geq 2$. Judging by the dimension of the space H_0 in relation to H_1, H_2, H_3 alone, these forms become less singular as n increases, so these might be useful test cases towards the triangular Hilbert transform. Finally, it would be interesting to gain a better understanding of the questions considered in this paper for higher degrees of multilinearity. While a classification in terms of direct summands is not possible, see Remark 2.2.3, there might be a different algebraic description of the properties of modules relevant for proving **p**-bounds.

2.1.6 Comparison with the literature

We point out that various related objects have been studied under the name singular Brascamp-Lieb forms. Some completely nondegenerate cases with higher degrees of multilinearity have been considered in [92, 38], using time frequency analysis. Some further multilinear cases with so-called 'cubical structure' are studied in [45, 46], using twisted techniques. However, there has been no attempt of a systematic study of all degenerate cases.

Our kernel K always satisfies the single parameter Mikhlin condition (2.1.3). This is in contrast to related multiparameter problems, which have been studied for example in [12, 89, 90, 93] in connection with fractional Leibniz-rules. In particular, we point out that the 'tensorization' of forms to obtain multiparameter forms, as for example in [12], is not the same as the procedure of taking direct sums of modules associated with singular Brascamp-Lieb forms. The former is a way of constructing multiparameter forms, while the latter constructs single parameter forms. However, as the Mikhlin condition (2.1.3) on $\mathbb{R}^n$ is implied by an n-parameter kernel condition, known multiparameter bounds imply some of the one parameter bounds. In particular, we note that the bounds obtained in [12, Theorem 6] for a tensor product of one bilinear Hilbert transform with n many paraproducts imply boundedness of the forms of type $\mathbf{N}_1 \oplus \mathbf{C}_1^n$.

Another question concerning the kernels is about the optimal regularity l, or more generally for optimal regularity conditions on the kernel K or the symbol. Classical results for linear singular integral operators giving such sharp regularity conditions have been generalized to Coiffman-Meyer type forms $\mathbf{C}_1^{\oplus n}$ in [114, 63, 88, 79], and to the bilinear Hilbert transform in [25].

Acknowledgment

We thank Christoph Thiele for numerous discussions and his support in facilitating collaboration among the authors. LB is supported by the Collaborative Research Center 1060 funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and the Hausdorff Center for Mathematics, funded by the DFG under Germany's Excellence Strategy - EXC-2047/1 - 390685813. PD is supported by the NSF grant DMS-2154356. FL is supported by the DAAD Graduate School Scholarship Programme - 57572629. This work was completed while the authors were in residence at the Hausdorff Research Institute for Mathematics in Bonn, during the trimester program "Boolean Analysis in Computer Science", funded as well by the DFG under Germany's Excellence Strategy - EXC-2047/1 - 390685813.

2.2 The classification: Proof of Theorem 2.1.14 and Theorem 2.1.15

Proof of Theorem 2.1.14. The first claim of Theorem 2.1.14 follows immediately from basic linear algebra. The remaining claims follow from the first, the classification of modules in

Theorem 2.1.11 and the facts that clearly $\mathbf{M}_{\mathbf{H}_\mathbf{M}} \cong \mathbf{M}$ and $\mathbf{H}_{\mathbf{M}_\mathbf{H}} \cong \mathbf{H}$. $\square$

Before proving Theorem 2.1.15, we recall some necessary conditions for a datum $\mathbf{H}$ to be $\mathbf{p}$-bounded for some $\mathbf{p} < \infty$. They were proven in [48], by adapting similar arguments for non-singular Brascamp-Lieb inequalities from [13].

Lemma 2.2.1. *Let $\mathbf{p} < \infty$ and suppose that $\mathbf{H}$ is $\mathbf{p}$-bounded. Then for each $i = 1, 2, 3$,*

$$\Pi_i \ker \Pi_0 = \Pi_i H \,. \tag{2.2.1}$$

Furthermore, for each subspace $H' \subseteq \ker \Pi_0$, it holds that

$$\dim H' \leq \sum_{i=1}^{3} \frac{\dim \Pi_i H'}{p_i} \,, \tag{2.2.2}$$

and if $H' = \ker \Pi_0$, then we have equality in (2.2.2).

Proof. Note that the Dirac δ distribution is a Calderón-Zygmund kernel. Hence, if $\mathbf{H}$ is $\mathbf{p}$-bounded, then the Brascamp-Lieb form

$$\int_H f_1(\Pi_1(x)) f_2(\Pi_2(x)) f_3(\Pi_3(x)) \delta(\Pi_0(x)) \, dx = c \int_{\ker \Pi_0} f_1(\Pi_1(x)) f_2(\Pi_2(x)) f_3(\Pi_3(x)) \, dx$$

is bounded on $L^{p_1} \times L^{p_2} \times L^{p_3}$. Theorem 1.13 in [13] then immediately gives (2.2.2). The first condition (2.2.1) follows from the fact that if $p_j < \infty$, then the Brascamp-Lieb form can only be bounded on L^{p_j} if $\Pi_i|_{\ker \Pi_0}$ is surjective. $\square$

Lemma 2.2.2. *Suppose that the datum $\mathbf{H}$ is of Hölder type. Then for each $i = 1, 2, 3$, we have that $H = \ker \Pi_0 \oplus \ker \Pi_i$.*

Proof. By (2.2.2) we have

$$\dim \ker \Pi_0 = \sum_{i=1}^{3} \frac{\dim \Pi_i \ker \Pi_0}{p_i} \leq \sum_{i=1}^{3} \frac{\dim \ker \Pi_0}{p_i} = \dim \ker \Pi_0 \,.$$

In the last step we used that $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$. So we must have equality in the middle, hence $\Pi_j|_{\ker \Pi_0}$ is injective, which gives that $\ker \Pi_j \cap \ker \Pi_0 = \{0\}$. Combining this with (2.2.1), we obtain

$$\dim \ker \Pi_0 = \dim \Pi_i \ker \Pi_0 = \dim \Pi_i H = \dim H - \dim \ker \Pi_i \,,$$

which completes the proof of the lemma. $\square$

Proof of Theorem 2.1.15. Let $\mathbf{p} < \infty$ and suppose that $\mathbf{H}$ is $\mathbf{p}$-bounded. To simplify some formulas, we will write below $q_i = p_i^{-1}$. By Theorem 2.1.11, we have that

$$\mathbf{M}_\mathbf{H} = \mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_k$$

for some modules $\mathbf{M}_k$ from Table 2.1. By Theorem 2.1.17, each $\mathbf{H}_j = \mathbf{H}_{\mathbf{M}_j}$, $j = 1, \ldots, k$, is $\mathbf{p}$-bounded.

Recall that Theorem 2.1.11 allows for permutations of the subspaces in the modules in Table 2.1, see Lemma 2.8.1 for an exact description of which permutations of the subspace give rise to nonisomorphic modules. We will denote modules from Table 2.1 by adding the permutation and the parameter n as subscripts.

We write

$$\mathbf{H}_j = (H_j, H_{j0}, H_{j1}, H_{j2}, H_{j3}, \Pi_{j0}, \Pi_{j1}, \Pi_{j2}, \Pi_{j3})\,.$$

By condition (2.2.1) of Lemma 2.2.1 and surjectivity of the maps Π_{ji}, we have for $i = 1, 2, 3$

$$\dim H_{ji} = \dim \Pi_{ji} H_j = \dim \Pi_{ji} \ker \Pi_{j0} \leq \dim H_j - \dim H_{j0}\,. \tag{2.2.3}$$

By comparing with Table 2.1, this immediately implies that $\mathbf{M}_j \ncong \mathbf{IV}^*_{n,\pi}$ and $\mathbf{M}_j \ncong \mathbf{V}^*_{n,\pi}$ for any n or π.

Suppose next that $\mathbf{M}_j \cong \mathbf{I}_{n,\pi}$ for some $n \geq 1$ and some permutation π. If H_{j0} corresponds to the second subspace in the block matrix in Table 2.1, then the map Π_{j3} corresponding to the fourth subspace is not surjective on $\ker \Pi_0$. Similarly, if H_{j0} corresponds to the fourth subspace, then the map corresponding to the second one is not surjective on $\ker \Pi_0$. Thus by (2.2.1) H_{j0} must correspond to the first or third subspace, and by Lemma 2.8.1 we can assume that it corresponds to the first. By (2.2.2), $\mathbf{p}$ must satisfy the Hölder condition $q_1 + q_2 + q_3 = 1$. By permuting the last three subspaces and using Lemma 2.8.1, we obtain three isomorphism classes of modules, which are exactly $\mathbf{J}_n^{(1)}, \mathbf{J}_n^{(2)}, \mathbf{J}_n^{(3)}$ in Table 2.2.

Suppose now that $\mathbf{M}_j \cong \mathbf{II}_{n,\pi}$ for some $n \geq 0$ and some permutation π. By (2.2.3), the permutation π must be so that H_{j0} corresponds to a subspace of dimension n. We will assume by permuting the functions that H_{j3} is the other subspace of dimension n. Applying (2.2.2) to the full space $H' = \ker \Pi_{j0}$ we obtain

$$(n+1)q_1 + (n+1)q_2 + nq_3 = n+1\,.$$

On the other hand, applying (2.2.2) to the one dimensional space $H' = \ker \Pi_{j3} \cap \ker \Pi_{j0}$ yields

$$q_1 + q_2 \geq 1\,.$$

Since $p_3 < \infty$ and hence $q_3 > 0$, it follows that $n = 0$. So in this case, we must have $\mathbf{M}_j \cong \mathbf{P}^{(3)}$ and $\frac{1}{p_1} + \frac{1}{p_2} = 1$. Note that swapping the two nonzero subspaces in $\mathbf{P}^{(j)}$ gives an isomorphic module, see Lemma 2.8.1. Permuting the functions yields the two additional possibilities $\mathbf{M}_j \cong \mathbf{P}^{(1)}$ or $\mathbf{M}_j \cong \mathbf{P}^{(2)}$, with the corresponding conditions on $\mathbf{p}$.

Next, assume that $\mathbf{M}_j \cong \mathbf{IV}_{n,\pi}$ for some $n \geq 0$ and π. Suppose first that H_{j0} corresponds to one of the first three subspaces in Table 2.1. We permute the subspaces so that $\dim H_{j3} = n$. Then we get from (2.2.2) that

$$(n+1)q_1 + (n+1)q_2 + nq_3 = n+1\,.$$

Taking $H' = \ker \Pi_{j0} \cap \ker \Pi_{j3}$, we also have

$$q_1 + q_2 \geq 1\,.$$

Since $p_3 < \infty$ it follows that $n = 0$ and hence $\mathbf{M}_j \cong \mathbf{K}^{(3)}$ and $\frac{1}{p_1} + \frac{1}{p_2} = 1$. Note that swapping the two nonzero subspaces in $\mathbf{P}^{(j)}$ gives an isomorphic module. Thus, permuting the subspaces, only yields the additional possibilities $\mathbf{M}_j \cong \mathbf{K}^{(1)}$ or $\mathbf{M}_j \cong \mathbf{K}^{(2)}$, with corresponding conditions on $\mathbf{p}$. It remains to consider the case where H_{j0} is the last subspace in Table 2.1. (2.2.2) applied to $H' = \ker \Pi_{j0}$ gives in this case

$$(n+1)(q_1 + q_2 + q_2) = n+2\,.$$

On the other hand, applying (2.2.2) to each of the one dimensional subspaces $\ker \Pi_{j0} \cap \ker \Pi_{ji}$, $i = 1, 2, 3$, and adding the resulting inequalities, yields

$$2(q_1 + q_2 + q_3) \geq 3\,.$$

Hence $n = 0$ and $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 2$ or $n = 1$ and $p_1 = p_2 = p_3 = 2$. This corresponds to $\mathbf{M}_j \cong \mathbf{Y}$ and $\mathbf{M}_j \cong \mathbf{L}$, respectively. Note that all permutations of the subspaces in these modules corresponding to functions yield isomorphic modules.

Finally assume that $\mathbf{M}_j \cong \mathbf{V}_{n,\pi}$ for some $n \geq 0$ and π. Note that all such modules for fixed n and different π are isomorphic. Applying condition (2.2.2) to $H' = \ker \Pi_{j0}$, we obtain

$$n(q_1 + q_2 + q_3) = n+1\,.$$

On the other hand, applying (2.2.2) to $H' = \ker \Pi_{j0} \cap \ker \Pi_{ji}$, $i = 1, 2, 3$, and adding the resulting inequalities, yields

$$2(q_1 + q_2 + q_3) \geq 3\,.$$

Hence, we must have either $n = 1$ and $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 2$ or $n = 2$ and $p_1 = p_2 = p_3 = 2$. This corresponds to $\mathbf{M}_j \cong \mathbf{Z}$ and $\mathbf{M}_j \cong \mathbf{B}$, respectively.

In the remaining cases, it follows immediately from (2.2.3) that $\mathbf{H}_{j0}$ must correspond to the first subspace. All permutations respecting this give rise to isomorphic modules, or for modules of type $\mathbf{0}$ to another module of type $\mathbf{0}$. From (2.2.2) it then follows that $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$. Case $\mathbf{0}$ then corresponds to $\mathbf{N}$, case $\mathbf{I}$ gives after permuting the subspaces rise to the Jordan block cases $\mathbf{J}^{(s)}$, case $\mathbf{III}$ corresponds to $\mathbf{C}$ and case $\mathbf{III}^*$ to $\mathbf{T}$.

Thus the possible choices of $\mathbf{p}$ are exactly as in case i) - iv) of Theorem 2.1.15. Collecting the possible summands for each choice of $\mathbf{p}$ completes the proof. $\square$

Remark 2.2.3. We now show that already the classification of n-linear singular Brascamp-Lieb forms of Hölder type is as hard as the classification of representations of the $n-1$-Kronecker quiver, i.e. of tuples of $n-1$ linear maps between two finite dimensional vector spaces, up to isomorphism. This classification problem is wild for $n > 3$, see for example Theorem 1 and 2 in [95]. Thus a classification as above is not possible for any $n > 3$, not even under the assumption that the forms are of Hölder type.

Note that the necessary conditions from both Lemma 2.2.1 and Lemma 2.2.2 continue to hold for more than three functions, with identical proofs. Suppose the $\mathbf{H}$ is a datum of Hölder type. Let $a = \dim H_0$ and $b = \dim H_1$. By Lemma 2.2.2, we have $\ker \Pi_0 \oplus \ker \Pi_1 = H$, so we can choose bases of H, H_0 and H_1 such that the matrices of Π_0, Π_1 are given by

$$\begin{pmatrix} I_a \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ I_b \end{pmatrix}\,.$$

Choosing in H_i, $2 \leq i \leq n$, the basis $\Pi_i(e_{a+1}), \dots \Pi_i(e_{a+b})$, the matrices of $\Pi_i, 2 \leq i \leq n$ are given by

$$\begin{pmatrix} A_i \\ I_b \end{pmatrix}\,.$$

for certain $a \times b$ matrices A_i. Let two singular Brascamp Lieb data $\mathbf{H}, \mathbf{H}'$ be given, and assume that they can be transformed into the above normal form with matrices A_i and A_i', $2 \leq i \leq n$, respectively. Then $\mathbf{H}, \mathbf{H}'$ are equivalent if and only if there exists an invertible $a \times a$ matrix P and an invertible $b \times b$ matrix Q such that for all $i = 2, \dots, n$

$$QA_iP = A_i' \,. \tag{2.2.4}$$

On the other hand, the datum $(A_2, \dots, A_n)$ determines $n-1$ linear maps from $\mathbb{R}^b \to \mathbb{R}^a$, so a representation of the $(n-1)$-Kronecker quiver. Two such representations are also isomorphic if and only if (2.2.4) holds. Thus classifying n-linear singular Brascamp-Lieb forms of Hölder type is as hard as classifying representations of the $n-1$-Kronecker quiver.

2.3 Bounds for forms of non-Hölder type: Proof of Theorem 2.1.16

We go through the cases one by one. For simplicity we omit the domain of integration from the notation. Here and in the following sections, we will find constants for various related inequalities, and by abuse of notation we will denote each of them by the letter C. In particular, the meaning of C may change from line to line.

2.3.1 Case i

Suppose first that $\mathbf{H}$ is as in the case i), that is, (2.1.9) holds for $i, j, k \in \{1, 2, 3\}$. We choose coordinates $x_1 \in \mathbb{R}^{n_1}$, $x_2, u \in \mathbb{R}^{n_2}$, $y_1 \in \mathbb{R}^{n_3}$, $y_2, v \in \mathbb{R}^{n_4}$. To prove bounds for $\Lambda_{\mathbf{H}}(f_1, f_2, f_3)$, it suffices after a change of variables to prove bounds for

$$\int f_j(x_1, x_2) f_k(y_1, y_2) f_i(x_1, y_1, x_2 + u, y_2 + v) K(u, v) \, dx_1 \, dy_1 \, dx_2 \, dy_2 \, du \, dv$$

$$= \int f_j(x_1, x_2) f_k(y_1, y_2) (f_i * \tilde{K})(x_1, y_1, x_2, y_2) \, dx_1 \, dy_1 \, dx_2 \, dy_2 \,,$$

where $\tilde{K}(u, v) = K(-u, -v)$ and the convolution is in the third and fourth argument only. Since $p_j < \infty$, we have $p_i > 1$. Thus we can further estimate, using Hölder's inequality for the exponents p_j, p_i with $\frac{1}{p_j} = 1 - \frac{1}{p_i}$ and a linear singular integral bound on f_i

$$\leq C \int \|f_j(x_1, x_2) f_k(y_1, y_2)\|_{L^{p_j}_{x_2, y_2}} \|f_i(x_1, y_1, x_2, y_2)\|_{L^{p_i}_{x_2, y_2}} \, dx_1 \, dy_1 \,.$$

By Hölder's inequality and the condition $p_j = p_k$, this is bounded by

$$\leq C \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3} \,,$$

which completes the proof. If $n_2 = n_4 = 0$, then the assumption $p_j < \infty$ is not needed, and the estimate follows just from Hölder's inequality.

2.3.2 Case ii

Suppose next that $\mathbf{H}$ is in the case ii) with $p_1, p_2, p_3 \neq 1$, so (2.1.10) holds. Again we choose coordinates $x_1, y_1 \in \mathbb{R}^{n_1}$, $x_2, y_2, z_2 \in \mathbb{R}^{n_2}$ and write the form $\Lambda_{\mathbf{H}}(f_1, f_2, f_3)$ after a change of variables up to a constant as

$$\int f_1(x_1 + y_1, x_2 + y_2) f_2(x_1, x_2 + z_2) f_3(y_1, y_2) K(z_2)\, dx_1\, dx_2\, dy_1\, dy_2\, dz_2 \,.$$

Using $*$ to denote convolution in the second argument only, we estimate this with Young's convolution inequality, and then a linear singular integral bound using that $p_3 > 1$, by

$$\leq \|f_1\|_{p_1} \|f_2 * \tilde{K}\|_{p_2} \|f_3\|_{p_3} \leq C \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3} \,.$$

Now suppose that there is some i with $p_i = 1$, so (2.1.11) holds. We may assume $i = 1$, because the conditions on $\mathbf{p}$ are otherwise symmetric. We choose coordinates $x_3 \in \mathbb{R}^{n_3}$, $x_4, z_4 \in \mathbb{R}^{n_4}$ and write the form $\Lambda_{\mathbf{H}}(f_1, f_2, f_3)$ up to a constant and a change of variables as

$$\int f_1(x_1 + y_1, x_2 + y_2) f_2(x_1, x_2, x_3, x_4) f_3(y_1, y_2 + z_2, x_3, x_4 + z_4) K(z_2, z_4)\, dx\, dy\, dz \,.$$

We recognize a convolution in the second and fourth coordinate of f_3 with $\tilde{K}$. Applying Hölder's inequality in x_3, x_4, using that $\frac{1}{p_2} + \frac{1}{p_3} = 1$, we bound the last display by

$$\int |f_1(x_1{+}y_1, x_2{+}y_2)|\, \|f_2(x_1, x_2, x_3, x_4)\|_{L^{p_2}_{x_3,x_4}}\, \|f_3 *_{2,4} \tilde{K}(y_1, y_2, x_3, x_4)\|_{L^{p_3}_{x_3,x_4}} dx_1 dx_2 dy_1 dy_2 \,.$$

By Young's convolution inequality and then a linear singular integral bound, using that $p_2 < \infty$ and hence $p_3 > 1$, this is again bounded by $C\|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3}$.

2.3.3 Case iii

Suppose finally that $\mathbf{H}$ is as in case iii), so (2.1.12) holds. We use Fourier inversion to express $\Lambda_{\mathbf{H}}(f_1, f_2, f_3)$ in terms of the Fourier transforms $\widehat{f_1}$, $\widehat{f_2}$ and $\widehat{f_3}$. Then we apply the triangle inequality to move absolute values inside and estimate $\widehat{K}$ by 1 using (2.1.3). The resulting expression is a non-singular Brascamp-Lieb form $\widehat{\Lambda}$ in $\widehat{f_1}$, $\widehat{f_2}$ and $\widehat{f_3}$. By Plancherel's theorem, the problem thus reduces to checking that this Brascamp-Lieb form is bounded at exponent $(p_1, p_2, p_3) = (2, 2, 2)$. Transferring a Brascamp-Lieb form to the Fourier side in this way commutes with taking direct sums of the associated modules. By Lemma 4.8 in [13], a Brascamp-Lieb form is $\mathbf{p}$-bounded if each direct summand is $\mathbf{p}$-bounded. Thus it suffices to verify $(2, 2, 2)$-boundedness of each possible direct summand of $\widehat{\Lambda}$.

The summand corresponding on the Fourier side to $\mathbf{L}$ is the Loomis-Whitney trilinear Brascamp-Lieb form, since we have

$$\Lambda_{\mathbf{L}}(f_1, f_2, f_3) = \int f_1(x, u) f_2(y, v) f_3(x + v, y + u) K(u + v)\, dx\, dy\, du\, dv$$

$$= \int \widehat{f_1}(\xi_1, \xi_2 + \xi_3) \widehat{f_2}(\xi_2, \xi_1 + \xi_3) \widehat{f_3}(-\xi_1, -\xi_2) \widehat{K}(-\xi_3)\, d\xi_1\, d\xi_2\, d\xi_3 \,.$$

Estimating $\|\widehat{K}\|_\infty \le 1$, changing variables $\xi_1 + \xi_3 + \xi_2 = \tau$ and shearing the functions $\widehat{f}_1, \widehat{f}_2$, this becomes exactly the Loomis-Whitney inequality trilinear form, which is then estimated by

$$\|\widehat{f}_1\|_2 \|\widehat{f}_2\|_2 \|\widehat{f}_3\|_2 = \|f_1\|_2 \|f_2\|_2 \|f_3\|_2 \,.$$

For the summands $\mathbf{B}$ we obtain similarly

$$\Lambda_{\mathbf{B}}(f_1, f_2, f_3) = \int f_1(x,y) f_2(x+z, y+u) f_3(x+v, z+u) K(u,v) \, dx \, dy \, dz \, du \, dv$$

$$= \int \widehat{f}_1(-\xi_2 - \xi_3, \xi_1) \widehat{f}_2(\xi_2, -\xi_1) \widehat{f}_3(\xi_3, -\xi_2) \widehat{K}(\xi_1 + \xi_2, -\xi_3) \, d\xi_1 \, d\xi_2 \, d\xi_3 \,,$$

which after estimating $\|\widehat{K}\|_\infty \le 1$ and changing variables is again bounded by the Loomis-Whitney trilinear form of $\widehat{f}_1, \widehat{f}_2$ and $\widehat{f}_3$. For the summands $\mathbf{M}_i$ from Table 2.4 boundedness of the summands in $\widehat{\Lambda}$ reduces similarly to the Cauchy-Schwarz inequality.

2.4 Proof of the projection theorems, Theorem 2.1.17 and Theorem 2.1.18

To prove Theorem 2.1.17 and Theorem 2.1.18, we will need to extend Calderón-Zygmund kernels K on some Hilbert space H_0 to kernels on a larger Hilbert space $H_0 \oplus H_0'$. The following lemma allows us to do that.

Lemma 2.4.1. *Let $d, d' \ge 1$ and let K be a Calderón-Zygmund kernel on $\mathbb{R}^d$. Define*

$$K'(x,y) = |x|^{-d'} \exp\left(-\pi \frac{|y|^2}{|x|^2}\right) K(x) \,.$$

For sufficiently small $c = c(d, d', l) > 0$, the kernel $cK'(x,y)$ is an l-Calderón-Zygmund kernel on $\mathbb{R}^{d+d'}$.

Proof. Using that the assumptions on K are invariant under dilations, it suffices by scaling to show that for $|\xi| = 1$, $|\eta| \le 1$ and for $|\eta| = 1$, $|\xi| \le 1$ and all $|\alpha| \le l$, we have

$$|\partial^\alpha \widehat{K}'(\xi, \eta)| \le 1/c \,.$$

Denote the heat kernel by $\Phi(\xi, t) = t^{-d/2} \exp(-\pi |\xi|^2 / t)$. By a direct computation, we find that

$$\widehat{K}'(\xi, \eta) = \int_{\mathbb{R}^d} \widehat{K}(u) \frac{1}{|\eta|^d} \exp\left(-\pi \frac{|\xi - u|^2}{|\eta|^2}\right) du = \int_{\mathbb{R}^d} \widehat{K}(u) \Phi(\xi - u, |\eta|^2) \, du \,. \tag{2.4.1}$$

First, suppose that $|\eta| = 1, |\xi| \le 1$. The derivatives of the heat kernel take the form

$$\partial^\alpha \Phi(\xi - u, |\eta|^2) = |\eta|^{-|\alpha|} p\left(\frac{\xi - u}{|\eta|}, \frac{\eta}{|\eta|}\right) \Phi(\xi - u, |\eta|^2) \,, \tag{2.4.2}$$

where p is a polynomial of degree $2|\alpha|$. Therefore, using (2.1.3), it holds for $|\alpha| \le l$

$$|\partial^\alpha \widehat{K}'(\xi, \eta)| \le C |\eta|^{-|\alpha|} \|\widehat{K}\|_\infty \int_{\mathbb{R}^d} (1 + |u|^{2m}) e^{-\pi |u|^2} du \le C \,.$$

Now suppose that $|\eta| \leq 1$ and $|\xi| = 1$. We split up the integral in (2.4.1). Pick a smooth function φ on H_0 with $\mathbf{1}_{B(0,1/4)} \leq \varphi \leq \mathbf{1}_{B(0,1/2)}$. Then

$$\begin{aligned}\widehat{K}'(\xi,\eta) &= \int (1-\varphi(u))\widehat{K}(u)\Phi(\xi-u,|\eta|^2)\,du + \int \varphi(u)\widehat{K}(u)\Phi(\xi-u,|\eta|^2)\,du \\ &= G_1(\xi,|\eta|^2) + G_2(\xi,|\eta|^2)\,.\end{aligned}$$

Note that the function $G_1(\xi,t)$ solves the heat equation

$$4\pi\partial_t G_1(\xi,t) = \Delta_\xi G_1(\xi,t)\,.$$

Using this to replace all derivatives in the second argument of G_1 by derivatives in ξ, we obtain

$$\partial_\xi^\beta\partial_\eta^\gamma(G_1(\xi,|\eta|^2)) = \sum_{j=1}^{|\gamma|} p_j(\eta)\partial_\xi^\beta\Delta_\xi^j G_1(\xi,|\eta|^2)\,,$$

for certain polynomials p_j that depend only on γ. It follows that

$$\begin{aligned}|\partial_\xi^\beta\partial_\eta^\gamma G_1(\xi,|\eta|^2)| &= \left|\sum_{j=1}^{|\gamma|}\int p_j(\eta)\left(\partial_\xi^\beta\Delta_\xi^j((1-\varphi(\xi-u))\widehat{K}(\xi-u))\right)\Phi(u,|\eta|^2)\,du\right| \\ &\leq \sum_{j=1}^{|\gamma|}\sup_{|\eta|\leq 1}|p_j(\eta)|\sup_u\left|\partial_u^\beta\Delta_u^j((1-\varphi(u))\widehat{K}(u))\right|\,.\end{aligned}$$

Since $1-\varphi$ is smooth and supported on the complement of $B(0,1/4)$ and since $|\beta|+2|\gamma| \leq 2l$, it follows from the Mikhlin condition (2.1.3) for the $2l$-Calderón-Zygmund kernel K that this is bounded by a constant depending only on l.

The derivatives of the second term G_2 are given by

$$\partial^\alpha G_2(\xi,|\eta|^2) = \int \varphi(u)\widehat{K}(u)\partial^\alpha\Phi(\xi-u,|\eta|^2)\,du\,.$$

On the support of the integrand, we have $|u| \leq 1/2$ and hence $1/2 \leq |\xi-u| \leq 3/2$. Further, we have $|\eta| \leq 1$. Using (2.4.2), we obtain

$$\begin{aligned}\left|\partial^\alpha\Phi(\xi-u,|\eta|^2)\right| &\leq C|\eta|^{-|\alpha|}p\left(\frac{\xi-u}{|\eta|},\frac{\eta}{|\eta|}\right)\Phi(\xi-u,|\eta|^2) \\ &\leq C|\eta|^{-d-|\alpha|}\left(\frac{1}{|\eta|}+1\right)^{2|\alpha|}\exp\left(-\frac{\pi}{|\eta|^2}\right) \leq C\,.\end{aligned}$$

Hence, we have

$$|\partial^\alpha G_2(\xi,\eta)| \leq C\|\widehat{K}\|_\infty\|\varphi\|_1 \leq C\,,$$

which completes the proof of the lemma. □

Next, we note that in proving Theorem 2.1.17 and Theorem 2.1.18, we may restrict attention to bounded Calderón-Zygmund kernels. Indeed, every Calderón-Zygmund kernel K is the weak limit as $R \to \infty$ of the bounded kernels K_R defined by

$$\widehat{K_R}(\xi) = \widehat{K}(\xi)(\varphi(R^{-1}\xi) - \varphi(R\xi))\,,$$

for smooth φ with $\mathbf{1}_{B(0,1)} \leq \varphi \leq \mathbf{1}_{B(0,2)}$. Letting $R \to \infty$ in the conclusion of either theorem with kernel K_R yields the conclusion for the general kernel K.

Proof of Theorem 2.1.17. We fix a bounded $2l$-Calderón-Zygmund kernel K on H_0. By assumption, for all l-Calderón-Zygmund kernels K' on $H_0 \oplus H_0'$ and all functions F_1, F_2, F_3

$$|\Lambda_{\mathbf{H}\oplus\mathbf{H}'}(K', F_1, F_2, F_3)| \leq C\|F_1\|_{p_1}\|F_2\|_{p_2}\|F_3\|_{p_3}\,. \tag{2.4.3}$$

Our goal is to show that there exists C' such that for all f_1, f_2, f_3

$$|\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3)| \leq C'\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}\,.$$

We choose orthonormal bases of the spaces H_0, H_0', so that the inner product becomes the standard inner product on $\mathbb{R}^n$, and we denote the corresponding norm by $|\cdot|$.

We will apply (2.4.3) to the kernel K' on $H_0 \oplus H_0'$ obtained by extending K as in Lemma 2.4.1, thus

$$K'(x,y) = |x|^{-\dim H_0'} \exp\left(-\pi\frac{|y|^2}{|x|^2}\right) K(x)\,.$$

We pick for $i = 1, 2, 3$, functions F_i^N on $H_i \oplus H_i'$ defined by

$$F_i^N(x,y) = f_i(x) N^{-\frac{\dim H_i'}{p_i}} \exp\left(-\pi\frac{|y|^2}{N^2}\right)\,.$$

Since the datum $\mathbf{H}\oplus\mathbf{H}'$ is by assumption $\mathbf{p}$-bounded, we can apply (2.2.1) and (2.2.2) to the subspaces $\ker \Pi_0 \subseteq \ker(\Pi_0 \oplus \Pi_0')$, $\ker \Pi_0' \subseteq \ker(\Pi_0 \oplus \Pi_0')$, and $\ker(\Pi_0 \oplus \Pi_0')$, to obtain respectively

$$\sum_{i=1}^{3} \frac{\dim H_i'}{p_i} \geq \dim H' - \dim H_0'\,, \tag{2.4.4}$$

$$\sum_{i=1}^{3} \frac{\dim H_i}{p_i} \geq \dim H - \dim H_0$$

and

$$\sum_{i=1}^{3} \frac{\dim H_i' + \dim H_i}{p_i} = \dim H' + \dim H - \dim H_0' - \dim H_0\,.$$

So we must have equality in (2.4.4). Evaluating $\Lambda_{\mathbf{H}\oplus\mathbf{H}'}$ with our choice of functions and kernel yields then

$$\Lambda_{\mathbf{H}\oplus\mathbf{H}'}(K', F_1^N, F_2^N, F_3^N) = \int \prod_{i=1}^{3} f_i(\Pi_i(x)) K(\Pi_0(x)) N^{-\dim H' + \dim H_0'} |\Pi_0(x)|^{-\dim H_0'}$$

$$\times \int_{H'} \exp\left(-\pi\frac{|\Pi_1'(y)|^2 + |\Pi_2'(y)|^2 + |\Pi_3'(y)|^2}{N^2} - \pi\frac{|\Pi_0'(y)|^2}{|\Pi_0(x)|^2}\right) dy\, dx\,. \tag{2.4.5}$$

Denote by $A(x)$ the matrix

$$A(x) = \frac{1}{|\Pi_0(x)|^2} {\Pi_0'}^t \Pi_0' + \frac{1}{N^2}({\Pi_1'}^t \Pi_1' + {\Pi_2'}^t \Pi_2' + {\Pi_3'}^t \Pi_3') = \frac{1}{|\Pi_0(x)|^2} A_0 + \frac{1}{N^2} A_1\,.$$

Then the Gaussian y-integral in (2.4.5) evaluates to $\det(A(x))^{-1/2}$, so

$$(2.4.5) = \int \prod_{i=1}^{3} f_i(\Pi_i(x))K(\Pi_0(x))N^{-\dim H'+\dim H_0'}|\Pi_0(x)|^{-\dim H_0'}\det(A(x))^{-1/2}\,dx\,. \tag{2.4.6}$$

We claim that there exists constants $c(\mathbf{H}'), C(\mathbf{H}')$ with

$$\det A(x) = c(\mathbf{H}')\cdot|\Pi_0(x)|^{-2\dim H_0'}N^{-2(\dim H'-\dim H_0')}\cdot(1+O(N^{-2}|\Pi_0(x)|^2)) \tag{2.4.7}$$

and

$$\det(A(x))^{-1/2} \le C(\mathbf{H}')\cdot|\Pi_0(x)|^{\dim H_0'}N^{(\dim H'-\dim H_0')}\,. \tag{2.4.8}$$

Suppose (2.4.7) and (2.4.8) hold. Because f_1, f_2 and f_3 are Schwartz functions and K is bounded, it follows that the integrand in (2.4.6) is uniformly in N controlled by an integrable function of x. Using (2.4.7) in (2.4.6) and sending $N\to\infty$, we obtain with the dominated convergence theorem

$$\Lambda_{\mathbf{H}}(K,f_1,f_2,f_3) = c(\mathbf{H}')\lim_{N\to\infty}\Lambda_{\mathbf{H}\oplus\mathbf{H}'}(K',F_1^N,F_2^N,F_3^N)\,.$$

With the boundedness assumption (2.4.3) on $\Lambda_{\mathbf{H}\oplus\mathbf{H}'}$, it follows that there exist constants $C', C'' > 0$ such that

$$|\Lambda_{\mathbf{H}}(K,f_1,f_2,f_3)| \le C'\limsup_{N\to\infty}\|F_1^N\|_{p_1}\|F_2^N\|_{p_2}\|F_3^N\|_{p_3} = C''\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}\,.$$

This completes the proof, up to verifying (2.4.7) and (2.4.8).

To show (2.4.7), we may assume by a base change that A_0 is a diagonal matrix, so that

$$A(x) = \frac{1}{|\Pi_0(x)|^2}\begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots\\ 0 & 0 & \dots & \lambda_{\dim H_0'} & 0 & \dots & 0\\ 0 & 0 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} + \frac{1}{N^2}A_1\,,$$

where $\lambda_1,\dots,\lambda_{\dim H_0'}$ are the nonzero eigenvalues of ${\Pi_0'}^t\Pi_0'$. Then (2.4.7) follows by expanding $\det A(x)$ using the Leibniz formula: The diagonal in A_0 contributes the first term, while the contribution of all other terms is controlled by the $O(N^{-2}|\Pi_0(x)|^2)$ term.

To show (2.4.8), we can assume $N \le C|\Pi_0(x)|$ for a sufficiently large constant C, since otherwise it already follows from (2.4.7). But then we have, since $A_0 \ge 0$

$$\begin{aligned}\det A(x) &\ge \det(\frac{1}{N^2}A_1) = N^{-2\dim H'}\det(A_1)\\ &\ge C^{-2\dim H_0'}|\Pi_0(x)|^{-2\dim H_0'}N^{-2(\dim H'-\dim H_0')}\det(A_1)\,.\end{aligned} \tag{2.4.9}$$

Note that $\det(A_1) > 0$ because $\ker\Pi_1'\cap\ker\Pi_2'\cap\ker\Pi_3' = \{0\}$, which follows from (2.2.2). Taking (2.4.9) to the power $-1/2$ then gives (2.4.8). □

Proof of Theorem 2.1.18. We proceed similarly as in the proof of Theorem 2.1.17. We fix again a bounded $2l$-Calderón-Zygmund kernel K. We further fix n and assume that there exists $C > 0$ such that for each l-Calderón-Zygmund kernel K'

$$|\Lambda_{\mathbf{H}_n}(K', F_1, F_2, F_3)| \le C\|F_1\|_{p_1}\|F_2\|_{p_2}\|F_3\|_{p_3} .$$

Our goal is to show that there exists C' such that for all f_1, f_2, f_3

$$|\Lambda_{\mathbf{H}_{n-1}}(K, f_1, f_2, f_3)| \le C'\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3} .$$

Suppose that the singular Brascamp-Lieb datum $\mathbf{H}_{n-1}$ associated with $\mathbf{M} \oplus \mathbf{J}^{(2)}_{n-1}$ is

$$(H \oplus \mathbb{R}^{n-1} \oplus \mathbb{R}^{n-1}, H_0 \oplus \mathbb{R}^{n-1}, H_1 \oplus \mathbb{R}^{n-1}, H_2 \oplus \mathbb{R}^{n-1}, H_3 \oplus \mathbb{R}^{n-1}, \Pi_0, \Pi_1, \Pi_2, \Pi_3) ,$$

for linear maps Π_i. Comparing the matrices associated with $\mathbf{J}^{(2)}_n$ and $\mathbf{J}^{(2)}_{n-1}$ in Table 2.2 shows that the datum associated with $\mathbf{H}_n$ is then given by

$$(H \oplus \mathbb{R}^n \oplus \mathbb{R}^n, H_0 \oplus \mathbb{R}^n, H_1 \oplus \mathbb{R}^n, H_2 \oplus \mathbb{R}^n , H_3 \oplus \mathbb{R}^n, \Pi'_0, \Pi'_1, \Pi'_2, \Pi'_3) ,$$

where we have, writing $x \in H$, $(y, y_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$, $(z, z_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$:

$$\Pi'_0(x, y, y_n, z, z_n) = (\Pi_0(x, y, z), y_n) ,$$

$$\Pi'_1(x, y, y_n, z, z_n) = (\Pi_1(x, y, z), z_n) ,$$

$$\Pi'_2(x, y, y_n, z, z_n) = (\Pi_2(x, y, z), z_n + y_n) ,$$

$$\Pi'_3(x, y, y_n, z, z_n) = (\Pi_3(x, y, z), z_n + y_{n-1}) .$$

We define for $i = 1, 2, 3$

$$F_i(x, z, z_n) = N^{-1/p_i} \exp\left(-\pi\frac{z_n^2}{N^2}\right) f_i(x, z) ,$$

and we set

$$K'(x, y, y_n) = |(x, y)|^{-1} \exp\left(-\pi\frac{y_n^2}{|(x, y)|^2}\right) K(x, y) .$$

By Lemma 2.4.1, the kernel K' is an l-Calderón-Zygmund kernel. We have that

$$\Lambda_{\mathbf{H}_n}(K', F_1, F_2, F_3) = \int_{H\oplus\mathbb{R}^{n-1}\oplus\mathbb{R}^{n-1}} \prod_{i=1}^{3} f_i(\Pi_i(x, y, z))K(x, y)$$

$$\times \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{|(x, y)|N} \exp\left(-\pi\frac{z_n^2 + (z_n + y_n)^2 + (z_n + y_{n-1})^2}{N^2} - \pi\frac{y_n^2}{|(x, y)|^2}\right) dy_n\, dz_n\, dx\, dy\, dz .$$

The z_n integral can be evaluated by first expanding $(z_n + y_n)^2$, $(z_n + y_{n-1})^2$ and then completing the square in z_n. One obtains that the inner two integrals equal

$$\frac{1}{|(x, y)|}\frac{1}{\sqrt{3}}\int_{\mathbb{R}} \exp\left(-\pi\frac{2(y_{n-1}^2 - y_{n-1}y_n + y_n^2)}{3N^2} - \pi\frac{y_n^2}{|(x, y)|^2}\right) dy_n .$$

This integral is bounded by, and converges by monotone convergence as $N \to \infty$, to

$$\frac{1}{|(x,y)|}\frac{1}{\sqrt{3}}\int_{\mathbb{R}} \exp\left(-\pi\frac{y_n^2}{|(x,y)|^2}\right)\,dy_n = \frac{1}{\sqrt{3}}\,.$$

Using that K is bounded and that f_1, f_2, f_3 are Schwartz functions, we obtain with the dominated convergence theorem

$$\Lambda_{\mathbf{H}_{n-1}}(K, f_1, f_2, f_3) = \lim_{N\to\infty} \sqrt{3}\Lambda_{\mathbf{H}_n}(K', F_1, F_2, F_3)\,.$$

Combined with boundedness of $\Lambda_{\mathbf{H}_n}$ this shows that there exist constants $C', C'' > 0$ with

$$|\Lambda_{\mathbf{H}_{n-1}}(K, f_1, f_2, f_3)| \le C' \limsup_{N\to\infty} \|F_1\|_{p_1}\|F_2\|_{p_2}\|F_3\|_{p_3} = C''\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}\,. \qquad \square$$

2.5 Method of rotations: Proof of Theorem 2.1.19

Fix the dimension $d \ge 3$. We denote $S^{d-1} = \{\theta \in \mathbb{R}^d \,:\, |\theta| = 1\}$, and we denote by σ the normalized $(d-1)$-dimensional Hausdorff probability measure on S^{d-1}. Further, if $\nu \in S^{d-1}$, then we denote by σ_ν the normalized $(d-2)$-dimensional Hausdorff probability measure on the great circle

$$(\operatorname{span}\nu)^\perp \cap S^{d-1}\,.$$

Recall that there is an orthogonal decomposition

$$L^2(S^{d-1}) = \overline{\bigoplus_{n=0}^{\infty}}\mathcal{H}_n\,,$$

where $\mathcal{H}_n$ is the space of spherical harmonics of degree n on S^{d-1}, see e.g. [111, Chapter IV]. Another way to characterize $\mathcal{H}_n$ is as the space of eigenfunctions of the spherical Laplacian corresponding to the eigenvalue $\lambda_n = -n(n+d-2)$.

We will use the spherical Sobolev spaces $H^s(S^{d-1})$ defined by

$$\mathcal{H}^s(S^{d-1}) = \{f \in L^2(S^{d-1}) \,:\, \|f\|^2_{H^s(S^{d-1})} = \sum_{n=0}^{\infty} \lambda_n^s \|\pi_n(f)\|^2_{L^2(S^{d-1})} < \infty\}\,,$$

where π_n denotes the orthogonal projection onto $\mathcal{H}_n$. We will also use the Funk transform, which is the operator T defined a priori on continuous functions F on S^{d-1} by

$$TF(\theta) = \int F(\nu)\,d\sigma_\theta(\nu)\,. \tag{2.5.1}$$

We will need the following properties of the Funk transform.

Lemma 2.5.1. *Let $H_0^s(S^{d-1})$ be the space of functions in the smoothness s Sobolev space on S^{d-1} of mean zero. For all $s \ge 0$, the Funk-transform T extends to a contraction*

$$T \,:\, H_0^s(S^{d-1}) \to H_0^s(S^{d-1})\,, \qquad \|T\|_{H_0^s \to H_0^s} = \frac{1}{d-1} < 1\,.$$

Moreover, for all $s \ge 0$, the operator T extends to a bounded operator

$$T \,:\, H_0^s(S^{d-1}) \to H_0^{s+\delta}(S^{d-1})\,,$$

where $\delta = \frac{d-2}{2}$.

The proof of Lemma 2.5.1 relies on the Funk-Hecke formula.

Lemma 2.5.2 (Funk-Hecke formula)**.** *Denote by ω_m the m-dimensional Hausdorff measure on S^m. Let $f : [-1,1] \to \mathbb{R}$ be a continuous function. Then for every spherical harmonic Y_n of degree n and $\theta \in S^{d-1}$,*

$$\int_{S^{d-1}} f(\nu \cdot \theta) Y_n(\nu)\, d\sigma(\nu) = \frac{\omega_{d-2}}{\omega_{d-1}} \lambda_n Y_n(\theta)\,,$$

where

$$\lambda_n = \int_{-1}^{1} \frac{C_n^{\frac{d-2}{2}}(t)}{C_n^{\frac{d-2}{2}}(1)} f(t)(1-t^2)^{\frac{d-3}{2}}\, dt\,.$$

Here $C_n^k(t)$ denotes the Gegenbauer polynomials, defined via the generating function

$$(1 - 2rt + r^2)^{-k} = \sum_{n \geq 0} C_n^k(t) r^n\,. \tag{2.5.2}$$

Proof. See for example [36], Theorem 1.2.9. □

Proof of Lemma 2.5.1. Let (f_k) be a sequence of continuous functions such that f_k is supported in $(-1/k, 1/k)$ and $\int_{-1/k}^{1/k} f_k(t)\, dt = 1$. A computation in coordinates shows that for every $\theta \in S^{d-1}$

$$\sigma_\theta(\nu) = \frac{\omega_{d-1}}{\omega_{d-2}} \lim_{k \to \infty} f_k(\nu \cdot \theta) \sigma(\nu)\,,$$

in the sense of weak convergence of measures. Applying Lemma 2.5.2 to the sequence f_k and taking limits, we obtain that

$$TY_n = \lambda_n Y_n$$

for every spherical harmonic Y_n of degree n, where

$$\lambda_n = \frac{C_n^{\frac{d-2}{2}}(0)}{C_n^{\frac{d-2}{2}}(1)}\,.$$

We compute the values of $C_n^{\frac{d-2}{2}}$ in 0 and 1 using (2.5.2). Note the identity

$$\frac{1}{(1-x)^k} = \sum_{n=0}^{\infty} \binom{n+k-1}{k-1} x^n\,. \tag{2.5.3}$$

Combining (2.5.2) and (2.5.3), we have

$$C_n^{\frac{d-2}{2}}(0) = \begin{cases} (-1)^{n/2} \binom{\frac{n}{2} + \frac{d-2}{2} - 1}{\frac{d-2}{2} - 1} & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd}\,, \end{cases}$$

and

$$C_n^{\frac{d-2}{2}}(1) = \binom{n+d-3}{d-3}\,.$$

Hence $|\lambda_n|$ clearly vanishes for odd n. For even n we obtain with the duplication formula $\Gamma(z)\Gamma(z+\frac{1}{2}) = \sqrt{\pi}2^{1-2z}\Gamma(2z)$ and Stirling's formula

$$|\lambda_n| = \frac{\Gamma(\frac{n}{2}+\frac{d}{2}-1)\Gamma(n+1)\Gamma(d-2)}{\Gamma(\frac{n}{2}+1)\Gamma(n+d-2)\Gamma(\frac{d}{2}-1)} = 2^{3-d}\frac{\Gamma(d-2)\Gamma(\frac{n+1}{2})}{\Gamma(\frac{d}{2}-1)\Gamma(\frac{n+d-1}{2})} = O(n^{\frac{2-d}{2}})\,. \tag{2.5.4}$$

Thus T maps H_0^s into $H_0^{s+\delta}$, for $\delta = \frac{d-2}{2} > 0$. Equation (2.5.4) combined with logarithmic convexity of the Γ-function also shows that $|\lambda_{2n}|$ is decreasing, so that

$$\|T\|_{H_0^s \to H_0^s} = |\lambda_2| = \frac{\Gamma(\frac{d}{2})\Gamma(3)\Gamma(d-2)}{\Gamma(2)\Gamma(d)\Gamma(\frac{d}{2}-1)} = \frac{1}{d-1}\,.$$

□

We define the manifold of all pairs of orthogonal vectors in S^{d-1}

$$\mathcal{M}_d = \{(\nu,\theta) \in S^{d-1} \times S^{d-1} \,:\, \theta\cdot\nu = 0\}\,.$$

Below we will make use of the fact that the normalized Hausdorff measure on $\mathcal{M}_d$ disintegrates as

$$d\sigma_\theta(\nu)d\sigma(\theta) = d\sigma_\nu(\theta)d\sigma(\nu)\,. \tag{2.5.5}$$

Theorem 2.1.19 is a consequence of the following key proposition.

Proposition 2.5.3. *Let $d \geq 3$ and $s > 1/2$. There exists a constant $C > 0$ such that the following holds. Let $\Omega \in H_0^s(S^{d-1})$. Then there exists a function $\Gamma : \mathcal{M}_d \to \mathbb{C}$ such that*

- *for all $\nu \in S^{d-1}$*

$$\int_{(\operatorname{span}\nu)^\perp \cap S^{d-1}} \Gamma(\nu,\theta)\, d\sigma_\nu(\theta) = 0\,. \tag{2.5.6}$$

  and

$$\|\Gamma(\nu,\cdot)\|_{H_0^{s-1/2}((\operatorname{span}\nu)^\perp\cap S^{d-1})} \leq C\|\Omega\|_{H_0^s(S^{d-1})}\,. \tag{2.5.7}$$

- *as measures, we have*

$$\Omega(\theta)\sigma(\theta) = \int_{S^{d-1}} \Gamma(\nu,\theta)\sigma_\nu(\theta)\, d\sigma(\nu)\,. \tag{2.5.8}$$

Moreover, Γ can be chosen so that the mapping $\Omega \mapsto \Gamma$ is continuous from $C^k(S^{d-1})$ into $C^k(\mathcal{M}_d)$, for every k.

Proposition 2.5.3 says that any mean zero function on S^{d-1} can be decomposed into mean zero functions on slices $S^{d-1} \cap (\operatorname{span}\nu)^\perp$. We will use this later to decompose Calderón-Zygmund kernels on $\mathbb{R}^d$ into kernels on $(\operatorname{span}\nu)^\perp$, $\nu \in S^{d-1}$.

Proof of Proposition 2.5.3. For $F : S^{d-1} \to \mathbb{C}$, we define a candidate solution to (2.5.6), (2.5.8) by

$$\Gamma[F] \,:\, \mathcal{M}_d \to \mathbb{C}, \qquad \Gamma[F](\nu,\theta) = F(\theta) - \int F(\gamma)\, d\sigma_\nu(\gamma)\,.$$

The function $\Gamma[F]$ satisfies (2.5.6) by construction. On the other hand, it satisfies (2.5.8) if and only if, as measures,

$$\Omega(\theta)\sigma(\theta) = \int \left[F(\theta) - \int F(\gamma)\,d\sigma_\nu(\gamma)\right] \sigma_\nu(\theta)\,d\sigma(\nu)\,.$$

Using (2.5.5) and the definition (2.5.1) of T to simplify the second summand, we obtain

$$= F(\theta)\sigma(\theta) - \int TF(\nu)\,d\sigma_\theta(\nu)\,\sigma(\theta) = (F(\theta) - T^2F(\theta))\sigma(\theta)\,.$$

Thus, (2.5.8) holds if

$$\Omega = (1 - T^2)F\,. \tag{2.5.9}$$

Lemma 2.5.1 now implies that $1 - T^2$ is invertible on $H_0^s(S^{d-1})$ for all $s > 1/2$ and hence (2.5.9) can be solved for F for every $\Omega \in H_0^s$, and the solution map is continuous. The function $\Gamma(\nu, \cdot)$ is up to a constant the restriction of F to the codimension one submanifold $S^{d-1} \cap (\operatorname{span}\nu)^\perp$. Since $F \in H_0^s$, we obtain with the trace theorem (2.5.7). Finally, we have for $\Omega \in C^k$ that

$$F = \sum_{l=0}^{\infty} T^{2l}\Omega = \sum_{l=0}^{2} T^{2l}\Omega + \sum_{l=3}^{\infty} T^{2l}\Omega\,.$$

The first three terms on the right hand side are in C^k, since T maps C^k into C^k. By Lemma 2.5.1 we have $T^6\Omega \in H^{k+d}$. Since $\|T\|_{H_0^{k+d}\to H_0^{k+d}} < 1$, the second sum converges in H^{k+d}, which embedds into C^k. Thus the solution map is continuous on C^k. □

We will apply Proposition 2.5.3 to the restriction of a homogeneous Calderón-Zygmund kernel to the sphere S^{d-1}. Since our defining assumptions (2.1.3) on Calderón-Zygmund kernels are formulated on the Fourier side, we need the following lemma to pass to kernels with prescribed smoothness in space.

Lemma 2.5.4 ([111, Chapter IV, Theorem 4.7])**.** *Let $s \geq d$. Let $\Omega \in H_0^s(S^{d-1})$ be a mean zero function on S^{d-1}. Then $m(\xi) = \Omega(\xi/|\xi|)$ defines a homogenous of degree 0 tempered distribution on $\mathbb{R}^d$. The inverse Fourier transform of m is a homogenous of degree $-d$ tempered distribution on $\mathbb{R}^d$ which can be written as*

$$\check{m}(x) = \Omega^*(x/|x|)|x|^{-d}\,.$$

The mapping $\Omega \mapsto \Omega^*$ is bounded with bounded inverse from $H_0^s(S^{d-1})$ into $H_0^{s-d}(S^{d-1})$.

Proof. See Theorem 4.7 in Chapter IV of [111]. □

Proof of Theorem 2.1.19. Let K be a homogenous l-Calderón-Zygmund kernel and let $\Omega : S^{d-1} \to \mathbb{C}$ be the function satisfying

$$\widehat{K}(\xi) = \Omega(\xi/|\xi|) + C_0\,, \qquad \int \Omega(\theta)\,d\sigma(\theta) = 0\,.$$

Since K is an l-Calderón-Zygmund kernel, $\Omega \in C^l(S^{d-1})$, so in particular $\Omega \in H_0^l(S^{d-1})$. By Lemma 2.5.4, the kernel K is then given by

$$K(x) = C_0\delta + \Omega^*(x/|x|)|x|^{-d}\,,$$

and Ω^* satisfies $\|\Omega^*\|_{H_0^{l-d}} \le C\|\Omega\|_{H_0^l}$. We apply Proposition 2.5.3 to Ω^*, using that $l \ge d+1$. We obtain for each $\nu \in S^{d-1}$ a function

$$\Omega^*_\nu(\theta) = \Gamma(\nu,\theta)$$

such that

$$\Omega^*(\theta)\sigma(\theta) = \int_{S^{d-1}} \Omega^*_\nu(\theta)\sigma_\nu(\theta)\,d\sigma(\nu)\,, \tag{2.5.10}$$

and such that

$$\|\Omega^*_\nu\|_{H_0^{l-d-1/2}} \le C\|\Omega^*\|_{H_0^{l-d}} \le C\|\Omega\|_{H_0^l}\,.$$

We define the kernel $K_\theta(x) = C_0\delta + \frac{\omega_{d-1}}{\omega_{d-2}}|x|^{1-d}\Omega^*_\theta(x/|x|)$ on $(\operatorname{span}\theta)^\perp$. By applying Lemma 2.5.4, in the opposite direction, to the function Ω^*_θ, we find that for all $\xi \in (\operatorname{span}\theta)^\perp$

$$\widehat{K}_\theta(\xi) = C_0 + \frac{\omega_{d-1}}{\omega_{d-2}}\Omega_\theta(\xi/|\xi|),$$

for a function Ω_θ with

$$\|\Omega_\theta\|_{H_0^{l-3/2}(S^{d-1}\cap(\operatorname{span}\theta)^\perp} \le C\|\Omega_0\|_{H_0^l(S^{d-1})}\,.$$

Applying finally the Sobolev embedding theorem, we find that there exists a constant $C > 0$ such that for each $\theta \in S^{d-1}$, the kernel $C^{-1}K_\theta$ is an $l - \lceil\frac{d+2}{2}\rceil$-Calderón-Zygmund kernel on $(\operatorname{span}\theta)^\perp$.

From (2.5.10), we obtain for each Schwartz function f on $\mathbb{R}^d$ by integration in polar coordinates:

$$\begin{aligned}
\int f(x)K(x)\,dx &= C_0 f(0) + \omega_{d-1}\int_0^\infty \int_{S^{d-1}} f(r\nu)\Omega^*(\nu)\,d\sigma(\nu)\,\frac{dr}{r}\\
&= C_0 f(0) + \omega_{d-1}\int_0^\infty \int_{S^{d-1}}\int_{S^{d-1}} f(r\nu)\Omega^*_\theta(\nu)\sigma_\theta(\nu)\,d\sigma(\theta)\,\frac{dr}{r}\\
&= \int_{S^{d-1}} C_0 f(0) + \omega_{d-1}\int_0^\infty \int_{S^{d-1}} f(r\nu)\Omega^*_\theta(\nu)\sigma_\theta(\nu)\,\frac{dr}{r}\,d\sigma(\theta)\\
&= \int_{S^{d-1}} C_0 f(0) + \frac{\omega_{d-1}}{\omega_{d-2}}\int_{(\operatorname{span}\theta)^\perp} f(x)\Omega^*_\theta(x/|x|)|x|^{1-d}\,d\mu_\theta(x)\,d\sigma(\theta)\\
&= \int_{S^{d-1}}\int_{(\operatorname{span}\theta)^\perp} f(x)K_\theta(x)\,d\mu_\theta(x)\,d\sigma(\theta)\,.
\end{aligned} \tag{2.5.11}$$

Here μ_θ is the $d-1$-dimensional Lebesgue measure on $(\operatorname{span}\theta)^\perp$. Combining (2.5.11) with Fubini's theorem, it follows that for all Schwartz functions f_1, f_2, f_3

$$\Lambda_{\mathbf{H}}(K, f_1, f_2, f_3) = \int_{S^{d-1}} \Lambda_{\mathbf{H}(\theta)}(K_\theta, f_1, f_2, f_3)\,d\sigma(\theta)\,.$$

Together with the triangle inequality and the assumption of integrable boundedness of the forms on the right hand side, this completes the proof. $\square$

2.6 Proof of Theorem 2.1.22

We will deduce Theorem 2.1.22 from the following multilinear multiplier bound from [56]. Note that the condition (2.6.1) on the multiplier is slightly more general than that obtained by translating the condition on the kernel of the corresponding singular Brascamp-Lieb forms. We will exploit this to deduce also bounds for forms with less singular multiplier.

Theorem 2.6.1 ([56, Theorem 1.1])**.** *Let $2 < p_1, p_2, p_3 < \infty$ with $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$ and fix a dimension $d \geq 1$. There exists $l \in \mathbb{N}$ such that the following holds. Let Γ be the $2d$-dimensional subspace $\{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^{3d} \ : \ \xi_1 + \xi_2 + \xi_3 = 0\}$ of $\mathbb{R}^{3d}$. Furthermore, let $\Gamma' \subseteq \Gamma$ be a d-dimensional subspace, which can be parametrized in terms of each ξ_1, ξ_2 and ξ_3. There exists a constant C such that the following holds. Let $M : \Gamma \to \mathbb{C}$ satisfy*

$$|\partial^\alpha M(\xi)| \leq (\mathrm{dist}(\xi, \Gamma'))^{-|\alpha|}\,, \qquad |\alpha| \leq l\,. \tag{2.6.1}$$

Then

$$\int_\Gamma \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3) M(\xi)\, d\mu_\Gamma(\xi) \leq C\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}\,.$$

Theorem 2.6.1 is a special case of the main result in [56]. It can also be deduced by following the proof of Theorem 2.1.1 in [104]: By a reduction similar to, but slightly more general than in Section 2.2 of [104], Theorem 2.6.1 reduces to a slightly more general version of Lemma 2.3.1 in [104]. The only difference to [104] is that the bump functions $\varphi_{y,\eta,t}$ are possibly different bump functions adapted to position y, frequency η and scale t, as opposed to dilation of a fixed bump function. This causes no issues, because the assumptions on M still guarantee the uniform estimates on $\varphi_{y,\eta,t}$ required in the proof.

Proof of Theorem 2.1.22. Choosing coordinates and swapping the role of f_1 and f_3, we may express the trilinear form $\Lambda_{\mathbf{H}}$ up to a constant as

$$\int f_1(x_1 + y_1, x_2 + C_1^T y_2) f_2(x_1 + B^T y_1, x_2 + C_2^T y_2) f_3(x_1, x_2) K(y_1, y_2)\, dy_1\, dy_2\, dx_1\, dx_2\,. \tag{2.6.2}$$

Here B is the direct sum of the matrices X in the modules $\mathbf{N}_{n_i}$ occuring as direct summands in $\mathbf{M_H}$. In particular, and that is all we will need, B and $I - B$ are invertible. The matrices C_1 and C_2 are direct sums of $I_{m_i}^{\uparrow}$ and $I_{m_i}^{\downarrow}$ respectively, from the direct summands $\mathbf{C}_{m_i}$.

By Fourier inversion, (2.6.2) equals a multiple of

$$\int_\Gamma \widehat{f_1}(\xi_{1,1}, \xi_{1,2})\widehat{f_2}(\xi_{2,1}, \xi_{2,2})\widehat{f_3}(\xi_{3,1}, \xi_{3,2})\widehat{K}(-\xi_{1,1} - B\xi_{2,1}, -C_1\xi_{1,2} - C_2\xi_{2,2})\, d\mu_\Gamma(\xi)\,.$$

Define the singular subspace

$$\Gamma_s = \{\xi \in \Gamma \ : \ \xi_{1,1} + B\xi_{2,1} = 0,\, C_1\xi_{1,2} + C_2\xi_{2,2} = 0\}\,.$$

Define D_1 and D_2 to be the direct sum of $Q_{m_i} I_{m_i}^{\uparrow}$ and $Q_{m_i} I_{m_i}^{\downarrow}$, respectively, where Q_m is the $m \times (m+1)$ matrix

$$Q_m = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ -1 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Then, the space Γ_s sits in the larger subspace

$$\Gamma' = \{\xi \in \Gamma \ : \ \xi_{1,1} + B\xi_{2,1} = 0,\, D_1\xi_{1,2} + D_2\xi_{2,2} = 0\}\,.$$

The space Γ' satisfies the nondegeneracy condition of Theorem 2.6.1, because B and $I - B$ are invertible and because D_1, D_2 and $D_1 - D_2$ are invertible. Indeed, this can be checked blockwise. The blocks in D_1 are of the form $Q_{m_i} I^{\uparrow}_{m_i}$, the blocks in D_2 are of the form $Q_{m_i} I^{\downarrow}_{m_i}$, and the blocks in $D_1 - D_2$ are of the form $Q_{m_i} I^{\uparrow}_{m_i} - Q_{m_i} I^{\downarrow}_{m_i}$, each of which are invertible. We choose

$$M(\xi) = \widehat{K}(-\xi_{1,1} - B\xi_{2,1}, -C_1\xi_{1,2} - C_2\xi_{2,2})\,.$$

Then we have for $|\alpha| \le m$, by (2.1.3)

$$\begin{aligned}
|\partial^\alpha M(\xi)| &\le C_m \sup_{|\beta| \le m} |\partial^\beta \widehat{K}(-\xi_{1,1} - B\xi_{2,1}, -C_1\xi_{1,2} - C_2\xi_{2,2})| \\
&\le C_m(|\xi_{1,1} + B\xi_{2,1}| + |C_1\xi_{1,2} + C_2\xi_{2,2}|)^{-|\alpha|} \\
&\le C(\mathrm{dist}(\xi, \Gamma_s))^{-|\alpha|} \le C(\mathrm{dist}(\xi, \Gamma'))^{-|\alpha|}\,.
\end{aligned}$$

Here the second to last line follows from the fact that both $\mathrm{dist}(\xi, \Gamma_s)$ and the expression

$$|\xi_{1,1} + B\xi_{2,1}| + |C_1\xi_{1,2} + C_2\xi_{2,2}|$$

define norms on the finite dimensional quotient space Γ/Γ_s, and are hence comparable.

Theorem 2.1.22 now follows from Theorem 2.6.1. $\square$

2.7 Proof of Theorem 2.1.23

For $x \in (\mathbb{R}^n)^6$ we write $x = (x^0, x^1)$, where $x^0 = (x^0_1, x^0_2, x^0_3) \in (\mathbb{R}^n)^3$, $x^1 = (x^1_1, x^1_2, x^1_3) \in (\mathbb{R}^n)^3$, and we write $z = (z_1, z_2, z_3) \in (\mathbb{R}^n)^3$. Reading off of Table 2.2, one finds that a singular Brascamp-Lieb datum associated with

$$\mathbf{M} = (J^{(1)}_1 \oplus J^{(2)}_1 \oplus J^{(3)}_1 \oplus C_1)^{\oplus n}$$

is $\mathbf{H}$ with projections $\Pi_1, \Pi_2, \Pi_3 : (\mathbb{R}^n)^9 \to (\mathbb{R}^n)^4$ and $\Pi_0 : (\mathbb{R}^n)^9 \to (\mathbb{R}^n)^5$, given by

$$\Pi_1(x, z) = (x^0_1, x^0_2, x^0_3, z_3)\,,$$

$$\begin{aligned}
\Pi_2(x, z) &= (x^0_1 + x^1_1, x^0_2 + x^1_2, x^0_3, z_2 + z_3)\,, \\
\Pi_3(x, z) &= (x^0_1 + x^1_1, x^0_2, x^0_3 + x^1_3, z_1 + z_3)\,, \\
\Pi_0(x, z) &= (x^1_1, x^1_2, x^1_3, z_1, z_2)\,.
\end{aligned}$$

A datum with the projections $\pi_1, \pi_2, \pi_3 : (\mathbb{R}^n)^9 \to (\mathbb{R}^n)^4$ and $\pi_0 : (\mathbb{R}^n)^9 \to (\mathbb{R}^n)^5$, where

$$\pi_1(x, z) = (x^1_1, x^0_2, x^0_3, z_3)\,,$$

$$\begin{aligned}
\pi_2(x, z) &= (x^0_1, x^1_2, x^0_3, z_2)\,, \\
\pi_3(x, z) &= (x^0_1, x^0_2, x^1_3, z_1)\,,
\end{aligned}$$

$$\pi_0(x,z) = (x^1 - x^0, z_1 - z_3, z_2 - z_3)\,,$$

is equivalent to the datum $\mathbf{H}$. This can be seen by first performing the change of variables $(x^1, z_1, z_2) \to (x^1 - x^0, z_1 - z_3, z_2 - z_3)$ in the singular Brascamp-Lieb form, relabeling (x_1^0, x_1^1) to (x_1^1, x_1^0), and then replacing K by its reflection in the first fiber $K(-\cdot,\cdot,\cdot,\cdot,\cdot)$. Therefore, to prove Theorem 2.1.23, it suffices to show

$$\Big| \int_{(\mathbb{R}^n)^9} \Big(\prod_{j=1}^3 f_j(\pi_j(x,z)) \Big) K(\pi_0(x,z))\, d(x,z) \Big| \le C \prod_{j=1}^3 \|f_j\|_{p_j}$$

whenever $2 < p_1, p_2, p_3 < \infty$, $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$.

We begin with a cone decomposition of $\widehat{K}$. For a function ϕ on $\mathbb{R}^d$, $d \ge 1$, and $t > 0$, let

$$\phi_t(x) = t^{-d}\phi(t^{-1}x)\,.$$

Let $B(0,R)$ denote the Euclidean open ball centered at 0 with radius R. Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a radial Schwartz function with

$$\operatorname{supp}(\widehat{\psi}) \subseteq B(0,1) \setminus B(0,1/4)\,,$$

satisfying $\int_0^\infty \widehat{\psi}(t\xi)\frac{dt}{t} = 1$ for each $\xi \neq 0$. Let $\xi = (\xi_1, \dots, \xi_5) \in (\mathbb{R}^n)^5$ and write

$$\widehat{K}(\xi) = \widehat{K}(\xi) \prod_{j=1}^5 \int_0^\infty \widehat{\psi}(t_j\xi_j)\, \frac{dt_j}{t_j}\,. \tag{2.7.1}$$

We rewrite this as a sum of integrals over five regions, depending on which parameter t_i is the smallest. That is, we write

$$K = \sum_{i=1}^5 K_i\,,$$

where

$$\widehat{K_i}(\xi) = \int_{T_i} \widehat{K}(\xi) \Big(\prod_{j=1}^5 \widehat{\psi}(t_j\xi_j) t_j^{-1} \Big)\, d(t_1, \dots, t_5)\,,$$

and

$$T_i = \{(t_1, t_2, t_3, t_4, t_5) \in (0,\infty)^5 : t_i \le t_j \text{ for } j \neq i\}\,.$$

We denote

$$\widehat{\varphi}(\eta) = \int_1^\infty \widehat{\psi}(s\eta)\, \frac{ds}{s}\,,$$

which is a radial smooth function supported in $B(0,1)$ and for $s_0 > 0$, it holds $\widehat{\varphi_{s_0}}(\eta) = \widehat{\varphi}(s_0\eta) = \int_{s_0}^\infty \widehat{\psi}(s\eta)\frac{ds}{s}$. On each T_i we then integrate over all larger parameters t_j, giving

$$\widehat{K_i}(\xi) = \int_0^\infty \widehat{K}(\xi)\widehat{\psi}(t\xi_i) \Big(\prod_{j=1, j\neq i}^5 \widehat{\varphi}(t\xi_j) \Big) \frac{dt}{t}\,.$$

We will proceed with additional decompositions of the kernels K_i. By symmetry in (2.7.1) and in the projections $\pi_0, \pi_1, \pi_2, \pi_3$, it suffices to consider $i = 1$ and $i = 4$ only.

We begin with $i = 1$. Here we will decompose the kernel further into Gaussian functions, which will be convenient later in the proof. Let g be the Gaussian $g(x) = e^{-\pi|x|^2}$, whose dimension should always be understood from the context. Then we write

$$\widehat{K_1}(\xi) = \int_0^\infty m_1(t\xi, t)|t\xi_1|^2 \widehat{g}_t(\xi_1, \xi_2, \xi_3)^2 \widehat{g}_t(\xi_4, \xi_5, \xi_4 + \xi_5) \, \frac{dt}{t} \, ,$$

where

$$m_1(\xi, t) = \widehat{K}(t^{-1}\xi)\widehat{\psi}(\xi_1)\Big(\prod_{j=2}^{5} \widehat{\varphi}(\xi_j)\Big)(|\xi_1|^2 g(\xi_1, \xi_2, \xi_3)^2 g(\xi_4, \xi_5, \xi_4 + \xi_5))^{-1} \, .$$

Since $|\xi_1|$ is bounded away from zero by the support assumption on $\widehat{\psi}$, the function $m_1(\xi, t)$ is smooth in the ξ variable. Moreover, on the support of $m_1(\xi, t)$, $\xi \in [-1, 1]^{5n}$. Denote

$$c_1(a, t) = (1 + |a|)^{6n} \widetilde{m_1}(a, t) \, ,$$

where the inverse Fourier transform is taken in the ξ variable. A standard integration by parts argument, together with the symbol estimates (2.1.3), gives

$$|c_1(a, t)| \le C_0(1 + |a|)^{-16n} \tag{2.7.2}$$

for an absolute constant C_0, provided $l \ge 22n$. The lower bound on l is chosen crudely such that the decay of the coefficients $c_1(a, t)$ suffices in all of the arguments below. We remark that we do not aim to optimize our arguments to minimize this bound.

Taking the inverse Fourier transform of m, we write

$$\widehat{K_1}(\xi) = \int_{(\mathbb{R}^n)^5} (1 + |a|)^{-6n} \int_0^\infty c_1(a, t) \, |t\xi_1|^2 \widehat{g}_t(\xi_1, \xi_2, \xi_3)^2 \widehat{g}_t(\xi_4, \xi_5, \xi_4 + \xi_5) e^{-2\pi i a \cdot t\xi} \, \frac{dt}{t} \, da \, . \tag{2.7.3}$$

By Fourier inversion,

$$K_1(\pi_0(x, z)) = \int_{(\mathbb{R}^n)^5} \widehat{K_1}(\xi) e^{2\pi i(\xi \cdot \pi_0(x,z))} \, d\xi \, .$$

Using the definition of π_0 and (2.7.3), we can therefore write $K_1(\pi_0(x, z))$ as a superposition of the kernels of the form

$$\sum_{i=1}^{n} \int_0^\infty \int_{(\mathbb{R}^n)^5} c_1(a, t) \widehat{(\partial_i g)_t}(-\xi_1, -\xi_2, -\xi_3) e^{2\pi i x^0 \cdot (-\xi_1, -\xi_2, -\xi_3)} \widehat{(\partial_i g)_t}(\xi_1, \xi_2, \xi_3)$$
$$\times e^{2\pi i(x^1 + t(a_1, a_2, a_3)) \cdot (\xi_1, \xi_2, \xi_3)} \widehat{g}_t(\xi_4, \xi_5, -\xi_4 - \xi_5) e^{2\pi i((z_1 + ta_4)\xi_4 + (z_2 + ta_5)\xi_5 - z_3(\xi_4 + \xi_5))} \, d\xi \, \frac{dt}{t} \, , \tag{2.7.4}$$

weighted by $(1 + |a|)^{-6n}$, where $a = (a_1, \ldots, a_5)$. Here we have also used for convenience that $\partial_i g$ is odd and that g is even, and replaced a by $-a$.

Fixing i and t, the integral in ξ can be viewed as the integral of the function

$$(\eta_1 \ldots, \eta_9) \mapsto \widehat{(\partial_i g)_t}(\eta_1, \eta_2, \eta_3) e^{2\pi i x^0 \cdot (\eta_1, \eta_2, \eta_3)} \widehat{(\partial_i g)_t}(\eta_4, \eta_5, \eta_6) e^{2\pi i(x^1 + t(a_1, a_2, a_3)) \cdot (\eta_4, \eta_5, \eta_6)}$$

$$\times \widehat{g_t}(\eta_7, \eta_8, \eta_9) e^{2\pi i (z_1 + ta_4, z_2 + ta_5, z_3)\cdot(\eta_7,\eta_8,\eta_9)}$$

over the five-dimensional subspace

$$\{(\eta_1 \dots, \eta_9) \in (\mathbb{R}^n)^9 : (\eta_1, \eta_2, \eta_3) = (-\eta_4, -\eta_5, -\eta_6), \eta_9 = -\eta_7 - \eta_8\}\,.$$

It equals the integral of the inverse Fourier transform of this function over the orthogonal complement of this subspace,

$$\{(r_1, \dots, r_9) \in (\mathbb{R}^n)^9 : (r_1, r_2, r_3) = (r_4, r_5, r_6), r_7 = r_8 = r_9\}\,.$$

Therefore, the term for a fixed i in (2.7.4) can be written, up to a constant, as

$$\int_0^\infty \int_{(\mathbb{R}^n)^4} c_1(a,t) (\partial_i \partial_{3n+i} g)_t((x,z) + r^\star + (0,0,0,ta,0))\, dr\, \frac{dt}{t}\,, \tag{2.7.5}$$

where $r = (r_1, r_2, r_3, r_4)$ and $r^\star = (r_1, r_2, r_3, r_1, r_2, r_3, r_4, r_4, r_4)$. Thus, it suffices to bound the form in which $K_1(\pi_0(x,z))$ is replaced by (2.7.5), with estimates uniform in a. Then it remains to sum over i and integrate in a. By symmetry, it suffices to prove bounds for $i = 1$. This will be done in Proposition 2.7.1.

Next we decompose the kernel K_4. We write it as $K_4 = K_6 + K_7$, where

$$\widehat{K_6}(\xi) = \int_0^\infty \widehat{K}(\xi) \widehat{\varphi}_t(\xi_1) \widehat{\varphi}_t(\xi_2) \widehat{\varphi}_t(\xi_3) \widehat{\psi}_t(\xi_4) \widehat{\varphi_{2^{10}t}}(\xi_5)\, \frac{dt}{t}\,,$$

$$\widehat{K_7}(\xi) = \int_0^\infty \widehat{K}(\xi) \widehat{\varphi}_t(\xi_1) \widehat{\varphi}_t(\xi_2) \widehat{\varphi}_t(\xi_3) \widehat{\psi}_t(\xi_4) (\widehat{\varphi}_t - \widehat{\varphi_{2^{10}t}})(\xi_5)\, \frac{dt}{t}\,.$$

Note that if ξ is in the support of the integrand in $\widehat{K_6}$ for a fixed t, then $2^{-3} < |t\xi_4 + t\xi_5| < 2^2$. On the other hand, if ξ is in the support of the integrand of $\widehat{K_7}$ for a fixed t, then $0 < |t\xi_4 + t\xi_5| < 2^2$, but $|t\xi_5| \geq 2^{-12}$ We will decompose these multiplier symbols further. To reduce the amount of notation we will use Gaussians for this decomposition as well, even though one could proceed with other Schwartz functions.

Now we write

$$\widehat{K_6}(\xi) = \int_{(\mathbb{R}^n)^5} (1+|a|)^{-6n} \int_0^\infty c_6(a,t) |t\xi_4|^2 \widehat{g_t}(\xi) |t\xi_4 + t\xi_5|^2 \widehat{g_t}(\xi_4 + \xi_5) e^{-2\pi i a\cdot t\xi}\, \frac{dt}{t}\, da\,,$$

where $c_6(a,t) = (1+|a|)^{6n} \widetilde{m_6}(a,t)$ and

$$m_6(\xi,t) = \widehat{K}(t^{-1}\xi) \widehat{\varphi}(\xi_1) \widehat{\varphi}(\xi_2) \widehat{\varphi}(\xi_3) \widehat{\psi}(\xi_4) \widehat{\varphi_{2^{10}}}(\xi_5) (|t\xi_4|^2 \widehat{g_t}(\xi) |t\xi_4 + t\xi_5|^2 \widehat{g_t}(\xi_4 + \xi_5))^{-1}\,.$$

Here, $c_6(a,t)$ satisfy the symbol estimate analogous to (2.7.2).

Thus, $K_6(\pi_0(x,z))$ can be written as a weighted superposition of integrals of the form

$$\int_0^\infty \int_{(\mathbb{R}^n)^5} c_6(a,t) \widehat{g_t}(\xi_1, \xi_2, \xi_3) \widehat{(\Delta g)_t}(\xi_4) \widehat{g_t}(\xi_5) \widehat{(\Delta g)_t}(\xi_4 + \xi_5)$$

$$\times e^{2\pi i (x^1 - x^0, z_1 - z_3, z_2 - z_3)\cdot\xi} e^{2\pi i t a\cdot\xi}\, d\xi\, \frac{dt}{t}$$

$$= \int_0^\infty \int_{(\mathbb{R}^n)^5} c_6(a,t) \widehat{g_t}(\xi_1, \xi_2, \xi_3) \widehat{(\Delta g)_t}(\xi_4) \widehat{g_t}(\xi_5) \widehat{(\Delta g)_t}(\xi_4 + \xi_5) e^{2\pi i (x^1 - x^0 + t(a_1,a_2,a_3))\cdot(\xi_1,\xi_2,\xi_3)}$$

$$\times e^{2\pi i(z_1+ta_4,z_2+ta_5,z_3)\cdot(\xi_4,\xi_5,-\xi_4-\xi_5)}\,d\xi\,\frac{dt}{t}\,. \tag{2.7.6}$$

Then, the integral in ξ_4,ξ_5 can be seen as the integral of the function

$$(\eta_1,\eta_2,\eta_3)\mapsto \widehat{(\Delta g)_t}(\eta_1)\widehat{g_t}(\eta_2)\widehat{(\Delta g)_t}(\eta_3)e^{2\pi i(z_1+ta_4,z_2+ta_5,z_3)\cdot(\eta_1,\eta_2,\eta_3)}$$

over the subspace

$$\{(\eta_1,\eta_2,\eta_3)\in(\mathbb{R}^n)^3:\eta_3=-\eta_1-\eta_2\}\,.$$

It equals the integral of the Fourier transform of this function over the orthogonal complement

$$\{(r_1,r_2,r_3)\in(\mathbb{R}^n)^3:r_1=r_2=r_3\}\,.$$

Using this and taking the inverse Fourier transform in ξ_1,ξ_2,ξ_3, the display (2.7.6) is up to a constant equal to

$$\int_0^\infty\int_{\mathbb{R}^n}c_6(a,t)g_t(x^1-x^0+t(a_1,a_2,a_3))(\Delta g)_t(z_1+r+ta_4)g_t(z_2+r+ta_5)(\Delta g)_t(z_3+r)\,dr\,\frac{dt}{t}\,.$$

Thus, it suffices to bound a form with $K_6(\pi_0(x,z))$ replaced by this kernel, with a constant uniform in a. This will follow from Proposition 2.7.2.

For the kernel K_7 we proceed with a similar decomposition but with a factor $|t\xi_5|^2$ instead of $|t\xi_4+t\xi_5|^2$. This leads to bounding a form with $K_7(\pi_0(x,z)))$ replaced by

$$\int_0^\infty\int_{\mathbb{R}^n}c_7(a,t)g_t(x^1-x^0+t(a_1,a_2,a_3))(\Delta g)_t(z_1+r+ta_4)(\Delta g)_t(z_2+r+ta_5)g_t(z_3+r)\,dr\,\frac{dt}{t}\,.$$

with a constant uniform in a, where $c_7(a,t)$ satisfies a bound analogous to (2.7.2). Note a symmetry between the last two displays, which can be seen by interchanging z_2 and z_3, translating $r\to r-ta_5$, and replacing a_4-a_5 by a_4 in the second display. Bounds for this form will also follow from Proposition 2.7.2.

To summarize, we have reduced Theorem 2.1.23 to the following two propositions.

Proposition 2.7.1. *Let $n\geq 1$. Let $2<p_1,p_2,p_3<\infty$ and $\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}=1$. There exists a constant $C>0$ such that for each $a\in(\mathbb{R}^n)^9$, $c(t)$ satisfying $|c(t)|\leq 1$ for each $t>0$, and all Schwartz functions $f_1,f_2,f_3:(\mathbb{R}^n)^4\to\mathbb{C}$,*

$$\Big|\int_0^\infty c(t)\int_{(\mathbb{R}^n)^{13}}\Big(\prod_{j=1}^3 f_j(\pi_j(x,z))\Big)(\partial_1\partial_{3n+1}g)_t((x,z)+r^\star+ta)\,d(x,z,r)\,\frac{dt}{t}\Big|$$

$$\leq C(1+|a|)^{8n}\prod_{j=1}^3\|f_j\|_{p_j}\,,$$

where $r=(r_1,r_2,r_3,r_4)$, $r^\star=(r_1,r_2,r_3,r_1,r_2,r_3,r_4,r_4,r_4)$.

Proposition 2.7.2. *Let $n\geq 1$. Let $1<p_1,p_2,p_3<\infty$ and $\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}=1$. There exists a constant $C>0$ such that for each $a\in(\mathbb{R}^n)^6$, $c(t)$ satisfying $|c(t)|\leq 1$ for each $t>0$, and all Schwartz functions $f_1,f_2,f_3:(\mathbb{R}^n)^4\to\mathbb{C}$,*

$$\Big|\int_0^\infty c(t)\int_{(\mathbb{R}^n)^{10}}\Big(\prod_{j=1}^3 f_j(\pi_j(x,z))\Big)g_t(x^1-x^0+t(a_1,a_2,a_3))(\Delta g)_t(z_1+r+ta_4)$$

$$\times g_t(z_2+r+ta_5)(\Delta g)_t(z_3+r+ta_6)\,d(x,z,r)\,\frac{dt}{t}\Big| \le C(1+|a|)^{16n}\prod_{j=1}^{3}\|f_j\|_{p_j}\,.$$

Note that in Proposition 2.7.2, one could decompose the Gaussian in x^1-x^0 into another integral of translated Gaussians. Also, the Laplacians could be split into sums of second order derivatives. This would yield a form that is, schematically, similar to the one in Proposition 2.7.1. However, we chose not to do this as, it will not be needed for the proof.

Proposition 2.7.1 will be proven using twisted techniques, while Proposition 2.7.2 will follow from the classical square and maximal function bounds, similarly as in the case of the Coifman-Meyer multipliers. We will prove these propositions in the following two sections.

2.7.1 Proof of Proposition 2.7.2

Denote $h=\Delta g$. Using the definition of the projections π_j and splitting the Gaussian into tensor products of three lower-dimensional Gaussians, the form we need to bound reads

$$\int_0^\infty c(t)\int_{(\mathbb{R}^n)^{10}} f_1(x_1^1,x_2^0,x_3^0,z_3)f_2(x_1^0,x_2^1,x_3^0,z_2)f_3(x_1^0,x_2^0,x_3^1,z_1)$$

$$\times\Big(\prod_{j=1}^{3}g_t(x_j^1-x_j^0+ta_j)\Big)h_t(z_1+r+ta_4)g_t(z_2+r+ta_5)h_t(z_3+r+a_6)\,d(x,z,r)\,\frac{dt}{t}\,.$$

Integrating in x_1 and z, using that Gaussians are even and replacing a_1,a_2,a_3 by $-a_1$, $-a_2$, $-a_3$, it suffices to estimate

$$\Big|\int_0^\infty c(t)\int_{(\mathbb{R}^n)^4}(f_1*_{1,4}(g_{t,a_1}\otimes h_{t,a_6}))(x^0,r)(f_2*_{2,4}(g_{t,a_2}\otimes g_{t,a_5}))(x^0,r)$$

$$\times(f_3*_{3,4}(g_{t,a_3}\otimes h_{t,a_4}))(x^0,r)\,d(x^0,r)\,\frac{dt}{t}\Big|\,,$$

where the subscript $*_{m_1,m_2}$ means that we take $2n$-dimensional convolutions with the functions f_j in the coordinates $m_1n,\dots,m_1(n+1)$ and $m_2n,\dots,m_2(n+1)$. Here, all functions g and h that appear in the tensor products are n-dimensional, we have denoted

$$g_{t,a_j}=t^{-n}g_t(\cdot+ta_j)\,,$$

and analogously for h_{t,a_j}. For two functions ϕ,ρ we also write $(\phi\otimes\rho)(u,v)=\phi(u)\rho(v)$. Applying Hölder's inequality in t for the exponents $(2,\infty,2)$ and using $|c(t)|\le 1$, we bound the last display by

$$\int_{(\mathbb{R}^n)^4}\Big(\int_0^\infty |(f_1*_{1,4}(g_{t,a_1}\otimes h_{t,a_6}))(x^0,r)|^2\frac{dt}{t}\Big)^{1/2}\sup_{t>0}|(f_2*_{2,4}(g_{t,a_2}\otimes g_{t,a_5}))(x^0,r)|$$

$$\times\Big(\int_0^\infty |(f_3*_{3,4}(g_{t,a_3}\otimes h_{t,a_4}))(x^0,r)|^2\frac{dt}{t}\Big)^{1/2}d(x^0,r)\,.$$

Applying Hölder's inequality in (x^0,p), we estimate this further by

$$\Big\|\Big(\int_0^\infty |f_1*_{1,4}(g_{t,a_1}\otimes h_{t,a_6})|^2\frac{dt}{t}\Big)^{1/2}\Big\|_{p_1}\Big\|\sup_{t>0}|f_2*_{2,4}(g_{t,a_2}\otimes g_{t,a_5})|\Big\|_{p_2}$$

$$\times \Big\| \Big(\int_0^\infty |f_3 *_{3,4} (g_{t,a_3} \otimes h_{t,a_4})|^2 \frac{dt}{t} \Big)^{1/2} \Big\|_{p_3}.$$

Using bounds for the two-dimensional fiber-wise maximal and square functions, the last display is bounded by an absolute constant times

$$(1+|a|)^{6n+2} \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3}.$$

For the maximal function, a polynomial loss in the shift follows by dominating the Gaussian $g_{t,a_j} \leq 10(2(1+|a_j|))^n g_{2t(1+|a_j|)}$. For the shifted square function, we have a uniform bound on L^2 in the shift a. We also have a weak L^1 bound with polynomial loss $(1+|a_j|)^{n+1}$, this follows from the standard proof of weak L^1 bounds of Calderón-Zygmund operators, but with an $L^2(\frac{dt}{t})$-vector valued kernel, as in [61, Section 5.6.1]. Interpolation and duality then give the polynomial loss $(1+|a_j|)^{n+1}$.

2.7.2 Proof of Proposition 2.7.1

In contrast with the previous section, now we cannot bound the form by the maximal and square functions of each of the functions f_j separately.

To prove Proposition 2.7.1, we will first prove an estimate for a local version of our form. Local estimates will then be combined into a global estimate using a stopping-time argument. A finite collection $\mathcal{T}$ of dyadic cubes in $\mathbb{R}^d$, $d \geq 1$, is called a convex tree if there exists $Q_{\mathcal{T}} \in \mathcal{T}$ such that $Q \subseteq Q_{\mathcal{T}}$ for every $Q \in \mathcal{T}$ and if $Q, Q'' \in \mathcal{T}$ and $Q \subseteq Q' \subseteq Q''$, then $Q' \in \mathcal{T}$. If $\ell(Q)$ denotes the side-length of a dyadic cube Q, we denote

$$\Omega_{\mathcal{T}} = \cup_{Q \in \mathcal{T}} Q \times (\ell(Q)/2, \ell(Q)).$$

For $f \in L^2_{\text{loc}}(\mathbb{R}^d)$ we also define a variant of a maximal operator on a tree $\mathcal{T}$

$$Mf(\mathcal{T}) = \sup_{Q \in \mathcal{T}} \sup_{Q' \supseteq Q} \Big(\frac{1}{|Q'|} \int_{Q'} |f|^2 \Big)^{1/2},$$

where the second supremum is over all cubes Q' with sides parallel to the coordinate axes, which contain the cube Q.

Let $c(t)$ and π_1, π_2, π_3 be as in Proposition 2.7.1. Let $\pi_4 : (\mathbb{R}^n)^9 \to (\mathbb{R}^n)^4$ be given by

$$\pi_4(x,z) = (x_1^1, x_2^1, x_3^1, z_2).$$

We will prove bounds for a more symmetric local quadrilinear form

$$\Lambda_{\mathcal{T},a}(f_1,f_2,f_3,f_4) = (1+|a|)^{-16n} \int_{\Omega_{\mathcal{T}}} c(t) \int_{(\mathbb{R}^n)^9} \Big(\prod_{j=1}^4 f_j(\pi_j(x,z)) \Big)$$

$$\times (\partial_1 \partial_{3n+1} g)_t((x,z) + r^\star + ta)\, d(x,z)\, dr\, \frac{dt}{t},$$

defined for bounded functions f_j on $\mathbb{R}^{4n}$ and a convex tree $\mathcal{T}$ in $\mathbb{R}^{4n}$. The main step in the proof of Proposition 2.7.1 is the following estimate, which will be applied with $f_4 = 1$.

Proposition 2.7.3. *There exists a constant $C > 0$ such that for any convex tree $\mathcal{T}$, any $a \in (\mathbb{R}^n)^9$, and bounded functions $f_1, f_2, f_3, f_4 : (\mathbb{R}^n)^4 \to \mathbb{C}$,*

$$|\Lambda_{\mathcal{T},a}(f_1, f_2, f_3, f_4)| \le C|Q_{\mathcal{T}}| \prod_{i=1}^{4} Mf_i(\mathcal{T}) .$$

Proof of Proposition 2.7.3. We may assume that the functions f_j are real-valued, as otherwise we split them into real and imaginary parts. Interchanging the order of integration, using the triangle inequality, and $|c(t)| \le 1$, we bound $|\Lambda_{\mathcal{T},a}(f_1, f_2, f_3, f_4)|$ by

$$(1+|a|)^{-16n} \int_{\Omega_{\mathcal{T}}} \int_{(\mathbb{R}^n)^7} \Big| \int_{\mathbb{R}^n} f_1(x_1^1, x_2^0, x_3^0, z_3) f_4(x_1^1, x_2^1, x_3^1, z_2)(\partial_1 g)_t(x_1^1 + r_1 + a_1 t)\, dx_1^1 \Big|$$

$$\times \Big| \int_{\mathbb{R}^n} f_2(x_1^0, x_2^1, x_3^0, z_2) f_3(x_1^0, x_2^0, x_3^1, z_1)(\partial_1 g)_t(x_1^0 + r_1 + a_4 t)\, dx_1^0 \Big|\, d\mu ,$$

where

$$d\mu = g_t((x_2^0, x_3^0, x_2^1, x_3^1) + (r_2, r_3, r_2, r_3) + t(a_2, a_3, a_5, a_6), z + (r_4, r_4, r_4) + t(a_7, a_8, a_9))$$

$$\times d(x_2^0, x_3^0, x_2^1, x_3^1, z)\, dr\, \frac{dt}{t} .$$

Applying the Cauchy-Schwarz inequality with respect to $d\mu$ bounds this form by $(1 + |a|)^{-16n}$ times the geometric mean of

$$\int_{\Omega_{\mathcal{T}}} \int_{(\mathbb{R}^n)^7} \Big| \int_{\mathbb{R}^n} f_1(x_1^1, x_2^0, x_3^0, z_3) f_4(x_1^1, x_2^1, x_3^1, z_2)(\partial_1 g)_t(x_1^1 + r_1 + a_1 t) dx_1^1 \Big|^2 d\mu . \tag{2.7.7}$$

and

$$\int_{\Omega_{\mathcal{T}}} \int_{(\mathbb{R}^n)^7} \Big| \int_{\mathbb{R}^n} f_2(x_1^0, x_2^1, x_3^0, z_2) f_3(x_1^0, x_2^0, x_3^1, z_1)(\partial_1 g)_t(x_1^0 + r_1 + a_4 t) dx_1^0 \Big|^2 d\mu .$$

These two terms are analogous, which can be seen by swapping the roles of x_3^0 and x_3^1 in the second term. Thus, it suffices to proceed with (2.7.7).

We integrate in z_1 and then expand out the square in (2.7.7). This gives

$$\int_{\Omega_{\mathcal{T}}} \int_{(\mathbb{R}^n)^8} f_1(x_1^0, x_2^0, x_3^0, z_3) f_4(x_1^0, x_2^1, x_3^1, z_2) f_1(x_1^1, x_2^0, x_3^0, z_3) f_4(x_1^1, x_2^1, x_3^1, z_2)$$

$$\times (\partial_1 g)_t((x^0, z_2) + r + u_1 t)(\partial_1 g)_t((x^1, z_3) + r + u_2 t) d(x, z_2, z_3)\, dr\, \frac{dt}{t} , \tag{2.7.8}$$

where $u_1 = (a_1, a_2, a_3, a_8)$, $u_2 = (a_1, a_5, a_6, a_9)$.

If $n = 1$, this expression corresponds to the local form $\widetilde{\Lambda}_{\mathcal{T}}$ from the dimension four case in [45]. More precisely, it can be interpreted as the local form applied to a 16-tuple of functions on $\mathbb{R}^4$, in the case of the identity matrix, and when all but four functions are set to the constant 1. Applying the result from [45] to this setup yields a variant of Proposition 2.7.3, where the maximal operators Mf_i are replaced by $(M|f_i|^4)^{1/4}$. However, this is insufficient to establish Proposition 2.7.1. It is therefore essential to view (2.7.8) as a variant of the quadrilinear form from the two-dimensional case in [45], but acting on functions defined on $\mathbb{R}^n \times \mathbb{R}^{3n}$ instead of $\mathbb{R} \times \mathbb{R}$. The variables are now

$$x_1^0, x_1^1 \in \mathbb{R}^n, \quad (x_2^0, x_3^0, z_3), (x_2^1, x_3^1, z_2) \in \mathbb{R}^{3n}\,.$$

This perspective leads to only one more application of the Cauchy-Schwarz inequality, which subsequently gives the maximal operators Mf_i and the desired Hölder estimate.

The paper [45] establishes an estimate for this quadrilinear form in the setting of functions on $\mathbb{R} \times \mathbb{R}$. The argument, however, extends to $\mathbb{R}^n \times \mathbb{R}^{3n}$ without significant complications. Since [45] additionally focuses on other matters, we nevertheless prove the estimate in our specific setting for the reader's convenience and to maintain a self-contained exposition.

We will rewrite the form (2.7.8) more concisely, for which we introduce slightly more general expressions. Let $d_1, d_2 \geq 1$. For $y \in (\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})^2$ we write $y = (y^0, y^1)$, where $y^0 = (y_1^0, y_2^0) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, $y^1 = (y_1^1, y_2^1) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, and let $q = (q_1, q_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. We write $m = d_1 + d_2$ and identify $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ with $\mathbb{R}^m$. Define the maps $\rho_j : (\mathbb{R}^m)^2 \to \mathbb{R}^m$ by

$$\rho_1(y) = (y_1^0, y_2^0)\,, \quad \rho_3(y) = (y_1^0, y_2^1)\,,$$
$$\rho_2(y) = (y_1^1, y_2^0)\,, \quad \rho_4(y) = (y_1^1, y_2^1)\,.$$

For bounded functions $F_1, \dots, F_4$ on $\mathbb{R}^m$, $v \in (\mathbb{R}^m)^2$, and $\mathcal{T}$ a convex tree in $\mathbb{R}^m$, we define

$$\Theta_{\mathcal{T},v}(F_1, F_2, F_3, F_4) = (1 + |v|)^{-4m} \int_{\Omega_{\mathcal{T}}} \int_{(\mathbb{R}^m)^2} \Big(\prod_{j=1}^4 F_j(\rho_j(y)) \Big)$$
$$\times (\partial_1 \partial_{m+1} g)_t(y + (q,q) + vt)\, dy\, dq\, \frac{dt}{t}\,.$$

Then, (2.7.8) can be recognized as

$$(1 + |u|)^{4m} \Theta_{\mathcal{T},u}(f_1, f_1, f_4, f_4)$$

with $u = (u_1, u_2)$, $d_1 = n$, $d_2 = 3n$. Recall that u consists of the components of a and satisfies $|u| \leq C|a|$. Thus, it will suffice to prove

$$|\Theta_{\mathcal{T},u}(f_1, f_1, f_4, f_4)| \leq C |Q_{\mathcal{T}}| Mf_1(\mathcal{T})^2 Mf_4(\mathcal{T})^2\,. \tag{2.7.9}$$

To show this estimate, we will remove the localization of the kernel and localize the functions, similarly as in [45]. This will allow for translation $q \to q - vt$ and global telescoping arguments. For a tree $\mathcal{T}$ in $\mathbb{R}^m$ we define a region in $\mathbb{R}^m$

$$T_k = \cup\{Q \in \mathcal{T} : \ell(Q) = 2^k\}.$$

For $F_j, \mathcal{T}, v$ as above, $\alpha \geq 1$, and $1 \leq i \leq m$, we define

$$\Theta^{(i)}_{\mathcal{T},v,\alpha}(F_1, F_2, F_3, F_4) = (\alpha + |v|)^{-4m} \sum_{k \in \mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{\mathbb{R}^m} \int_{(\mathbb{R}^m)^2} \Big(\prod_{j=1}^4 (F_j 1_{T_k})(\rho_j(y)) \Big)$$
$$\times (\partial_i \partial_{m+i} g)_{Dt}(y + (q,q) + vt)\, dy\, dq\, \frac{dt}{t}\,,$$

where $D = D(\alpha)$ is a $2m \times 2m$ diagonal matrix with diagonal entries $d_{ll} = 1$ if $1 \leq l \leq d_1$ or $d_1 + d_2 < l \leq 2d_1 + d_2$, and $d_{ll} = \alpha$ otherwise. Here, $g_{Dt} = (\det D)^{-1} g_t(D^{-1} \cdot)$. We will only use this expression when either $\alpha = 1$ or $v = 0$.

The following lemma will reduce the problem to proving a bound for $\Theta^{(1)}_{\mathcal{T},u,1}$ instead.

Lemma 2.7.4. *There exists a constant $C > 0$ such that for any $F_j, \mathcal{T}, v$ as above,*

$$|\Theta_{\mathcal{T},v}(F_1, F_2, F_3, F_4) - \Theta^{(1)}_{\mathcal{T},v,1}(F_1, F_2, F_3, F_4)| \le C|Q_{\mathcal{T}}| \prod_{i=1}^{4} MF_i(\mathcal{T})\,.$$

We will use another lemma, which will reduce the problem of bounding $\Theta^{(1)}_{\mathcal{T},0,\alpha}$ to bounding the sum of $\Theta^{(i)}_{\mathcal{T},0,\alpha}$, $i \neq 1$, instead.

Lemma 2.7.5. *There exists a constant $C > 0$ such that for any $F_j, \mathcal{T}, \alpha$ be as above,*

$$\Big|\sum_{i=1}^{m} \Theta^{(i)}_{\mathcal{T},0,\alpha}(F_1, F_2, F_3, F_4)\Big| \le C|Q_{\mathcal{T}}| \prod_{i=1}^{4} MF_i(\mathcal{T})\,.$$

We postpone the proofs of these two lemmas until the end of this section and return to proving (2.7.9).

By Lemma 2.7.4, it thus suffices to show

$$|\Theta^{(1)}_{\mathcal{T},u,1}(f_1, f_1, f_4, f_4)| \le C|Q_{\mathcal{T}}| Mf_1(\mathcal{T})^2 Mf_4(\mathcal{T})^2\,.$$

Note that we can write $\Theta^{(1)}_{\mathcal{T},u,1}(f_1, f_1, f_4, f_4)$ in an analogous way as in (2.7.7), by writing the product of all non-positive terms as a square. Indeed, we can write is as

$$(1+|u|)^{-4m} \sum_{k\in\mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{\mathbb{R}^{d_1+3d_2}} \Big| \int_{\mathbb{R}^{d_1}} (f_1 1_{T_k})(\rho_1(y))(f_4 1_{T_k})(\rho_3(y))(\partial_1 g)_t(y_1^0 + q_1 + a_1 t)\, dy_1^0 \Big|^2$$

$$\times g_t((y_2^0, y_2^1) + (q_2, q_2) + (a_2, a_3, a_8, a_5, a_6, a_9)t)\, d(y_2^0, y_2^1, q)\, \frac{dt}{t}\,,$$

where we have unravelled the definition of u inside the Gaussian. We estimate a non-centered Gaussian by a centered Gaussian as

$$g_t(\cdot + vt) \le 10(2(1+|v|))^{2d_2} g_{2t(1+|v|)}\,,$$

and apply this to the Gaussian outside of the squared term. We also change variables $q_1 \to q_1 - u_1 t$. This gives a constant multiple of the form

$$\alpha^{2d_2}(1+|u|)^{-4m}\Theta^{(1)}_{\mathcal{T},0,\alpha}(f_1, f_1, f_4, f_4)$$

with $\alpha = 2(1 + |(a_2, a_3, a_8, a_5, a_6, a_9)|)$. Since $\alpha^{2d_2}(1+|u|)^{-4m} \le C$, it will suffice to show

$$\Theta^{(1)}_{\mathcal{T},0,\alpha}(f_1, f_1, f_4, f_4) \le C|Q_{\mathcal{T}}| Mf_1(\mathcal{T})^2 Mf_4(\mathcal{T})^2\,.$$

Note that by symmetry, $\Theta^{(i)}_{\mathcal{T},0,\alpha}(f_1, f_1, f_4, f_4) \ge 0$ for each $1 \le i \le d_1$. By Lemma 2.7.5, it will thus suffice to prove

$$\Big|\sum_{i=d_1+1}^{m} \Theta^{(i)}_{\mathcal{T},0,\alpha}(f_1, f_1, f_4, f_4)\Big| \le C|Q_{\mathcal{T}}| Mf_1(\mathcal{T})^2 Mf_4(\mathcal{T})^2\,.$$

To show this inequality, we bound the left-hand side of the last display by

$$\sum_{i=d_1+1}^{m} \sum_{k\in\mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{\mathbb{R}^{3d_1+d_2}} \Big| \int_{\mathbb{R}^{d_2}} (f_1 1_{T_k})(y_1^0, y_2^0)(f_1 1_{T_k})(y_1^1, y_2^0)(\partial_i g)_{\alpha t}(y_2^0 + q_2)\, dy_2^0 \Big|$$

$$\times \Big| \int_{\mathbb{R}^{d_2}} (f_4 1_{T_k})(y_1^0, y_2^1)(f_4 1_{T_k})(y_1^1, y_2^1)(\partial_i g)_{\alpha t}(y_2^1 + q_2) dy_2^1 \Big| g_t((y_1^0, y_1^1) + (q_1, q_1))\, d(y_1^0, y_1^1, q)\, \frac{dt}{t}\,.$$

We apply the Cauchy-Schwarz inequality in y_1^0, y_1^1, q, t and in the sums, and then expand out the square, similarly as we did in (2.7.8). This gives an estimate by

$$\prod_{j\in\{1,4\}} \Big(\sum_{i=d_1+1}^{m} \Theta^{(i)}_{\mathcal{T},0,\alpha}(f_j, f_j, f_j, f_j)\Big)^{1/2}.$$

For each $1 \le i \le m$, we have $\Theta^{(i)}_{\mathcal{T},0,\alpha}(f_j, f_j, f_j, f_j) \ge 0$. Therefore, Lemma 2.7.5 gives

$$\Theta^{(i)}_{\mathcal{T},0,\alpha}(f_j, f_j, f_j, f_j) \le C|Q_{\mathcal{T}}| M f_j(\mathcal{T})^4$$

for each $1 \le i \le m$ and $1 \le j \le 4$. This finishes the proof of Proposition 2.7.3, up to verification of Lemmas 2.7.4 and 2.7.5. □

To prove Lemmas 2.7.4 and 2.7.5 we will need the following Brascamp-Lieb inequality.

Lemma 2.7.6. *For any measurable functions $F_1, F_2, F_3, F_4 : \mathbb{R}^m \to \mathbb{C}$,*

$$\Big| \int_{(\mathbb{R}^m)^2} \Big(\prod_{j=1}^{4} F_j(\rho_j(y))\Big) dy \Big| \le \prod_{j=1}^{4} \|F_j\|_2\,.$$

This lemma was proven in the case $d = 1$ by repeated applications of the Cauchy-Schwarz inequality in [45, Lemma 3.2]. The proof when the variables are in higher dimensions follows in the analogous way and we omit it.

We will also need an estimate on the boundary of a convex tree from [45].

Lemma 2.7.7 ([45, Lemma 4.1])**.** *There exists a constant $C > 0$ such that for any convex tree $\mathcal{T}$ in $\mathbb{R}^m$,*

$$\sum_{k\in\mathbb{Z}} 2^{mk} \#(\partial T_k \cap (2^k\mathbb{Z})^m) \le C|Q_{\mathcal{T}}|\,.$$

Now we are ready to prove Lemmas 2.7.4 and 2.7.5.

Proof of Lemma 2.7.4. We proceed along the lines of the argument in [45, Section 4] in the case of the identity matrix. First we use [45, Lemma 3.4], which in the particular case of the identity matrix gives

$$\big|(1 + |v|)^{-4m} (\partial_1 \partial_{m+1} g)_t(y + (q, q) + vt)\big| \le C t^{-2m} \sum_{n\ge 0} 2^{-4nm} \chi(|y + (q, q)| \le 2^n t)\,,$$

where $\chi(\mathcal{A})$ equals 1 if the condition $\mathcal{A}$ is satisfied and 0 otherwise. With this and (2.7.2), we estimate

$$|\Theta_{\mathcal{T},v}(F_1, F_2, F_3, F_4) - \Theta^{(1)}_{\mathcal{T},v,1}(F_1, F_2, F_3, F_4)|$$

by an absolute constant times

$$\sum_{n\geq 0} 2^{-4mn} \sum_{k\in\mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{T_k} \int_{S_k^c} \Big(\prod_{j=1}^4 |F_j(\rho_j(y))|\Big) t^{-2m} \chi(|y-(q,q)|\leq 2^n t)\, dy\, dq\, \frac{dt}{t} \quad (2.7.10)$$

$$+\sum_{n\geq 0} 2^{-4mn} \sum_{k\in\mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{T_k^c} \int_{S_k} \Big(\prod_{j=1}^4 |F_j(\rho_j(y))|\Big) t^{-2m} \chi(|y-(q,q)|\leq 2^n t)\, dy\, dq\, \frac{dt}{t}\,, \quad (2.7.11)$$

where $S_k = \{y \in \mathbb{R}^{2m} : \rho_j(y) \in T_k \text{ for all } j = 1,2,3,4\}$.

First we estimate the summand in (2.7.10) for fixed n and k by

$$C2^{-2mk} \int_{T_k} \int_{S_k^c} \Big(\prod_{j=1}^4 |F_j(\rho_j(y))| \chi(|\rho_j(y) - q| \leq 2^{n+k})\Big)\, dy\, dq\,,$$

where we used $\rho_j(q,q) = q$. Let E be the set of $q \in T_k$ such that the inner integral of the last display does not vanish. We estimate the last display using Lemma 2.7.6 by

$$C2^{-2mk} \int_E \prod_{j=1}^4 \|F_j(w)\chi(|w-q|\leq 2^{n+k})\|_{L^2(w)}\, dq \leq C2^{2mn}|E| \prod_{j=1}^4 MF_j(\mathcal{T})\,. \quad (2.7.12)$$

We proceed by estimating $|E|$. If $q \in E$, then there is $y \in S_k^c$ such that for all j,

$$|\rho_j(y) - q| \leq 2^{n+k}\,.$$

By definition of S_k, $\rho_{j_0}(y) \notin T_k$ for some $1 \leq j_0 \leq 4$. Let Q_q be a dyadic cube of side length 2^k containing q and let Q_y be a dyadic cube of side length 2^k such that $\rho_{j_0}(y) \in Q_y$. Then $Q_q \subseteq T_k$ and $Q_y \not\subseteq T_k$. But both Q_q and Q_y are contained in the ball B of radius $C2^{n+k}$ about q for sufficiently large $C > 1$. Therefore, there is $w \in \partial T_k \cap (2^k\mathbb{Z})^m$ such that $w \in B$. But then q is contained in the ball of radius $C2^{n+k}$ about w. This implies

$$|E| \leq C2^{nm+km} \#(\partial T_k \cap (2^k\mathbb{Z})^m)\,.$$

Applying this estimate to (2.7.12) and summing in n and k, we obtain for (2.7.10) a upper bound by a constant times

$$\Big(\sum_{k\in\mathbb{Z}} 2^{km} \#(\partial T_k \cap (2^k\mathbb{Z})^m)\Big) \prod_{j=1}^4 MF_j(\mathcal{T})\,.$$

Lemma 2.7.7 then yields the desired bound for (2.7.10).

It remains to estimate (2.7.11). Fix n and k and estimate the corresponding summand by

$$C2^{-2mk} \int_{T_k^c} \int_{S_k} \Big(\prod_{j=1}^4 |F_j(\rho_j(y))| \chi(|\rho_j(y) - q| \leq 2^{n+k})\Big)\, dy\, dq\,.$$

Let E be the set of $q \in T_k^c$ such that the inner integral of the last display is not zero. We estimate the last display with Lemma 2.7.6 by

$$C2^{-2mk} \int_E \prod_{j=1}^4 \|F_j(w)\chi(|w-q|\leq 2^{n+k})\|_{L^2(w)}\, dq\,. \quad (2.7.13)$$

If $p \in E$, then there is $y \in S_k$ such that for all j, $|\rho_j(y) - q| \le 2^{n+k}$. By definition of S_k, for every j there is $q^{(j)} \in T_k$ such that $\rho_j(y) = q^{(j)}$. Using the triangle inequality, we estimate (2.7.13) by

$$C2^{-2mk} \int_E \prod_{j=1}^4 \|F_j(w)\chi(|w - q^{(j)}| \le 2^{n+k+1})\|_{L^2(w)}\, dq \le C2^{2mn}|E| \prod_{j=1}^4 MF_j(\mathcal{T})\,.$$

To obtain the last inequality we may argue as for (2.7.12), because $q^{(j)} \in T_k$. Similarly as before, the ball of radius $C2^{n+k+1}$ about p contains p_j and as before we see that it also contains a point in $\partial T_k \cap (2^k\mathbb{Z})^m$. We estimate

$$|E| \le C2^{nm+km}\#(\partial T_k \cap (2^k\mathbb{Z})^m)\,,$$

sum in n and k, and use Lemma 2.7.7 to conclude the desired bound for (2.7.11). This finishes the proof of the lemma. $\square$

Proof of Lemma 2.7.5. We proceed along the lines of the argument in [45, Section 5.2] in the case of the identity matrix, and streamline the proof in our setting.

Integrating by parts in q, we see that

$$-2\sum_{i=1}^m \int_{\mathbb{R}^m} (\partial_i\partial_{i+m} g)_{Dt}(y + (q,q))\, dq = \int_{\mathbb{R}^m} (\Delta g)_{Dt}(y + (q,q))\, dq\,.$$

Using the heat equation $(\Delta g)_{tD} = 2\pi t\partial_t(g_{tD})$, we thus obtain

$$-\alpha^{4m} 4\pi \sum_{i=1}^m \Theta^{(i)}_{\mathcal{T},0,\alpha}(F_1,F_2,F_3,F_4)$$

$$= \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}^m}\int_{\mathbb{R}^{2m}} \Big(\prod_{j=1}^4 (F_j 1_{T_k})(\rho_j(y))\Big) \int_{2^{k-1}}^{2^k} t\partial_t(g_{tD})(y + (q,q))\,\frac{dt}{t}\, dy\, dq\,.$$

Let $k_{\mathcal{T}}$ be defined by $\ell(Q_{\mathcal{T}}) = 2^{k_{\mathcal{T}}}$. By the fundamental theorem of calculus in t, the last display equals

$$\sum_{k\in\mathbb{Z}} \int_{\mathbb{R}^m}\int_{\mathbb{R}^{2m}} \Big(\prod_{j=1}^4 (F_j 1_{T_k})(\rho_j(y))\Big)(g_{2^kD} - g_{2^{k-1}D})(y + (q,q))\, dy\, dq$$

$$= \int_{\mathbb{R}^m}\int_{\mathbb{R}^{2m}} \Big(\prod_{j=1}^4 (F_j 1_{Q_{\mathcal{T}}})(\rho_j(y))\Big) g_{2^{k_{\mathcal{T}}}D}(y + (q,q))\, dy\, dq \tag{2.7.14}$$

$$+ \sum_{k<k_{\mathcal{T}}} \int_{\mathbb{R}^m}\int_{\mathbb{R}^{2m}} \Big(\prod_{j=1}^4 (F_j 1_{T_k})(\rho_j(y)) - \prod_{j=1}^4 (F_j 1_{T_{k+1}})(\rho_j(y))\Big) g_{2^kD}(y + (q,q))\, dy\, dq\,. \tag{2.7.15}$$

We estimate the two terms (2.7.14) and (2.7.15) separately.

First we estimate (2.7.15). Let χ be the characteristic function of $[-1,1]^{2m}$. We bound

$$g \le C\sum_{n\ge 0} e^{-2^n}\chi_{2^n}\,.$$

We fix $k < k_{\mathcal{T}}$ and $n \geq 0$, and consider

$$\int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} \Big(\prod_{j=1}^4 (F_j 1_{T_k})(\rho_j(y)) - \prod_{j=1}^4 (F_j 1_{T_{k+1}})(\rho_j(y)) \Big) \chi_{2^{n+k}D}(y + (q,q)) \, dy \, dq \, .$$

Using the distributive law and $T_k \subseteq T_{k+1}$, we estimate the integrand as

$$\Big| \prod_{j=1}^4 (F_j 1_{T_k})(\rho_j(y)) - \prod_{j=1}^4 (F_j 1_{T_{k+1}})(\rho_j(y)) \Big|$$

$$\leq \sum_{j_0=1}^4 |F_{j_0} 1_{T_{k+1} \setminus T_k}|(\rho_{j_0}(y)) \prod_{j \neq j_0} |F_j 1_{T_{k+1}}|(\rho_j(y)) \, .$$

We fix j_0. For simplicity of notation we set $j_0 = 1$, the other values of j_0 will be analogous.

Let Q be a cube of side length 2^k contained in $T_{k+1} \setminus T_k$ and consider

$$\int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} |F_1 1_Q|(\rho_1(y)) \Big(\prod_{j=2}^4 |F_j 1_{T_{k+1}}|(\rho_j(y)) \Big) \chi_{2^{n+k}D}(y + (q,q)) \, dy \, dq \, . \tag{2.7.16}$$

Since $\alpha \geq 1$, we have $y + (q,q) \in 2^{n+k}\alpha[-1,1]^{2m}$. Applying ρ_j, we obtain for $1 \leq j \leq 4$,

$$\rho_j(y) + q \in 2^{n+k}\alpha[-1,1]^m \, .$$

We also have $\rho_1(y) \in Q$, so $q \in P$, where

$$P = 2^{n+k}\alpha[-1,1]^m - Q \, .$$

Thus, for each $j = 2,3,4$, we have $\rho_j(y) \in S$, where

$$S = Q + 2^{n+k+1}\alpha[-1,1]^m \, .$$

Thus, we can bound (2.7.16) by

$$2^{-2m(n+k)} \, (\det D)^{-1} |P| \int_{\mathbb{R}^{2m}} |F_1 1_Q|(\rho_1(y)) \Big(\prod_{j=2}^4 |F_j 1_{T_{k+1} \cap S}|(\rho_j(y)) \Big) \, dy \, .$$

By Lemma 2.7.6, we estimate this by

$$C 2^{-2m(n+k)} (\det D)^{-1} |P| \|F_1 1_Q\|_2 \prod_{j=2}^4 \|F_j 1_{T_{k+1} \cap S}\|_2$$

$$= C 2^{-2m(n+k)} (\det D)^{-1} |P| |Q|^{1/2} |S|^{3/2} \Big(\frac{1}{|Q|} \int_Q F_1^2 \Big)^{1/2} \prod_{j=2}^4 \Big(\frac{1}{|S|} \int_S F_j^2 \Big)^{1/2} .$$

Next, we crudely estimate $(\det D)^{-1} = \alpha^{-2d_2} \leq 1$ and $|S|^{3/2} \leq C 2^{2mn} \alpha^{2m} |Q|^{3/2}$. We also use that and $|Q| = C 2^{mk}$, $|P| \leq C 2^{m(n+k)} \alpha^m$, and that S covers Q. This bounds the last display by

$$C |Q| 2^{mn} \alpha^{3m} \prod_{j=1}^4 M F_j(\mathcal{T}) \, .$$

Summing over the disjoint cubes Q in $T_{k+1} \setminus T_k$, summing over $k < k_{\mathcal{T}}$, and using that the regions $T_{k+1} \setminus T_k$ are disjoint in $Q_{\mathcal{T}}$, we then estimate (2.7.15) by

$$C\Big(\sum_{n\geq 0} e^{-2^n} 2^{mn}\Big)\alpha^{3m}|Q_{\mathcal{T}}| \prod_{j=1}^{4} MF_j(\mathcal{T}) \,.$$

Then it remains to sum in n.

It remains to estimate (2.7.14), which is done similarly as (2.7.15) but simpler. Estimating the Gaussian by a superposition of characteristic functions of cubes, we consider

$$\int_{\mathbb{R}^m}\int_{\mathbb{R}^{2m}} \Big(\prod_{j=1}^{4}(|F_j|1_{Q_{\mathcal{T}}})(\rho_j(y))\Big)\chi_{2^{n+k}D}(y+(q,q))\,dy\,dq\,.$$

This is then estimated analogously to (2.7.16). $\square$

To finish the proof of Proposition 2.7.2 it remains to do a stopping time argument, similarly as in [71, 45]. Denote by $\Lambda(f_1, f_2, f_3)$ the form in the statement of Proposition 2.7.2. Let $\mathcal{Q}$ denote the collection of all dyadic cubes in $\mathbb{R}^{4n}$ contained in $[-2^N, 2^N]^{4n}$, with side-lengths in $[2^{-N}, 2^N]$ for a large $N > 0$. By the monotone convergence theorem, we may assume that in the integral defining Λ one has $(t,p) \in \Omega_{\mathcal{Q}}$. By homogeneity we may also normalize

$$\|f_j\|_{p_j} = 1$$

for each $j = 1, 2, 3$. Thus, it suffices to prove

$$|\Lambda(f_1, f_2, f_3)| \leq C(1+|a|)^{16n}\,.$$

For every triple of integers $k = (k_1, k_2, k_3)$, we define

$$\mathcal{P}_k = \{Q \in \mathcal{Q} : 2^{k_j - 1} < \sup_{Q' \supseteq Q}\Big(\frac{1}{|Q'|}\int_{Q'}|f_j|^2\Big)^{1/2} \leq 2^{k_j} \text{ for } j = 1,2,3\}\,,$$

where the supremum is over all cubes Q' in $\mathbb{R}^{4n}$ with sides parallel to the coordinate axes. Let $\mathcal{P}_k^{\max}$ be the collection of all maximal dyadic cubes in $\mathcal{P}_k$ with respect to set inclusion. For every $Q \in \mathcal{P}_k^{\max}$, the collection

$$\mathcal{T}_Q = \{Q' \in \mathcal{P}_k : Q' \subseteq Q\}$$

is a convex tree and for different $Q \in \mathcal{P}_k^{\max}$, the corresponding trees are disjoint. Proposition 2.7.3 gives

$$|\Lambda_{\mathcal{T}_Q, a}(f_1, f_2, f_3, 1)| \leq |Q| \sum_{j=1}^{3}\Big(\frac{1}{|Q'|}\int_{Q'}|f_j|^2\Big)^{1/2} \leq |Q| 2^{k_1+k_2+k_3}\,.$$

Therefore,

$$(1+|a|)^{-16n}|\Lambda(f_1,f_2,f_3)| \leq \sum_{k\in\mathbb{Z}^3}\sum_{Q\in\mathcal{P}_k^{\max}}|\Lambda_{\mathcal{T}_Q,a}(f_1,f_2,f_3,1)| \leq C\sum_{k\in\mathbb{Z}^3}2^{k_1+k_2+k_3}\sum_{Q\in\mathcal{P}_k^{\max}}|Q|\,.$$

By disjointness of the maximal cubes, for each $j = 1, 2, 3$,

$$\sum_{Q \in \mathcal{P}_k^{\max}} |Q| = \Big| \bigcup_{Q \in \mathcal{P}_k^{\max}} Q \Big| \subseteq |\{Mf_j > C2^{k_j}\}| ,$$

where Mf_j denotes a "quadratic" variant of the Hardy-Littlewood maximal function

$$Mf_j(x) = \sup_{Q' \ni x} \Big(\frac{1}{|Q'|} \int_{Q'} |f_j(x)|^2 \, dx \Big)^{1/2} ,$$

with supremum is over all cubes Q' with sides parallel to the coordinate axes. We split $\mathbb{Z}^3 = \mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3$, where $\mathcal{K}_j = \{(k_1, k_2, k_3) : k_j p_j \geq k_{j'} p_{j'}$ for $j' = 1, 2, 3\}$. Thus,

$$(1 + |a|)^{-16n} |\Lambda(f_1, f_2, f_3)| \leq \sum_{j=1}^{3} \sum_{(k_1,k_2,k_3) \in \mathcal{K}_j} 2^{k_1+k_2+k_3} |\{Mf_j > C2^{k_j}\}|$$

$$= \sum_{j=1}^{3} \sum_{k_j \in \mathbb{Z}} 2^{p_j k_j} |\{Mf_j > C2^{k_j}\}| \prod_{j' \neq j} \sum_{k_{j'} : k_{j'} \leq p_j k_j / p_{j'}} 2^{k_{j'} - \frac{p_j k_j}{p_{j'}}}$$

$$\leq C \sum_{j=1}^{3} \|Mf_j\|_{p_j}^{p_j} \leq C \sum_{j=1}^{3} \|f_j\|_{p_j}^{p_j} \leq C .$$

This finishes the proof of Proposition 2.7.2.

2.8 List of indecomposable modules

In Tables 2.1–2.4 below we list the modules used in the classification Theorems 2.1.11 and 2.1.15. We specify the modules using block matrices

A_{10}	A_{11}	A_{12}	A_{13}
A_{20}	A_{21}	A_{22}	A_{23}

.

We define M to be $\mathbb{R}^n$ for some n, and identify each subspace M_i with $\mathbb{R}^{n_i}$ as well. The block columns

$$\begin{pmatrix} A_{1i} \\ A_{2i} \end{pmatrix}$$

for $i = 0, 1, 2, 3$ then specify the matrices of the embeddings $M_i \to M$ defining the module $\mathbf{M}$, which fixes implicitly also the dimensions of the subspaces M_i and of M. Two modules defined like this are isomorphic if the corresponding block matrices can be transformed into each other by row operations on the whole matrix and column operations on each block column. In terms of the corresponding Brascamp-Lieb data, the transposes of the block columns are the matrices of the maps Π_i.

Following the notation of [87], we write I_n for the $n \times n$ identity matrix. We denote by $J_n(\lambda)$ an $n \times n$ Jordan block with eigenvalue λ. An arrow in the superscript of a matrix indicates that a row or column of zeros is to be added in the direction the arrow points, for example $I_n^{\uparrow}$ is the $(n+1) \times n$ matrix with one row of zeros, followed by the $n \times n$ identity matrix in the rows below.

Finally, the matrix $X = X(P, s)$ in modules $\mathbf{0}$ and $\mathbf{N}_n$ denotes the companion matrix of the polynomial $(P(t))^s$, for some $s \geq 1$ and an irreducible polynomial $P \in \mathbb{R}[t]$ with $P(t) \neq t$ and $P(t) \neq t - 1$. The companion matrix of a polynomial $Q(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0$ is the matrix

$$\begin{pmatrix} 0 & \dots & 0 & -a_0 \\ 1 & \dots & 0 & -a_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & -a_{n-1} \end{pmatrix},$$

with characteristic polynomial Q. Note that the conditions on P imply that $P(t) = t - \lambda$ with $\lambda \neq 0, 1$ or $P(t) = t^2 - 2\lambda t + \mu$ with $\mu > \lambda^2$.

The indecomposable modules in Table 2.1 are only listed up to permutation of the subspaces. The additional information which permutations give rise to non- isomorphic modules is given by the following lemma, which is Remark 1 in [87].

Lemma 2.8.1 ([87, Remark 1])**.** *For the modules $\mathbf{II}, \mathbf{III}, \mathbf{III}^*, \mathbf{IV}, \mathbf{IV}^*, \mathbf{V}, \mathbf{V}^*$, each permutation of the subspaces that leaves their dimensions invariant gives rise to an isomorphic module. For the modules of type $\mathbf{0}$, all permutations of the subspaces give rise to another module of type $\mathbf{0}$, but possibly with different X. For module $\mathbf{I}$, swapping the columns $1, 3$ or swapping columns $2, 4$ gives rise to an isomorphic module. Thus there are 6 isomorphism classes of modules that can be obtained by permuting the subspaces in type $\mathbf{I}$.*

Lemma 2.8.1 can be directly checked by transforming the corresponding block matrices into each other using the allowed row and column transformations. We omit this and refer to [87].

M	$\dim(M, M_0, M_1, M_2, M_3)$	block matrix			
0	$(2n, n, n, n, n)$	I_n	0	I_n	X
		0	I_n	I_n	I_n
I	$(2n, n, n, n, n)$	I_n	0	I_n	$J_n(0)$
		0	I_n	I_n	I_n
II	$(2n+1, n+1, n+1, n, n)$	I_{n+1}	I_{n+1}	$I_n^{\downarrow}$	0
		0	$I_n^{\rightarrow}$	I_n	I_n
III	$(2n+1, n, n, n, n+1)$	I_{n+1}	0	$I_n^{\uparrow}$	$I_n^{\downarrow}$
		0	I_n	I_n	I_n
III*	$(2n+1, n+1, n+1, n+1, n)$	I_n	0	$I_n^{\leftarrow}$	$I_n^{\rightarrow}$
		0	I_{n+1}	I_{n+1}	I_{n+1}
IV	$(2n+2, n+1, n+1, n+1, n)$	I_{n+1}	0	I_{n+1}	$I_n^{\uparrow}$
		0	I_{n+1}	I_{n+1}	$I_n^{\downarrow}$
IV*	$(2n+2, n+1, n+1, n+1, n+2)$	I_{n+1}	0	I_{n+1}	$I_{n+1}^{\leftarrow}$
		0	I_{n+1}	I_{n+1}	$I_{n+1}^{\rightarrow}$
V	$(2n+1, n, n, n, n)$	I_n	0	$J_n(0)$	I_n
		0	I_n	I_n	$J_n(0)$
		0..0	0..0	10..0	10..0
V*	$(2n+1, n+1, n+1, n+1, n+1)$	$I_n^{\leftarrow}$	$I_n^{\leftarrow}$	$I_n^{\rightarrow}$	0
		0	$I_n^{\rightarrow}$	$I_n^{\leftarrow}$	$I_n^{\leftarrow}$
		10..0	10..0	10..0	10..0

Table 2.1: Indecomposable modules of the four subspace quiver, up to permutation of the subspaces. The following list is a direct result from the diagrams in [87].

M	$\dim(M, M_0, M_1, M_2, M_3)$	block matrix			
$\mathbf{N}_n$	$(2n, n, n, n, n)$	I_n	0	I_n	X
		0	I_n	I_n	I_n
$\mathbf{J}_n^{(1)}$	$(2n, n, n, n, n)$	I_n	0	I_n	$J_n(1)$
		0	I_n	I_n	I_n
$\mathbf{J}_n^{(2)}$	$(2n, n, n, n, n)$	I_n	0	I_n	$J_n(0)$
		0	I_n	I_n	I_n
$\mathbf{J}_n^{(3)}$	$(2n, n, n, n, n)$	I_n	0	$J_n(0)$	I_n
		0	I_n	I_n	I_n
$\mathbf{C}_n$	$(2n+1, n+1, n, n, n)$	I_{n+1}	0	$I_n^{\uparrow}$	$I_n^{\downarrow}$
		0	I_n	I_n	I_n
$\mathbf{T}_n$	$(2n+1, n, n+1, n+1, n+1)$	I_n	0	$I_n^{\leftarrow}$	$I_n^{\rightarrow}$
		0	I_{n+1}	I_{n+1}	I_{n+1}

Table 2.2: Indecomposable modules corresponding to data of Hölder type

M	$\dim(M, M_0, M_1, M_2, M_3)$	block matrix			
Y	$(2,0,1,1,1)$	0	1	0	1
		0	0	1	1
Z	$(3,1,1,1,1)$	1	0	0	1
		0	1	1	0
		0	0	1	1
L	$(4,1,2,2,2)$	$I_1^{\uparrow}$	I_2	0	I_2
		$I_1^{\downarrow}$	0	I_2	I_2
B	$(5,2,2,2,2)$	I_2	0	$J_2(0)$	I_2
		0	I_2	I_2	$J_2(0)$
		00	00	10	10

Table 2.3: Indecomposable modules corresponding to Young's convolution inequality and to Loomis-Whitney type inequalities

M	$\dim(M, M_0, M_1, M_2, M_3)$	block matrix			
$\mathbf{P}^{(1)}$	$(1,0,0,1,1)$	0	0	I_1	I_1
$\mathbf{P}^{(2)}$	$(1,0,1,0,1)$	0	I_1	0	I_1
$\mathbf{P}^{(3)}$	$(1,0,1,1,0)$	0	I_1	I_1	0
$\mathbf{K}^{(1)}$	$(2,1,0,1,1)$	I_1	0	0	I_1
		0	0	I_1	I_1
$\mathbf{K}^{(2)}$	$(2,1,1,0,1)$	I_1	0	0	I_1
		0	I_1	0	I_1
$\mathbf{K}^{(3)}$	$(2,1,1,1,0)$	I_1	0	I_1	0
		0	I_1	I_1	0

Table 2.4: Indecomposable modules corresponding to Hölder's inequality or Hölder's inequality combined with boundedness of a linear singular integral operator

Chapter 3

Sharp Fourier extension for functions with localized support on the circle

This chapter consists of the publication [7] that will appear in Revista Matemática Iberoamericana. The copyright is held by the EMS Publishing House GmbH and the article is reproduced here in accordance with their copyright policy.

3.1 Introduction

We are interested in the conjecture that constant functions are extremizers for the Tomas-Stein Fourier extension inequality for the circle

$$\|\widehat{f\sigma}\|_{L^6(\mathbb{R}^2)} \leq C\|f\|_{L^2(\sigma)}. \tag{3.1.1}$$

Here σ is the arc length measure on the unit circle $S^1 \subset \mathbb{R}^2$ and $\hat{\mu}(\xi) = \int e^{-ix\xi}\,\mathrm{d}\mu(\xi)$ is the Fourier transform.

The corresponding conjecture for S^2 was proven by Foschi [54], and in [23] Foschi's argument is adapted to S^1, and the conjecture of interest is reduced to the following.

Conjecture 3.1.1. *The quadratic form*

$$Q(f) := \int_{(S^1)^6} (|\omega_1 + \omega_2 + \omega_3|^2 - 1)(f(\omega_1, \omega_2, \omega_3)^2 - f(\omega_1, \omega_2, \omega_3)f(\omega_4, \omega_5, \omega_6))\,\mathrm{d}\Sigma$$

is positive semi-definite on the subspace V of all antipodal functions in $L^2((S^1)^3, \mathbb{R})$. Here we denote

$$\mathrm{d}\Sigma = \mathrm{d}\Sigma(\omega) = \delta(\sum_{j=1}^{6} \omega_k) \prod_{j=1}^{6} \mathrm{d}\sigma(\omega_j),$$

and a function f is antipodal if $f(\pm\omega_1, \pm\omega_2, \pm\omega_3)$ does not depend on the choice of signs.

Conjecture 3.1.1 has been verified for all functions with Fourier modes up to degree 120 in [99] and [3], via a numerical computation of the eigenvalues of Q on the finite dimensional space of such functions. Further, using different methods, in [29] the conjectured sharp form of inequality (3.1.1) has been established for certain infinite dimensional subspaces of $L^2(\sigma)$ with constrained Fourier support. Our main result establishes Conjecture 3.1.1 for functions with localized spatial support.

Let C_ε be the cylinder of radius ε centered at the line $\mathbb{R}(1,1,1)$, and define

$$V_\varepsilon := \left\{ f \in V \ : \ \operatorname{supp} f(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3}) \subset \bigcup_{k \in \pi\mathbb{Z}^3} k + C_\varepsilon \right\}.$$

Theorem 3.1.2. *Let $\varepsilon = 1/20$. Then for all $f \in V_\varepsilon$ it holds that $Q(f) \geq 0$.*

Note that since constant functions are in the kernel of Q, the same result holds for $V_\varepsilon \oplus \langle \mathbf{1} \rangle$, where $\mathbf{1}$ is the constant 1 function.

As a corollary, functions with support sufficiently close to a pair of antipodal points satisfy (3.1.1) with the conjectured sharp constant. Define

$$\Phi(g) := \frac{\|\widehat{g\sigma}\|_{L^6(\mathbb{R}^2)}}{\|g\|_{L^2(\sigma)}}.$$

Corollary 3.1.3. *Let $\varepsilon' = \sqrt{3/8}\varepsilon$. Suppose that $g \in L^2(\sigma)$ is such that $g(e^{i\theta})$ is supported in $(-\varepsilon', \varepsilon') + \pi\mathbb{Z}$. Then $\Phi(g) \leq \Phi(\mathbf{1})$, where $\mathbf{1}$ is the constant 1 function on S^1.*

Note that by rotation symmetry, the same holds when $g(e^{i\theta})$ is supported in $I + \pi\mathbb{Z}$ for any interval I of length $2\varepsilon'$.

The constants ε and ε' in Theorem 3.1.2 and Corollary 3.1.3 are not optimal. Numerical computations suggest that with our method ε can be improved up to about 0.104 and ε' up to about 0.063, see Section 3.7.

The numerical results in [3] suggest that eigenfunctions of Q on the subspace of functions with Fourier modes up to degree N corresponding to small eigenvalues concentrate in space. Theorem 3.1.2 shows that Q is positive on all such sufficiently concentrated functions, thus it should be a useful partial result in establishing positive semi-definiteness of Q on the full space of antipodal functions. A more precise observation by Jiaxi Cheng, a graduate student in Bonn, is that the smallest eigenvalue is of size $\sim N^{-2}\log(N)$, see Section 2 of [96]. The existence of such an eigenvalue is also explained by the asymptotic formula for the multiplier m in Lemma 3.4.1, which looks like $c|\log|x|||x|^2$ near 0. Unfortunately, we cannot prove that this is the smallest eigenvalue.

More generally, the topic of sharp Fourier extension inequalities has attracted a lot of interest in recent years. In the following we consider general dimensions $d \geq 2$. Then the Tomas-Stein extension inequality states that for every

$$q \geq q_d := \frac{2(d+1)}{d-1},$$

there exists $C(d,q) > 0$ such that for all $f \in L^2(S^{d-1}, \sigma^{d-1})$

$$\|\widehat{f\sigma}\|_{L^q(\mathbb{R}^d)} \leq C(d,q)\|f\|_{L^2(\sigma)}. \tag{3.1.2}$$

Here σ^{d-1} denotes the $d-1$-dimensional Hausdorff measure on S^{d-1}.

It is known that extremizers for (3.1.2) exist when $q > q_d$, for all d, see [49]. At the endpoint $q = q_d$, existence and smoothness of extremizers have been shown for $d = 3$ in [27], [28] and for $d = 2$ in [106], [105]. For higher dimensional spheres $d \geq 4$, existence of extremizers for $q = q_d$ is known conditional on the conjecture that Gaussians maximize the corresponding extension inequality for the paraboloid, see [57].

For certain specific choices of (d, q), a full characterization of the extremizers of (3.1.2) is known. Most such results grew out of the work of Foschi [54], who showed that constant functions maximize (3.1.2) for $(d, q) = (2, 4)$, and gave a full characterization of all complex valued maximizers. His method can be adapted for some non-endpoint extension inequalities on higher dimensional spheres, see [24]. Using different methods, maximizers of (3.1.2) for some choices of (d, q) with even $q > 4$ are characterized in [98]. In some further cases it is known that constant functions are local maximizers. This was shown in [23] for $(d, q) = (2, 6)$, and in [60] for (d, q_d) with $3 \leq d \leq 60$. For further background and references on sharp Fourier extension inequalities we refer to [53] and [96].

Acknowledgement

I am grateful to Christoph Thiele for introducing me to this problem and for many helpful discussions, and to Jan Holstermann for pointing out the short proof of Lemma 3.6.7. The author was supported by the Collaborative Research Center 1060 funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and the Hausdorff Center for Mathematics, funded by the DFG under Germany's Excellence Strategy - GZ 2047/1, ProjectID 390685813.

3.2 Proof of Corollary 3.1.3

Corollary 3.1.3 is a direct consequence of Theorem 3.1.2 and the program formulated in [23]. We give a brief sketch of the implication here; for the details of the program and proofs we refer the reader to [23].

Proof of Corollary 3.1.3. Let $g \in L^2(\sigma)$ be such that $g(e^{i\theta})$ is supported in $(-\varepsilon', \varepsilon') + \pi\mathbb{Z}$. Define $\tilde{g}(x) = g(-x)$ and

$$g_\# = \sqrt{\frac{|g|^2 + |\tilde{g}|^2}{2}}.$$

As shown in [23], Step 1 and 2, it holds that $\Phi(g) \leq \Phi(g_\#)$, and $g_\#$ is antipodal and $g_\#(e^{i\theta})$ is supported in $(-\varepsilon', \varepsilon') + \pi\mathbb{Z}$. Define $f(\omega_1, \omega_2, \omega_3) := g_\#(\omega_1)g_\#(\omega_2)g_\#(\omega_3)$. Then the function $f(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3})$ is supported in $\bigcup_{k\in\pi\mathbb{Z}^3} k + (-\varepsilon', \varepsilon')^3$. Since $(-\varepsilon', \varepsilon')^3$ is a subset of the cylinder $C_{\sqrt{8/3}\epsilon'}$, it follows that $f \in V_{\sqrt{8/3}\varepsilon'} = V_\varepsilon$, hence $Q(f) \geq 0$, by Theorem 3.1.2. This verifies Conjecture 1.4 in [23] for $g_\#$. Using Step 3, 4 and 5 in [23], we conclude that $\Phi(g) \leq \Phi(1)$. $\square$

3.3 Proof of Theorem 3.1.2

3.3.1 Orthogonal decomposition

We consider the sesquilinear form

$$\begin{aligned} B(f,g) = \int_{(S^1)^6} &(|\omega_1+\omega_2+\omega_3|^2-1) \\ &\cdot \left(f(\omega_1,\omega_2,\omega_3)\overline{g(\omega_1,\omega_2,\omega_3)} - f(\omega_1,\omega_2,\omega_3)\overline{g(\omega_4,\omega_5,\omega_6)}\right) \mathrm{d}\Sigma(\omega). \end{aligned}$$

By a change of variables, it holds that $B(f,g) = B(Rf,Rg)$, where $Rf(\omega_1,\omega_2,\omega_3) = f(e^i\omega_1, e^i\omega_2, e^i\omega_3)$. Define

$$Z_d = \{(k_1,k_2,k_3) \in (2\mathbb{Z})^3 \,:\, k_1+k_2+k_3 = d\}$$

and

$$X_d = \Big\{ \sum_{k\in Z_d} a_k \omega_1^{k_1}\omega_2^{k_2}\omega_3^{k_3} \;:\; (a_k) \in \ell^2(Z_d) \Big\} \subset L^2((S^1)^3).$$

For $d \neq d'$, the spaces X_d and $X_{d'}$ are eigenspaces of R with different eigenvalues e^{id} and $e^{id'}$, and hence are orthogonal with respect to B. Note that the orthogonal projection π_d onto X_d can be expressed as

$$\pi_d(f)(\omega_1,\omega_2,\omega_3) = \int_0^1 e^{-2\pi i d t} f(e^{2\pi i t}\omega_1, e^{2\pi i t}\omega_2, e^{2\pi i t}\omega_3)\,\mathrm{d}t,$$

which implies that $\pi_d(V_\varepsilon) \subset V_\varepsilon$. Therefore, we have that

$$V_\varepsilon = \overline{\bigoplus_{d\in\mathbb{Z}} \pi_d(V_\varepsilon)} = \overline{\bigoplus_{d\in\mathbb{Z}} (V_\varepsilon \cap X_d)}.$$

Hence, it suffices to show positive semi-definiteness of B on each of the spaces

$$X_{d,\varepsilon} := V_\varepsilon \cap X_d.$$

3.3.2 Reducing the dimension

From now on, we use the convention that

$$\omega_i = (\cos(\theta_i), \sin(\theta_i)), \tag{3.3.1}$$

and abuse notation by writing $f(\omega(\theta)) = f(\theta)$. We also define

$$\begin{aligned} a(\theta_1,\theta_2,\theta_3) &:= (\cos(\theta_1)+\cos(\theta_2)+\cos(\theta_3))^2 + (\sin(\theta_1)+\sin(\theta_2)+\sin(\theta_3))^2 \\ &= |\omega_1+\omega_2+\omega_3|^2, \end{aligned}$$

so that the weight in the bilinear form B is given by $a-1$, and record the useful identity

$$a(\theta_1,\theta_2,\theta_3) = 3 + 2\cos(\theta_1-\theta_2) + 2\cos(\theta_2-\theta_3) + 2\cos(\theta_3-\theta_1). \tag{3.3.2}$$

The domain of integration $\omega \in (S^1)^6$ in the bilinear form B becomes $\theta \in \mathbb{R}^6/(2\pi\mathbb{Z})^6$. As we assume that $f \in X_d$ for some d, we fully understand how f transforms under simultaneous rotations of $\omega_1, \omega_2, \omega_3$ by the same angle. We will use this to integrate out such simultaneous rotations of $\omega_1, \omega_2, \omega_3$ and of $\omega_4, \omega_5, \omega_6$. These rotations correspond to shifts of $(\theta_1, \theta_2, \theta_3)$ and $(\theta_4, \theta_5, \theta_6)$ in direction $(1, 1, 1)$, which makes it natural to choose the following fundamental domain of $\mathbb{R}^3/(2\pi\mathbb{Z})^3$ as our domain of integration in θ.

Lemma 3.3.1. *Let C be the rhombus with corners*

$$(\pi, -\pi, 0), \quad (-\frac{\pi}{3}, -\frac{\pi}{3}, \frac{2\pi}{3}), \quad (-\pi, \pi, 0) \quad \textit{and} \quad (\frac{\pi}{3}, \frac{\pi}{3}, -\frac{2\pi}{3}).$$

Then the prism $P := C + \{(t, t, t) \,:\, t \in [0, 2\pi)\}$ over C of height $2\pi\sqrt{3}$ is a fundamental domain for $\mathbb{R}^3/(2\pi\mathbb{Z})^3$.

Proof. Denote by p the orthogonal projection onto the hyperplane

$$H := \{(\theta_1, \theta_2, \theta_3) \,:\, \theta_1 + \theta_2 + \theta_3 = 0\}.$$

The image of $(2\pi\mathbb{Z})^3$ under p is the hexagonal lattice

$$\Lambda := \mathbb{Z}v_1 \oplus \mathbb{Z}v_2 \subset H,$$

where

$$v_1 := (\frac{4\pi}{3}, -\frac{2\pi}{3}, -\frac{2\pi}{3}) \quad \text{and} \quad v_2 := (-\frac{2\pi}{3}, \frac{4\pi}{3}, -\frac{2\pi}{3}).$$

It is easy to see that the rhombus C is a fundamental domain of H modulo the lattice Λ. Thus for every x, there exists y with $x - y \in (2\pi\mathbb{Z})^3$ and $p(y) \in C$. Then for an appropriate choice of $k \in \mathbb{Z}$, the point $z = y + 2\pi k(1, 1, 1)$ lies in P, and $x - z \in (2\pi\mathbb{Z})^3$.

Conversely, let $z, z' \in P$ be such that $z - z' \in (2\pi\mathbb{Z})^3$. Then $p(z) - p(z')$ lies in $p((2\pi\mathbb{Z})^3) = \Lambda$, and $p(z), p(z') \in C$. It follows that $p(z) = p(z')$. Thus $z - z' \in 2\pi\mathbb{Z}\cdot(1, 1, 1)$, and from $z, z' \in P$ it follows that $z = z'$. $\square$

In the next lemma, we perform integrations in direction $(1, 1, 1)$ in $(\theta_1, \theta_2, \theta_3)$ and $(\theta_4, \theta_5, \theta_6)$, thereby reducing to a quadratic form depending only on the restriction $f|_C$. We define the function $\lambda_d : C \times C \to S^1$ by

$$\lambda_d(\theta'_1, \theta'_2, \theta'_3, \theta_1, \theta_2, \theta_3) = \exp(id \cdot (\arg(e^{i\theta'_1} + e^{i\theta'_2} + e^{i\theta'_3}) - \arg(e^{i\theta_1} + e^{i\theta_2} + e^{i\theta_3}))).$$

The only property of λ_d that will be used in the proof below is that $|\lambda_d| = 1$.

Lemma 3.3.2. *For all $d \in \mathbb{Z}$ and all $f \in X_d$, we have that $B(f, f)$ equals*

$$12\pi \int_{C^2} \delta(a(\theta) - a(\theta'))(a(\theta) - 1)(|f(\theta)|^2 - \lambda_d(\theta', \theta)f(\theta)\overline{f(\theta')})\, \mathrm{d}\mathcal{H}^2_C(\theta)\, \mathrm{d}\mathcal{H}^2_C(\theta').$$

Here $\mathcal{H}^2_C$ denotes the 2-dimensional Hausdorff measure on C.

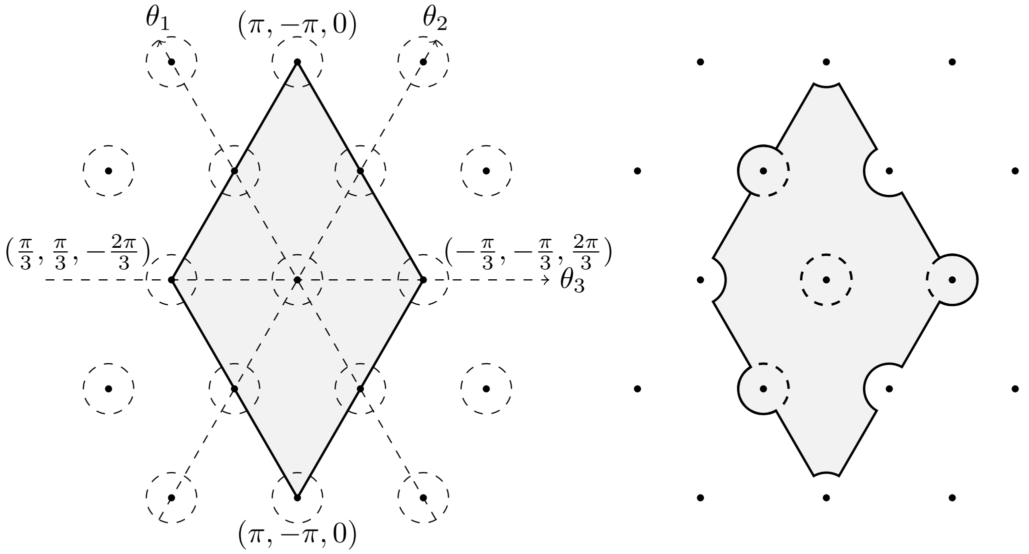

Figure 3.1: Left: The lattice $\frac{1}{2}\Lambda$ in the hyperplane H and the fundamental domain C (gray) of Λ. The restriction $|f||_H$ is supported in the union of the dashed balls and periodic with respect to $\frac{1}{2}\Lambda$. Right: One possible choice of a fundamental domain C' such that $f|_{C'}$ is supported in the union of the balls (dashed) B_1, B_2, B_3 and B_4.

Proof. By Lemma 3.3.1, we have

$$\begin{aligned}B(f,f) = \int_{P\times P} &\delta(\omega_1+\omega_2+\omega_3-\omega_4-\omega_5-\omega_6)(|\omega_1+\omega_2+\omega_3|^2-1)\\ &\times (|f(\omega_1,\omega_2,\omega_3)|^2 - f(\omega_1,\omega_2,\omega_3)\overline{f(\omega_4,\omega_5,\omega_6)}) \prod_{j=1}^{6} \mathrm{d}\theta_j\end{aligned}$$

$$\begin{aligned}= 2\pi\sqrt{3}\int_{C\times P} &\delta(\omega_1+\omega_2+\omega_3-\omega_4-\omega_5-\omega_6)(|\omega_1+\omega_2+\omega_3|^2-1)\\ &\times (|f(\omega_1,\omega_2,\omega_3)|^2 - f(\omega_1,\omega_2,\omega_3)\overline{f(\omega_4,\omega_5,\omega_6)})\,\mathrm{d}\mathcal{H}^2_C(\theta_1,\theta_2,\theta_3) \prod_{j=4}^{6} \mathrm{d}\theta_j.\end{aligned}$$

Here we have used that $f \in X_d$, to integrate out simultaneous rotations of all 6 points ω_j by the same angle. For $x, y \in \mathbb{R}^2$, it holds that

$$\delta(x-y) = 2\delta(|x|^2-|y|^2)\delta(\arg(x)-\arg(y)).$$

Hence, we can rewrite the last expression as

$$\begin{aligned}= 4\pi\sqrt{3}\int_{C\times P} &\delta(|\omega_1+\omega_2+\omega_3|^2-|\omega_4+\omega_5+\omega_6|^2)\\ &\times \delta(\arg(\omega_1+\omega_2+\omega_3)-\arg(\omega_4+\omega_5+\omega_6))(|\omega_1+\omega_2+\omega_3|^2-1)\\ &\times (|f(\omega_1,\omega_2,\omega_3)|^2 - f(\omega_1,\omega_2,\omega_3)\overline{f(\omega_4,\omega_5,\omega_6)})\,\mathrm{d}\mathcal{H}^2_C(\theta_1,\theta_2,\theta_3) \prod_{j=4}^{6} \mathrm{d}\theta_j\end{aligned}$$

$$= 12\pi \int_{C\times C} \int_0^{2\pi} \delta(a(\theta_1,\theta_2,\theta_3) - a(\theta_4,\theta_5,\theta_6))$$
$$\times\, \delta(\arg(\omega_1+\omega_2+\omega_3) - \arg(\omega_4+\omega_5+\omega_6) - t)(a(\theta_1,\theta_2,\theta_3) - 1)$$
$$\times\, (|f(\omega_1,\omega_2,\omega_3)|^2 - f(\omega_1,\omega_2,\omega_3)\overline{f(e^{it}\omega_4, e^{it}\omega_5, e^{it}\omega_6)})\,\mathrm{d}t\,\mathrm{d}\mathcal{H}^4_{C\times C}(\theta).$$

Since $f \in X_d$, we have

$$f(e^{it}\omega_4, e^{it}\omega_5, e^{it}\omega_6) = e^{itd} f(\omega_4,\omega_5,\omega_6).$$

Thus, we can integrate out t and obtain the claimed identity. □

3.3.3 Completing the proof

By Lemma 3.3.2, we have for all d and all $f \in X_d$:

$$\begin{aligned} B(f,f) &\geq 12\pi \int_C |f(\theta)|^2 (a(\theta)-1) \int_C \delta(a(\theta)-a(\theta'))\,\mathrm{d}\mathcal{H}^2_C(\theta')\,\mathrm{d}\mathcal{H}^2_C(\theta) \\ &\quad - 12\pi \int_{C^2} \delta(a(\theta)-a(\theta'))|a(\theta)-1||f(\theta)||f(\theta')|\,\mathrm{d}\mathcal{H}^2_C(\theta)\,\mathrm{d}\mathcal{H}^2_C(\theta') \\ &=: 12\pi(I - II). \end{aligned} \tag{3.3.3}$$

If $f \in X_{d,\varepsilon}$, then the restriction of f onto the hyperplane $H = \{(\theta_1,\theta_2,\theta_3) : \theta_1+\theta_2+\theta_3 = 0\}$ is supported in $\frac{1}{2}\Lambda + B_\varepsilon(0)$. Furthermore, the function $|f|$ is periodic with respect to $\frac{1}{2}\Lambda$, since it is periodic with respect to $\pi\mathbb{Z}^3$ and invariant under all translations in direction $(1,1,1)$. Thus it suffices to show the following.

Lemma 3.3.3. *Suppose that $\varepsilon \leq 1/20$. Then for all functions $f : H \to [0,\infty)$ that are periodic with respect to $\frac{1}{2}\Lambda$ and supported in $\frac{1}{2}\Lambda + B_\varepsilon(0)$, it holds that $I \geq II$.*

Proof. Recall that C is a fundamental domain of the lattice Λ. The expressions in the integrals for the terms I and II are Λ periodic, so we may replace C by any other fundamental domain C'. Since f is supported in $\frac{1}{2}\Lambda + B(0,\varepsilon)$, there exists a fundamental domain C' such that $f|_{C'}$ is supported in

$$\begin{aligned} &B_\varepsilon(0,0,0) \cup B_\varepsilon(\tfrac{2\pi}{3}, -\tfrac{\pi}{3}, -\tfrac{\pi}{3}) \cup B_\varepsilon(-\tfrac{\pi}{3}, \tfrac{2\pi}{3}, -\tfrac{\pi}{3}) \cup B_\varepsilon(-\tfrac{\pi}{3}, -\tfrac{\pi}{3}, \tfrac{2\pi}{3}) \\ &\qquad =: B_1 \cup B_2 \cup B_3 \cup B_4. \end{aligned}$$

We decompose

$$I = \sum_{i=1}^4 \int_{B_i} |f(\theta)|^2 (a(\theta)-1) \int_C \delta(a(\theta)-a(\theta'))\,\mathrm{d}\mathcal{H}^2_C(\theta')\,\mathrm{d}\mathcal{H}^2_C(\theta) =: \sum_{i=1}^4 I_i, \tag{3.3.4}$$

$$\begin{aligned} II &= \sum_{1\leq i,j\leq 4} \int_{B_i\times B_j} \delta(a(\theta)-a(\theta'))|a(\theta)-1||f(\theta)||f(\theta')|\,\mathrm{d}\mathcal{H}^2_C(\theta)\,\mathrm{d}\mathcal{H}^2_C(\theta') \\ &=: \sum_{1\leq i,j\leq 4} II_{ij}. \end{aligned} \tag{3.3.5}$$

Note that $|\theta| < \pi/6$ implies, by (3.3.2), that $a(\theta) \geq 3 + 6\cos(\pi/3) = 6$, and that similarly $|\theta - (2\pi/3, -\pi/3, -\pi/3)| < \pi/6$ implies that $a(\theta) \leq 3$. Therefore, for $j = 2, 3, 4$ the measure $\delta(a(\theta) - a(\theta'))$ vanishes on $B_1 \times B_j$, thus $I_{1j} = I_{j1} = 0$.

Next, we record that $II_{11} \leq I_1$, by Cauchy-Schwarz and since $a(\theta) \geq 6$ on B_1:

$$\begin{aligned} II_{11} &= \int_{B_1^2} \delta(a(\theta) - a(\theta'))|a(\theta) - 1||f(\theta)||f(\theta')| \mathrm{d}\mathcal{H}^2_C(\theta)\, \mathrm{d}\mathcal{H}^2_C(\theta') \\ &\leq \frac{1}{2} \int_{B_1^2} \delta(\tilde{a}(\theta) - \tilde{a}(\theta'))(\tilde{a}(\theta) - 1)(|f(\theta)|^2 + |f(\theta')|^2)\, \mathrm{d}\mathcal{H}^2_C(\theta)\, \mathrm{d}\mathcal{H}^2_C(\theta') \\ &\leq \int_{B_1} |f(\theta)|^2 (\tilde{a}(\theta) - 1) \int_C \delta(\tilde{a}(\theta) - \tilde{a}(\theta'))\, \mathrm{d}\mathcal{H}^2_C(\theta)\, \mathrm{d}\mathcal{H}^2_C(\theta') = I_1. \end{aligned}$$

The remaining terms are estimated in the next two sections. By Lemmas 3.4.1 and 3.5.1, we have

$$\begin{aligned} I_2 + I_3 + I_4 &\geq 30 \int_{B_1} |\theta|^2 |f(\theta)|^2\, \mathrm{d}\mathcal{H}^2_H(\theta) > 9\frac{101}{100}\pi \int_{B_1} |\theta|^2 |f(\theta)|^2\, \mathrm{d}\mathcal{H}^2_H(\theta) \\ &\geq \sum_{2 \leq i,j \leq 4} II_{ij}, \end{aligned}$$

which completes the proof. $\square$

3.4 Estimating term I

Lemma 3.4.1. *It holds that*

$$I_2 + I_3 + I_4 = \int_{B_1} m(\theta) |f(\theta)|^2\, \mathrm{d}\mathcal{H}^2_H(\theta), \tag{3.4.1}$$

where I_j is defined in (3.3.4)*, and $m(\theta) \geq 30|\theta|^2$.*

Proof. By definition of the I_j, equation (3.4.1) holds with

$$m(\theta) = \sum_{j=2}^{4} (a(\theta + c_j) - 1) \int_C \delta(a(\theta + c_j) - a(\theta'))\, \mathrm{d}\mathcal{H}^2_C(\theta'),$$

where c_j is the center of the ball B_j. Reversing the argument in the proof of Lemma 3.3.2, it follows that for $x \in \mathbb{R}^2$

$$\begin{aligned} &\int_C \delta(|x|^2 - a(\theta'))\, \mathrm{d}\mathcal{H}^2_C(\theta') \\ &= \frac{1}{\sqrt{3}} \int_P \delta(|x|^2 - |\omega_1 + \omega_2 + \omega_3|^2) \delta(\arg(x) - \arg(\omega_1 + \omega_2 + \omega_3)) \prod_{j=1}^{3} \mathrm{d}\theta'_j \\ &= \frac{1}{2\sqrt{3}} \int_{(S^1)^3} \delta(x - (\omega_1 + \omega_2 + \omega_3)) \prod_{j=1}^{3} \mathrm{d}\sigma(\omega_j) \\ &= \frac{1}{2\sqrt{3}} \sigma * \sigma * \sigma(x). \end{aligned}$$

The convolution $\sigma * \sigma * \sigma$ is radial. We set $\sigma * \sigma * \sigma(x) = \rho(|x|)$, giving

$$m(\theta) = \frac{1}{2\sqrt{3}} \sum_{j=2}^{4} (a(\theta + c_j) - 1)\rho(\sqrt{a(\theta + c_j)}). \tag{3.4.2}$$

In polar coordinates

$$\begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} = s\cos(\alpha)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + s\sin(\alpha)\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \tag{3.4.3}$$

we compute in Lemma 3.6.6 the asymptotic expansion

$$\begin{aligned}
&(a(\theta + c_4) - 1)\rho(\sqrt{a(\theta + c_4)}) && (3.4.4)\\
&\qquad = -12s^2(3\sin^2(\alpha) - \cos^2(\alpha))\log(s) && (3.4.5)\\
&\qquad\quad - 6s^2(3\sin^2(\alpha) - \cos^2(\alpha))\log|3\sin^2(\alpha) - \cos^2(\alpha)| && (3.4.6)\\
&\qquad\quad + 18\log 2\, s^2(3\sin^2(\alpha) - \cos^2(\alpha)) && (3.4.7)\\
&\qquad\quad + E,
\end{aligned}$$

with

$$|E| \leq -180s^4 \log s + 71s^4 \quad \text{when } s \leq 1/20.$$

As the function a is invariant under permutation of its arguments and constant in direction $(1, 1, 1)$, it is invariant under the rotation T by $2\pi/3$ about the line $\mathbb{R}(1, 1, 1)$. Since

$$c_2 + \theta(\alpha, s) = T(c_4 + \theta(\alpha + 4\pi/3, s)) \quad \text{and} \quad c_3 + \theta(\alpha, s) = T^2(c_4 + \theta(\alpha + 2\pi/3, s)),$$

we obtain the same asymptotic expansion for $a(\theta + c_j)\rho(\sqrt{a(\theta + c_j)})$, $j = 2, 3$, but with α replaced by $\alpha + 4\pi/3$ and $\alpha + 2\pi/3$.

We now consider (3.4.2). The term (3.4.5) contributes $-6\sqrt{3}s^2\log(s)$ to m and the term (3.4.7) contributes $9\sqrt{3}\log(2)s^2$, since for all α

$$\sum_{j=1}^{3} (3\sin^2(\alpha + \frac{2\pi j}{3}) - \cos^2(\alpha + \frac{2\pi j}{3})) = 3.$$

For term (3.4.6) we use the sharp estimate

$$\sum_{j=1}^{3} (3\sin^2(\alpha + \frac{2\pi j}{3}) - \cos^2(\alpha + \frac{2\pi j}{3}))\log|3\sin^2(\alpha + \frac{2\pi j}{3}) - \cos^2(\alpha + \frac{2\pi j}{3})| \leq 3\log(3),$$

which we prove in Lemma 3.6.7. Hence, for $s \leq 1/20$,

$$\begin{aligned}
m(\theta) &\geq -6\sqrt{3}s^2\log(s) + (9\sqrt{3}\log(2) - 3\sqrt{3}\log(3))s^2 + 90\sqrt{3}s^4\log s - 62s^4\\
&\geq (6\sqrt{3}\log(20) + 9\sqrt{3}\log(2) - 3\sqrt{3}\log(3) - \frac{90\sqrt{3}}{400}\log(20) - 62/400)s^2\\
&\approx 34.906\, s^2,
\end{aligned}$$

as claimed. $\square$

3.5 Estimating term II

Lemma 3.5.1. *For all $2 \le i,j \le 4$ and all f, it holds that*

$$II_{ij} \le \frac{101}{100}\pi \int_{B_1} |\theta|^2 |f(\theta)|^2 \,\mathrm{d}\mathcal{H}^2_H(\theta).$$

Proof. We first treat the term II_{44}, and later explain the changes for the other terms. We have

$$\begin{aligned} II_{44} &= \int_{B_4\times B_4} \delta\Big(1 - \frac{1-a(\theta')}{1-a(\theta)}\Big)|f(\theta)||f(\theta')| \,\mathrm{d}\mathcal{H}^2_H(\theta)\,\mathrm{d}\mathcal{H}^2_H(\theta') \\ &= \int_{B_1\times B_1} \delta\Big(1 - \frac{1-a(c_4+\theta')}{1-a(c_4+\theta)}\Big)|f(\theta)||f(\theta')| \,\mathrm{d}\mathcal{H}^2_H(\theta)\,\mathrm{d}\mathcal{H}^2_H(\theta'). \end{aligned}$$

We introduce polar coordinates $\theta = \theta(s,\alpha)$ as in (3.4.3) and write also $\theta' = \theta(t,\beta)$. With the definitions

$$h(s,t,\alpha,\beta) := \frac{1-a(c_4+\theta')}{1-a(c_4+\theta)} \quad \text{and} \quad g(s,\alpha) := |\theta|^2|f(\theta)|,$$

we obtain by changing variables

$$II_{44} = \int_0^{2\pi}\int_0^{2\pi}\int_0^{\varepsilon}\int_0^{\varepsilon} \delta(1-h(s,t,\alpha,\beta))g(s,\alpha)g(t,\beta)\,\frac{\mathrm{d}s}{s}\,\frac{\mathrm{d}t}{t}\,\mathrm{d}\alpha\,\mathrm{d}\beta. \tag{3.5.1}$$

Doing a Taylor expansion of $1 - a(c_4+\theta)$ at 0 yields (see Lemma 3.6.5)

$$h(s,t,\alpha,\beta) = \frac{t^2}{s^2}\frac{3\sin^2(\beta)-\cos^2(\beta)}{3\sin^2(\alpha)-\cos^2(\alpha)}\frac{1+\psi(t,\beta)}{1+\psi(s,\alpha)}, \tag{3.5.2}$$

where $\psi(s,\alpha)$ is a smooth function of s and α, and $\psi(s,\alpha) = O(s^2)$. If the last factor in (3.5.2) were equal to 1, then the inner two integrals in (3.5.1) would simplify to

$$\int_0^\infty g(s,\alpha)g(c(\alpha,\beta)s,\beta)\,\frac{\mathrm{d}s}{s},$$

for some constant $c(\alpha,\beta)$, which is easily estimated using Cauchy-Schwarz. The following is a perturbed version of this argument.

Fix α, β and write $h(s,t) = h(s,t,\alpha,\beta)$. Let $s(t)$ be defined implicitly by $h(s(t),t) = 1$ (note that s also depends on α and β). Then

$$\begin{aligned} \int_0^{\varepsilon}\int_0^{\varepsilon} \delta(1-h(s,t))g(s,\alpha)g(t,\beta)\,\frac{\mathrm{d}s}{s}\,\frac{\mathrm{d}t}{t} &= \int_0^{\varepsilon} g(t,\beta)g(s(t),\alpha)\frac{1}{|\partial_s h(s(t),t)|}\,\frac{1}{s(t)t}\,\mathrm{d}t \\ &= \int_0^{\varepsilon} g(t,\beta)g(s(t),\alpha)\frac{1}{2+s(t)\frac{\psi'(s(t),\alpha)}{1+\psi(s(t),\alpha)}}\,\frac{1}{t}\,\mathrm{d}t. \end{aligned} \tag{3.5.3}$$

Here we used that

$$\partial_s h(s,t) = -h(s,t)\Big(\frac{2}{s} + \frac{\psi'(s,\alpha)}{1+\psi(s,\alpha)}\Big),$$

and hence

$$-\partial_s h(s(t),t) = \frac{2}{s(t)} + \frac{\psi'(s(t),\alpha)}{1+\psi(s(t),\alpha)}.$$

Applying Cauchy-Schwarz, we obtain that (3.5.3) is bounded by

$$\left(\int_0^\varepsilon g(t,\beta)^2 \frac{1}{2 + s(t)\frac{\psi'(s(t),\alpha)}{1+\psi(s(t),\alpha)}} \frac{1}{t}\,\mathrm{d}t\right)^{1/2}$$

$$\cdot \left(\int_0^\varepsilon g(s(t),\alpha)^2 \frac{1}{2 + s(t)\frac{\psi'(s(t),\alpha)}{1+\psi(s(t),\alpha)}} \frac{1}{t}\,\mathrm{d}t\right)^{1/2}. \quad (3.5.4)$$

After substituting $s = s(t)$ in the second integral, its integrand becomes the same as in the first one, but with the roles of (s,α) and (t,β) interchanged. By Lemma 3.6.5, it holds for $s \le 1/20$ that

$$|\psi(s,\alpha)| < \frac{1}{100} \quad \text{and} \quad |\psi'(s,\alpha)| \le \frac{1}{10},$$

giving

$$\left| s\frac{\psi'(s,\alpha)}{1+\psi(s,\alpha)} \right| \le \frac{1}{198}.$$

Thus, the factor in the integrals in (3.5.4) is bounded above by $198/395 < 101/200$. It follows that

$$\begin{aligned} II_{44} &\le \frac{101}{200} \int_0^{2\pi}\int_0^{2\pi} \left(\int_0^\varepsilon g(t,\beta)^2 \frac{\mathrm{d}t}{t}\right)^{1/2} \left(\int_0^\varepsilon g(s,\alpha)^2 \frac{\mathrm{d}s}{s}\right)^{1/2} \mathrm{d}\alpha\,\mathrm{d}\beta \\ &\le \frac{101}{100}\pi \int_0^{2\pi}\int_0^\varepsilon |g(s,\alpha)|^2 \frac{\mathrm{d}t}{t}\,\mathrm{d}\alpha = \frac{101}{100}\pi \int_{B_1} |\theta|^2 |f(\theta)|^2 \,\mathrm{d}\mathcal{H}^2_H(\theta). \end{aligned}$$

For the other eight integrals the same estimate holds: By the argument in the proof of Lemma 3.4.1, changing c_4 to some other c_j only changes the expansion in (3.5.2) by a translation in α and β. Then the rest of the argument goes through exactly as for II_{44}. □

3.6 Technical estimates

Here we prove the computational lemmas that were used in the main argument.

We have the following explicit formula for ρ (see [23], Lemma 8):

$$\rho(r) = \frac{4}{r}\int_{A(r)}^1 \frac{\mathrm{d}u}{\sqrt{1-u^2}\sqrt{\frac{(1-r)^2}{2r} + 1 - u}\sqrt{\frac{(3+r)(1-r)}{2r} + 1 + u}} \quad (3.6.1)$$

with

$$A(r) = -1 + \max\left\{0, \frac{(3+r)(r-1)}{2r}\right\}.$$

From this, we obtain the following asymptotic formula.

Lemma 3.6.1. *Let ρ be defined by $\rho(|x|) = \sigma * \sigma * \sigma(x)$. Then we have for all r with $|r-1| \le 1/10$*

$$|\rho(r) + 6\log|1-r| - 12\log 2| \le -22|r-1|\log|r-1| + 23|r-1|.$$

We have not tried to optimize the error in this estimate. We give an elementary, self-contained proof below. For an alternative proof one can use the identity (see [100, p. 17] or [16, eq. (1.2)])

$$\rho(x) = \begin{cases} \frac{16}{\sqrt{(x+1)^3(3-x)}} K\left(\sqrt{\frac{16x}{(x+1)^3(3-x)}}\right) & \text{if } 0 \le x < 1 \\ \frac{4}{\sqrt{x}} K\left(\sqrt{\frac{(x+1)^3(3-x)}{16x}}\right) & \text{if } 1 < x \le 3 \\ 0 & \text{if } x > 3 \end{cases} \tag{3.6.2}$$

where

$$K(k) = \int_0^1 \frac{1}{\sqrt{1-x^2}\sqrt{1-k^2x^2}}\,\mathrm{d}x$$

is the complete elliptic integral of the first kind, together with known asymptotics for $K(k)$ as $k \nearrow 1$.

We first prove some auxiliary lemmas.

Lemma 3.6.2. *For all $\delta > 0$ it holds that*

$$0 \le \int_0^1 \frac{1}{\sqrt{u}\sqrt{u+\delta}}\,\mathrm{d}u - \log\left(\frac{4}{\delta}\right) \le \frac{1}{2}\delta.$$

Proof. We have

$$\int_0^1 \frac{1}{\sqrt{u}\sqrt{u+\delta}}\,\mathrm{d}u = -\log(\delta) + 2\log\left(1+\sqrt{1+\delta}\right).$$

Furthermore, by the mean value theorem, there exists $0 < \delta' < \delta$ such that

$$\log\left(1+\sqrt{1+\delta}\right) = \log(2) + \delta g(\delta')$$

where

$$0 < g(\delta) = \frac{1}{2(1+\sqrt{1+\delta})\sqrt{1+\delta}} \le \frac{1}{4}$$

is the derivative of $\log(1+\sqrt{1+\delta})$. □

Lemma 3.6.3. *For all $0 < a, b < 1$, we have:*

$$\left|\int_0^1 \frac{1}{\sqrt{1-x^2}\sqrt{a+1-x}\sqrt{b+1+x}}\,\mathrm{d}x - \int_0^1 \frac{1}{\sqrt{1-x^2}\sqrt{a+1-x}\sqrt{1+x}}\,\mathrm{d}x\right|$$
$$\le \frac{b}{2}\left(\log\left(\frac{4}{a}\right) + \frac{a}{2}\right).$$

Proof. By the mean value theorem, we have for all $x \geq 0$

$$|(b+1+x)^{-1/2} - (1+x)^{-1/2}| \leq \frac{1}{2}b.$$

Hence the left hand side of the claimed inequality is estimated by

$$\frac{b}{2}\int_0^1 \frac{1}{\sqrt{1-x}\sqrt{a+1-x}}\,\mathrm{d}x \leq \frac{b}{2}\Big(\log\Big(\frac{4}{a}\Big) + \frac{a}{2}\Big),$$

where we applied Lemma 3.6.2. □

Lemma 3.6.4. *For all $1 > a > 0$, we have*

$$\Big|\int_0^1 \frac{1}{(1+x)\sqrt{1-x}\sqrt{a+1-x}}\,\mathrm{d}x - \frac{1}{2}\log\Big(\frac{8}{a}\Big)\Big| \leq \frac{1}{2}a\log\Big(1+\frac{1}{a}\Big).$$

Proof. We have with $v = 1-x$

$$\int_0^1 \frac{1}{(1+x)\sqrt{1-x}\sqrt{a+1-x}}\,\mathrm{d}x = \int_0^1 \frac{1}{(2-v)\sqrt{v}\sqrt{a+v}}\,\mathrm{d}v$$

which can be expanded to equal

$$\frac{1}{2}\int_0^1 \frac{1}{\sqrt{v}\sqrt{a+v}}\,\mathrm{d}v + \frac{1}{2}\int_0^1 \frac{1}{2-v}\,\mathrm{d}v - \frac{a}{2}\int_0^1 \frac{1}{(2-v)\sqrt{a+v}(\sqrt{v}+\sqrt{a+v})}\,\mathrm{d}v.$$

Computing the second integral and using Lemma 3.6.2 for the first one yields the main term $\log(8/a)/2$. For the error estimate we combine Lemma 3.6.2 and the bound

$$\int_0^1 \frac{1}{(2-v)\sqrt{a+v}(\sqrt{v}+\sqrt{a+v})}\,\mathrm{d}v \leq \int_0^1 \frac{1}{v+a}\,\mathrm{d}v = \log\Big(1+\frac{1}{a}\Big),$$

and note that the errors have opposite signs. □

Proof of Lemma 3.6.1. We start with the case $r = 1-\varepsilon < 1$. By (3.6.1), we have

$$\frac{1-\varepsilon}{4}\rho(1-\varepsilon) = \int_{-1}^1 \frac{1}{\sqrt{1-u^2}\sqrt{\frac{\varepsilon^2}{2-2\varepsilon}+1-u}\sqrt{\frac{(4-\varepsilon)\varepsilon}{2-2\varepsilon}+1+u}}\,\mathrm{d}u.$$

Combining Lemma 3.6.3 and Lemma 3.6.4 with $a = \varepsilon^2/(2-2\varepsilon)$ and $b = (4-\varepsilon)\varepsilon/(2-2\varepsilon)$, we obtain that this integral equals

$$\frac{1}{2}\Big(\log\Big(\frac{8}{a}\Big) + \log\Big(\frac{8}{b}\Big)\Big) + E = 3\log(2) - \frac{3}{2}\log(\varepsilon) - \log(2-2\varepsilon) + \frac{1}{2}\log(4-\varepsilon) + E$$

with

$$|E| \leq \frac{1}{2}\Big(b\log\Big(\frac{4}{a}\Big) + a\log\Big(\frac{4}{b}\Big) + ab + a\log\Big(1+\frac{1}{a}\Big) + b\log\Big(1+\frac{1}{b}\Big)\Big). \tag{3.6.3}$$

It is easy to see that

$$\Big|\frac{1}{2}\log(4-\varepsilon) - \log(2-2\varepsilon)\Big| \leq \frac{\varepsilon}{2}.$$

Further, one verifies that, when $0 < \varepsilon \le 1/10$,

$$a \le \frac{1}{18}\varepsilon, \quad b \le \frac{19}{9}\varepsilon, \quad \log\left(\frac{4}{a}\right) \le 3\log(2) - 2\log(\varepsilon), \quad \log\left(\frac{4}{b}\right) \le \log(2) - \log(\varepsilon)$$

and

$$\log\left(1 + \frac{1}{a}\right) \le \log(2) - 2\log(\varepsilon), \quad \log\left(1 + \frac{1}{b}\right) \le -\log(\varepsilon).$$

Using this, one can check that

$$|E| \le \frac{13}{4}\varepsilon\log\left(\frac{1}{\varepsilon}\right) + \frac{5}{2}\varepsilon.$$

To summarize, we have shown that

$$\left|\frac{1-\varepsilon}{4}\rho(1-\varepsilon) - 3\log(2) + \frac{3}{2}\log(\varepsilon)\right| \le \frac{13}{4}\varepsilon\log\left(\frac{1}{\varepsilon}\right) + 3\varepsilon.$$

We multiply by $4/(1-\varepsilon)$, and use that $|4/(1-\varepsilon) - 4| \le 40\varepsilon/9$ to obtain

$$|\rho(1-\varepsilon) - 12\log(2) + 6\log(\varepsilon)| \le 22\varepsilon\log\left(\frac{1}{\varepsilon}\right) + 23\varepsilon.$$

Now we turn to the case $r = 1 + \varepsilon > 1$. There we have

$$\begin{aligned}\rho(1+\varepsilon) &= \frac{4}{1+\varepsilon}\int_{-1+\frac{(4+\varepsilon)\epsilon}{2+2\varepsilon}}^{1} \frac{1}{\sqrt{1-u^2}\sqrt{\frac{\varepsilon^2}{2+2\varepsilon}+1-u}\sqrt{-\frac{(4+\varepsilon)\varepsilon}{2+2\varepsilon}+1+u}}\,\mathrm{d}u \\ &= \frac{16}{4-\varepsilon^2}\int_{-1}^{1} \frac{1}{\sqrt{1-v^2}\sqrt{\frac{2\varepsilon^2}{4-\varepsilon^2}+1-v}\sqrt{\frac{8\varepsilon}{4-\varepsilon^2}+1+v}}\,\mathrm{d}v.\end{aligned}$$

We first approximate the integral. We can argue as in the case $r < 1$, now with $a = 2\varepsilon^2/(4-\varepsilon^2)$ and $b = 8\varepsilon/(4-\varepsilon^2)$. The main term is easily seen to be the same as in the case $r < 1$, and the error is bounded by

$$-\log\left(1 - \frac{\varepsilon^2}{4}\right) + E \le \frac{\varepsilon}{40} + E,$$

with E satisfying (3.6.3). Now we have

$$a \le \frac{1}{15}\varepsilon, \quad b \le \frac{800}{399}\varepsilon, \quad \log\left(\frac{4}{a}\right) \le 3\log(2) - 2\log(\varepsilon), \quad \log\left(\frac{4}{b}\right) \le \log(2) - \log(\varepsilon)$$

and

$$\log\left(1 + \frac{1}{a}\right) \le \log\left(\frac{201}{100}\right) - 2\log(\varepsilon), \quad \log\left(1 + \frac{1}{b}\right) \le -\log(\varepsilon).$$

Using this, we obtain

$$|E| + \frac{\varepsilon}{40} \le \frac{13}{4}\varepsilon\log\left(\frac{1}{\varepsilon}\right) + \frac{9}{4}\varepsilon.$$

In other words, it holds that

$$\left|\frac{4-\varepsilon^2}{16}\rho(1+\varepsilon) - 3\log(2) + \frac{3}{2}\log(\varepsilon)\right| \le \frac{13}{4}\varepsilon\log\left(\frac{1}{\varepsilon}\right) + \frac{9}{4}\varepsilon.$$

We multiply by $16/(4-\varepsilon^2)$ and use that $|16/(4-\varepsilon^2) - 4| \le 40\varepsilon/399$ to obtain

$$|\rho(1+\varepsilon) - 12\log(2) + 6\log(\varepsilon)| \le 14\varepsilon\log\left(\frac{1}{\varepsilon}\right) + 9\varepsilon.$$

This completes the proof. $\square$

Lemma 3.6.5. *Let θ be given by* (3.4.3)*. Then it holds that*

$$a(c_4+\theta)-1 = s^2(3\sin^2(\alpha)-\cos^2(\alpha))(1+\psi(s,\alpha)),$$

where $\psi(s,\alpha)$ is a smooth function satisfying the following estimates:

$$|\psi(s,\alpha)| \leq \frac{7}{24}s^2 + \frac{17}{720}s^4 + s^6 e^{\sqrt{2}s}$$
$$|\psi'(s,\alpha)| \leq \frac{14}{24}s + \frac{17}{180}s^3 + 2s^5 e^{\sqrt{2}s}.$$

Proof. By the definition of h, the trigonometric identities and the Taylor expansion of cos, we have

$$\begin{aligned}
a(c_4+\theta)-1 &= a((0,0,\pi)+\theta)-1\\
&= (\cos(\theta_1)+\cos(\theta_2)-\cos(\theta_3))^2 + (\sin(\theta_1)+\sin(\theta_2)-\sin(\theta_3))^2 - 1\\
&= 2+2\cos(\theta_1-\theta_2)-2\cos(\theta_1-\theta_3)-2\cos(\theta_2-\theta_3)\\
&= 2\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k)!}((\theta_1-\theta_2)^{2k}-(\theta_1-\theta_3)^{2k}-(\theta_2-\theta_3)^{2k}) \qquad (3.6.4)\\
&=: \sum_{k=1}^{\infty} s^{2k}\frac{(-1)^k}{(2k)!}P_{2k}(\sin(\alpha),\cos(\alpha)).
\end{aligned}$$

It follows from (3.6.4) that each P_{2k} vanishes when $\theta_1=\theta_3$ and when $\theta_2=\theta_3$, which is equivalent to $\alpha=\pm\pi/6$, or to $\cos(\alpha)=\pm\sqrt{3}\sin(\alpha)$. Hence, the homogeneous polynomial $P_{2k}(X,Y)$ vanishes on the lines $\sqrt{3}X+Y=0$ and $\sqrt{3}X-Y=0$. We conclude that for all k, the factor $3X^2-Y^2$ divides $P_{2k}(X,Y)$. Define Q_{2k} by

$$Q_{2k}(X,Y)(3X^2-Y^2) = (-1)^k P_{2k}(X,Y).$$

Then we have, using that $Q_2=1$:

$$a(c_4+\theta)-1 = s^2(3\sin^2(\alpha)-\cos^2(\alpha))(1+\psi(s,\alpha))$$

where ψ is defined by

$$\psi(s,\alpha) = \sum_{k=2}^{\infty} s^{2k-2}\frac{1}{(2k)!}Q_{2k}(\cos(\alpha),\sin(\alpha)).$$

Now we fix k and estimate

$$p(\alpha) := P_{2k}(\sin(\alpha),\cos(\alpha)) \quad \text{and} \quad q(\alpha) := Q_{2k}(\sin(\alpha),\cos(\alpha)).$$

By (3.4.3), we have that

$$\begin{aligned}
\theta_1-\theta_2 &= \sqrt{2}\cos(\alpha),\\
\theta_3-\theta_1 &= -\frac{1}{\sqrt{2}}\cos(\alpha)-\frac{\sqrt{3}}{\sqrt{2}}\sin(\alpha) = \sqrt{2}\cos(\alpha+\frac{2\pi}{3}),\\
\theta_2-\theta_3 &= -\frac{1}{\sqrt{2}}\cos(\alpha)+\frac{\sqrt{3}}{\sqrt{2}}\sin(\alpha) = \sqrt{2}\cos(\alpha-\frac{2\pi}{3}).
\end{aligned}$$

k	$P_{2k}(X,Y)$	$Q_{2k}(X,Y)$
1	$-3X^2+Y^2$	1
2	$-9X^4-18X^2Y^2+7Y^4$	$-3X^2-7Y^2$
3	$-\frac{1}{2}(27X^6+135X^4Y^2+45X^2Y^4-31Y^6)$	$\frac{1}{2}(9X^4+48X^2Y^2+31Y^4)$

Table 3.1: The polynomials P_{2k} and Q_{2k} for small values of k.

Thus, by (3.6.4),

$$p(\alpha) = 2^{k+1}(-1)^k(\cos(\alpha)^{2k} - \cos(\alpha+\frac{2\pi}{3})^{2k} - \cos(\alpha-\frac{2\pi}{3})^{2k}),$$

Taking derivatives, and noting that the terms inside the brackets are each at most 1, we obtain:

$$|p(\alpha)| \leq 6\cdot 2^k, \quad |p'(\alpha)| \leq 12k2^k \quad \text{and} \quad |p''(\alpha)| \leq 24k^22^k.$$

Denote

$$q(\alpha) := \frac{p(\alpha)}{3\sin^2(\alpha)-\cos^2(\alpha)} = \frac{p(\alpha)}{(\sqrt{3}\sin(\alpha)-\cos(\alpha))(\sqrt{3}\sin(\alpha)+\cos(\alpha))}.$$

If both factors $|\sqrt{3}\sin(\alpha)\pm\cos(\alpha)|$ are at least $1/2$, we have that

$$q(\alpha) \leq 24\cdot 2^k.$$

If not, then $|\alpha-\pi/6| < 1/5$ or $|\alpha+\pi/6| < 1/5$. Without loss of generality we are in the first case. Then, by Taylor's formula:

$$\Big|\frac{p(\alpha)}{\alpha-\pi/6} - p'(\pi/6)\Big| \leq \frac{1}{2}|\alpha-\pi/6|\sup|p''| \leq \frac{1}{10}24k^22^k,$$

hence

$$\Big|\frac{p(\alpha)}{\alpha-\pi/6}\Big| \leq 15k^22^k.$$

Furthermore, since $|\alpha-\pi/6| \leq 1/5$,

$$\begin{aligned}\Big|\frac{\alpha-\pi/6}{(\sqrt{3}\sin(\alpha)-\cos(\alpha))(\sqrt{3}\sin(\alpha)+\cos(\alpha))}\Big| &\leq 2\Big|\frac{\alpha-\pi/6}{\sqrt{3}\sin(\alpha)-\cos(\alpha)}\Big| \\ &\leq \frac{1/5}{\sin(1/5)} < 2.\end{aligned}$$

Multiplying the last two estimates, we conclude that $|q| \leq 30k^22^k$. We also directly compute for small k:

$$|Q_4(\sin(\alpha),\cos(\alpha))| = |-7\cos^2(\alpha) - 3\sin^2(\alpha)| \leq 7$$

and

$$|Q_6(\sin(\alpha),\cos(\alpha))| = \frac{1}{2}|9\sin^4(\alpha) + 48\sin^2(\alpha)\cos^2(\alpha) + 31\cos^2(\alpha)| \leq \frac{5125}{312} < 17.$$

Plugging in these estimates, we obtain

$$|\psi(s,\alpha)| \le \frac{7}{24}s^2 + \frac{17}{720}s^4 + \sum_{k=4}^{\infty} \frac{60k^2}{(2k)!}(\sqrt{2}s)^{2k-2} \le \frac{7}{24}s^2 + \frac{17}{720}s^4 + s^6 e^{\sqrt{2}s}$$

and

$$|\psi'(s,\alpha)| \le \frac{14}{24}s + \frac{17}{180}s^3 + \sqrt{2}\sum_{k=4}^{\infty} \frac{60k^2}{(2k-1)!}(\sqrt{2}s)^{2k-3} \le \frac{14}{24}s + \frac{17}{180}s^4 + 2s^5 e^{\sqrt{2}s},$$

as claimed. □

Lemma 3.6.6. *Let θ be given by* (3.4.3). *Then for all* $0 \le s \le 1/20$, *we have*

$$\begin{aligned}(a(c_4+\theta)-1)\rho(\sqrt{a(c_4+\theta)}) &= -12s^2(3\sin^2(\alpha)-\cos^2(\alpha))\log(s)\\ &\quad - 6s^2(3\sin^2(\alpha)-\cos^2(\alpha))\log|3\sin^2(\alpha)-\cos^2(\alpha)|\\ &\quad + 18\log 2\, s^2(3\sin^2(\alpha)-\cos^2(\alpha)) + E,\end{aligned}$$

with

$$|E| \le -180s^4\log s + 71s^4.$$

Proof. By Lemma 3.6.1, it holds that

$$\begin{aligned}(x^2-1)\rho(x) &= -6(x^2-1)\log|x-1| + 12\log(2)(x^2-1) + (x^2-1)E_1\\ &= -6(x^2-1)\log|x^2-1| + 18\log(2)(x^2-1) + (x^2-1)E_1\\ &\quad + 6(x^2-1)\log(1+\frac{1}{2}(x-1)),\end{aligned}\tag{3.6.5}$$

where

$$|E_1| \le -22|x-1|\log|x-1| + 23|x-1|.$$

Denote also the last term in (3.6.5) by E_2. We set

$$x = \sqrt{a(c_4+\theta)}.$$

Lemma 3.6.5 implies that $|x-1| \le |x^2-1| \le 2s^2$. Using this and monotonicity of $r\log r$, we obtain

$$|(x^2-1)E_1| \le |x^2-1|(-22|x-1|\log|x-1| + 23|x-1|) \le -176s^4\log(s) + 32s^4 \tag{3.6.6}$$

and

$$|E_2| \le 6|x^2-1|\left|\log\left(1+\frac{1}{2}(x-1)\right)\right| \le 24s^4. \tag{3.6.7}$$

By Lemma 3.6.5, it holds that

$$\begin{aligned}&-6(x^2-1)\log|x^2-1|\\ &= -6s^2(3\sin^2(\alpha)-\cos^2(\alpha))(1+\psi(s,\alpha))(2\log(s)+\log(3\sin^2(\alpha)-\cos^2(\alpha))\\ &\quad + \log(1+\psi(s,\alpha)))\\ &= -12s^2\log(s)(3\sin^2(\alpha)-\cos^2(\alpha)) &(3.6.8)\\ &\quad - 6s^2(3\sin^2(\alpha)-\cos^2(\alpha))\log|3\sin^2(\alpha)-\cos^2(\alpha)| &(3.6.9)\\ &\quad - 6s^2(3\sin^2(\alpha)-\cos^2(\alpha))\log(1+\psi(s,\alpha))) &(3.6.10)\\ &\quad - 6s^2\psi(s,\alpha)(3\sin^2(\alpha)-\cos^2(\alpha))\\ &\qquad \times (2\log(s)+\log|3\sin^2(\alpha)-\cos^2(\alpha)|+\log(1+\psi(s,\alpha))). &(3.6.11)\end{aligned}$$

The term (3.6.10) is bounded by $18s^2|\psi(s,\alpha)| \leq 6s^4$. The term (3.6.11) is bounded by

$$-18s^2\log(s)|\psi(s,\alpha)| + 4s^2|\psi(s,\alpha)| + 18s^2\psi(s,\alpha)^2 \leq -6s^4\log(s) + 2s^4.$$

For the second term in (3.6.5), we have

$$\begin{aligned}&18\log(2)(x^2-1)\\&= 18\log(2)s^2(3\sin^2(\alpha)-\cos^2(\alpha)) + 18\log(2)s^2(3\sin^2(\alpha)-\cos^2(\alpha))\psi(s,\alpha),\end{aligned} \tag{3.6.12}$$

with the second term bounded by

$$27\log(2)s^2|\psi(s,\alpha)| \leq 9\log(2)s^4.$$

Putting together the main terms (3.6.8), (3.6.9) and (3.6.12), and the estimates for the error terms in (3.6.6), (3.6.7), (3.6.10), (3.6.11) and in (3.6.12), one obtains the lemma. □

Lemma 3.6.7. *For all α, it holds that*

$$\sum_{j=1}^{3}\left(3\sin^2\left(\alpha+\frac{2\pi j}{3}\right)-\cos^2\left(\alpha+\frac{2\pi j}{3}\right)\right)\log\left|3\sin^2\left(\alpha+\frac{2\pi j}{3}\right)-\cos^2\left(\alpha+\frac{2\pi j}{3}\right)\right|$$
$$\leq 3\log(3).$$

Proof. Let

$$a_j = \sin^2(\alpha+\frac{2\pi j}{3}) - \frac{1}{3}\cos^2(\alpha+\frac{2\pi j}{3}) = \frac{1}{3}-\frac{2}{3}\cos(2\alpha+\frac{4\pi j}{3}).$$

It is easy to check that

$$a_1+a_2+a_3 = 1 \quad \text{and} \quad a_1^2+a_2^2+a_3^2 = 1. \tag{3.6.13}$$

Defining

$$b_j = \frac{a_j+a_{j-1}}{2}$$

(note that $a_{j+3}=a_j$), it follows that

$$b_1+b_2+b_3 = 1 \quad \text{and} \quad b_1^2+b_2^2+b_3^2 = 1/2,$$

hence $b_1, b_2, b_3 \geq 0$. Using Jensen's inequality, we deduce

$$\sum_{j=1}^{3} a_j\log(|a_j|) = \sum_{j=1}^{3} b_j\log(\frac{|a_j||a_{j-1}|}{|a_{j-2}|}) \leq \log(\sum_{j=1}^{3} b_j\frac{|a_j||a_{j-1}|}{|a_{j-2}|}).$$

By (3.6.13), we have that

$$2a_ja_{j-1} = (a_j+a_{j-1})^2-(a_j^2+a_{j-1}^2) = (1-a_{j-2})^2-(1-a_{j-2}^2) = 2a_{j-2}(a_{j-2}-1).$$

Thus, using again (3.6.13)

$$\sum_{j=1}^{3} b_j\frac{|a_j||a_{j-1}|}{|a_{j-2}|} = \sum_{j=1}^{3} b_j(1-a_{j-2}) = 1.$$

We conclude that

$$\sum_{j=1}^{3} 3a_j\log(|3a_j|) = 3\log(3) + 3\sum_{j=1}^{3} a_j\log|a_j| \leq 3\log 3.$$

□

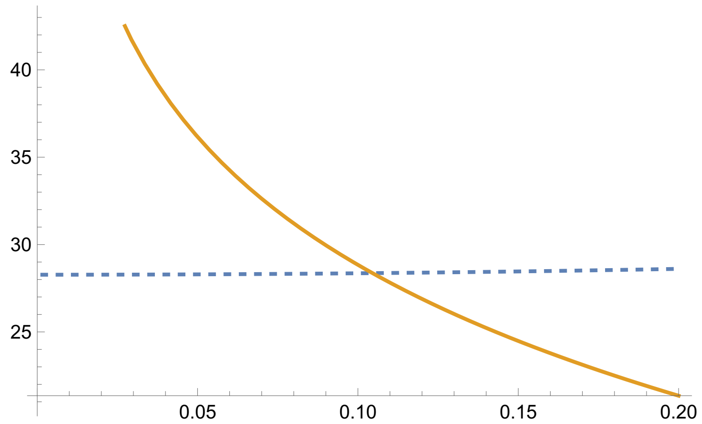

Figure 3.2: The left hand side (solid) and the right hand side (dashed) of (3.7.1).

3.7 Discussion

3.7.1 Optimal value of ε

An inspection of the above argument shows that $Q(f) \geq 0$ for all $f \in V_\varepsilon$ as long as

$$\inf_{\theta \in H, |\theta| \leq \varepsilon} \frac{1}{2} \sum_{j=2}^{4} (a(\theta + c_j) - 1)\rho(\sqrt{a(\theta + c_j)}) \geq 18\pi \sup_{s \leq \varepsilon, \alpha \in [0, 2\pi]} \frac{1}{2 + s\frac{\psi'(s,\alpha)}{1+\psi(s,\alpha)}}. \tag{3.7.1}$$

(Non-rigorous) numerical computations suggest that this inequality holds up to $\varepsilon = 0.104$. The constant ε' in Corollary 3.1.3 could then be increased to 0.063.

3.7.2 Fourier coefficients of Q

In [3], some numerical observations on the Fourier coefficients

$$\hat{B}(k, l) := B(\omega_1^{k_1}\omega_2^{k_2}\omega_3^{k_3}, \omega_4^{l_1}\omega_5^{l_2}\omega_6^{l_3})$$

of B with $k_1 + k_2 + k_3 = l_1 + l_2 + l_3 = 0$ are discussed. Namely, they are very large only when k is very close to l and when $k_1^2 + k_2^2 + k_3^2 \approx l_1^2 + l_2^2 + l_3^2$. We can explain this using Lemma 3.3.2 as follows.

By Lemma 3.3.2 and since $\lambda_0 = 1$, for all $f \in X_0$ the form $B(f, f)$ can be expressed as

$$\int_C m(\theta)|f(\theta)|^2 \,\mathrm{d}\mathcal{H}_C^2(\theta) + \int_{C^2} n(\theta)\delta(a(\theta) - a(\theta'))f(\theta)\overline{f(\theta')} \,\mathrm{d}\mathcal{H}_C^2(\theta) \,\mathrm{d}\mathcal{H}_C^2(\theta')$$

for certain functions m and n.

The first term is a multiplier, hence it acts on the Fourier side by convolution with a fixed bump function. This bump function decays at least like $|k - l|^{-3}$, because the third derivative of m is still integrable. This explains the large coefficients when k is close to l.

The Fourier coefficients of the second term are the Fourier coefficients of the measure $\mu := n(\theta)\delta(a(\theta) - a(\theta'))$ supported on the 3-manifold

$$M := \{(x, y) \in C^2 \,:\, a(x) = a(y)\} \subset \mathbb{R}^6.$$

The measure μ has a smooth, bounded density with respect to the Hausdorff measure on this manifold, except in the critical points of a. The Fourier transform of the parts where the measure has a smooth, bounded density can be estimated using the method of stationary phase and are of lower order than the contribution of the critical points. To explain what happens at a critical point (where $\det D^2 a \neq 0$), we choose coordinates x_1, x_2, y_1, y_2 for C^2, such that the critical point of a is at 0. After a scaling in a and a linear change of variables, either

$$a(x) = x_1^2 + x_2^2 + O(|x|^3) \quad \text{or} \quad a(x) = x_1^2 - x_2^2 + O(|x|^3). \tag{3.7.2}$$

Thus, ignoring higher order terms,

$$\delta(a(x) - a(y)) \approx \delta(|x|^2 - |y|^2) \quad \text{or} \quad \delta(a(x) - a(y)) \approx \delta(x_1^2 - x_2^2 - y_1^2 + y_2^2).$$

The Fourier transforms of these measures can be explicitly computed, in fact, they are up to a constant factor their own Fourier transform. Now, a has one local maximum and two local minima, which together with the above discussion explain why $\hat{B}(k, l)$ is very large on the cone $|k|^2 = |l|^2$. The contribution of all other critical points is of smaller order, since the weight n vanishes there.

This discussion can be turned into a rigorous proof that the Fourier coefficients of μ concentrate near the cone $|k|^2 = |l|^2$. However, we can only show that they concentrate in e.g.

$$\{(k, l) : ||k| - |l|| \leq C|k|^{1/2}\},$$

and not in an $O(1)$ neighborhood of the cone, because of the higher order terms in (3.7.2).

Bibliography

[1] T. C. Anderson, D. Maldague, L. B. Pierce, and P.-L. Yung. "On polynomial Carleson operators along quadratic hypersurfaces". In: *J. Geom. Anal.* 34.10 (2024), Paper No. 321, 47. DOI: `10.1007/s12220-024-01676-9` (cit. on pp. 5, 22).

[2] N. Y. Antonov. "Convergence of Fourier series". In: *Proceedings of the XX Workshop on Function Theory (Moscow, 1995).* Vol. 2. 2. 1996, pp. 187–196 (cit. on pp. 3, 22).

[3] J. Barker, C. Thiele, and P. Zorin-Kranich. "Band-limited maximizers for a Fourier extension inequality on the circle, II". In: *Exp. Math.* 32.2 (2023), pp. 280–293. DOI: `10.1080/10586458.2021.1926011` (cit. on pp. 19, 101, 102, 119).

[4] M. Bateman and C. Thiele. "L^p estimates for the Hilbert transforms along a one-variable vector field". In: *Anal. PDE* 6.7 (2013), pp. 1577–1600. DOI: `10.2140/apde.2013.6.1577` (cit. on p. 54).

[5] L. Becker. *A degree one Carleson operator along the paraboloid.* 2023. arXiv: `2312.01134 [math.CA]` (cit. on pp. i, 6, 21).

[6] L. Becker. "Maximal polynomial modulations of singular Radon transforms". In: *J. Funct. Anal.* 286.6 (2024), Paper No. 110299, 46. DOI: `10.1016/j.jfa.2023.110299` (cit. on pp. 5, 23).

[7] L. Becker. "Sharp Fourier extension for functions with localized support on the circle". In: *Rev. Mat. Iberoam.* (2024) (cit. on pp. i, 20, 101).

[8] L. Becker, P. Durcik, and F. Y.-H. Lin. *On trilinear singular Brascamp-Lieb integrals.* 2024. arXiv: `2411.00141 [math.CA]` (cit. on pp. i, 13–15, 55).

[9] L. Becker, F. van Doorn, A. Jamneshan, R. Srivastava, and C. Thiele. *Carleson Operators on Doubling Metric Measure Spaces.* 2024. arXiv: `2405.06423 [math.CA]` (cit. on p. 4).

[10] D. Beltran, S. Guo, and J. Hickman. *On a planar Pierce–Yung operator.* 2024. arXiv: `2407.07563 [math.CA]` (cit. on pp. 5, 22).

[11] D. Beltran, J. Roos, and A. Seeger. "Multi-scale sparse domination". In: *Mem. Amer. Math. Soc.* 298.1491 (2024), pp. v+104. DOI: `10.1090/memo/1491` (cit. on p. 7).

[12] C. Benea and C. Muscalu. "Multiple vector-valued inequalities via the helicoidal method". In: *Anal. PDE* 9.8 (2016), pp. 1931–1988. DOI: `10.2140/apde.2016.9.1931` (cit. on p. 65).

[13] J. Bennett, A. Carbery, M. Christ, and T. Tao. "The Brascamp-Lieb inequalities: finiteness, structure and extremals". In: *Geom. Funct. Anal.* 17.5 (2008), pp. 1343–1415. DOI: 10.1007/s00039-007-0619-6 (cit. on pp. 9, 13, 57, 60, 66, 70).

[14] F. Bernicot. "Fiber-wise Calderón-Zygmund decomposition and application to a bi-dimensional paraproduct". In: *Illinois J. Math.* 56.2 (2012), pp. 415–422 (cit. on p. 64).

[15] G. D. Birkhoff. "Proof of the ergodic theorem." In: *Proc. Natl. Acad. Sci. USA* 17 (1931), pp. 656–660. DOI: 10.1073/pnas.17.12.656 (cit. on p. 11).

[16] J. M. Borwein, A. Straub, J. Wan, and W. Zudilin. "Densities of short uniform random walks". In: *Canad. J. Math.* 64.5 (2012). With an appendix by Don Zagier, pp. 961–990. DOI: 10.4153/CJM-2011-079-2 (cit. on p. 112).

[17] J. Bourgain. "Pointwise ergodic theorems for arithmetic sets". In: *Inst. Hautes Études Sci. Publ. Math.* 69 (1989). With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein, pp. 5–45 (cit. on p. 11).

[18] A. P. Calderón. "Cauchy integrals on Lipschitz curves and related operators". In: *Proc. Nat. Acad. Sci. U.S.A.* 74.4 (1977), pp. 1324–1327. DOI: 10.1073/pnas.74.4.1324 (cit. on p. 9).

[19] A. P. Calderón. "Ergodic theory and translation-invariant operators". In: *Proc. Nat. Acad. Sci. U.S.A.* 59 (1968), pp. 349–353. DOI: 10.1073/pnas.59.2.349 (cit. on p. 11).

[20] A. P. Calderón and A. Zygmund. "On singular integrals". In: *Amer. J. Math.* 78 (1956), pp. 289–309. DOI: 10.2307/2372517 (cit. on p. 61).

[21] L. Carleson. "On convergence and growth of partial sums of Fourier series". In: *Acta Math.* 116 (1966), pp. 135–157. DOI: 10.1007/BF02392815 (cit. on pp. 2, 3, 22, 63).

[22] L. Carleson and P. Sjölin. "Oscillatory integrals and a multiplier problem for the disc". In: *Studia Math.* 44 (1972), pp. 287–299. DOI: 10.4064/sm-44-3-287-299 (cit. on p. 17).

[23] E. Carneiro, D. Foschi, D. Oliveira e Silva, and C. Thiele. "A sharp trilinear inequality related to Fourier restriction on the circle". In: *Rev. Mat. Iberoam.* 33.4 (2017), pp. 1463–1486. DOI: 10.4171/RMI/978 (cit. on pp. 18, 19, 101, 103, 111).

[24] E. Carneiro and D. Oliveira e Silva. "Some sharp restriction inequalities on the sphere". In: *Int. Math. Res. Not. IMRN* 17 (2015), pp. 8233–8267. DOI: 10.1093/imrn/rnu194 (cit. on p. 103).

[25] J. Chen, M. Hsu, and F. Y.-H. Lin. *A sharp Hörmander condition for bilinear Fourier multipliers with Lipschitz singularities.* 2024. arXiv: 2403.04721 [math.CA] (cit. on p. 65).

[26] M. Christ, A. Nagel, E. M. Stein, and S. Wainger. "Singular and maximal Radon transforms: analysis and geometry". In: *Ann. of Math. (2)* 150.2 (1999), pp. 489–577. DOI: 10.2307/121088 (cit. on p. 5).

[27] M. Christ and S. Shao. "Existence of extremals for a Fourier restriction inequality". In: *Anal. PDE* 5.2 (2012), pp. 261–312. DOI: 10.2140/apde.2012.5.261 (cit. on p. 103).

[28] M. Christ and S. Shao. "On the extremizers of an adjoint Fourier restriction inequality". In: *Adv. Math.* 230.3 (2012), pp. 957–977. DOI: 10.1016/j.aim.2012.03.020 (cit. on p. 103).

[29] V. Ciccone and F. Gonçalves. "Sharp Fourier extension on the circle under arithmetic constraints". In: *J. Funct. Anal.* 286.2 (2024), Paper No. 110219, 21. DOI: 10.1016/j.jfa.2023.110219 (cit. on p. 102).

[30] L. Cladek and B. Krause. "Improved endpoint bounds for the lacunary spherical maximal operator". In: *Anal. PDE* 17.6 (2024), pp. 2011–2032. DOI: 10.2140/apde.2024.17.2011 (cit. on p. 5).

[31] L. Cladek and Y. Ou. "Sparse domination of Hilbert transforms along curves". In: *Math. Res. Lett.* 25.2 (2018), pp. 415–436. DOI: 10.4310/MRL.2018.v25.n2.a4 (cit. on p. 7).

[32] R. R. Coifman and Y. Meyer. "Commutateurs d'intégrales singulières et opérateurs multilinéaires". In: *Ann. Inst. Fourier (Grenoble)* 28.3 (1978), pp. xi, 177–202. DOI: 10.5802/aif.708 (cit. on p. 62).

[33] R. R. Coifman and Y. Meyer. *Au delà des opérateurs pseudo-différentiels.* Vol. 57. Astérisque. With an English summary. Société Mathématique de France, Paris, 1978, pp. i+185 (cit. on p. 62).

[34] J.-P. Conze and E. Lesigne. "Théorèmes ergodiques pour des mesures diagonales". In: *Bull. Soc. Math. France* 112.2 (1984), pp. 143–175 (cit. on p. 10).

[35] W. W. Crawley-Boevey. "On tame algebras and bocses". In: *Proc. London Math. Soc. (3)* 56.3 (1988), pp. 451–483. DOI: 10.1112/plms/s3-56.3.451 (cit. on p. 58).

[36] F. Dai and Y. Xu. *Approximation theory and harmonic analysis on spheres and balls.* Springer Monographs in Mathematics. Springer, New York, 2013, pp. xviii+440. DOI: 10.1007/978-1-4614-6660-4 (cit. on p. 77).

[37] C. Demeter. "Pointwise convergence of the ergodic bilinear Hilbert transform". In: *Illinois J. Math.* 51.4 (2007), pp. 1123–1158 (cit. on p. 11).

[38] C. Demeter, M. Pramanik, and C. Thiele. "Multilinear singular operators with fractional rank". In: *Pacific J. Math.* 246.2 (2010), pp. 293–324. DOI: 10.2140/pjm.2010.246.293 (cit. on pp. 64, 65).

[39] C. Demeter and C. Thiele. "On the two-dimensional bilinear Hilbert transform". In: *Amer. J. Math.* 132.1 (2010), pp. 201–256. DOI: 10.1353/ajm.0.0101 (cit. on pp. 11, 12, 55, 56, 62–64).

[40] H. Derksen and J. Weyman. "Quiver representations". In: *Notices Amer. Math. Soc.* 52.2 (2005), pp. 200–206 (cit. on p. 57).

[41] F. Di Plinio and A. Fragkos. "The weak-type Carleson theorem via wave packet estimates". In: *Trans. Amer. Math. Soc.* 378.3 (2025), pp. 1551–1592. DOI: 10.1090/tran/9303 (cit. on pp. 3, 22).

[42] Y. Do, R. Oberlin, and E. A. Palsson. "Variation-norm and fluctuation estimates for ergodic bilinear averages". In: *Indiana Univ. Math. J.* 66.1 (2017), pp. 55–99. DOI: 10.1512/iumj.2017.66.5983 (cit. on p. 11).

[43] J. A. Drozd. “Tame and wild matrix problems”. In: *Representation theory, II (Proc. Second Internat. Conf., Carleton Univ., Ottawa, Ont., 1979).* Vol. 832. Lecture Notes in Math. Springer, Berlin, 1980, pp. 242–258 (cit. on p. 58).

[44] J. Duoandikoetxea and J. L. Rubio de Francia. “Maximal and singular integral operators via Fourier transform estimates”. In: *Invent. Math.* 84.3 (1986), pp. 541–561. DOI: 10.1007/BF01388746 (cit. on p. 25).

[45] P. Durcik, L. Slavíková, and C. Thiele. “Local bounds for singular Brascamp-Lieb forms with cubical structure”. In: *Math. Z.* 302.4 (2022), pp. 2375–2405. DOI: 10.1007/s00209-022-03148-8 (cit. on pp. 63, 65, 89, 90, 92, 94, 96).

[46] P. Durcik, L. Slavíková, and C. Thiele. *Norm-variation of triple ergodic averages for commuting transformations.* 2023. arXiv: 2307.07372 [math.CA] (cit. on p. 65).

[47] P. Durcik and C. Thiele. “Singular Brascamp-Lieb inequalities with cubical structure”. In: *Bull. Lond. Math. Soc.* 52.2 (2020), pp. 283–298. DOI: 10.1112/blms.12310 (cit. on p. 9).

[48] P. Durcik and C. Thiele. “Singular Brascamp-Lieb: a survey”. In: *Geometric aspects of harmonic analysis.* Vol. 45. Springer INdAM Ser. Springer, Cham, 2021, pp. 321–349 (cit. on pp. 9, 60, 66).

[49] L. Fanelli, L. Vega, and N. Visciglia. “On the existence of maximizers for a family of restriction theorems”. In: *Bull. Lond. Math. Soc.* 43.4 (2011), pp. 811–817. DOI: 10.1112/blms/bdr014 (cit. on p. 103).

[50] C. Fefferman. *Inequalities for strongly singular convolution operators.* Thesis (Ph.D.)–Princeton University. ProQuest LLC, Ann Arbor, MI, 1969 (cit. on p. 16).

[51] C. Fefferman. “Pointwise convergence of Fourier series”. In: *Ann. of Math. (2)* 98 (1973), pp. 551–571. DOI: 10.2307/1970917 (cit. on pp. 3, 6, 8, 22, 24, 25, 27).

[52] C. Fefferman. “The multiplier problem for the ball”. In: *Ann. of Math. (2)* 94 (1971), pp. 330–336. DOI: 10.2307/1970864 (cit. on p. 17).

[53] D. Foschi and D. Oliveira e Silva. “Some recent progress on sharp Fourier restriction theory”. In: *Anal. Math.* 43.2 (2017), pp. 241–265. DOI: 10.1007/s10476-017-0306-2 (cit. on p. 103).

[54] D. Foschi. “Global maximizers for the sphere adjoint Fourier restriction inequality”. In: *J. Funct. Anal.* 268.3 (2015), pp. 690–702. DOI: 10.1016/j.jfa.2014.10.015 (cit. on pp. 18, 101, 103).

[55] D. Foschi. “Maximizers for the Strichartz inequality”. In: *J. Eur. Math. Soc. (JEMS)* 9.4 (2007), pp. 739–774. DOI: 10.4171/JEMS/95 (cit. on p. 18).

[56] M. Fraccaroli, O. Saari, and C. Thiele. *Uniform bounds for bilinear symbols with linear K-quasiconformally embedded singularity.* 2024. arXiv: 2402.11661 [math.CA] (cit. on pp. 14, 62, 81).

[57] R. L. Frank, E. H. Lieb, and J. Sabin. “Maximizers for the Stein-Tomas inequality”. In: *Geom. Funct. Anal.* 26.4 (2016), pp. 1095–1134. DOI: 10.1007/s00039-016-0380-9 (cit. on p. 103).

[58] I. M. Gelfand and V. A. Ponomarev. "Problems of linear algebra and classification of quadruples of subspaces in a finite-dimensional vector space". In: *Hilbert space operators and operator algebras (Proc. Internat. Conf., Tihany, 1970)*. Vol. 5. Colloq. Math. Soc. János Bolyai. North-Holland, Amsterdam-London, 1972, pp. 163–237 (cit. on pp. 12, 13, 58).

[59] J. E. Gilbert and A. R. Nahmod. "Boundedness of bilinear operators with nonsmooth symbols". In: *Math. Res. Lett.* 7.5-6 (2000), pp. 767–778. DOI: `10.4310/MRL.2000.v7.n6.a9` (cit. on p. 62).

[60] F. Gonçalves and G. Negro. "Local maximizers of adjoint Fourier restriction estimates for the cone, paraboloid and sphere". In: *Anal. PDE* 15.4 (2022), pp. 1097–1130. DOI: `10.2140/apde.2022.15.1097` (cit. on p. 103).

[61] L. Grafakos. *Classical Fourier analysis*. Third. Vol. 249. Graduate Texts in Mathematics. Springer, New York, 2014, pp. xviii+638. DOI: `10.1007/978-1-4939-1194-3` (cit. on p. 88).

[62] L. Grafakos and X. Li. "Uniform bounds for the bilinear Hilbert transforms. I". In: *Ann. of Math. (2)* 159.3 (2004), pp. 889–933. DOI: `10.4007/annals.2004.159.889` (cit. on p. 10).

[63] L. Grafakos and Z. Si. "The Hörmander multiplier theorem for multilinear operators". In: *J. Reine Angew. Math.* 668 (2012), pp. 133–147. DOI: `10.1515/crelle.2011.137` (cit. on p. 65).

[64] S. Guo, L. B. Pierce, J. Roos, and P.-L. Yung. "Polynomial Carleson operators along monomial curves in the plane". In: *J. Geom. Anal.* 27.4 (2017), pp. 2977–3012. DOI: `10.1007/s12220-017-9790-7` (cit. on pp. 23, 24).

[65] L. Hörmander. *The analysis of linear partial differential operators. I.* Classics in Mathematics. Springer-Verlag, Berlin, 2003. DOI: `10.1007/978-3-642-61497-2` (cit. on pp. 37–39).

[66] M. Hsu and V. Lie. *On a Carleson-Radon Transform (the non-resonant setting)*. 2024. arXiv: `2411.01660 [math.CA]` (cit. on p. 5).

[67] D. Hundertmark and V. Zharnitsky. "On sharp Strichartz inequalities in low dimensions". In: *Int. Math. Res. Not.* (2006), Art. ID 34080, 18. DOI: `10.1155/IMRN/2006/34080` (cit. on p. 18).

[68] R. A. Hunt. "On the convergence of Fourier series". In: *Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967)*. Southern Illinois Univ. Press, Carbondale, Ill., 1968, pp. 235–255 (cit. on pp. 3, 22).

[69] P. Jones. "Bilinear singular integrals and maximal functions". In: *Linear and Complex Analysis Problem Book 3, Part I.* Springer-Verlag, New York, 1994 (cit. on p. 55).

[70] Y. Katznelson. "Sur les ensembles de divergence des séries trigonométriques". In: *Studia Math.* 26 (1966), pp. 301–304. DOI: `10.4064/sm-26-3-305-306` (cit. on p. 2).

[71] V. Kovač. "Boundedness of the twisted paraproduct". In: *Rev. Mat. Iberoam.* 28.4 (2012), pp. 1143–1164. DOI: `10.4171/RMI/707` (cit. on pp. 12, 55, 56, 63, 64, 96).

[72] V. Kovač, C. Thiele, and P. Zorin-Kranich. "Dyadic triangular Hilbert transform of two general functions and one not too general function". In: *Forum Math. Sigma* 3 (2015), Paper No. e25, 27. DOI: `10.1017/fms.2015.25` (cit. on pp. 9, 64).

[73] L. Kronecker. *Algebraische Reduction der Schaaren bilinearer Formen.* Sitzungsberichte der Preussischen Akademie der Wissenschaften. 1890 (cit. on p. 13).

[74] M. Lacey. "Sparse bounds for spherical maximal functions". In: *J. Anal. Math.* 139.2 (2019), pp. 613–635. DOI: `10.1007/s11854-019-0070-2` (cit. on pp. 7, 25).

[75] M. Lacey. "The bilinear maximal functions map into L^p for $2/3 < p \leq 1$". In: *Ann. of Math. (2)* 151.1 (2000), pp. 35–57. DOI: `10.2307/121111` (cit. on p. 11).

[76] M. Lacey and C. Thiele. "L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$". In: *Ann. of Math. (2)* 146.3 (1997), pp. 693–724. DOI: `10.2307/2952458` (cit. on pp. 10, 55, 62).

[77] M. Lacey and C. Thiele. "A proof of boundedness of the Carleson operator". In: *Math. Res. Lett.* 7.4 (2000), pp. 361–370. DOI: `10.4310/MRL.2000.v7.n4.a1` (cit. on pp. 3, 22).

[78] M. Lacey and C. Thiele. "On Calderón's conjecture". In: *Ann. of Math. (2)* 149.2 (1999), pp. 475–496. DOI: `10.2307/120971` (cit. on pp. 10, 55, 62).

[79] J. Lee et al. "The Hörmander multiplier theorem for n-linear operators". In: *Math. Ann.* 381.1-2 (2021), pp. 499–555. DOI: `10.1007/s00208-021-02162-1` (cit. on p. 65).

[80] A. K. Lerner. "A simple proof of the A_2 conjecture". In: *Int. Math. Res. Not. IMRN* 14 (2013), pp. 3159–3170. DOI: `10.1093/imrn/rns145` (cit. on p. 7).

[81] A. K. Lerner. "On an estimate of Calderón-Zygmund operators by dyadic positive operators". In: *J. Anal. Math.* 121 (2013), pp. 141–161. DOI: `10.1007/s11854-013-0030-1` (cit. on p. 7).

[82] V. Lie. "On the boundedness of the Carleson operator near L^1". In: *Rev. Mat. Iberoam.* 29.4 (2013), pp. 1239–1262. DOI: `10.4171/RMI/755` (cit. on pp. 3, 22).

[83] V. Lie. "The (weak-L^2) boundedness of the quadratic Carleson operator". In: *Geom. Funct. Anal.* 19.2 (2009), pp. 457–497. DOI: `10.1007/s00039-009-0010-x` (cit. on pp. 4, 22).

[84] V. Lie. "The polynomial Carleson operator". In: *Ann. of Math. (2)* 192.1 (2020), pp. 47–163. DOI: `10.4007/annals.2020.192.1.2` (cit. on pp. 4, 22).

[85] W. Littman. "$L^p - L^q$-estimates for singular integral operators arising from hyperbolic equations". In: *Partial differential equations.* Vol. XXIII. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, 1973, pp. 479–481 (cit. on p. 39).

[86] N. N. Luzin. "The integral and trigonometric series". PhD thesis. 1915 (cit. on p. 2).

[87] G. Medina and A. Zavadskij. "The four subspace problem: an elementary solution". In: *Linear Algebra Appl.* 392 (2004), pp. 11–23. DOI: `10.1016/j.laa.2004.05.010` (cit. on pp. 58, 97–99).

[88] C. Muscalu. "Calderón commutators and the Cauchy integral on Lipschitz curves revisited: I. First commutator and generalizations". In: *Rev. Mat. Iberoam.* 30.2 (2014), pp. 727–750. DOI: `10.4171/RMI/798` (cit. on p. 65).

[89] C. Muscalu, J. Pipher, T. Tao, and C. Thiele. "Bi-parameter paraproducts". In: *Acta Math.* 193.2 (2004), pp. 269–296. DOI: 10.1007/BF02392566 (cit. on p. 65).

[90] C. Muscalu, J. Pipher, T. Tao, and C. Thiele. "Multi-parameter paraproducts". In: *Rev. Mat. Iberoam.* 22.3 (2006), pp. 963–976. DOI: 10.4171/RMI/480 (cit. on p. 65).

[91] C. Muscalu and W. Schlag. *Classical and multilinear harmonic analysis. Vol. II.* Vol. 138. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2013, pp. xvi+324 (cit. on p. 9).

[92] C. Muscalu, T. Tao, and C. Thiele. "Multi-linear operators given by singular multipliers". In: *J. Amer. Math. Soc.* 15.2 (2002), pp. 469–496. DOI: 10.1090/S0894-0347-01-00379-4 (cit. on pp. 62, 65).

[93] C. Muscalu and Y. Zhai. "Five-linear singular integral estimates of Brascamp-Lieb-type". In: *Anal. PDE* 15.4 (2022), pp. 1011–1095. DOI: 10.2140/apde.2022.15.1011 (cit. on p. 65).

[94] L. A. Nazarova. "Representations of a tetrad". In: *Izv. Akad. Nauk SSSR Ser. Mat.* 31 (1967), pp. 1361–1378 (cit. on pp. 12, 58).

[95] L. A. Nazarova. "Representations of quivers of infinite type". In: *Izv. Akad. Nauk SSSR Ser. Mat.* 37 (1973), pp. 752–791 (cit. on pp. 12, 58, 68).

[96] G. Negro, D. Oliveira e Silva, and C. Thiele. "When does $e^{-|\tau|}$ maximize Fourier extension for a conic section?" In: *Harmonic analysis and convexity.* Vol. 9. Adv. Anal. Geom. De Gruyter, Berlin, 2023, pp. 391–426 (cit. on pp. 18, 102, 103).

[97] R. Oberlin. "Sparse bounds for a prototypical singular Radon transform". In: *Canad. Math. Bull.* 62.2 (2019), pp. 405–415. DOI: 10.4153/cmb-2018-007-5 (cit. on pp. 7, 25, 41, 42).

[98] D. Oliveira e Silva and R. Quilodrán. "Global maximizers for adjoint Fourier restriction inequalities on low dimensional spheres". In: *J. Funct. Anal.* 280.7 (2021), Paper No. 108825, 73. DOI: 10.1016/j.jfa.2020.108825 (cit. on p. 103).

[99] D. Oliveira e Silva, C. Thiele, and P. Zorin-Kranich. "Band-limited maximizers for a Fourier extension inequality on the circle". In: *Exp. Math.* 31.1 (2022), pp. 192–198. DOI: 10.1080/10586458.2019.1596847 (cit. on pp. 19, 101).

[100] K. Pearson. *A mathematical theory of random migration.* Draper's Company Research Memoirs, Biometric Series III. Dulau and Co., London, 1906, p. 68 (cit. on p. 112).

[101] L. B. Pierce and P.-L. Yung. "A polynomial Carleson operator along the paraboloid". In: *Rev. Mat. Iberoam.* 35.2 (2019), pp. 339–422. DOI: 10.4171/rmi/1057 (cit. on pp. 5, 21, 22).

[102] J. P. G. Ramos. "The Hilbert transform along the parabola, the polynomial Carleson theorem and oscillatory singular integrals". In: *Math. Ann.* 379.1-2 (2021), pp. 159–185. DOI: 10.1007/s00208-020-02075-5 (cit. on pp. 5, 23).

[103] J. Roos. "Bounds for anisotropic Carleson operators". In: *J. Fourier Anal. Appl.* 25.5 (2019), pp. 2324–2355. DOI: 10.1007/s00041-018-09657-7 (cit. on pp. 5, 23).

[104] J. Roos. “Singular integrals and maximal operators related to Carleson’s theorem and curves in the plane”. PhD thesis. Rheinische Friedrich-Wilhelms-Universität Bonn, Aug. 2017 (cit. on pp. 5, 62, 81).

[105] S. Shao. “On existence of extremizers for the Tomas-Stein inequality for S^1”. In: *J. Funct. Anal.* 270.10 (2016), pp. 3996–4038. DOI: 10.1016/j.jfa.2016.02.019 (cit. on pp. 18, 103).

[106] S. Shao. *On smoothness of extremizers of the Tomas-Stein inequality for S^1*. 2018. arXiv: 1601.07119 [math.AP] (cit. on pp. 18, 103).

[107] P. Sjölin. “Convergence almost everywhere of certain singular integrals and multiple Fourier series”. In: *Ark. Mat.* 9 (1971), pp. 65–90. DOI: 10.1007/BF02383638 (cit. on pp. 4, 22, 24).

[108] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Vol. 43. Princeton Mathematical Series. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993, pp. xiv+695 (cit. on p. 17).

[109] E. M. Stein and S. Wainger. “Oscillatory integrals related to Carleson’s theorem”. In: *Math. Res. Lett.* 8.5-6 (2001), pp. 789–800. DOI: 10.4310/MRL.2001.v8.n6.a9 (cit. on pp. 22, 23).

[110] E. M. Stein and S. Wainger. “Problems in harmonic analysis related to curvature”. In: *Bull. Amer. Math. Soc.* 84.6 (1978), pp. 1239–1295. DOI: 10.1090/S0002-9904-1978-14554-6 (cit. on pp. 5, 8, 24, 25, 40).

[111] E. M. Stein and G. Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Vol. No. 32. Princeton Mathematical Series. Princeton University Press, Princeton, NJ, 1971, pp. x+297 (cit. on pp. 76, 79).

[112] C. Thiele. “A uniform estimate”. In: *Ann. of Math. (2)* 156.2 (2002), pp. 519–563. DOI: 10.2307/3597197 (cit. on p. 10).

[113] P. A. Tomas. “Restriction theorems for the Fourier transform”. In: *Harmonic analysis in Euclidean spaces*. Vol. XXXV, Part 1. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, 1979, pp. 111–114 (cit. on p. 16).

[114] N. Tomita. “A Hörmander type multiplier theorem for multilinear operators”. In: *J. Funct. Anal.* 259.8 (2010), pp. 2028–2044. DOI: 10.1016/j.jfa.2010.06.010 (cit. on p. 65).

[115] G. Uraltsev and M. Warchalski. *The full range of uniform bounds for the bilinear Hilbert transform*. 2022. arXiv: 2205.09851 [math.CA] (cit. on p. 62).

[116] P. Zorin-Kranich. “Maximal polynomial modulations of singular integrals”. In: *Adv. Math.* 386 (2021), Paper No. 107832, 40. DOI: 10.1016/j.aim.2021.107832 (cit. on pp. 4, 22, 24, 25, 52, 54).