Fourier Restriction and Decoupling

Griffin Pinney

October 2021

A thesis submitted in partial fulfilment of the requirements
for the degree of Bachelor of Philosophy (Honours) in Mathematics
of the Australian National University

For all those who have fostered in me a love of mathematics, problem solving, and learning, from childhood to now.

Declaration

The work in this thesis is my own except where otherwise stated.

Griffin Pinney

Acknowledgements

First and foremost, thank you Po Lam for all your patience and dedication throughout the year. Every time I visited your office (or your Zoom room), I left as a far better mathematician than when I entered. I cannot overstate how much I appreciate all your guidance.

Thank you Wenqi for reading my drafts; your keen eye prevented many mistakes from slipping through the cracks.

Thank you to my fellow Honours students for making the environment in the Hanna Neumann building so enjoyable this year; I'll forever cherish our fun times and pranks.

Last but not least, thank you to my family for supporting me in everything I do.

Abstract

The Fourier restriction problem has a rich history, and with broad applications including geometric measure theory, combinatorics, number theory, and PDE, Fourier restriction has evolved into one of the most active areas of research in modern harmonic analysis. On the other hand, Fourier decoupling is a new and powerful tool which has led to recent breakthroughs in number theory and PDE. We explore both Fourier restriction and decoupling, our analysis culminating in a new proof of a classic restriction theorem based on decoupling techniques. In addition to providing a new perspective on a classic theorem, our decoupling method highlights insightful connections between the fields of Fourier restriction and decoupling.

Contents

Notation and Conventions

We will strive to explain all notation as it is introduced; however, some notation and conventions are of such importance that we wish to implement them immediately and without further comment. We record them here for reference:
If $(X, \mathcal{X}, \mu)$ is a measure space, we say that a property holds *μ-a.e*, or *for μ-a.e.* $x \in X$, if it holds for all $x \in X$ except possibly on a set of μ-measure 0.
If $1 \leq p < \infty$ and $f : X \to \mathbb{C}$ is μ-measurable, we define

$$\|f\|_{L^p(X,\mu)} = \Big(\int_X |f(x)|^p \, d\mu(x) \Big)^{1/p};$$

for $p = \infty$, we define

$$\|f\|_{L^\infty(X,\mu)} = \operatorname*{ess\,sup}_{x \in X} |f(x)| = \sup\{c \in \mathbb{R} : |f(x)| \leq c \text{ for } \mu\text{-a.e. } x \in X\}.$$

Given $1 \leq p \leq \infty$, we define $L^p(X, \mu)$ to be the set of μ-measurable functions $f : X \to \mathbb{C}$ such that $\|f\|_{L^p(X,\mu)} < \infty$, modulo μ-a.e. equivalence.
The measure μ will often be clear from context, in which case we tidy our notation by writing $\|f\|_{L^p(X)}$ in place of $\|f\|_{L^p(X,\mu)}$ and $L^p(X)$ in place of $L^p(X, \mu)$. A particularly common case is when the measure space is $\mathbb{R}^n$, which we always assume to be equipped with the Lebesgue σ-algebra and the Lebesgue measure. Consequently, on $\mathbb{R}^n$, the terms *a.e.* and *for a.e.* $x \in \mathbb{R}^n$ are always assumed to be with respect to the Lebesgue measure.
If $A \subset \mathbb{R}^n$ is measurable, we denote its Lebesgue measure by $|A|$, not to be confused with the absolute value of a real or complex number; the difference should always be clear from context. A common case is the ball of radius $R > 0$ centred at $x \in \mathbb{R}^n$, which we denote by $B(x, R)$.
Finally, if $A \subset X$ is measurable, we denote by χ_A the *characteristic function of* A, defined by

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

Chapter 1

Introduction

First considered by Stein in 1967, the problem of *Fourier restriction* remains a highly active area of contemporary research in harmonic analysis. With connections to numerous fields as disparate as combinatorics and PDE, insight into the restriction problem is valuable not only to harmonic analysis, but to a wide range of fields of modern research.

In contrast to the long history of Fourier restriction, the study of *Fourier decoupling* began to flourish as recently as 2015, leading to exciting breakthroughs in number theory and new results in PDE. Our goal is to give a brief overview of both Fourier restriction and decoupling, before demonstrating how decoupling techniques may be used to give a new proof of a classic restriction result which we will refer to as the *Tomas restriction theorem*.

We begin in Chapter 2 by reviewing the basic tools required to engage with Fourier restriction and decoupling. Of greatest importance is the theory of Fourier analysis on $\mathbb{R}^n$, and we review the standard results following a series of lecture notes by Wolff [Wol03]. We then give a brief overview of interpolation – one of the most useful tools in harmonic analysis, which we will have frequent need to employ. Our treatment of this subject is drawn from a text of Stein and Weiss [SW71] and an exposition of Tao [Tao09]. We conclude this chapter by recalling some useful analytic inequalities, as well as some miscellaneous results mostly regarding the surface measure $d\sigma$ (Definition 2.42) associated to the truncated paraboloid in $\mathbb{R}^n$ (Definition 2.40).

We dedicate Chapter 3 to an overview of Fourier restriction, beginning by summarising the initial observations of Stein which mark the origins of the restriction problem. We then introduce the equivalent dual "extension" formulation of the restriction problem (which will become our main focus), before using the dual formulation to motivate and state the *restriction conjecture* for the truncated paraboloid in $\mathbb{R}^n$ (Conjecture 3.13). Significant early progress on the restriction conjecture was made by Tomas [Tom75] with the first proof the Tomas restriction theorem. Since our main result in Chapter 5 is a new proof of the Tomas restriction theorem based on decoupling methods, we provide a careful analysis of the original methods of Tomas, drawing inspiration from an elaboration of Tao [Tao20b]. Finally, we introduce and explore standard "local" variants of restriction and extension estimates (Definitions 3.15 and 3.16) which will be essential for the application of decoupling to the Tomas restriction theorem.

In Chapter 4, we give a similar overview of Fourier decoupling. We begin with a brief summary of the considerations of Wolff [Wol00] which first led to the study of decoupling inequalities, before defining the notion of a decoupling inequality itself (Equation (4.4)). We give some simple well-known examples, following which we state the more sophisticated results of *decoupling for the paraboloid* and *decoupling for the moment curve* (Theorems 4.5 and 4.8), due to Bourgain-Demeter [BD15] and Bourgain-Demeter-Guth [BDG16] respectively. Decoupling for the moment curve was famously used by Bourgain-Demeter-Guth to prove a long-standing conjecture in number theory known as the *Vinogradov main conjecture* (or sometimes, the *main conjecture of Vinogradov's mean value theorem*); we give a short demonstration of how the main conjecture follows from decoupling for the moment curve, adapting an exposition of Tao [Tao15] to make a more suitable statement of the decoupling theorem applicable. We conclude this chapter by recounting a proof of the model $n = 2$ case of decoupling for the paraboloid (Theorem 4.10), following an article of Li [Li21] and an elaboration of Tao [Tao20a], making some small optimisations where possible.

In Chapter 5, we unite Fourier restriction and decoupling by giving a new proof of the Tomas restriction theorem based on the decoupling theorem for the paraboloid. We find that a local variant of the Tomas restriction theorem (Theorem 5.2) follows without too much difficulty from the decoupling theorem; the main difficulty lies in upgrading the local extension estimates thus obtained to a suitable family of global extension estimates. Our approach to this undertaking, known as *ε-removal*, is based on that taken by Tao [Tao99] to prove a related result. Finally, we conclude with some speculation on how similar decoupling methods could be used to prove further restriction estimates.

Chapter 2

Background

We will find it necessary in subsequent chapters to make frequent use of certain analytic tools which, whilst not being our main focus, are essential for us to understand. It is assumed that the reader is already familiar with most of these tools, but we take this opportunity to recall the most important of them nonetheless. Our emphasis will be on stating results rather than proving them (with some exceptions), and on summarising the relevant theory.

2.1 Fourier Analysis on $\mathbb{R}^n$

We review here the basic theory of Fourier analysis on $\mathbb{R}^n$. Unless indicated otherwise, the results and definitions stated herein are drawn from the first several chapters of [Wol03]. There are an abundance of texts on this topic, with [SS03] and [Gra04] being just two examples from a myriad of good alternatives.

2.1.1 The L^1 Fourier Transform

Definition 2.1 (The Fourier Transform)**.** Given $f \in L^1(\mathbb{R}^n)$, we define its *Fourier transform* $\hat{f}$ by the pointwise formula

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \cdot \xi} \, dx. \tag{2.1}$$

More generally, if μ is a complex measure on $\mathbb{R}^n$, we define its Fourier transform $\hat{\mu}$ by

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x \cdot \xi} \, d\mu(x). \tag{2.2}$$

If $f \in L^1(\mathbb{R}^n)$ and μ is a measure, we refer to the supports of $\hat{f}$ and $\hat{\mu}$ as the *Fourier supports* of f and μ respectively.

Equation (2.2) indeed generalises (2.1) since for any $f \in L^1(\mathbb{R}^n)$, we may identify f with the complex measure fdx.
If we are given a function g which is the Fourier transform of a pre-existing function and we wish to emphasise this, we informally say that g takes inputs from the *frequency domain*, and we use Greek letters such as ξ and ω as the variables. Conversely, if we are given a function f which we do not wish to think of as the Fourier transform of a pre-existing function, we informally say that f takes inputs from the *spatial domain*, and we use Latin letters such as x and y as the variables.
We have the following basic formulas describing how the Fourier transform of an L^1 function is affected by various operations. Each of these formulas may be proved by recalling the definitions and applying a simple change of variables.

Proposition 2.2 (Basic Fourier Transform Formulas). *Let $f \in L^1(\mathbb{R}^n)$, $\tau \in \mathbb{R}^n$, and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear map. Then,*

1. *Let $f_\tau(x) = f(x - \tau)$. We have $\widehat{f_\tau}(\xi) = e^{-2\pi i\tau\cdot\xi}\hat{f}(\xi)$.*
2. *Let $e_\tau(x) = e^{2\pi i\tau\cdot x}$. We have $\widehat{e_\tau f}(\xi) = \hat{f}(\xi - \tau)$.*
3. *We have $\widehat{f \circ T} = |\det T|^{-1}(\hat{f} \circ T^{-t})$.*
4. *Let $\tilde{f}(x) = \overline{f(-x)}$. Then, $\widehat{\tilde{f}} = \overline{\hat{f}}$.*

Loosely, we refer to the interplay between formulas one and two as *translation invariance*. A particularly common case of the third formula is when T is scalar multiplication by some factor $r \in \mathbb{R}$. In particular, letting $f_r(x) = f(rx)$, we get $\widehat{f_r}(\xi) = |r|^{-n}\hat{f}_{r^{-1}}$. We will tend to avoid the case $n = 1$, so there will be no risk of confusing the notations f_τ and f_r as introduced above.

Notation 2.3 (Multiindex Notation). Fix $n \geq 2$. Given $\alpha = (\alpha_1, \cdots, \alpha_n)$, where each α_i is a nonnegative integer, we say α is a *multiindex*, and define its *order* to be $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Given $f : \mathbb{R}^n \to \mathbb{C}$, we define

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}},$$

and for any $x \in \mathbb{R}^n$, we define

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

The next two propositions can be summarised by the general principle that the faster f decays, the smoother $\hat{f}$ should be, and conversely the smoother f is, the faster $\hat{f}$ should decay.

Proposition 2.4. *If* $\int_{\mathbb{R}^n} |x|^N |f(x)|\,dx < \infty$ *for all* $N \geq 0$*, then* $\hat{f} \in C^\infty(\mathbb{R}^n)$*, with*

$$D^\alpha \hat{f}(\xi) = ((-2\pi i\xi)^\alpha f)\widehat{\ }.$$

Here and throughout, we use the notation $(\cdot)\widehat{\ }$ to denote the Fourier transform of the expression inside the parentheses.

Proposition 2.5. *If* $D^\alpha f$ *exists and is in* $L^1(\mathbb{R}^n)$ *for all* $0 \leq |\alpha| \leq N$*, then*

$$\widehat{D^\alpha f}(\xi) = (2\pi i\xi)^\alpha \hat{f}(\xi) \qquad \forall 0 \leq |\alpha| \leq N,$$

and there is a constant $C > 0$ *depending on* N *and* f *such that*

$$|\hat{f}(\xi)| \leq C(1 + |\xi|)^{-N}. \tag{2.3}$$

If $f \in L^1(\mathbb{R}^n)$, then heuristically, $\hat{f}(\xi)$ measures how much the frequency $e^{2\pi i x\cdot\xi}$ contributes to f. As such, we would like to have a *Fourier inversion formula* of the form $f(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi i x\cdot\xi}\,d\xi$ for a.e. $x \in \mathbb{R}^n$. However, a problem lies in the fact that the Fourier transform of an L^1 function is not necessarily also in L^1. To remedy this, it is convenient to introduce a certain class of functions which behaves well under the Fourier transform.

Definition 2.6 (Schwartz Space)**.** The *Schwartz space*, denoted $\mathcal{S}(\mathbb{R}^n)$, is the subspace of $C^\infty(\mathbb{R}^n)$ consisting of all functions f for which $x^\alpha D^\beta f$ is bounded for all multiindices α and β.

Thus, $f \in \mathcal{S}(\mathbb{R}^n)$ if and only if f has derivatives of rapid decrease of all orders. Clearly, $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, and a typical example showing this containment is strict is the Gaussian $\Gamma(x) = e^{-\pi|x|^2}$. Using the elementary fact that there exists some $C_N > 0$ such that $C_N^{-1}(1 + |x|)^N \leq \sum_{0\leq|\alpha|\leq N} |x^\alpha| \leq C_N(1 + |x|)^N$, we may also obtain a useful alternative characterisation of the Schwartz space as the subspace of $C^\infty(\mathbb{R}^n)$ consisting of all functions f for which $(1 + |x|)^N D^\beta f$ is bounded for all $N \geq 0$ and all multiindices β.
Propositions 2.4 and 2.5 lead to the following:

Theorem 2.7. *If* $f \in \mathcal{S}(\mathbb{R}^n)$*, then* $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$*.*

Thus, the Fourier transform defines an operator $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ upon restriction to $\mathcal{S}(\mathbb{R}^n)$. Notably, the Gaussian $\Gamma \in \mathcal{S}(\mathbb{R}^n)$ as defined above is a fixed point for this operator; that is, $\hat{\Gamma} = \Gamma$.

2.1.2 The Fourier Inversion Formula

Recall that we define the *convolution* of two functions $f, g : \mathbb{R}^n \to \mathbb{C}$ by

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy = \int_{\mathbb{R}^n} f(y)g(x-y)\,dy$$

(the two being equal by a change of variables), and note that $\operatorname{supp}(f*g) \subset \overline{\operatorname{supp} f + \operatorname{supp} g}$. The following result is standard:

Lemma 2.8. *Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \phi\,dx = 1$, and for each $\varepsilon > 0$, define $\phi^\varepsilon(x) = \varepsilon^{-n}\phi(\varepsilon^{-1}x)$. Then,*

1. *If f is a continuous function which limits to zero at infinity, then $\phi^\varepsilon * f \to f$ uniformly as $\varepsilon \to 0$.*

2. *If $f \in L^p(\mathbb{R}^n)$ for $1 \le p < \infty$, then $\phi^\varepsilon * f \to f$ in $L^p(\mathbb{R}^n)$ as $\varepsilon \to \infty$.*

The following *duality relation* for the Fourier transform follows by a direct application of Fubini's theorem:

Lemma 2.9 (Duality Relation)**.** *If $f, g \in L^1(\mathbb{R}^n)$, then*

$$\int_{\mathbb{R}^n} \hat{f}(x)g(x)\,dx = \int_{\mathbb{R}^n} f(x)\hat{g}(x)\,dx.$$

Combining Lemmas 2.8 and 2.9 with the observation that $\hat{\Gamma} = \Gamma$, we may obtain the desired Fourier inversion formula under the additional assumption that $\hat{f} \in L^1$:

Theorem 2.10 (Fourier Inversion)**.** *If $f \in L^1(\mathbb{R}^n)$ and $\hat{f} \in L^1(\mathbb{R}^n)$, then for a.e. $x \in \mathbb{R}^n$,*

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\,d\xi.$$

Proof. For each $\varepsilon > 0$, define a function I_ε by

$$I_\varepsilon(x) = \int_{\mathbb{R}^n} \hat{f}(\xi)\Gamma_\varepsilon(\xi)e^{2\pi i x\cdot\xi}\,d\xi.$$

Since $\hat{f} \in L^1(\mathbb{R}^n)$, the Lebesgue dominated convergence theorem applies, giving

$$\lim_{\varepsilon\to 0} I_\varepsilon(x) = \int_{\mathbb{R}^n} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\,d\xi \tag{2.4}$$

for all $x \in \mathbb{R}^n$. On the other hand, Lemma 2.9 and the properties of the Fourier transform from Proposition 2.2 give

$$I_\varepsilon(x) = \int_{\mathbb{R}^n} f(\xi)\widehat{\Gamma_\varepsilon}(\xi - x)\,d\xi = \int_{\mathbb{R}^n} f(\xi)\Gamma^\varepsilon(x-\xi)\,d\xi$$
$$= \Gamma^\varepsilon * f(x),$$

where we have used the property $\hat{\Gamma} = \Gamma$ and the observation that Γ is even. Lemma 2.8 now gives $I_\varepsilon \to f$ in $L^1(\mathbb{R}^n)$ as $\varepsilon \to 0$, which together with equation (2.4) implies the result. □

In light of the Fourier inversion formula, we make the following definition:

Definition 2.11 (The Inverse Fourier Transform)**.** Given $f \in L^1(\mathbb{R}^n)$, we define its *inverse Fourier transform* $\check{f}$ by the pointwise formula

$$\check{f}(x) = \int_{\mathbb{R}^n} f(\xi)e^{2\pi i x\cdot\xi}\,d\xi.$$

More generally, if μ is a complex measure on $\mathbb{R}^n$, we define its inverse Fourier transform $\check{\mu}$ by

$$\check{\mu}(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\,d\mu(\xi).$$

We note that for $f \in L^1(\mathbb{R}^n)$, $\check{f}(x) = \hat{f}(-x)$, and similarly $\hat{f}(\xi) = \check{f}(-\xi)$. That is, $\check{f} = (\hat{f})_{-1}$ and $\hat{f} = (\check{f})_{-1}$. Using these identities, it is clear that any statement for the Fourier transform has a simple analogue for the inverse Fourier transform. We will tend not to distinguish such analogues from their original counterparts, and will simply cite the relevant result for the Fourier transform when we wish to invoke them; a particularly common case of this will be analogues for the inverse Fourier transform of the formulas given in Proposition 2.4. As is the case for the Fourier transform, we will also use the notation $(\cdot)^\vee$ to denote the inverse Fourier transform of the expression inside the parentheses.
Recalling that the Fourier transform of an L^1 function is not necessarily itself in L^1, the terminology *inverse Fourier transform* is potentially misleading, since it is does not define an inverse of the map $f \mapsto \hat{f}$ on $L^1(\mathbb{R}^n)$. We use this terminology nonetheless, though some care must be taken to avoid confusion. If we restrict to the Schwartz space, however, and recall from Theorem 2.7 that the Fourier transform maps $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$ (which is clearly contained in $L^1(\mathbb{R}^n)$), the Fourier inversion formula implies that the inverse Fourier transform indeed defines an inverse of the map $f \mapsto \hat{f}$ on $\mathcal{S}(\mathbb{R}^n)$. That is, for $f \in \mathcal{S}(\mathbb{R}^n)$, we have

$\check{\hat{f}} = f$ and $\hat{\check{f}} = f$ (the second of these following from the first by the identities $\check{f} = (\hat{f})_{-1}$ and $\hat{f} = (\check{f})_{-1}$). We will tend to use these properties without further comment in subsequent chapters.
The following result regarding the interplay between convolution and the Fourier transform will be frequently useful. To make sense of the proposition, we note that both $L^1(\mathbb{R}^n)$ and $\mathcal{S}(\mathbb{R}^n)$ are closed under taking convolutions.

Proposition 2.12. *We have*

$$\widehat{f * g} = \hat{f}\hat{g} \qquad \forall f, g \in L^1(\mathbb{R}^n),$$

and

$$\widehat{fg} = \hat{f} * \hat{g} \qquad \forall f, g \in \mathcal{S}(\mathbb{R}^n).$$

The first of these formulas follows by a simple application of Fubini's theorem, and the second follows from the first by Fourier inversion. Like Fourier inversion, we will tend to use these properties without further comment in subsequent chapters.
We will find it useful in many instances to produce a function $\psi \in \mathcal{S}(\mathbb{R}^n)$ which is bounded away from zero on some bounded set, and whose Fourier support is particularly small. The existence of such functions is a result that is often used but rarely proved, so we outline in the following proposition a method for their construction. The proof is an exercise in understanding Proposition 2.2, as the construction is based on the simple observation that dilating a function on the spatial domain results in a contraction on the frequency domain (with some scaling), and conversely.

Proposition 2.13. *Let $U \subset \mathbb{R}^n$ be bounded. Then, there exists $\psi \in \mathcal{S}(\mathbb{R}^n)$ such that $|\psi| \geq C > 0$ on U, and $\hat{\psi}$ is supported in $B(0,1)$. Moreover, ψ can be chosen so that both ψ and $\hat{\psi}$ are nonnegative.*

Proof. Let $\varphi \in C_c^\infty(B(0,1))$ be a nonnegative, nonzero bump function supported in $B(0,1)$. Then, $\check{\varphi} \in \mathcal{S}(\mathbb{R}^n)$ is such that $\check{\varphi}(0) = \int_{\mathbb{R}^n} \varphi \, dx > 0$, so by continuity, there exists some $0 < R < 1$ and some $C > 0$ such that $|\check{\varphi}| \geq C$ on $B(0,R)$. Since U is bounded, we may choose some $M \geq 1$ such that $U \subset B(0,M)$, and it follows that $|\check{\varphi}_{R/M}| \geq C$ on U. Since $(\check{\varphi}_{R/M})\widehat{\ } = (M/R)^n \varphi_{M/R}$ is supported in $B(0, R/M) \subset B(0,1)$, we may take $\psi = \check{\varphi}_{R/M}$.
To ensure that ψ and $\hat{\psi}$ are both nonnegative, we in addition ask for φ to be radial and supported in $B(0,1/2)$. Then, using the notation from formula four

of Proposition 2.2 we have $\tilde{\varphi} = \varphi$, from which the formula itself gives $\check{\varphi} = \overline{\check{\varphi}}$. Then, $\varphi * \varphi$ is a nonnegative, nonzero bump function supported in $B(0,1)$, and Proposition 2.12 combined with our observation that $\check{\varphi} = \overline{\check{\varphi}}$ implies that $(\varphi*\varphi)\check{} = \check{\varphi}\overline{\check{\varphi}} = |\check{\varphi}|^2$ is also nonnegative. Repeating our previous argument with $\varphi * \varphi$ in place of φ gives the result. □

2.1.3 Plancherel and Hausdorff-Young

The $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ Fourier transform is particularly well-behaved with respect to the $L^2(\mathbb{R}^n)$ norm, as the following important theorem shows. It states that the Fourier transform is an isometry on $\mathcal{S}(\mathbb{R}^n)$ when equipped with the $L^2(\mathbb{R}^n)$ norm.

Theorem 2.14 (Plancherel). *If $f, g \in \mathcal{S}(\mathbb{R}^n)$, we have*

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}\,d\xi.$$

Proof. By Fourier inversion, we have $f(x) = \check{\hat{f}}(x) = \hat{\hat{f}}(-x)$. A change of variables therefore gives

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n} \hat{\hat{f}}(-x)\overline{g(x)}\,dx = \int_{\mathbb{R}^n} \hat{\hat{f}}(x)\tilde{g}(x)\,dx,$$

from which the duality relation (Lemma 2.9) and formula four of Proposition 2.2 give the result. □

We will most often use Plancherel when $f = g$, in which case the theorem states $\|f\|_{L^2(\mathbb{R}^n)} = \|\hat{f}\|_{L^2(\mathbb{R}^n)}$.
Noting that $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for all $1 \leq p < \infty$ (since this is true of $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$), we obtain as a corollary of Plancherel that the $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ may be uniquely extended to an isometry $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ (in fact, $\mathcal{F}$ is unitary). Indeed, given $f \in L^2(\mathbb{R}^n)$, choose some sequence of Schwartz functions $(f_n)_{n\in\mathbb{N}}$ limiting to f in $L^2(\mathbb{R}^n)$, and note that the sequence $(\hat{f})_{n\in\mathbb{N}}$ is Cauchy in $L^2(\mathbb{R}^n)$ by Plancherel. This sequence therefore has a unique limit in $L^2(\mathbb{R}^n)$, which we define to be $\mathcal{F}(f)$. The fact that this operator is a well-defined isometry follows by a standard argument.
Given $f \in \mathcal{S}(\mathbb{R}^n)$, it is easy to see by the triangle inequality and the definition of the Fourier transform that $\hat{f}$ is bounded, with $\|\hat{f}\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)}$. Since Plancherel gives $\|\hat{f}\|_{L^2(\mathbb{R}^n)} = \|f\|_{L^2(\mathbb{R}^n)}$, we see that we have $\|\hat{f}\|_{L^{p'}(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)}$

in the cases $p = 1$ and $p = 2$ (recalling that for $1 \leq p \leq \infty$, p' is defined to be the unique exponent $1 \leq p' \leq \infty$ for which $1/p + 1/p' = 1$). This can be extended to a family of intermediate estimates for $p \in (1, 2)$:

Theorem 2.15 (Hausdorff-Young)**.** *If* $1 \leq p \leq 2$*, then*

$$\|\hat{f}\|_{L^{p'}(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)} \qquad \forall f \in \mathcal{S}(\mathbb{R}^n). \tag{2.5}$$

The idea of extending $L^{p_1} \to L^{q_1}$ and $L^{p_2} \to L^{q_2}$ bounds for an operator to a family of bounds for intermediate p and q is known as *interpolation*, which we will explore in the following section. We will see a proof of the Hausdorff-Young theorem as a specific example.
As was the case when $p = 2$, we note that equation (2.5) allows us to extend the $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ Fourier transform to a bounded operator $L^p(\mathbb{R}^n) \to L^{p'}(\mathbb{R}^n)$ for any $1 \leq p \leq 2$. This idea of converting estimates of the form (2.5) for Schwartz functions to bounded operators between L^p spaces will contribute to the motivation of the *restriction problem* in the following chapter.

2.2 Interpolation

We dedicate this section to a brief overview of one of the most useful tools in harmonic analysis: interpolation. We fix two σ-finite measure spaces $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ to be used throughout our discussion, noting that for our purposes, the most relevant measure spaces will be be $\mathbb{R}^n$ with the Lebesgue σ-algebra and Lebesgue measure, and discrete sets with the counting measure (and products of such measure spaces). In what follows, the results and definitions are all drawn from either [Tao09] or Chapter 5 of [SW71]. The subject of interpolation is vast, and there are more advanced texts such as [BL76] dedicated to its study.

2.2.1 Strong-Type Bounds and Complex Interpolation

Let T be an operator mapping $L^p(X)$ to measurable functions on Y. Recall that if $1 \leq p, q \leq \infty$ and there exists some constant $C > 0$ such that

$$\|Tf\|_{L^q(Y)} \leq C\|f\|_{L^p(X)} \qquad \forall f \in L^p(X), \tag{2.6}$$

we say that T is *bounded* $L^p(X) \to L^q(Y)$. In this case, we define the $L^p \to L^q$ *operator norm* of T, denoted $\|T\|_{L^p \to L^q}$, to be the least constant C such that (2.6) holds. There is an alternative terminology that is convenient for the purposes of interpolation:

Definition 2.16. If T is bounded $L^p(X) \to L^q(Y)$, we say that T is of *strong-type* (p, q), and refer to estimates of the form (2.6) as *strong-type* $L^p \to L^q$ *bounds.*

Given an operator T, we will often be interested in determining for which pairs of exponents T is of strong-type (p, q). To make this easier to discuss, we make the following definition:

Definition 2.17. The *strong-type diagram* of T is the set of all pairs $(1/p, 1/q) \in [0, 1]^2$ for which T is of strong-type (p, q).

Note that by decomposing a function f as a sum $\chi_{\{|f|\geq 1\}} f + \chi_{\{|f|<1\}} f$, it is clear that we have $L^p(X) \subset L^{p_1}(X) + L^{p_2}(X)$ for all $p_1 \leq p \leq p_2$. It follows that if T is defined on $L^{p_1}(X) + L^{p_2}(X)$, then it is also defined on $L^p(X)$ for all $p_1 \leq p \leq p_2$. For the purposes of interpolation, operators will always be defined on a space of the form $L^{p_1}(X) + L^{p_2}(X)$.
We are now ready to state the first interpolation theorem which will be of use: the Riesz-Thorin interpolation theorem. Perhaps the simplest interpolation theorem, this allows us to convert a pair of strong-type bounds for $L^{p_1} \to L^{q_1}$ and $L^{p_2} \to L^{q_2}$ to a family of strong-type bounds for intermediate p and q, provided the operator T is linear.
Riesz-Thorin interpolation and related theorems are broadly referred to as *complex interpolation* in reference to the methods used to prove them, which rely upon techniques from complex analysis such as the maximum modulus principle.

Theorem 2.18 (Riesz-Thorin Interpolation)**.** *Suppose T is a linear operator on $L^{p_1}(X)+L^{p_2}(X)$ of strong-type (p_1, q_1) and (p_2, q_2) for some $1 \leq p_1, p_2, q_1, q_1 \leq \infty$. Then T is of strong-type (p_θ, q_θ) for all $0 \leq \theta \leq 1$, where p_θ and q_θ are defined by*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}, \quad \frac{1}{q_\theta} = \frac{1-\theta}{q_1} + \frac{\theta}{q_2}.$$

Moreover, we have $\|T\|_{L^{p_\theta} \to L^{q_\theta}} \leq \|T\|_{L^{p_1} \to L^{q_1}}^{1-\theta} \|T\|_{L^{p_2} \to L^{q_2}}^{\theta}$ for all $0 \leq \theta \leq 1$.

Remark 2.19. In the case where one of the exponents is ∞, we interpret $1/\infty = 0$.

We note that as θ varies from 0 to 1, $(1/p_\theta, 1/q_\theta)$ traces a straight line from $(1/p_1, 1/q_1)$ to $(1/p_2, 1/q_2)$. This is the motivation behind defining the strong-type diagram as a set of pairs of the form $(1/p, 1/q)$ rather than pairs of the form (p, q); the Riesz-Thorin interpolation theorem can be summarised by the statement that the strong-type diagram of a linear operator is convex.

Example 2.20. Let T be the linear operator on $L^1(\mathbb{R}^n) + L^2(\mathbb{R}^n)$ obtained by extending the L^1 and L^2 Fourier transforms to $L^1 + L^2$. Our observations from the previous section show that T is of strong-type $(1, \infty)$ with $\|T\|_{L^1 \to L^\infty} \leq 1$, and of strong-type $(2,2)$ with $\|T\|_{L^2 \to L^2} = 1$. Riesz-Thorin interpolation therefore implies that T is of strong-type (p_θ, q_θ) for all $0 \leq \theta \leq 1$, with $\|T\|_{L^{p_\theta} \to L^{q_\theta}} \leq \|T\|_{L^1 \to L^\infty}^{1-\theta} \|T\|_{L^2 \to L^2}^{\theta} = 1$, where

$$\frac{1}{p_\theta} = (1-\theta) + \frac{\theta}{2}, \quad \frac{1}{q_\theta} = \frac{\theta}{2}.$$

Noting that $\frac{1}{p_\theta} + \frac{1}{q_\theta} = 1$, we have $q_\theta = p'_\theta$, so T is of strong-type (p, p') for all $1 \leq p \leq 2$, with $\|T\|_{L^p \to L^{p'}} \leq 1$. This is implies the Hausdorff-Young theorem since $Tf = \hat{f}$ for $f \in \mathcal{S}(\mathbb{R}^n)$.

2.2.2 Weak-Type Bounds and Real Interpolation

We now discuss an alternative form of interpolation involving "weak" counterparts $L^{p,\infty}$ to the "strong" L^p spaces.

Definition 2.21. Let $f : X \to \mathbb{C}$ be measurable. We define the *distribution function of* f by

$$\lambda_f(t) = \mu(\{x \in X : |f(x)| \geq t\}) \qquad \forall t > 0.$$

Given any $t > 0$, we have

$$\|f\|_{L^p(X)}^p = \int_X |f(x)|^p \, d\mu(x) \geq \int_{\{x \in X : |f(x)| \geq t\}} t^p \, d\mu(x) = t^p \lambda_f(t).$$

This short computation gives the well-known *Chebyshev's inequality*:

Theorem 2.22 (Chebyshev's Inequality)**.** *Given* $1 \leq p < \infty$ *and a measurable function* $f : X \to \mathbb{C}$*, we have*

$$\lambda_f(t) \leq \frac{1}{t^p} \|f\|_{L^p(X)}^p \qquad \forall t > 0.$$

This leads to the following definition:

Definition 2.23 (Weak L^p Norms)**.** Given $1 \leq p < \infty$ and a measurable function $f : X \to \mathbb{C}$, we define the *weak* L^p *norm* of f by

$$\|f\|_{L^{p,\infty}(X)} = \sup_{t > 0} t \lambda_f(t)^{1/p}.$$

For $p = \infty$, we simply define $\|f\|_{L^{\infty,\infty}(X)} = \|f\|_{L^\infty(X)}$.

Analogous to the definition of L^p spaces, we also make the following definition:

Definition 2.24 (Weak L^p Spaces)**.** Given $1 \leq p \leq \infty$, we define the *weak* L^p *space* $L^{p,\infty}(X)$ to be the space of all measurable functions $f : X \to \mathbb{C}$ such that $\|f\|_{L^{p,\infty}(X)} < \infty$ (modulo almost-everywhere equivalence).

Remark 2.25. We have $L^{\infty,\infty}(X) = L^{\infty}(X)$.

It is evident that Chebyshev's inequality is equivalent to the statement that $\|f\|_{L^{p,\infty}(X)} \leq \|f\|_{L^p(X)}$; this is the motivation behind the terminologies *weak* L^p *norm* and *weak* L^p *space*, since we have $L^p(X) \subset L^{p,\infty}(X)$. To further contrast the spaces $L^p(X)$ and $L^{p,\infty}(X)$, we sometimes refer to the former as *strong* L^p *spaces*.
Let T be an operator mapping $L^p(X)$ to measurable functions on Y. In analogy to equation (2.6), if $1 \leq p, q \leq \infty$ and there exists some constant $C > 0$ such that

$$\|Tf\|_{L^{q,\infty}(Y)} \leq C\|f\|_{L^p(X)} \qquad \forall f \in L^p(X), \tag{2.7}$$

we say that T is *bounded* $L^p(X) \to L^{q,\infty}(Y)$. As was the case for strong L^p spaces, there is an alternative terminology that is convenient for the purposes of interpolation:

Definition 2.26. If T is bounded $L^p(X) \to L^{q,\infty}(Y)$, we say that T is of *weak-type* (p, q), and refer to estimates of the form (2.7) as *weak-type* $L^p \to L^q$ bounds.

Analogous to the definition of the strong-type diagram, we make the following definition:

Definition 2.27. The *weak-type diagram* of T is the set of all pairs $(1/p, 1/q) \in [0, 1]^2$ for which T is of weak-type (p, q).

Note that due to Chebyshev's inequality, if T is of strong-type (p, q) then it must also be of weak-type (p, q). It follows that the strong-type diagram is contained in the weak-type diagram.

Definition 2.28. We say that the operator T is *sublinear* if for all f, g in the domain of T and all $c \in \mathbb{C}$, we have $|T(f + g)(y)| \leq |Tf(y)| + |Tg(y)|$ and $|T(cf)(y)| \leq |c||Tf(y)|$ for ν-a.e. $y \in Y$.

Remark 2.29. Clearly, linear operators are also sublinear.

We are now ready to state the second interpolation theorem which will be of use: the Marcinkiewicz interpolation theorem. This is similar to Riesz-Thorin interpolation, but has some advantages and disadvantages. The main advantages are that Marcinkiewicz interpolation allows one to convert a pair of weak-type bounds for $L^{p_1} \to L^{q_1}$ and $L^{p_2} \to L^{q_2}$ to a family of *strong-type* bounds for intermediate p and q, provided the operator T is sublinear. The main disadvantages are that Marcinkiewicz interpolation requires further restrictions on the exponents p_i and q_i (namely, we need $q_1 \neq q_2$ and $p_i \leq q_i$ for $i = 1, 2$); moreover, Marcinkiewicz interpolation does not give the nice strong-type bounds $\|T\|_{L^{p_\theta} \to L^{q_\theta}} \leq \|T\|^{1-\theta}_{L^{p_1} \to L^{q_1}} \|T\|^{\theta}_{L^{p_2} \to L^{q_2}}$ as was the case with Riesz-Thorin interpolation.
Marcinkiewicz interpolation and related theorems are broadly referred to as *real interpolation* in reference to the methods used to prove them, which rely on techniques from real analysis.

Theorem 2.30 (Marcinkiewicz Interpolation)**.** *Suppose T is a sublinear operator on $L^{p_1}(X)+L^{p_2}(X)$ of weak-type (p_1, q_1) and (p_2, q_2) for some exponents $1 \leq p_1 \leq q_1 \leq \infty$, $1 \leq p_2 \leq q_2 \leq \infty$, where $q_1 \neq q_2$. Then T is of strong-type (p_θ, q_θ) for all $0 < \theta < 1$, where p_θ and q_θ are defined by*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}, \quad \frac{1}{q_\theta} = \frac{1-\theta}{q_1} + \frac{\theta}{q_2}.$$

The Marcinkiewicz interpolation theorem can be summarised by the statement that the weak-type diagram of a sublinear operator is convex, and moreover, the the interior of the weak-type diagram is contained in the strong-type diagram.

2.3 Some Useful Inequalities

We now briefly review some inequalities that will be of use. As before, we let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-finite measure spaces, and given an exponent $1 \leq p \leq \infty$, we let p' denote the unique exponent $1 \leq p' \leq \infty$ for which $1/p + 1/p' = 1$. The first is one of the most-well known inequalities in analysis, though we include it for the sake of completeness:

Theorem 2.31 (Hölder's Inequality)**.** *Let $1 \leq p \leq \infty$. If $f \in L^p(X)$ and $g \in L^{p'}(X)$, we have*

$$\int_X |fg| \, d\mu \leq \|f\|_{L^p(X)} \|g\|_{L^q(X)}.$$

The next inequality, not quite as well-known as Hölder's inequality, is useful for estimating the L^p norm of a function on X given by integrating a function on $X \times Y$ with respect to the Y variable.

Theorem 2.32 (Minkowski's Integral Inequality). *Let $1 \le p < \infty$, and let $F : X \times Y \to \mathbb{C}$ be measurable. Then,*

$$\Big(\int_X \Big|\int_Y F(x,y)\,d\nu(y)\Big|^p\,d\mu(x)\Big)^{1/p} \le \int_Y \Big(\int_X |F(x,y)|^p\,d\mu(x)\Big)^{1/p}\,d\nu(y).$$

Remark 2.33. Minkowski's integral inequality also holds with standard modifications when $p = \infty$.

The following well-known inequality describes when the convolution of two functions $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$ lies in some other $L^r(\mathbb{R}^n)$:

Theorem 2.34 (Young's Convolution Inequality). *Let $1 \le p, q \le \infty$ satisfy $1 \le 1/p + 1/q \le 2$, and let $1 \le r \le \infty$ be such that $1 + 1/r = 1/p + 1/q$. If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then $f * g \in L^r(\mathbb{R}^n)$ with*

$$\|f * g\|_{L^r(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^q(\mathbb{R}^n)}.$$

Young's convolution inequality can be proved by elementary means, but a far simpler proof is afforded to us by interpolation. Fixing $g \in L^q(\mathbb{R}^n)$, Minkowski's integral inequality gives the bound $\|f*g\|_{L^q(\mathbb{R}^n)} \le \|f\|_{L^1(\mathbb{R}^n)}\|g\|_{L^q(\mathbb{R}^n)}$, and the triangle inequality followed by Hölder's inequality yields the bound $\|f * g\|_{L^\infty(\mathbb{R}^n)} \le \|f\|_{L^{q'}(\mathbb{R}^n)}\|g\|_{L^q(\mathbb{R}^n)}$. Considering the linear operator T defined by $Tf = f * g$, we see that T is of strong-type $(1, q)$ and (q', ∞) with operator norm at most $\|g\|_{L^q(\mathbb{R}^n)}$ in both cases. Applying Riesz-Thorin interpolation yields Young's convolution inequality in full generality.
As a special case of Theorem 2.6 of [O'N63], Young's convolution inequality remains true if each norm is replaced by its weak counterpart.
The final inequality which will be of use is *Schur's test*, which gives a bound on the $L^2 \to L^2$ operator norm of an integral operator based on its integral kernel.

Theorem 2.35 (Schur's Test). *Let T be an integral operator with kernel $K : X \times Y \to \mathbb{C}$; that is,*

$$Tf(x) = \int_Y K(x,y)f(y)\,d\nu(y).$$

Then,

$$\|T\|^2_{L^2\to L^2} \le \sup_{x\in X}\Big(\int_Y |K(x,y)|\,d\nu(y)\Big)\cdot \sup_{y\in Y}\Big(\int_X |K(x,y)|\,d\mu(x)\Big).$$

Remark 2.36. Schur's test is actually more general; we have stated it in the specific form that we will need later.

Finally, we highlight some important notation related to inequalities:

Notation 2.37. If two quantities X and Y satisfy $|X| \leq CY$ for some constant $C > 0$ depending on some finite number of parameters $p_1, \cdots, p_n$, we write $X \lesssim_{p_1,\cdots,p_n} Y$ or equivalently $X = O_{p_1,\cdots,p_n}(Y)$, and call C the *implied constant.* We allow the implied constant to change each line when using this notation.
We understand $Y \gtrsim_{p_1,\cdots,p_n} X$ to mean $X \lesssim_{p_1,\cdots,p_n} Y$, and we write $X \sim_{p_1,\cdots,p_n} Y$ if $Y \lesssim_{p_1,\cdots,p_n} X \lesssim_{p_1,\cdots,p_n} Y$.

If we are treating a certain parameter as fixed during a given argument, we will tend not to record the dependence of implicit constants on that parameter. For example, we will often treat the dimension n as fixed, and thus we do not tend to emphasise the dependence of implicit constants on n. Otherwise, the omission of a non-fixed parameter from the above notations should be understood to mean that the implied constants do not depend on that parameter.

Example 2.38. Equation (2.3) could be more concisely written as

$$|\hat{f}(\xi)| \lesssim_{N,f} (1 + |\xi|)^{-N}.$$

The parameter ξ is not considered fixed, so its omission from the notation $\lesssim_{N,f}$ is understood to mean that the implied constant is independent of ξ. Though the implied constant depends on the function f, if we were treating f as fixed for the purposes of a given argument, we would simply write $|\hat{f}(\xi)| \lesssim_N (1 + |\xi|)^{-N}$.

2.4 Miscellaneous Results

We record here some results that will be of use but which do not fit into any of the previous sections.

Theorem 2.39 (Partitions of Unity). *Let $\mathcal{U} = (U_\alpha)_{\alpha \in A}$ be a collection of open sets in $\mathbb{R}^n$ indexed by some set A. Then, there exists a smooth partition of unity subordinate to $\mathcal{U}$, which is a family $(\eta_\alpha)_{\alpha \in A}$ of smooth functions $\eta_\alpha : \bigcup_{\alpha \in A} U_\alpha \to \mathbb{R}$ such that*

1. $0 \leq \eta_\alpha \leq 1$ *for all* $\alpha \in A$.
2. $\operatorname{supp} \eta_\alpha \subset U_\alpha$ *for all* $\alpha \in A$.

3. *Each* $x \in \bigcup_{\alpha \in A} U_\alpha$ *lies in* $\operatorname{supp} \eta_\alpha$ *for only finitely many* $\alpha \in A$.

4. $\sum_{\alpha \in A} \eta_\alpha(x) = 1$ *for all* $x \in \bigcup_{\alpha \in A} U_\alpha$ *(noting that the sum has finitely many nonzero terms for each* x *in light of the previous property).*

We will have frequent need to consider submanifolds of $\mathbb{R}^n$, which we assume without further qualification to be smooth and compact. In fact, for the sake of concreteness, we will primarily consider the *truncated paraboloid* in $\mathbb{R}^n$. To avoid trivialities, we henceforth assume $n \geq 2$ throughout.

Definition 2.40. The *truncated paraboloid* in $\mathbb{R}^n$ is the hypersurface

$$P^{n-1} := \{(\xi', \xi_n) \in \mathbb{R}^n : \xi_n = |\xi'|^2 \,;\, \xi' \in [-1,1]^{n-1}\}.$$

Remark 2.41. As indicated by the use of the Greek letter ξ in the above definition, we will tend to think of submanifolds as being situated in the frequency domain.

For the sake of brevity, we will simply refer to the truncated paraboloid as defined above as the *paraboloid*, though care should be taken not to confuse this with its non-truncated counterpart, $\{(\xi', \xi_n) \in \mathbb{R}^n : \xi_n = |\xi'|^2 \,;\, \xi' \in \mathbb{R}^{n-1}\}$.
Note that the paraboloid admits a global parametrisation $\phi : [-1,1]^{n-1} \to P^{n-1}$ defined by $\phi(\xi') = (\xi', |\xi'|^2)$. Using this, we may define the surface measure $d\sigma$ on P^{n-1}, motivated by considerations from Riemannian geometry:

Definition 2.42 (Surface Measure for the Paraboloid)**.** Define $h : [-1,1]^{n-1} \to \mathbb{R}$ by $h(\xi') = \sqrt{(\det D\phi^T D\phi)(\xi')}$. Given a function $g : P^{n-1} \to \mathbb{C}$, we define

$$\int_{P^{n-1}} g(\xi)\, d\sigma(\xi) = \int_{[-1,1]^{n-1}} (g \circ \phi)(\xi') h(\xi')\, d\xi' = \int_{[-1,1]^{n-1}} g(\xi', |\xi'|^2) h(\xi')\, d\xi'.$$

Remark 2.43. We have $h \sim 1$ on $[-1,1]^{n-1}$.

A standard differential geometry computation shows that the foregoing definition is independent of the parametrisation ϕ. Leveraging this fact, a similar definition can be made for an arbitrary submanifold (if the submanifold does not admit a global parametrisation, one uses a more general version of a partition of unity to "stitch together" local definitions of $d\sigma$).
The next result, based on the theory of *oscillatory integrals* and *stationary phase*, gives a bound on the decay of the Fourier transform $\widehat{d\sigma}$:

Theorem 2.44. *Let* $n \geq 2$*, and let* $d\sigma$ *be the surface measure on* P^{n-1}*. Then,* $|\widehat{d\sigma}(\xi)| \lesssim (1 + |\xi|)^{-(n-1)/2}$.

Remark 2.45. Theorem 2.44 remains true if $d\sigma$ is replaced by the complex measure $f d\sigma$ for any $f \in C^\infty(S)$; moreover, if ξ is in a direction normal to P^{n-1} at some point, then we have the stronger result $\widehat{d\sigma}(\xi) \sim (1+|\xi|)^{-(n-1)/2}$.

There will be several instances in which we would like to estimate the surface measure of a ball in $\mathbb{R}^n$. The following lemma provides a simple bound:

Lemma 2.46. *For all $\omega \in \mathbb{R}^n$ and all $R > 0$, we have*

$$\sigma(B(\omega, R)) \lesssim R^{n-1}.$$

Proof. We have

$$\sigma(B(\omega, R)) = \int_{P^{n-1}} \chi_{B(\omega,R)}(\xi)\, d\sigma(\xi),$$

so by definition of the surface measure, we have

$$\begin{aligned} \sigma(B(\omega, R)) &= \int_{[-1,1]^{n-1}} \chi_{B(\omega,R)}(\xi', |\xi'|^2) h(\xi')\, d\xi' \\ &\lesssim \int_{[-1,1]^{n-1}} \chi_{B(\omega,R)}(\xi', |\xi'|)\, d\xi', \end{aligned} \tag{2.8}$$

where we have used our observation from Remark 2.43. Clearly, $\chi_{B(\omega,R)}(\xi', |\xi'|)$ is supported in the projection of $B(\omega, R)$ onto $\mathbb{R}^{n-1} \cong \{\xi \in \mathbb{R}^n : \xi_n = 0\}$, which is a ball of radius R in $\mathbb{R}^{n-1}$. This has $(n-1)$-dimensional Lebesgue measure $\sim R^{n-1}$, from which (2.8) gives $\sigma(B(\omega, R)) \lesssim R^{n-1}$. $\square$

Finally, we recall a useful covering lemma:

Lemma 2.47 (Vitali Covering Lemma)**.** *Let $\{B(x_\alpha, R_\alpha) : \alpha \in A\}$ be a collection of balls in $\mathbb{R}^n$ indexed by some set A. If $\sup_{\alpha \in A} R_\alpha < \infty$, then there exists a countable subset $B \subset A$ such that the balls $\{B(x_\beta, R_\beta) : \beta \in B\}$ are pairwise disjoint and satisfy*

$$\bigcup_{\alpha \in A} B(x_\alpha, R_\alpha) \subset \bigcup_{\beta \in B} B(x_\beta, 5R_\beta).$$

Chapter 3

Fourier Restriction and Extension

The study of Fourier restriction was initiated by Stein in the 1960's, and with applications as broad as geometric measure theory, combinatorics, number theory, and PDE, Fourier restriction remains a highly active field of contemporary research. In this chapter, we aim to give a brief overview of some key aspects of the field.

We begin by summarising the observations of Stein which led to the advent of the restriction problem, before introducing the dual "extension" formulation of the problem which will become our main focus. We then state the restriction conjecture for the truncated paraboloid, motivated by certain necessary conditions, following which we recount an early breakthrough of Tomas [Tom75] on the restriction conjecture, a new proof of which will be our main result in Chapter 5. Finally, we introduce and explore a "local" variant of restriction and extension estimates which will play an important role in the aforementioned proof.

3.1 Initial Considerations

When $f \in L^1(\mathbb{R}^n)$, the Fourier transform $\hat{f}$ is defined pointwise by the formula (2.1), and its restriction to an arbitrary subset of $\mathbb{R}^n$ is therefore well-defined. However, as was observed in Section 2.1, when $f \in L^p(\mathbb{R}^n)$ for $p \in (1, 2]$, $\hat{f}$ is an element of $L^{p'}(\mathbb{R}^n)$ and is therefore only defined up to almost-everywhere equivalence. In particular, $\hat{f}$ may take arbitrary values on a set of measure zero, which is to say that its restriction to a set of measure zero is not *a priori* well-defined. Despite this, it was observed by Stein in 1967 that for certain submanifolds $S \subset \mathbb{R}^n$ of appropriate curvature and for certain exponents $p \in (1,2]$, we may meaningfully restrict the Fourier transform of a function $f \in L^p(\mathbb{R}^n)$ to S, up to $d\sigma$-a.e. equivalence. Stein was driven by the observation that it is the Hausdorff-Young estimate $\|\hat{f}\|_{p'} \lesssim \|f\|_p$ which allows us to extend the $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ Fourier transform to a bounded operator $L^p(\mathbb{R}^n) \to L^{p'}(\mathbb{R}^n)$, and therefore allows us to define $\hat{f}$ up to a.e. equivalence on $\mathbb{R}^n$ when $f \in L^p(\mathbb{R}^n)$. Analogously, an estimate of the form

$$\|\hat{f}|_S\|_{L^q(S,\,d\sigma)} \lesssim \|f\|_{L^p(\mathbb{R}^n)} \qquad \forall f \in \mathcal{S}(\mathbb{R}^n) \tag{3.1}$$

would allow us, after an appropriate approximation argument, to realise $\hat{f}|_S$ as an element of $L^q(S, d\sigma)$ for any $f \in L^p(\mathbb{R}^n)$, so that $\hat{f}$ could be uniquely restricted to S up to $d\sigma$-a.e. equivalence. Loosely, we refer to estimates of the form (3.1) as *restriction estimates.*

Remark 3.1. We will tend to write $\|\hat{f}\|_{L^q(S,\,d\sigma)}$ in place of $\|\hat{f}|_S\|_{L^q(S,\,d\sigma)}$, since the restriction to S is implicit in the $L^q(S, d\sigma)$ norm.

Restriction estimates are so important that they are worthy of their own notation:

Definition 3.2. Given a submanifold S and a pair of exponents $1 \leq p, q \leq \infty$, we denote by $R_S(p \to q)$ the statement that equation (3.1) holds.

Broadly speaking, the goal of Fourier restriction is to determine, for a given submanifold S (or family of submanifolds), for which pairs of exponents the restriction estimate $R_S(p \to q)$ holds.
We record in the next few lemmas some standard elementary observations about restriction estimates for later reference.

Lemma 3.3. $R_S(1 \to \infty)$ *holds for any submanifold* S.

Proof. By the triangle inequality and the definition of the Fourier transform, we have $\|\hat{f}\|_{L^\infty(S,\,d\sigma)} \leq \|f\|_{L^1(\mathbb{R}^n)}$ for all $f \in \mathcal{S}(\mathbb{R}^n)$. □

Due to its simplicity, we will sometimes refer to $R_S(1 \to \infty)$ as the *trivial estimate.* As we will later see, the trivial estimate is often useful for the purposes of interpolation; by combining the trivial estimate with some other non-trivial restriction estimate $R_S(p \to q)$, we may obtain a family of "intermediate" restriction estimates for a range of exponents p and q.

Lemma 3.4. *Given a submanifold S and a pair of exponents $1 \leq p, q \leq \infty$, $R_S(p \to q)$ implies $R_S(p \to r)$ for all $1 \leq r < q$.*

Proof. Since S is compact by assumption, the $d\sigma$ measure of S is finite. Hölder's inequality with exponent $q/r > 1$ therefore gives that for any $f \in \mathcal{S}(\mathbb{R}^n)$,

$$\int_S |\hat{f}(\xi)|^r \, d\sigma(\xi) \leq \sigma(S)^{1-r/q} \Big(\int_S |\hat{f}(\xi)|^q \, d\sigma(\xi) \Big)^{r/q}.$$

Raising both sides to the power of $1/r$, we find that

$$\|\hat{f}\|_{L^r(S, d\sigma)} \lesssim \|\hat{f}\|_{L^q(S, d\sigma)} \lesssim \|f\|_{L^p(\mathbb{R}^n)},$$

where the second inequality follows by $R_S(p \to q)$. □

An analogous result is also true for the first exponent:

Lemma 3.5. *Given a submanifold S and a pair of exponents $1 \leq p, q \leq \infty$, $R_S(p \to q)$ implies $R_S(r \to q)$ for all $1 \leq r < p$.*

Proof. Since S is compact by assumption, we may extend the function $\varphi \equiv 1$ on S to a function $\varphi \in C_c^\infty(\mathbb{R}^n)$. Defining $\psi \in \mathcal{S}(\mathbb{R}^n)$ by $\psi = \check{\varphi}$, we have $\hat{\psi} \equiv 1$ on S by Fourier inversion, hence $\hat{f}|_S = (\hat{f}\hat{\psi})|_S = (f * \psi)\widehat{\ }|_S$. It follows that

$$\|\hat{f}\|_{L^q(S, d\sigma)} = \|(f * \psi)\widehat{\ }\|_{L^q(S, d\sigma)} \lesssim \|f * \psi\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^r(\mathbb{R}^n)},$$

where the first inequality follows by $R_S(p \to q)$ and the second follows by Young's convolution inequality. □

We note that in light of Lemmas 3.4 and 3.5, the general goal is to prove restriction estimates for the largest possible exponents

3.2 The Dual Formulation

Given a submanifold $S \subset \mathbb{R}^n$, define an operator R_S on $\mathcal{S}(\mathbb{R}^n)$ by $R_S f = \hat{f}|_S$ (not to be confused with the R_S appearing in the notation $R_S(p \to q)$).

Definition 3.6. We call R_S the *restriction operator* associated to S.

Clearly, $R_S(p \to q)$ is equivalent to the statement that R_S can be extended to a bounded operator $L^p(\mathbb{R}^n) \to L^q(S, d\sigma)$. When expressed in the language of operators, it becomes evident that there should be a dual statement that is equivalent to $R_S(p \to q)$, and this is indeed the case. To discover the dual statement, let $f \in \mathcal{S}(\mathbb{R}^n)$ and $g \in L^1(S, d\sigma)$ be given. Fubini's theorem gives

$$\begin{aligned}\int_S R_S f(\xi)\overline{g(\xi)}\, d\sigma(\xi) &= \int_S \Big(\int_{\mathbb{R}^n} f(x)e^{-2\pi i x\cdot\xi}\, dx\Big)\overline{g(\xi)}\, d\sigma(\xi) \\ &= \int_{\mathbb{R}^n} f(x)\overline{(g\, d\sigma)^{\vee}(x)}\, dx.\end{aligned}$$

Letting E_S be the operator acting on $L^1(S, d\sigma)$ by $E_S g = (g\, d\sigma)^{\vee}$, the above can be summarised by the formula

$$\int_S R_S f(\xi)\overline{g(\xi)}\, d\sigma(\xi) = \int_{\mathbb{R}^n} f(x)\overline{E_S g(x)}\, dx. \tag{3.2}$$

Definition 3.7. We call E_S the *extension operator* associated to S.

Formally, equation (3.2) can be interpreted to mean that the extension operator is the adjoint of the restriction operator.
Alternative to equation (3.1), one might consider whether we have an estimate of the form

$$\|E_S g\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|g\|_{L^{q'}(S, d\sigma)} \qquad \forall g \in L^{q'}(S, d\sigma) \tag{3.3}$$

(noting that $L^{q'}(S, d\sigma) \subset L^1(S, d\sigma)$ by a similar application of Hölder's inequality as in Lemma 3.4). Loosely, we refer to estimates of this form as *extension estimates*. Such estimates are also important enough to deserve their own notation:

Definition 3.8. Given a submanifold S and a pair of exponents $1 \le p', q' \le \infty$, we denote by $R_S^*(q' \to p')$ the statement that equation (3.3) holds.

As one might expect, the statements $R_S(p \to q)$ and $R_S^*(q' \to p')$ are in fact equivalent.

Theorem 3.9. *Given a submanifold S and a pair of exponents $1 \le p, q \le \infty$, the statements $R_S(p \to q)$ and $R_S^*(q' \to p')$ are equivalent (with the same implied constants).*

Proof. Suppose $\|R_S f\|_{L^q(S, d\sigma)} \le A\|f\|_{L^p(\mathbb{R}^n)}$ for all $f \in \mathcal{S}(\mathbb{R}^n)$, and let $g \in L^{q'}(S, d\sigma)$. A standard duality formula gives

$$\begin{aligned}\|E_S g\|_{L^{p'}(\mathbb{R}^n)} &= \sup\Big\{\Big|\int_{\mathbb{R}^n} f(x)\overline{E_S g(x)}\, dx\Big| : f \in \mathcal{S}(\mathbb{R}^n),\ \|f\|_{L^p(\mathbb{R}^n)} = 1\Big\} \\ &= \sup\Big\{\Big|\int_S R_S f(\xi)\overline{g(\xi)}\, d\sigma(\xi)\Big| : f \in \mathcal{S}(\mathbb{R}^n), \|f\|_{L^p(\mathbb{R}^n)} = 1\Big\},\end{aligned} \tag{3.4}$$

where the second line follows by (3.2). But for any $f \in \mathcal{S}(\mathbb{R}^n)$ such that $\|f\|_{L^p(\mathbb{R}^n)} = 1$, the hypothesis $\|R_S f\|_{L^q(S,\,d\sigma)} \leq A\|f\|_{L^p(\mathbb{R}^n)}$ and Hölder's inequality give

$$\Big|\int_S R_S f(\xi)\overline{g(\xi)}\,d\sigma(\xi)\Big| \leq A\|g\|_{L^{q'}(S,\,d\sigma)}.$$

The supremum in Equation (3.4) is therefore at most $A\|g\|_{L^{q'}(S,\,d\sigma)}$, so we have $\|E_S g\|_{L^{p'}(\mathbb{R}^n)} \leq A\|g\|_{L^{q'}(S\,d\sigma)}$ for all $g \in L^{q'}(S, d\sigma)$.
The converse follows by a similar argument, using the duality formula

$$\|R_S f\|_{L^q(S,\,d\sigma)} = \sup\Big\{\Big|\int_S R_S f(\xi)\overline{g(\xi)}\,d\sigma(\xi)\Big| : g \in L^{q'}(S, d\sigma),\ \|g\|_{L^{q'}(S,\,d\sigma)} = 1\Big\}.$$

□

In light of Theorem 3.9, it is clear that we lose no generality by focusing only on estimates of the form $R_S^*(q' \to p')$. Indeed, despite the origins of the restriction problem lying in the non-dual formulation, there has been a trend in the literature to favour the dual formulation. We adopt this approach, henceforth focusing mainly on estimates of the form $R_S^*(q' \to p')$ (with some exceptions). As such, we will drop the use of dual exponents for such estimates, and simply write them in the form $R_S^*(p \to q)$. One advantage of this approach is that the dual formulation can be used more readily to establish necessary conditions for restriction or extension estimates to hold, as we will see in the following section. We will also see later that extension operators and estimates are intimately connected to the study of PDE.
In addition to focusing on extension estimates, we will henceforth adopt the convention of using variants of g to denote functions on $\mathbb{R}^n$, preserving variants of f to denote functions on a submanifold (reversing an implicit convention held prior to this point). We also note that we will tend to blur the distinction between restriction and extension estimates, the two being equivalent.
We pause momentarily to record the dual counterparts of our elementary observations in Lemmas 3.3, 3.4, and 3.5:

Lemma 3.10. *$R_S^*(1 \to \infty)$ holds for any submanifold S.* □

Lemma 3.11. *Given a submanifold S and a pair of exponents $1 \leq p, q \leq \infty$, $R_S^*(p \to q)$ implies $R_S^*(r \to q)$ for all $r > p$.* □

Lemma 3.12. *Given a submanifold S and a pair of exponents $1 \leq p, q \leq \infty$, $R_S^*(p \to q)$ implies $R_S^*(p \to r)$ for all $r > q$.* □

Each of these follows by a direct application of Theorem 3.9 and the corresponding restriction lemma. Alternatively, elementary proofs of these results can be given along the same lines as their restriction counterparts.
We note that in light of Lemmas 3.11 and 3.12, the general goal is to prove extension estimates for the smallest possible exponents

3.3 Necessary Conditions and the Restriction Conjecture

For a given submanifold S, it is natural to ask for which pairs of exponents (p, q) $R_S^*(p \to q)$ holds. The answer depends on the submanifold S, with characteristics such as curvature and dimension being important. We will limit our attention to compact hypersurfaces, which is to say compact submanifolds of $\mathbb{R}^n$ of dimension $n-1$. In fact, as mentioned in Section 2.4, we will primarily consider the truncated paraboloid P^{n-1} in $\mathbb{R}^n$, though most of our results may be readily generalised to any compact hypersurface with positive definite second fundamental form.

Since we are focusing our attention on a particular submanifold for each n, we will sometimes tidy our notation by omitting the subscript from the operator $E_{P^{n-1}}$. Thus, we understand all instances of the operator E to mean $E_{P^{n-1}}$ for whatever dimension n is implicitly fixed in the given argument.
Recall that for $f \in L^p(S, d\sigma)$, $Ef = (fd\sigma)^{\vee}$ is just the inverse Fourier transform of the surface-supported measure $fd\sigma$. We can therefore use our understanding of the asymptotics of the Fourier transforms of such measures from Section 2.4 to derive necessary conditions for extension estimates to hold. Indeed, suppose $R^*_{P^{n-1}}(p \to q)$ holds for some $1 \leq p, q \leq \infty$, and consider the function $f \equiv 1$ on P^{n-1}. Then, $Ef = (d\sigma)^{\vee}$, so $R^*_{P^{n-1}}(p \to q)$ dictates that $\|(d\sigma)^{\vee}\|_{L^q(\mathbb{R}^n)} \lesssim \|1\|_{L^p(P^{n-1}, d\sigma)} = O(1)$. But by Remark 2.45, $(d\sigma)^{\vee}$ decays asymptotically like $(1 + |x|)^{-\frac{n-1}{2}}$ in directions normal to P^{n-1}, from which we see by integrating in polar coordinates that $\|(d\sigma)^{\vee}\|_{L^q(\mathbb{R}^n)}$ can only be finite if $q\frac{n-1}{2} > n$. It follows that for $R^*_{P^{n-1}}(p \to q)$ to hold, we must have

$$q > \frac{2n}{n-1}. \tag{3.5}$$

We may establish further necessary conditions on the exponents p and q by considering the scalings $(\xi', \xi_n) \mapsto (\lambda\xi', \lambda^2\xi_n)$ for $\lambda \geq 1$. Indeed, let $f \in L^p(P^{n-1}, d\sigma)$, and extend f to an L^p function $\tilde{f}$ on the non-truncated paraboloid by defining $\tilde{f}(\xi) = 0$ for $\xi \notin P^{n-1}$. Given $\lambda \geq 1$, define $T_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ by $T(\xi', \xi_n) =$

$(\lambda\xi', \lambda^2\xi_n)$, and define $f^\lambda \in L^p(P^{n-1}, d\sigma)$ by $f^\lambda = \tilde{f} \circ T_\lambda$ (noting that T_λ preserves the paraboloid). $R^*_{P^{n-1}}(p \to q)$ gives

$$\|Ef^\lambda\|_{L^q(\mathbb{R}^n)} \lesssim \|f^\lambda\|_{L^p(P^{n-1}, d\sigma)}, \tag{3.6}$$

but a change of variables reveals that

$$\|f^\lambda\|_{L^p(P^{n-1}, d\sigma)} \sim \lambda^{-\frac{n-1}{p}} \|f\|_{L^p(P^{n-1}, d\sigma)}, \tag{3.7}$$

and another change of variables gives $Ef^\lambda(x) = \lambda^{-(n-1)}(Ef)^{\lambda^{-1}}(x)$. Hence,

$$\begin{aligned} \|Ef^\lambda\|_{L^q(\mathbb{R}^n)} &= \lambda^{-(n-1)} \|(Ef)^{\lambda^{-1}}\|_{L^q(\mathbb{R}^n)} \\ &= \lambda^{-(n-1)+\frac{n+1}{q}} \|Ef\|_{L^q(\mathbb{R}^n)}. \end{aligned} \tag{3.8}$$

Combining equations (3.6), (3.7) and (3.8), we see that for all $\lambda \geq 1$,

$$\lambda^{-(n-1)+\frac{n+1}{q}} \|Ef\|_{L^q(\mathbb{R}^n)} \lesssim \lambda^{-\frac{n-1}{p}} \|f\|_{L^p(P^{n-1}, d\sigma)},$$

hence

$$\|Ef\|_{L^q(\mathbb{R}^n)} \lesssim \lambda^{\frac{n-1}{p'} - \frac{n+1}{q}} \|f\|_{L^p(P^{n-1}, d\sigma)}. \tag{3.9}$$

If $\frac{n-1}{p'} - \frac{n+1}{q} < 0$, then by letting $\lambda \to \infty$ in (3.9), we find that $Ef = 0$ for all $f \in L^p(P^{n-1}, d\sigma)$. But this is a contradiction, giving the additional necessary condition

$$\frac{n+1}{q} \leq \frac{n-1}{p'}. \tag{3.10}$$

We note that there is an alternative derivation of the necessary condition (3.10) along similar lines to that of (3.5) which is applicable to other submanifolds; the idea is to let f^R be a "smoothed out" version of the characteristic function of a small cap of radius R, and consider the behaviour of $\|Ef^R\|_{L^q(\mathbb{R}^n)}$ and $\|f^R\|_{L^p(P^{n-1}, d\sigma)}$ as $R \to 0$. This is known as the *Knapp example*, originally published by Strichartz [Str77]; an accessible heuristic explanation can be found in [Tao04].

The famous restriction conjecture (in the case of the paraboloid) simply states that the necessary conditions (3.5) and (3.10) on the exponents p and q are also sufficient:

Conjecture 3.13 (The Restriction Conjecture)**.** $R^*_{P^{n-1}}(p \to q)$ *holds if and only if the conditions* (3.5) *and* (3.10) *are satisfied.*

This conjecture can be easily visualised using the strong-type diagram for the extension operator $E_{P^{n-1}}$, as shown in the following figure:

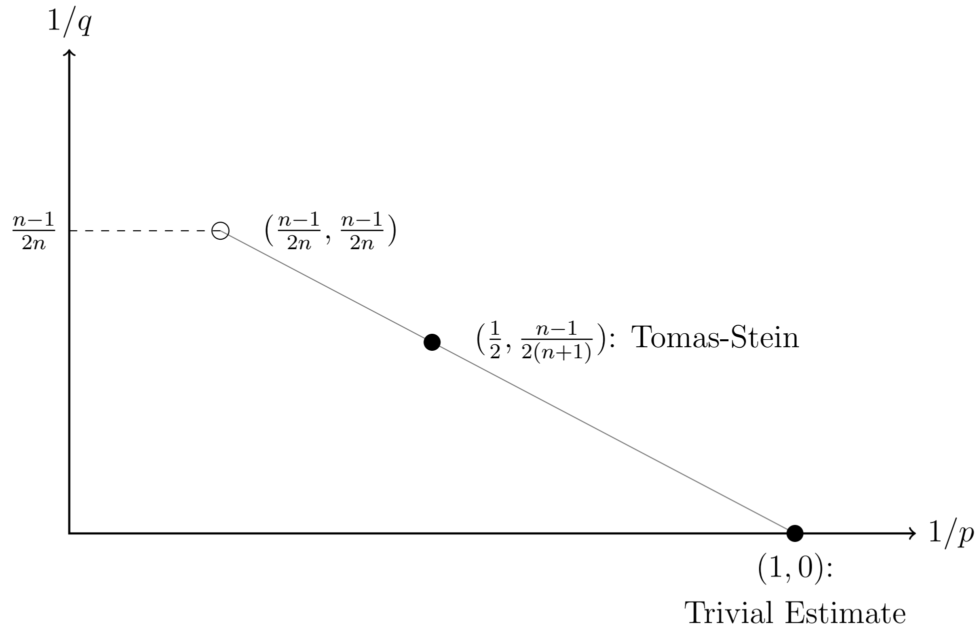

Figure 3.1: The conjectured strong-type diagram for $E_{P^{n-1}}$.

The restriction conjecture states that $E_{P^{n-1}}$ is of strong-type (p,q) for all $(\frac{1}{p}, \frac{1}{q})$ lying inside the convex trapezoidal region. This is the intersection of the region $0 \le \frac{1}{p}, \frac{1}{q} \le 1$ (corresponding to the fact that $1 \le p, q \le \infty$) with the regions $\frac{1}{q} < \frac{n-1}{2n}$ and $\frac{n+1}{q} \le \frac{n-1}{p'}$ (corresponding to the necessary conditions (3.5) and (3.10) respectively). Thus, the trapezoidal region is bounded above by the line $\frac{1}{q} = \frac{n-1}{2n}$, and on the diagonal by the line $\frac{n+1}{q} = \frac{n-1}{p'}$ (or equivalently, $\frac{1}{q} = \frac{n-1}{n+1}(1 - \frac{1}{p})$). The latter of these is known as the *critical line*, since proving the conjecture along this line would imply the conjecture in full. Indeed, Lemma 3.11 states that if the conjecture is known at some point in the above region, then it is automatically true for all points to the left and on the same horizontal line (we note that Lemma 3.10, on the other hand, simply states that the conjecture is already known at the point $(1,0)$, as marked in the figure). Moreover, by interpolation, if the conjecture is known for two pairs of exponents (p_1, q_1) and (p_2, q_2), then it is also true for all pairs of exponents (p,q) for which $(\frac{1}{p}, \frac{1}{q})$ lies on the straight line joining $(\frac{1}{p_1}, \frac{1}{q_1})$ and $(\frac{1}{p_2}, \frac{1}{q_2})$. It follows by Lemma 3.11 and interpolation with the trivial estimate that if the conjecture is known at some point $(\frac{1}{p_1}, \frac{1}{q_1})$ on the critical line, then it must also be true for all points $(\frac{1}{p_2}, \frac{1}{q_2})$ in the trapezoidal region with $\frac{1}{q_2} \le \frac{1}{q_1}$. As

such, progress on the restriction conjecture is measured by how close one can get on the critical line to the endpoint $(\frac{n-1}{2n}, \frac{n-1}{2n})$.
The restriction conjecture was resolved in full for the case $n = 2$ by Zygmund [Zyg74], but remains open in all other dimensions. The first major development in higher dimensions was made by Tomas [Tom75], by proving the conjecture along the line $\frac{1}{p} = \frac{1}{2}$ up to but not including the endpoint $(\frac{1}{2}, \frac{n-1}{2(n+1)})$ for all $n \geq 2$, a statement which we will call the *Tomas restriction theorem*. This was soon strengthened by Stein to include the endpoint, and we therefore refer to the estimate $R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1})$ as the *Tomas-Stein estimate*, with $\frac{2(n+1)}{n-1}$ being referred to as the *Tomas-Stein exponent*. Our main result in Chapter 5 will be a new proof of the Tomas restriction theorem, so we dedicate the next section to a careful analysis of the original methods of Tomas, as well as some applications of the Tomas-Stein estimate to PDE.
The restriction conjecture remains a very active field of research, with the best results to date given in the recent paper [HZ20].

3.4 The Tomas Restriction Theorem

A significant early development on the restriction conjecture was made by Tomas [Tom75], with a proof that $R^*_{S^{n-1}}(2 \to q)$ holds for all $q > \frac{2(n+1)}{n-1}$. In fact, Tomas proved the equivalent restriction statement that $R_{S^{n-1}}(p \to 2)$ holds for all $1 \leq p < \frac{2(n+1)}{n+3}$. Tomas' proof utilised the decay of the Fourier transform of the surface measure of S^{n-1}, and as such, generalises easily to other hypersurfaces including P^{n-1}. In keeping with our focus on the paraboloid, we present here an elaboration of Tomas' original proof, adapted to the paraboloid. Some inspiration has been drawn from an exposition of Tao [Tao20b].

Theorem 3.14 (The Tomas Restriction Theorem). *$R_{P^{n-1}}(p \to 2)$ holds for all $1 \leq p < \frac{2(n+1)}{n+3}$.*

Proof. Let $1 \leq p < \frac{2(n+1)}{n+3}$. We must show that

$$\int_{P^{n-1}} |\hat{g}(\xi)|^2 \, d\sigma(\xi) \lesssim \|g\|^2_{L^p(\mathbb{R}^n)}$$

for all $g \in \mathcal{S}(\mathbb{R}^n)$. Using the identity $|\hat{g}(\xi)|^2 = \hat{g}(\xi)\overline{\hat{g}(\xi)}$ and Fubini's theorem, we find that

$$\begin{aligned}\int_{P^{n-1}} |\hat{g}(\xi)|^2\, d\sigma(\xi) &= \int_{P^{n-1}} \Big(\int_{\mathbb{R}^n} g(x)e^{-2\pi i x\cdot\xi}\, dx\Big)\Big(\int_{\mathbb{R}^n} \overline{g(y)}e^{2\pi i y\cdot\xi}\, dx\Big)\, d\sigma(\xi)\\ &= \int_{\mathbb{R}^n} g(x)\overline{(g * (d\sigma)\check{\ })(x)}\, dx.\end{aligned}$$

By Hölder's inequality, it follows that

$$\int_{P^{n-1}} |\hat{g}(\xi)|^2\, d\sigma(\xi) \le \|g\|_{L^p(\mathbb{R}^n)}\|g * (d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)},$$

and it therefore suffices to prove

$$\|g * (d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|g\|_{L^p(\mathbb{R}^n)}.$$

We use a variant of a standard technique from harmonic analysis known as *dyadic decomposition*: fix some $\varphi \in C_c^\infty(\mathbb{R}^n)$ with compact support in $B(0,1)$ such that $\varphi \equiv 1$ on $B(0,1/2)$, and define $\psi(x) = \varphi(x) - \varphi(2x)$ so that $\psi \in C_c^\infty(\mathbb{R}^n)$ is supported in the annulus $B(0,1) \setminus B(0,1/4)$. By a telescoping series argument, we then have

$$1 = \varphi(x) + \sum_{k=1}^{\infty} \psi(x/2^k),$$

so we may write

$$(d\sigma)\check{\ }(x) = \varphi(x)(d\sigma)\check{\ }(x) + \sum_{k=1}^{\infty} \psi(x/2^k)(d\sigma)\check{\ }(x).$$

The triangle inequality then gives

$$\|g * (d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)} \le \|g * \varphi(d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)} + \sum_{k=1}^{\infty} \|g * \psi(\cdot/2^k)(d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)},$$

so it suffices to prove the estimates

$$\|g * \varphi(d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|g\|_{L^p(\mathbb{R}^n)} \tag{3.11}$$

and

$$\|g * \psi(\cdot/2^k)(d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)} \lesssim 2^{-ck}\|g\|_{L^p(\mathbb{R}^n)} \qquad \forall k \ge 1, \tag{3.12}$$

for some constant $c > 0$ independent of k and g.

To prove (3.11), we note that $p < 2$ implies $p' > 2$, hence $p'/2 > 1$. Letting $r = p'/2$, we then have $1 + \frac{1}{p'} = \frac{1}{p} + \frac{1}{r}$, so Young's convolution inequality gives

$$\|g * \varphi(d\sigma)\check{\ }\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|g\|_{L^p(\mathbb{R}^n)}\|\varphi(d\sigma)\check{\ }\|_{L^r(\mathbb{R}^n)}.$$

This is (3.11) since $\varphi(d\sigma)\check{}$ is compactly supported, hence $\|\varphi(d\sigma)\check{}\|_{L^r(\mathbb{R}^n)}$ is finite. To prove (3.12), define for each integer $k \geq 1$ an operator T_k mapping functions on $\mathbb{R}^n$ to functions on $\mathbb{R}^n$ by $T_k g = g * \psi(\cdot/2^k)(d\sigma)\check{}$. Then, equation (3.12) is equivalent to a strong-type (p, p') bound for T_k for all $k \geq 1$, with $\|T_k\|_{L^p \to L^{p'}} = O(2^{-ck})$. We will establish this by interpolating between strong-type $(1, \infty)$ and $(2, 2)$ bounds for T_k.

Fix some $k \geq 1$. For the strong-type $(1, \infty)$ estimate, we use Young's convolution inequality to find

$$\begin{aligned} \|g * \psi(\cdot/2^k)(d\sigma)\check{}\|_{L^\infty(\mathbb{R}^n)} &\lesssim \|g\|_{L^1(\mathbb{R}^n)} \|\psi(\cdot/2^k)(d\sigma)\check{}\|_{L^\infty(\mathbb{R}^n)} \\ &\lesssim 2^{-\frac{k(n-1)}{2}} \|g\|_{L^1(\mathbb{R}^n)}, \end{aligned} \tag{3.13}$$

where the last inequality follows since $\psi(\cdot/2^k) \lesssim 1$ is supported on the annulus $B(0, 2^k) \setminus B(0, 2^{k-2})$, on which $(d\sigma)\check{} \lesssim 2^{-\frac{k(n-1)}{2}}$ by Theorem 2.44.

For the strong-type $(2, 2)$ estimate, we use Plancherel's theorem to obtain

$$\begin{aligned} \|g * \psi(\cdot/2^k)(d\sigma)\check{}\|_{L^2(\mathbb{R}^n)} &= \|\hat{g}(\psi(\cdot/2^k)(d\sigma)\check{})\hat{}\|_{L^2(\mathbb{R}^n)} \\ &\leq \|g\|_{L^2(\mathbb{R}^n)} \|(\psi(\cdot/2^k)(d\sigma)\check{})\hat{}\|_{L^\infty(\mathbb{R}^n)}, \end{aligned} \tag{3.14}$$

so we must estimate the Fourier transform $(\psi(\cdot/2^k)(d\sigma)\check{})\hat{}$. Fixing some $\xi \in \mathbb{R}^n$, Fubini's theorem gives

$$\begin{aligned} (\psi(\cdot/2^k)(d\sigma)\check{})\hat{}(\xi) &= \int_{\mathbb{R}^n} \psi(x/2^k) \Big(\int_{P^{n-1}} e^{2\pi i x \cdot \omega} \, d\sigma(\omega) \Big) e^{-2\pi i x \cdot \xi} \, dx \\ &= \int_{P^{n-1}} \Big(\int_{\mathbb{R}^n} \psi(x/2^k) e^{-2\pi i x \cdot (\xi - \omega)} \, dx \Big) \, d\sigma(\omega) \\ &= \int_{P^{n-1}} \widehat{\psi_{2^{-k}}}(\xi - \omega) \, d\sigma(\omega), \end{aligned} \tag{3.15}$$

so by Proposition 2.2, we have

$$(\psi(\cdot/2^k)(d\sigma)\check{})\hat{}(\xi) = 2^{kn} \int_{P^{n-1}} \hat{\psi}(2^k(\xi - \omega)) \, d\sigma(\omega).$$

Hence, by the triangle inequality and the rapid decay of the Schwartz function $\hat{\psi}$, we have

$$|(\psi(\cdot/2^k)(d\sigma)\check{})\hat{}(\xi)| \lesssim 2^{kn} \int_{P^{n-1}} (1 + 2^k|\xi - \omega|)^{-100n} \, d\sigma(\omega). \tag{3.16}$$

We now use yet another variant of dyadic decomposition: letting $A_j = B(\xi, 2^{j-k}) \setminus B(\xi, 2^{j-k-1})$ for all integers $j \geq 1$, observe that we may express $\mathbb{R}^n$ as the disjoint union

$$\mathbb{R}^n = B(\xi, 2^{-k}) \cup \bigcup_{j=1}^{\infty} A_j.$$

We may therefore decompose the right-hand side of equation (3.16) using the Lebesgue monotone convergence theorem to obtain

$$\begin{aligned}
|(\psi(\cdot/2^k)(d\sigma)\check{\ })\hat{\ }(\xi)| &\lesssim 2^{kn}\Big(\int_{P^{n-1}\cap B(\xi,2^{-k})} (1+2^k|\xi-\omega|)^{-100n}\, d\sigma(\omega) \\
&\qquad + \sum_{j=1}^{\infty}\int_{P^{n-1}\cap A_j} (1+2^k|\xi-\omega|)^{-100n}\, d\sigma(\omega)\Big) \\
&= 2^{kn}(I+II).
\end{aligned} \tag{3.17}$$

To estimate I, we note that the integrand is at most 1 and Lemma 2.46 gives $\sigma(B(\xi,2^{-k})) \lesssim 2^{-k(n-1)}$, from which we see $I \lesssim 2^{-k(n-1)}$. To estimate II, we observe that for $\omega \in A_j$, we have $|\xi-\omega| \geq 2^{j-k-1}$, and it follows that $(1+2^k|\xi-\omega|)^{-100n} \leq (2^{j-1})^{-100n} \lesssim 2^{-100jn}$. Since $A_j \subset B(0,2^{j-k})$, we also have $\sigma(A_j) \leq \sigma(B(0,2^{j-k})) \lesssim 2^{(j-k)(n-1)}$. It follows that we also have

$$\begin{aligned}
II &\lesssim \sum_{j=1}^{\infty} 2^{-100jn+(j-k)(n-1)} \\
&= 2^{-k(n-1)}\sum_{j=1}^{\infty} 2^{-(99n+1)j} \lesssim 2^{-k(n-1)}.
\end{aligned}$$

Equation (3.17) therefore gives $\|(\psi(\cdot/2^k)(d\sigma)\check{\ })\hat{\ }\|_{L^\infty(\mathbb{R}^n)} \lesssim 2^k$, so equation (3.14) gives

$$\|g * \psi(\cdot/2^k)(d\sigma)\check{\ }\|_{L^2(\mathbb{R}^n)} \lesssim 2^k\|g\|_{L^2(\mathbb{R}^n)} \tag{3.18}$$

Interpolating between the strong-type bounds (3.13) and (3.18) for the operator T_k using Riesz-Thorin interpolation, we find that for all $0 \leq \theta \leq 1$, T_k is of strong-type (p_θ, q_θ) with

$$\begin{aligned}
\|T_k\|_{L^{p_\theta}\to L^{q_\theta}} &\leq \|T_k\|_{L^1\to L^\infty}^{1-\theta}\|T_k\|_{L^2\to L^2}^{\theta} \\
&\lesssim 2^{-k\frac{(n-1)}{2}(1-\theta)+k\theta},
\end{aligned}$$

where $\frac{1}{p_\theta} = 1-\frac{\theta}{2}$ and $\frac{1}{q_\theta} = \frac{\theta}{2}$. Setting $\theta = \frac{2}{p'}$ (noting that $0 < \frac{2}{p'} < 1$ since $p < 2$, implying $p' > 2$), we obtain

$$\|T\|_{L^p\to L^{p'}} \lesssim 2^{-2k(1+\frac{n-1}{2})(\frac{n-1}{2(n+1)}-\frac{1}{p'})}.$$

This proves (3.12) with $c = 2(1+\frac{n-1}{2})(\frac{n-1}{2(n+1)}-\frac{1}{p'})$, where $c > 0$ since $1 \leq p < \frac{2(n+1)}{n+3}$ implies $p' > \frac{2(n+1)}{n-1}$. □

By using a more delicate interpolation method, Stein was able to extend the Tomas restriction theorem to the endpoint, proving the Tomas-Stein estimate $R_{P^{n-1}}(\frac{2(n+1)}{n+3} \to 2)$, or equivalently, $R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1})$. One approach to this estimate, different to the original method of Stein, is given in [Tao20b].
Some connections between PDE and extension estimates can be seen by considering the Schrödinger equation $2\pi i \partial_{x_n} u = \Delta_{x'} u$ on $\mathbb{R}^{n-1} \times [0, \infty)$ with initial data $u(x', 0) = g(x')$. By a standard argument in which one converts the PDE into a family of ODE by taking the Fourier transform with respect to the x' variable and using Proposition 2.5, one obtains that a solution is given by

$$u(x', x_n) = \int_{\mathbb{R}^{n-1}} \hat{g}(\xi') e^{2\pi i (x', x_n)\cdot(\xi', |\xi'|^2)} \, d\xi'. \tag{3.19}$$

In particular, if $\hat{g} \in C_c^\infty([-1,1]^{n-1})$, equation (3.19) gives that $u(x', x_n) = (E_{P^{n-1}} G)(x', x_n)$, where we define $G : P^{n-1} \to \mathbb{C}$ by $G(\xi', |\xi'|^2) = \hat{g}(\xi')/h(\xi')$ for h as defined in Definition 2.42. The Tomas-Stein estimate $R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1})$ therefore gives

$$\begin{aligned} \|u\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} &\lesssim \|G\|_{L^2(P^{n-1}, d\sigma)} \\ &\lesssim \|g\|_{L^2(\mathbb{R}^{n-1})}, \end{aligned}$$

where we have used Plancherel and the fact that $h \sim 1$ on $[-1,1]^{n-1}$. Similar estimates may be derived for general $g \in \mathcal{S}(\mathbb{R}^{n-1})$ by proving the Tomas-Stein estimate for the non-truncated paraboloid.

3.5 Local Restriction and Extension Estimates

Let $S \subset \mathbb{R}^n$ be a submanifold. We now introduce "local" variants of the statements $R_S(p \to q)$ and $R^*_S(p \to q)$:

Definition 3.15. Let $\varepsilon > 0$. If for all $R \geq 1$ and all $g \in \mathcal{S}(\mathbb{R}^n)$ supported in $B(0, R)$ we have

$$\|\hat{g}\|_{L^q(S, d\sigma)} \lesssim R^\varepsilon \|g\|_{L^p(B(0,R))}, \tag{3.20}$$

we say that $R_S(p \to q\,; \varepsilon)$ holds.

Loosely, we refer to estimates of the form (3.20) as *local restriction estimates.*

Definition 3.16. Let $\varepsilon > 0$. If for all $R \geq 1$ and all $f \in L^{q'}(S, d\sigma)$ we have

$$\|E_S f\|_{L^{p'}(B(0,R))} \lesssim R^\varepsilon \|f\|_{L^{q'}(S, d\sigma)}, \tag{3.21}$$

we say that $R^*_S(q' \to p'\,; \varepsilon)$ holds.

Loosely, we refer to estimates of the form (3.21) as *local extension estimates.* We note that by translation invariance, estimate (3.21) is equivalent to the same estimate in which the $L^{p'}$ norm is taken over any ball $B(x, R)$ of radius R. Indeed, given any $x \in \mathbb{R}^n$, an analogue of Proposition 2.2 gives $(E_S f)_{-x} = E_S(e_x f)$. It follows that if estimate (3.21) holds, then for all $f \in L^{q'}(S, d\sigma)$, we have

$$\begin{aligned}
\|E_S f\|_{L^{p'}(B(x,R))} &= \|(E_S f)_{-x}\|_{L^{p'}(B(0,R))} \\
&\lesssim R^{\varepsilon} \|e_x f\|_{L^{q'}(S,\, d\sigma)} \\
&= R^{\varepsilon} \|f\|_{L^{q'}(S,\, d\sigma)}.
\end{aligned}$$

Analogously, estimate (3.20) is equivalent to the same estimate in which g is supported in any ball $B(x, R)$ of radius R, and where the L^p norm is taken over the same ball.

Without further qualification, we understand the terms *restriction estimates* and *extension estimates* to refer to those of the form (3.1) and (3.3) respectively; when referring to estimates of the form (3.20) and (3.21), we will always include the description *local*. On the other hand, when we wish to emphasise that we are not referring to local restriction or extension estimates, we will refer to estimates of the form (3.1) and (3.3) as *global* restriction and extension estimates respectively.

It is not hard to see by an analogous argument to Theorem 3.9 that $R_S(p \to q\,;\, \varepsilon)$ and $R^*_S(q' \to p'\,;\, \varepsilon)$ are equivalent. We therefore lose no generality by focusing on local extension estimates, and as such, we will adopt the same convention for local estimates as we do for global estimates; namely, we will tend to focus on local extension estimates, which we will write without dual exponents in the form $R^*_S(p \to q\,;\, \varepsilon)$.

Clearly, $R^*_S(p \to q\,;\, \varepsilon)$ implies $R^*_S(p \to q\,;\, \varepsilon')$ for all $\varepsilon' > \varepsilon$. Moreover, it is easy to see that $R^*_S(p \to q)$ implies $R^*_S(p \to q\,;\, \varepsilon)$ for all $\varepsilon > 0$. The utility of local estimates lies in the question of whether the converse, or a partial converse, of this statement holds. In particular, if we knew that $R^*_S(p \to q\,;\, \varepsilon)$ for all $\varepsilon > 0$ implied $R^*_S(p \to q)$, then we could use local extension estimates to prove a corresponding global extension estimate. Unfortunately, the converse as stated here is not true in general, but under some circumstances something almost as good is true: knowing $R^*_S(p \to q\,;\, \varepsilon)$ for some $\varepsilon > 0$ implies a global extension estimate where the ε loss is transferred to one or both of the exponents p, q. That is, it can be shown under some circumstances that $R^*_S(p \to q\,;\, \varepsilon)$ implies $R^*_S(p_\varepsilon \to q_\varepsilon)$ for some $p_\varepsilon \geq p$ and $q_\varepsilon \geq q$. In particular, we would like to have $p_\varepsilon \to p$ and $q_\varepsilon \to q$ as $\varepsilon \to 0$, in which case knowing $R^*_S(p \to q\,;\, \varepsilon)$ for all $\varepsilon > 0$ implies a family of global extension estimates having $R^*_S(p \to q)$ as one of its endpoints. This will be our approach to proving the Tomas restriction theorem in Chapter 5 based on the decoupling methods introduced in Chapter 4; decoupling will provide a relatively short proof of the local extension estimates $R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1}\,;\,\varepsilon)$ for all $\varepsilon > 0$, after which we must prove a concrete *ε-removal theorem* allowing us to deduce the Tomas restriction theorem based on this family of local extension estimates.

We now prove a well-known equivalent condition for $R^*_S(p \to q\,;\,\varepsilon)$ to hold. Heuristically, the idea is that upon localising to scale R on the spatial side in equation (3.21),we may safely "blur" to scale R^{-1} on the frequency side (this heuristic is motivated by a general principle in Fourier analysis known as the *uncertainty principle*). This result holds for more general compact submanifolds, but we will once again adhere to the case of the paraboloid P^{n-1} for the sake of concreteness. In what follows, given $\delta > 0$, we let $\mathcal{N}_\delta(P^{n-1})$ denote the δ-neighbourhood of P^{n-1}; that is,

$$\mathcal{N}_\delta(P^{n-1}) = \{\xi \in \mathbb{R}^n : |\xi - \omega| < \delta \text{ for some } \omega \in P^{n-1}\}.$$

We divide the necessary and sufficient conditions into two separate propositions:

Proposition 3.17. *Let $1 \le p, q \le \infty$, let $\varepsilon > 0$, and suppose that for all $R \ge 1$ and all $g \in \mathcal{S}(\mathbb{R}^n)$ with Fourier support in $\mathcal{N}_{R^{-1}}(P^{n-1})$, we have $\|g\|_{L^q(B(0,R))} \lesssim R^{\varepsilon - 1/p'}\|\hat{g}\|_{L^p(\mathbb{R}^n)}$. Then, $R^*_{P^{n-1}}(p \to q\,;\,\varepsilon)$ holds.*

Proof. Let $f \in L^p(P^{n-1}, d\sigma)$ and choose by Proposition 2.13 some $\psi \in \mathcal{S}(\mathbb{R}^n)$ such that $\psi \sim 1$ on $B(0,1)$, and $\hat{\psi}$ is nonnegative and supported in $B(0,1)$. Given $R \ge 1$, we have

$$\begin{aligned} \|Ef\|_{L^q(B(0,R))} &= \Big(\int_{B(0,R)} |Ef(x)|^q\,dx\Big)^{1/q} \\ &\sim \Big(\int_{B(0,R)} |\psi(x/R)Ef(x)|^q\,dx\Big)^{1/q} \\ &= \|\psi_{R^{-1}}Ef\|_{L^q(B(0,R))}. \end{aligned} \tag{3.22}$$

Now, applying Fubini's theorem as in the derivation of equation (3.15) gives

$$(\psi_{R^{-1}}Ef)\widehat{\ }(\xi) = \int_{P^{n-1}} \widehat{\psi_{R^{-1}}}(\xi - \omega) f(\omega)\,d\sigma(\omega); \tag{3.23}$$

we therefore say that $(\psi_{R^{-1}}Ef)\widehat{\ } = \widehat{\psi_{R^{-1}}} * (f d\sigma)$ – a statement which can be made precise using the theory of *distributions*. Using the result regarding the support of a convolution and the fact that $\psi_{R^{-1}}$ has Fourier support in $B(0, R^{-1})$, we see that $(\psi_{R^{-1}}Ef)\widehat{\ }$ is supported in $P^{n-1} + B(0, R^{-1}) = \mathcal{N}_{R^{-1}}(P^{n-1})$ (alternatively, this can be shown by an elementary argument). Since $\psi_{R^{-1}}Ef$ is also Schwartz, the hypothesis gives

$$\|\psi_{R^{-1}}Ef\|_{L^q(B(0,R))} \lesssim R^{\varepsilon-1/p'}\|(\psi_{R^{-1}}Ef)\widehat{\ }\|_{L^p(\mathbb{R}^n)}.$$

By equation (3.22), we see that it therefore suffices to prove

$$\|(\psi_{R^{-1}}Ef)\widehat{\ }\|_{L^p(\mathbb{R}^n)} \lesssim R^{1/p'}\|f\|_{L^p(P^{n-1},\,d\sigma)}.$$

To do so, we write $\widehat{\psi_{R^{-1}}} = R^n\hat{\psi}_R$ in equation (3.23), and apply Hölder's inequality with the splitting

$$\hat{\psi}_R(\xi-\omega)f(\omega) = \hat{\psi}_R(\xi-\omega)^{1/p'+1/p}f(\omega)$$

to obtain

$$|(\psi_{R^{-1}}Ef)\widehat{\ }(\xi)| \le R^n\Big(\int_{P^{n-1}} |\hat{\psi}_R(\xi-\omega)|\,d\sigma(\omega)\Big)^{1/p'} \Big(\int_{P^{n-1}} |\hat{\psi}_R(\xi-\omega)||f(\omega)|^p\,d\sigma(\omega)\Big)^{1/p}. \tag{3.24}$$

Noting that

$$\begin{aligned}\int_{P^{n-1}} |\hat{\psi}_R(\xi-\omega)|\,d\sigma(\omega) &\le \|\hat{\psi}\|_{L^\infty(\mathbb{R}^n)}\int_{P^{n-1}} \chi_{\xi-\operatorname{supp}\hat{\psi}_R}(\omega)\,d\sigma(\omega)\\ &= \|\hat{\psi}\|_{L^\infty(\mathbb{R}^n)}\int_{P^{n-1}} \chi_{B(\xi,R^{-1})}(\omega)\,d\sigma(\omega)\\ &\lesssim R^{-(n-1)}\end{aligned}$$

(by Lemma 2.46), equation (3.24) gives that for all $\xi \in \mathbb{R}^n$,

$$|(\psi_{R^{-1}}Ef)\widehat{\ }(\xi)| \lesssim R^{n-(n-1)/p'}\Big(\int_{P^{n-1}} |\hat{\psi}_R(\xi-\omega)||f(\omega)|^p\,d\sigma(\omega)\Big)^{1/p}.$$

We therefore have

$$\|(\psi_{R^{-1}}Ef)\widehat{\ }\|_{L^p(\mathbb{R}^n)} \lesssim R^{n-(n-1)/p'}\Big(\int_{\mathbb{R}^n}\Big(\int_{P^{n-1}} |\hat{\psi}_R(\xi-\omega)||f(\omega)|^p\,d\sigma(\omega)\Big)\,d\xi\Big)^{1/p},$$

so upon applying Fubini's theorem and observing that $\int_{\mathbb{R}^n} |\hat{\psi}_R(\xi)|\,d\xi \sim R^{-n}$, we find that

$$\begin{aligned}\|(\psi_{R^{-1}}Ef)\widehat{\ }\|_{L^p(\mathbb{R}^n)} &\lesssim R^{n-(n-1)/p'-n/p}\|f\|_{L^p(P^{n-1},\,d\sigma)}\\ &= R^{1/p'}\|f\|_{L^p(P^{n-1},\,d\sigma)}.\end{aligned} \tag{3.25}$$

□

The converse is as follows:

Proposition 3.18. *Let $1 \leq p, q \leq \infty$, let $\varepsilon > 0$, and suppose that $R^*_{P^{n-1}}(p \to q\,;\varepsilon)$ holds. Then, for all $R \geq 1$ and all $g \in \mathcal{S}(\mathbb{R}^n)$ with Fourier support in $\mathcal{N}_{R^{-1}}(P^{n-1})$, we have $\|g\|_{L^q(B(0,R))} \lesssim R^{\varepsilon - 1/p'} \|\hat{g}\|_{L^p(\mathbb{R}^n)}$.*

Proof. Let $R \geq 1$, and let $g \in \mathcal{S}(\mathbb{R}^n)$ have Fourier support in $\mathcal{N}_{R^{-1}}(P^{n-1})$. Fix some constant $C > 0$, and for each $\delta > 0$ define

$$\tilde{\mathcal{N}_\delta}(P^{n-1}) := \{\xi = (\xi', \xi_n) \in \mathbb{R}^n : \xi' \in [-(1+\delta), 1+\delta]^{n-1}\,;\, |\xi_n - |\xi'|^2| \leq C\delta\}.$$

It is clear that by choosing C large enough, we can ensure $\mathcal{N}_\delta(P^{n-1}) \subset \tilde{\mathcal{N}_\delta}(P^{n-1})$ for all $\delta > 0$. It follows by Fourier inversion that for all $x \in \mathbb{R}^n$, we have

$$g(x) = \int_{\tilde{\mathcal{N}}_{R^{-1}}} \hat{g}(\xi) e^{2\pi i x\cdot\xi}\, d\xi.$$

Define a map $\Phi : [-1,1]^{n-1} \times [-CR^{-1}, CR^{-1}] \to \tilde{\mathcal{N}}_{R^{-1}}$ by

$$\Phi(\omega', \omega_n) = ((1+R^{-1})\omega', |(1+R^{-1})\omega'|^2 + \omega_n),$$

and define a linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ by $T(x', x_n) = ((1+R^{-1})x', (1+R^{-1})^2 x_n)$. Since $1 + R^{-1} \sim 1$ for $R \geq 1$, we have

$$|\det D\Phi(\omega)| = (1+R^{-1})^{n-1} \sim 1$$

for all $\omega \in [-1,1]^{n-1} \times [-CR^{-1}, CR^{-1}]$. It follows by a change of variables and Fubini's theorem that

$$\begin{aligned} |g(x)| &\sim \Big| \int_{-CR^{-1}}^{CR^{-1}} \Big(\int_{[-1,1]^{n-1}} \hat{g}(\Phi(\omega', \omega_n)) e^{2\pi i x\cdot(\Phi(\omega',\omega_n))}\, d\omega' \Big)\, d\omega_n \Big| \\ &\leq \int_{-CR^{-1}}^{CR^{-1}} \Big| \int_{[-1,1]^{n-1}} \hat{g}(\Phi(\omega', \omega_n)) e^{2\pi i (Tx)\cdot(\omega', |\omega'|^2)}\, d\omega' \Big|\, d\omega_n. \end{aligned} \tag{3.26}$$

For each $\omega_n \in [-CR^{-1}, CR^{-1}]$, define a function $f_{\omega_n} : P^{n-1} \to \mathbb{C}$ by

$$f_{\omega_n}(\omega', |\omega'|^2) = \frac{\hat{g}(\Phi(\omega', \omega_n))}{h(\omega')} \qquad \forall \omega' \in [-1,1]^{n-1},$$

where h is as defined in Definition 2.42. Equation (3.26) then gives

$$\begin{aligned} |g(x)| &\lesssim \int_{-CR^{-1}}^{CR^{-1}} \Big| \int_{P^{n-1}} f_{\omega_n}(\xi) e^{2\pi i (Tx)\cdot\xi}\, d\xi \Big|\, d\omega_n \\ &= \int_{-CR^{-1}}^{CR^{-1}} |Ef_{\omega_n}(Tx)|\, d\omega_n, \end{aligned}$$

so upon applying Minkowski's integral inequality, we get

$$\begin{aligned}\|g\|_{L^q(B(0,R))} &\lesssim \int_{-CR^{-1}}^{CR^{-1}} \Big(\int_{B(0,R)} |Ef_{\omega_n}(Tx)|^q \, dx \Big)^{1/q} d\omega_n \\ &\lesssim \int_{-CR^{-1}}^{CR^{-1}} \Big(\int_{T(B(0,R))} |Ef_{\omega_n}(x)|^q \, dx \Big)^{1/q} d\omega_n, \qquad (3.27)\end{aligned}$$

where we have applied a change of variables on the second line, again using the fact that $1+R^{-1} \sim 1$. Now, it is clear from the definition of T that $T(B(0,R)) \subset B(0, R(1+R^{-1})^2) \subset B(0,4R)$. Hence, the hypothesis $R^*_{P^{n-1}}(p \to q\,;\,\varepsilon)$ gives

$$\begin{aligned}\Big(\int_{T(B(0,R))} |Ef_{\omega_n}(x)|^q \, dx \Big)^{1/q} &\leq \|Ef_{\omega_n}\|_{L^q(B(0,4R))} \\ &\lesssim R^\varepsilon \|f_{\omega_n}\|_{L^p(P^{n-1},\, d\sigma)}.\end{aligned}$$

Substituting this into equation (3.27) and using Hölder's inequality, we get

$$\begin{aligned}\|g\|_{L^q(B(0,R))} &\lesssim R^\varepsilon \int_{-CR^{-1}}^{CR^{-1}} \|f_{\omega_n}\|_{L^p(P^{n-1},\, d\sigma)} \, d\omega_n \\ &\lesssim R^{\varepsilon - 1/p'} \Big(\int_{-CR^{-1}}^{CR^{-1}} \|f_{\omega_n}\|^p_{L^p(P^{n-1},\, d\sigma)} \, d\omega_n \Big)^{1/p}. \qquad (3.28)\end{aligned}$$

Now, we have

$$\|f_{\omega_n}\|^p_{L^p(P^{n-1},\, d\sigma)} \sim \int_{[-1,1]^{n-1}} |\hat{g}(\Phi(\omega', \omega_n))|^p \, d\omega',$$

hence,

$$\begin{aligned}\Big(\int_{-CR^{-1}}^{CR^{-1}} \|f_{\omega_n}\|^p_{L^p(P^{n-1},\, d\sigma)} \, d\omega_n \Big)^{1/p} &\sim \Big(\int_{-CR^{-1}}^{CR^{-1}} \Big(\int_{[-1,1]^{n-1}} |\hat{g}(\Phi(\omega', \omega_n))|^p \, d\omega' \Big) d\omega_n \Big)^{1/p} \\ &\sim \Big(\int_{\tilde{\mathcal{N}}_{R^{-1}}} |\hat{g}(\xi)|^p \, d\xi \Big)^{1/p} = \|\hat{g}\|_{L^p(\mathbb{R}^n)},\end{aligned}$$

where the last line follows by undoing the change of variables which led to equation (3.26). Inserting this into equation (3.28) gives the result. □

Chapter 4

Fourier Decoupling

The *local smoothing conjecture for the wave equation*, first stated in [Sog91], is a well-known conjecture from PDE which seeks to bound the $L^p(\mathbb{R}^n \times [1,2])$ norm of solutions to the wave equation in $n+1$ dimensions by Sobolev norms of the initial data. It is well-known that the conjecture would follow from a certain *square function estimate* of the form

$$\Big\| \sum_{j \in \mathcal{I}} f_j \Big\|_{L^p(\mathbb{R}^{n+1})} \lesssim_\varepsilon |\mathcal{I}|^\varepsilon \Big\| \Big(\sum_{j \in \mathcal{I}} |f_j|^2 \Big)^{1/2} \Big\|_{L^p(\mathbb{R}^{n+1})}, \tag{4.1}$$

where $\varepsilon > 0$. When working on the local smoothing conjecture, Wolff [Wol00] was compelled to consider a weaker form of the estimate (4.1) in which the roles of the ℓ^2 norm and L^p norm on the right-hand side were interchanged:

$$\Big\| \sum_{j \in \mathcal{I}} f_j \Big\|_{L^p(\mathbb{R}^{n+1})} \lesssim_\varepsilon |\mathcal{I}|^\varepsilon \Big(\sum_{j \in \mathcal{I}} \|f_j\|^2_{L^p(\mathbb{R}^{n+1})} \Big)^{1/2} \tag{4.2}$$

(in fact, Wolff considered a slightly different estimate with an ℓ^p norm on the right-hand side). Initially named in honor of Wolff, estimates of the form (4.2) are now known as *decoupling inequalities.*

The study of decoupling inequalities has recently flourished, and the implications have been far-reaching. In this chapter, we give a brief overview of some of the developments, beginning with some motivation and elementary examples. We then state some of the landmark results, namely decoupling for the paraboloid [BD15] and decoupling for the moment curve [BDG16], before seeing how decoupling for the moment curve was used to resolve a long-standing conjecture from number theory known as the *Vinogradov main conjecture.* Finally, we demonstrate a proof of the model two-dimensional case of decoupling for the paraboloid.

4.1 Motivation and Simple Examples

Let $(X, \mathcal{X}, \mu)$ be a measure space, and let $1 \leq p \leq \infty$. Given a finite collection of functions $\{f_j\}_{j\in\mathcal{I}} \subset L^p(X)$, the triangle inequality followed by the Cauchy-Schwarz inequality gives the bound

$$\Big\|\sum_{j\in\mathcal{I}} f_j\Big\|_{L^p(X)} \leq |\mathcal{I}|^{1/2}\Big(\sum_{j\in\mathcal{I}} \|f_j\|^2_{L^p(X)}\Big)^{1/2}. \tag{4.3}$$

However, it is clear that under some circumstances, this trivial bound can be improved upon.

Example 4.1. Consider the case when the measure space is $\mathbb{R}^n$ with the Lebesgue measure, $p = 2$, and the functions f_j are pairwise orthogonal. Then, equation (4.3) holds without the factor of $|\mathcal{I}|^{1/2}$, and we say that the functions f_j exhibit *square-root cancellation*. Thus, by Plancherel's theorem, the functions f_j having pairwise disjoint Fourier supports is sufficient to guarantee square-root cancellation

Example 4.2. Adhering to the case $p = 2$, a similar setting in which it is possible to improve upon equation (4.3) is when the Fourier supports of f_j, rather than being disjoint, have bounded overlap. That is to say, when there exists some $C > 0$ such that every $\xi \in \mathbb{R}^n$ lies in at most C of the sets $\operatorname{supp} f_j$. In this case, the Cauchy-Schwarz inequality implies the pointwise bound

$$\Big|\sum_{j\in\mathcal{I}} \hat{f}_j\Big|^2 \leq C\sum_{j\in\mathcal{I}} |\hat{f}_j|^2.$$

Upon integrating and taking square roots, we have

$$\Big\|\sum_{j\in\mathcal{I}} f_j\Big\|_{L^2(\mathbb{R}^n)} \leq C^{1/2}\Big(\sum_{j\in\mathcal{I}} \|f_j\|^2_{L^2(\mathbb{R}^n)}\Big)^{1/2},$$

which is an improvement upon (4.3) as long as $C < |\mathcal{I}|$.

Examples 4.1 and 4.2 are specific case of a more general type of inequality known as a *decoupling inequality.* Informally, we say that a collection of functions $\{f_j\}_{j\in\mathcal{I}}$ *exhibits ℓ^2 decoupling in L^p* if

$$\Big\|\sum_{j\in\mathcal{I}} f_j\Big\|_{L^p(X)} \lesssim_\varepsilon |\mathcal{I}|^\varepsilon\Big(\sum_{j\in\mathcal{I}} \|f_j\|^2_{L^p(X)}\Big)^{1/2} \tag{4.4}$$

for all $\varepsilon > 0$. The ℓ^2 refers to the occurrence of the ℓ^2 norm of the sequence $(\|f_j\|_{L^p(X)})_{j\in\mathcal{I}}$ on the right-hand side of equation (4.4), and we could similarly consider analogous ℓ^q decoupling inequalities for other exponents q (though we will have no need to do so). The name "decoupling" itself is due to the fact that the functions f_j can be thought of as being "coupled" on the left-hand side of (4.4) into the sum $\sum_{j\in\mathcal{I}} f_j$, but they have been "decoupled" on the right-hand side into individual components. This terminology is most clearly motivated by examples in which the f_j are complex exponentials of different frequencies. Such functions are so common in this field that we are compelled to give them a more convenient notation:

Notation 4.3. We define $e(x) = e^{2\pi i x}$ for all $x \in \mathbb{R}$.

One typical example is the following (given, for example, in [Pie19]):

Example 4.4. For each $N \in \mathbb{N}$ and $\varepsilon > 0$, we have the decoupling inequality

$$\Big\|\sum_{j=1}^{N} e(j^2x)\Big\|_{L^4([0,1])} \lesssim_\varepsilon N^\varepsilon\Big(\sum_{j=1}^{N}\|e(j^2x)\|^2_{L^4([0,1])}\Big)^{1/2}.$$

Indeed, the right-hand side is equal to $N^{\varepsilon+1/2}$, and for the left-hand side, we see that

$$\begin{aligned}\Big\|\sum_{j=1}^{N} e(j^2x)\Big\|^4_{L^4([0,1])} &= \int_0^1 \Big|\sum_{j=1}^{N} e(j^2x)\Big|^4\,dx\\ &= \sum_{1\le j_1,j_2,j_3,j_4\le N}\int_0^1 e((j_1^2+j_2^2-j_3^2-j_4^2)x)\,dx.\end{aligned}$$

Now, recall that an integral of the form $\int_0^1 e(kx)\,dx$ for $k \in \mathbb{Z}$ is equal to 1 if $k = 0$ and is equal to 0 otherwise. It follows that

$$\Big\|\sum_{j=1}^{N} e(j^2x)\Big\|^4_{L^4([0,1])} = \#\{1 \le j_1, j_2, j_3, j_4 \le N : j_1^2 + j_2^2 = j_3^2 + j_4^2\},$$

and number-theoretic considerations show that the right-hand side is $\lesssim_\varepsilon N^{2+\varepsilon}$ for all $\varepsilon > 0$; indeed, there are N^2 ways to choose j_2 and j_4, and for each choice, there are $\lesssim_\varepsilon N^\varepsilon$ solutions to the diophantine equation $j_1^2 - j_3^2 = j_4^2 - j_2^2$ in the variables j_1 and j_3 [Tao08]. Taking fourth roots gives the result.

This example is a good illustration of an important technique connecting number theory and analysis. This technique exploits the simple identity

$$\int_{[0,1]^n} e(k \cdot x)\, dx = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k \in \mathbb{Z}^n \setminus \{0\} \end{cases}$$

to see that estimating $L^p([0,1]^n)$ norms of exponential sums (when p is an even integer) is equivalent to estimating the number of solutions to a corresponding system of n diophantine equations. In Example 4.4, we saw that an *a priori* estimate on the number of solutions to the diophantine equation $j_1^2 + j_2^2 = j_3^2 + j_4^2$ for $1 \leq j_1, \cdots, j_4 \leq N$ led to an ℓ^2 decoupling inequality in $L^4([0,1])$ for the complex exponentials $\{e(j^2 x)\}_{j\in\mathbb{N}}$. Of more interest is the converse situation, in which an *a priori* decoupling inequality provides the means to estimate the number of solutions to a diophantine system. We will see later a powerful example of how a decoupling inequality was used to prove a long-standing conjecture of Vinogradov on the number of solutions to diophantine systems of a particular form.

4.2 Some Landmark Results

One situation in which decoupling has been found to hold is when a collection of Schwartz functions $\{f_j\}_{j\in\mathcal{I}}$ have Fourier supports in regions of moderate separation adapted to a submanifold of appropriate curvature. Two notable examples of appropriate submanifolds are the paraboloid and the moment curve. We take this opportunity to state these landmark results.

4.2.1 The Paraboloid

For each $0 < \delta \leq 1$ and each $\xi \in \mathbb{R}^{n-1}$, let $P_{\xi,\delta} \subset \mathbb{R}^n$ denote the region

$$P_{\xi,\delta} = \{(\xi', \xi_n) : \xi' \in \xi + (-\delta, \delta)^{n-1} \, ; \, |\xi_n - |\xi|^2 - 2\xi \cdot (\xi' - \xi)| < \delta^2\}.$$

This can be thought of as the tangent hyperplane to the paraboloid at $(\xi, |\xi|^2)$, thickened to scale δ^2 in the ξ_n direction and lying above a cube of scale δ about ξ. To aid with the visualisation of the regions $P_{\xi,\delta}$, the following diagram depicts one such region in the case $n = 2$:

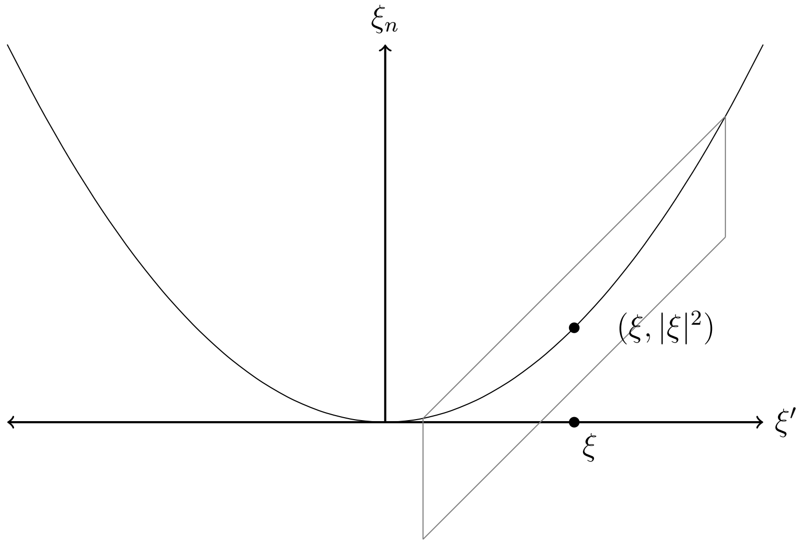

Figure 4.1: One of the regions $P_{\xi,\delta}$ visualised when $n = 2$, in which case they are narrow parallelograms "tangent" to the parabola.

The landmark result for the paraboloid, due to Bourgain and Demeter, is as follows:

Theorem 4.5 (Decoupling for the Paraboloid: Theorem 1.1, [BD15])**.** *Let* $0 < \delta \leq 1$*, and let* $\Sigma \subset [-1,1]^{n-1}$ *be* δ*-separated. If for each* $\xi \in \Sigma$*,* $f_\xi \in \mathcal{S}(\mathbb{R}^n)$ *has Fourier support in* $P_{\xi,\delta}$*, then*

$$\Big\| \sum_{\xi \in \Sigma} f_\xi \Big\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \lesssim_\varepsilon \delta^{-\varepsilon} \Big(\sum_{\xi \in \Sigma} \|f_\xi\|^2_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \Big)^{1/2}$$

for all $\varepsilon > 0$.

Remark 4.6. The decoupling exponent $\frac{2(n+1)}{n-1}$ afforded to us by Theorem 4.5 is the same as the Tomas-Stein exponent. This is the primary fact we will exploit to give a new proof of the Tomas restriction theorem based on decoupling in Chapter 5.

We note that the decoupling theorem for the paraboloid remains true if we replace the regions $P_{\xi,\delta}$ by the vertically scaled regions

$$\tilde{P}_{\xi,\delta} = \{(\xi', \xi_n) : \xi' \in \xi + (-\delta, \delta)^{n-1} \, ; \, |\xi_n - |\xi|^2 - 2\xi \cdot (\xi' - \xi)| < C^2\delta^2\}$$

for some fixed $C \geq 1$. Indeed, let $\Sigma \subset [-1,1]^{n-1}$ be δ-separated, and let $f_\xi \in \mathcal{S}(\mathbb{R}^n)$ have Fourier support in $\tilde{P}_{\xi,\delta}$ for all $\xi \in \Sigma$. Observe that $\tilde{P}_{\xi,\delta} \subset P_{\xi,C\delta}$ for all $\xi \in \Sigma$, so each f_ξ has Fourier support in $P_{\xi,C\delta}$; moreover, we may partition Σ into $O(1)$ $C\delta$-separated subsets $\Sigma_1, \cdots, \Sigma_m$. The triangle inequality followed by the decoupling theorem therefore gives

$$\Big\| \sum_{\xi\in\Sigma} f_\xi \Big\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \leq \sum_{j=1}^{m} \Big\| \sum_{\xi\in\Sigma_j} f_\xi \Big\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)}$$
$$\lesssim_\varepsilon \delta^{-\varepsilon} \sum_{j=1}^{m} \Big(\sum_{\xi\in\Sigma_j} \|f_\xi\|^2_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \Big)^{1/2}$$

for all $\varepsilon > 0$. Applying the Cauchy-Schwarz inequality on the last line and absorbing $m^{1/2} = O(1)$ into the implied constant gives

$$\Big\| \sum_{\xi\in\Sigma} f_\xi \Big\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \lesssim_\varepsilon \delta^{-\varepsilon} \Big(\sum_{\xi\in\Sigma} \|f_\xi\|^2_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \Big)^{1/2}.$$

4.2.2 The Moment Curve

For each $n \geq 2$, define $\gamma_n : [0,1] \to \mathbb{R}^n$ by $\gamma_n(t) = (t, t^2, \cdots, t^n)$.

Definition 4.7. The *moment curve* in $\mathbb{R}^n$ is the curve $\Gamma^n = \gamma_n([0,1])$. That is, $\Gamma^n := \{(t, t^2, \cdots, t^n) : t \in [0,1]\}$.

For each $0 < \delta \leq 1$ and each $\xi \in \mathbb{R}$, let $\theta_{\xi,\delta}$ denote the region

$$\theta_{\xi,\delta} = \{\gamma_n(\xi) + s_1\gamma_n'(\xi) + \cdots + s_n\gamma_n^{(n)}(\xi) : |s_k| < \delta^k \text{ for all } k = 1, \cdots, n\},$$

where $\gamma_n^{(k)}$ denotes the k^{th} derivative of γ_n. Thus, $\theta_{\xi,\delta}$ is a box of dimensions $\sim \delta \times \delta^2 \times \cdots \times \delta^n$ centred at $\gamma_n(\xi)$, with sides running parallel to $\gamma_n'(\xi), \gamma_n''(\xi), \cdots, \gamma_n^{(n)}(\xi)$ respectively.
The landmark result for the moment curve, due to Bourgain, Demeter, and Guth, is as follows:

Theorem 4.8 (Decoupling for the Moment Curve: Theorem 1.2, [BDG16]; Theorem 1.2, [GLY21])**.** *Let $0 < \delta \leq 1$, and let $\Sigma \subset [0,1]$ be δ-separated. If for each $\xi \in \Sigma$, $f_\xi \in \mathcal{S}(\mathbb{R}^n)$ has Fourier support in $\theta_{\xi,\delta}$, then*

$$\Big\| \sum_{\xi\in\Sigma} f_\xi \Big\|_{L^{n(n+1)}(\mathbb{R}^n)} \lesssim_\varepsilon \delta^{-\varepsilon} \Big(\sum_{\xi\in\Sigma} \|f_\xi\|^2_{L^{n(n+1)}(\mathbb{R}^n)} \Big)^{1/2}$$

for all $\varepsilon > 0$.

Theorem 4.8 was famously used by Bourgain, Demeter and Guth [BDG16] to prove a long-standing conjecture in number theory known as the Vinogradov main conjecture; we explore this application in the following section.

4.3 The Vinogradov Main Conjecture

Given integers $s, n, N \geq 1$, let $J_{s,n}(N)$ denote the number of integral solutions to the system of diophantine equations

$$x_1^j + \cdots + x_s^j = x_{s+1}^j + \cdots + x_{2s}^j \quad \text{for all } 1 \leq j \leq n, \tag{4.5}$$

with $1 \leq x_i \leq N$ for all $i = 1, \cdots, 2s$. There are N^s choices for $x_1, \cdots, x_s$, so by taking $x_{s+k} = x_k$ for all $1 \leq k \leq s$, we see that there are at least N^s solutions to the system (4.5) (these are the so-called "diagonal solutions"). On the other hand, there are N^{2s} choices for $x_1, \cdots, x_{2s}$, and given one choice, one may heuristically consider $x_1^j + \cdots + x_s^j$ and $x_{s+1}^j + \cdots + x_{2s}^j$ as random integers in the range $[1, sN^j]$. The probability that these are equal are equal would then be $1/sN^j$, so the probability that we have equality for each $j = 1, \cdots, n$ would be $\prod_{j=1}^n 1/sN^j \sim_{s,n} N^{-\frac{n(n+1)}{2}}$. This heuristic, plus the existence of the diagonal solutions, leads to the following conjecture:

Conjecture 4.9 (The Vinogradov Main Conjecture)**.** *For all $s, n \geq 1$ and all $\varepsilon > 0$, we have*

$$J_{s,n}(N) \lesssim_{s,n,\varepsilon} N^{\varepsilon}(N^s + N^{2s-\frac{n(n+1)}{2}}).$$

It is not hard to see that

$$J_{s,n}(N) = \int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{2s} dx,$$

so the conjecture is equivalent to the statement

$$\int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{2s} dx \lesssim_{s,n,\varepsilon} N^{\varepsilon}(N^s + N^{2s-\frac{n(n+1)}{2}}) \tag{4.6}$$

for all $\varepsilon > 0$. Heuristically, the Fourier transform of $e(\gamma_n(j) \cdot x)$ should be supported at the point $\gamma_n(j)$ (an idea which can be made precise by the theory of distributions), from which it is evident that the decoupling theorem for the moment curve might be applicable. This is indeed the case, as was first shown in [BDG16]. We recount a variant of the argument here, inspired by an exposition of Tao [Tao15]. Tao cites a different version of the decoupling theorem for the moment curve, so some extra steps must be taken to make Theorem 4.8 applicable.

Proof (The Vinogradov Main Conjecture): We begin with some standard reductions. The first is that it suffices to prove (4.6) for the decoupling exponent $2s = n(n+1)$. Indeed, suppose it is known that

$$\int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{n(n+1)} dx \lesssim_{n,\varepsilon} N^{\frac{n(n+1)}{2}+\varepsilon}. \tag{4.7}$$

Then, given any $s > \frac{n(n+1)}{2}$, we may write $2s = n(n+1) + \eta$ and use the triangle inequality to obtain

$$\begin{aligned}\int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{2s} dx &\le N^{\eta} \int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{n(n+1)} dx \\ &\lesssim_{n,\varepsilon} N^{\frac{n(n+1)}{2}+\eta+\varepsilon} = N^{2s-\frac{n(n+1)}{2}+\varepsilon}.\end{aligned}$$

Hence, (4.6) holds in this case. On the other hand, if $s < \frac{n(n+1)}{2}$, we apply Hölder's inequality with exponent $p = \frac{n(n+1)}{2s} > 1$ to obtain

$$\begin{aligned}\int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{2s} dx &\le \Big(\int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{n(n+1)} dx \Big)^{\frac{2s}{n(n+1)}} \\ &\lesssim_{n,\varepsilon} N^{s+\varepsilon},\end{aligned}$$

so that (4.6) also holds.
To prove (4.7) using Theorem 4.8, we must first change variables so that the frequencies lie on the arc of the moment curve corresponding to the unit interval. In particular, letting $T_N : \mathbb{R}^n \to \mathbb{R}^n$ be the map $T_N(x_1, \cdots, x_n) = (Nx_1, \cdots, N^n x_n)$, we see that

$$\begin{aligned}\int_{[0,1]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot x) \Big|^{n(n+1)} dx &= N^{-\frac{n(n+1)}{2}} \int_{[0,N]\times\cdots\times[0,N^n]} \Big| \sum_{j=1}^{N} e(\gamma_n(j) \cdot T_N^{-1} x) \Big|^{n(n+1)} dx \\ &= N^{-\frac{n(n+1)}{2}} \int_{[0,N]\times\cdots\times[0,N^n]} \Big| \sum_{j=1}^{N} e(\gamma_n(j/N) \cdot x) \Big|^{n(n+1)} dx.\end{aligned}$$

We now note that $[0, N^n]^n$ can be written as the union of $N^{\frac{n(n-1)}{2}}$ translates of $[0, N]\times\cdots\times[0, N^n]$ by elements of $N\mathbb{Z}\times\cdots\times N^n\mathbb{Z}$, and by periodicity, the integral over such a translate is unchanged. The right-hand side is therefore equal to

$$N^{-n^2} \int_{[0,N^n]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j/N) \cdot x) \Big|^{n(n+1)} dx,$$

from which we see it suffices to prove

$$\int_{[0,N^n]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j/N)\cdot x) \Big|^{n(n+1)} dx \lesssim_{n,\varepsilon} N^{n^2+\frac{n(n+1)}{2}+\varepsilon}.$$

Theorem 4.8 is now almost applicable. The last step is to replace the integral over $[0,N^n]^n$ by an integral over $\mathbb{R}^n$, and to do so, we choose by Proposition 2.13 some $\psi \in \mathcal{S}(\mathbb{R}^n)$ such that $\psi \sim 1$ on $[0,1]^n$, with $\hat{\psi}$ supported in $B(0,1)$. Then, given any $N \in \mathbb{N}$, we have

$$\int_{[0,N^n]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j/N)\cdot x) \Big|^{n(n+1)} dx \lesssim_n \int_{\mathbb{R}^n} \Big| \sum_{j=1}^{N} \psi(x/N^n) e(\gamma_n(j/N)\cdot x) \Big|^{n(n+1)} dx.$$

But $\psi(x/N^n)$ has Fourier support in $B(0,N^{-n})$, from which it follows that the Fourier support of $\psi(x/N^n)e(\gamma_n(j/N)\cdot x)$ is contained in $B(\gamma_n(j/N),N^{-n}) \subset \theta_{j/N,N^{-1}}$ (by similar computations which led to equation (3.23)). Theorem 4.8 therefore applies, giving

$$\int_{[0,N^n]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j/N)\cdot x) \Big|^{n(n+1)} dx \lesssim_{n,\varepsilon} N^{\varepsilon} \Big(\sum_{j=1}^{N} \|\psi(x/N^n)\|^2_{L^{n(n+1)}(\mathbb{R}^n)} \Big)^{\frac{n(n+1)}{2}}.$$

But a change of variables gives

$$\|\psi(x/N^n)\|^{n(n+1)}_{L^{n(n+1)}(\mathbb{R}^n)} = N^{n^2} \|\psi\|^{n(n+1)}_{L^{n(n+1)}(\mathbb{R}^n)},$$

hence $\|\psi(x/N^n)\|^2_{L^{n(n+1)}(\mathbb{R}^n)} \sim N^{\frac{2n}{n+1}}$. It follows that

$$\begin{aligned} \int_{[0,N^n]^n} \Big| \sum_{j=1}^{N} e(\gamma_n(j/N)\cdot x) \Big|^{n(n+1)} dx &\lesssim_{n,\varepsilon} N^{\varepsilon} \Big(\sum_{j=1}^{N} N^{\frac{2n}{n+1}} \Big)^{\frac{n(n+1)}{2}} \\ &= N^{n^2+\frac{n(n+1)}{2}+\varepsilon}. \end{aligned}$$

□

4.4 The Parabola: A Model Case

We have now seen one example of the power of decoupling theorems, but we are yet to provide any insight into how they may be proven. We therefore dedicate this section to proving the model $n = 2$ case of Theorem 4.5.

We will closely follow an argument of Tao [Tao20a], which is in turn based on an article of Li [Li21]. We streamline Tao's argument by rescaling two of the parameters in the definition of the bilinear decoupling constant, with the effect

of rendering Lemma 14 of [Tao20a] superfluous. We also opt for our final proof to more closely resemble the argument of Li than that of Tao.
Recall that when $n = 2$, given $0 < \delta \leq 1$ and $\xi \in \mathbb{R}$, we let $P_{\xi,\delta}$ denote the parallelogram

$$P_{\xi,\delta} = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1 - \xi| < \delta\,;\, |\xi_2 - \xi^2 - 2\xi(\xi_1 - \xi)| < \delta^2\}.$$

To reiterate, the theorem is as follows:

Theorem 4.10 (Decoupling for the Parabola)**.** *Let $0 < \delta \leq 1$, and let $\Sigma \subset [-1, 1]$ be δ-separated. If for each $\xi \in \Sigma$, $f_\xi \in \mathcal{S}(\mathbb{R}^2)$ has Fourier support in $P_{\xi,\delta}$, then*

$$\Big\| \sum_{\xi \in \Sigma} f_\xi \Big\|_{L^6(\mathbb{R}^2)} \lesssim_\varepsilon \delta^{-\varepsilon} \Big(\sum_{\xi \in \Sigma} \|f_\xi\|^2_{L^6(\mathbb{R}^2)} \Big)^{1/2}$$

for all $\varepsilon > 0$.

For the sake of convenience, we introduce the notion of a *decoupling constant*:

Definition 4.11. Let $\mathcal{U} = \{U_1, \cdots, U_m\}$ be a finite multiset of nonempty open sets $U_i \subset \mathbb{R}^n$, and let $1 \leq p \leq \infty$. Then, the L^p *decoupling constant* $\mathrm{Dec}_p(\mathcal{U})$ associated to $\mathcal{U}$ and p is the smallest constant for which

$$\Big\| \sum_{i=1}^m f_i \Big\|_{L^p(\mathbb{R}^n)} \leq \mathrm{Dec}_p(\mathcal{U}) \Big(\sum_{i=1}^m \|f_i\|^2_{L^p(\mathbb{R}^n)} \Big)^{1/2}$$

whenever $f_i \in \mathcal{S}(\mathbb{R}^n)$ has Fourier support in U_i for each i.

Thus, letting $D(\delta)$ denote the supremum of the L^6 decoupling constants $\mathrm{Dec}_6(\{P_{\xi,\delta} : \xi \in \Sigma\})$ across all δ-separated subsets $\Sigma \subset [-1, 1]$, Theorem 4.10 is equivalent to the statement that $D(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}$ for all $\varepsilon > 0$.
The following property of decoupling constants will be useful throughout the arguments which follow:

Proposition 4.12 (Affine Invariance)**.** *Let $1 \leq p \leq \infty$. Then*

$$\mathrm{Dec}_p(\{LU_1, \cdots, LU_m\}) = \mathrm{Dec}_p(\{U_1, \cdots, U_m\})$$

whenever the $U_i \subset \mathbb{R}^n$ are nonempty and open, and $L : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible affine transformation.

Proof. Since any invertible affine transformation L is a composition of an invertible linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ and a translation $x \mapsto x+\tau$, it suffices to prove the proposition for such maps.
Let $f_i \in \mathcal{S}(\mathbb{R}^n)$ have Fourier support in U_i for each i. Given an invertible linear map $T : \mathbb{R}^n \to \mathbb{R}^n$, Proposition 2.2 implies that $(f_i \circ T^t)\widehat{\ } = |\det T^{-t}|\hat{f}_i \circ T^{-1}$ is supported in TU_i. It follows by a change of variables that

$$\begin{aligned}\Big\| \sum_{i=1}^m f_i \Big\|_{L^p(\mathbb{R}^n)} &= |\det T^t|^{1/p} \Big\| \sum_{i=1}^m f_i \circ T^t \Big\|_{L^p(\mathbb{R}^n)} \\ &\leq |\det T^t|^{1/p}\,\mathrm{Dec}_p(\{TU_1, \cdots, TU_m\}) \Big(\sum_{i=1}^m \|f_i \circ T^t\|_{L^p(\mathbb{R}^n)}^2 \Big)^{1/2} \\ &= \mathrm{Dec}_p(\{TU_1, \cdots, TU_m\}) \Big(\sum_{i=1}^m \|f_i\|_{L^p(\mathbb{R}^n)}^2 \Big)^{1/2},\end{aligned}$$

hence $\mathrm{Dec}_p(\{U_1, \cdots, U_m\}) \leq \mathrm{Dec}_p(\{TU_1, \cdots, TU_m\})$. Applying this with TU_i in place of U_i and T^{-1} in place of T gives the reverse inequality.
Given a translation defined by $x \mapsto x+\tau$, Proposition 2.2 gives that $(e^{2\pi i \tau \cdot x} f_i)\widehat{\ } = (\widehat{f_i})_\tau$ is supported in $U_i + \tau$. A similar computation as above shows that we have $\mathrm{Dec}_p(\{U_1, \cdots, U_m\}) \leq \mathrm{Dec}_p(\{U_1+\tau, \cdots, U_m+\tau\})$, and applying this with $U_i+\tau$ in place of U_i and $-\tau$ in place of τ gives the reverse inequality. □

The first important property we must prove specific to the parabola is the following rescaling lemma:

Lemma 4.13 (Parabolic Rescaling)**.** *Let $0 < \delta \leq \delta_0 \leq 1$, and let Σ be a δ-seperated subset of an interval $I \subset [-1,1]$ of length $2\delta_0$. Then,*

$$\mathrm{Dec}_6(\{P_{\xi,\delta} : \xi \in \Sigma\}) \leq D(\delta/\delta_0).$$

Proof. Write $I = [\xi_0 - \delta_0, \xi_0 + \delta_0]$. It is routine to check that the invertible affine transformation $G_{\xi_0} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$G_{\xi_0}(\xi_1, \xi_2) = (\xi_1 - \xi_0, \xi_2 - 2\xi_0\xi_1 + \xi_0^2)$$

maps the parabola to the parabola, and parallelograms $P_{\xi,\delta}$ to parallelograms $P_{\xi-\xi_0,\delta}$. Since G_{ξ_0} maps I to $[-\delta_0, \delta_0]$ and preserves the separation between each of the $\xi \in \Sigma$, we may assume without loss of generality by affine invariance that $I = [-\delta_0, \delta_0]$.
It is also routine to check that the invertible linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T(\xi_1, \xi_2) = (\xi_1/\delta_0, \xi_2/\delta_0^2)$$

maps the parabola to the parabola, and parallelograms $P_{\xi,\delta}$ to parallelograms $P_{\xi/\delta_0,\delta/\delta_0}$. Since the set $\Sigma/\delta_0 \subset [-1,1]$ is δ/δ_0-separated, affine invariance again gives

$$\mathrm{Dec}_6(\{P_{\xi,\delta} : \xi \in \Sigma\}) = \mathrm{Dec}_6(\{P_{\xi/\delta_0,\delta/\delta_0} : \xi \in \Sigma\}) \leq D(\delta/\delta_0),$$

where the last inequality is by the definition of $D(\delta/\delta_0)$. □

Using parabolic rescaling, it is not too hard to show that $D(\delta_1\delta_2) \lesssim D(\delta_1)D(\delta_2)$ for all $0 < \delta_1, \delta_2 \leq 1$. This property suggests that an "induction on scales" approach could be used to prove $D(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}$; indeed, starting with $D(1/2) \lesssim_\varepsilon 2^\varepsilon$, we could iterate $D(\delta_1\delta_2) \lesssim D(\delta_1)D(\delta_2)$ to obtain $D(2^{-k}) \lesssim_{k,\varepsilon} (2^k)^\varepsilon$ for all positive integers k. The problem with this approach is that the implied constant depends on k, and grows without bound as $k \to \infty$. To proceed with an induction on scales approach, a more complicated bilinear variant of $D(\delta)$ must be introduced.

Definition 4.14 (Bilinear Decoupling Constant)**.** Let $0 < \delta \leq 2\rho_1, 2\rho_2 \leq \nu \leq 1$. Define $M_{2,4}(\delta, \nu, \rho_1, \rho_2)$ to be the infimum over all constants C such that

$$\int_{\mathbb{R}^2} \Big| \sum_{\xi_1 \in \Sigma_1} f_{\xi_1} \Big|^2 \Big| \sum_{\xi_2 \in \Sigma_2} g_{\xi_2} \Big|^4 dx \leq C^6 \Big(\sum_{\xi_1 \in \Sigma_1} \|f_{\xi_1}\|^2_{L^6(\mathbb{R}^2)} \Big) \Big(\sum_{\xi_2 \in \Sigma_2} \|g_{\xi_2}\|^2_{L^6(\mathbb{R}^2)} \Big)^2$$

whenever $\Sigma_1, \Sigma_2 \subset [-1,1]$ are δ-separated subsets of intervals $I_1, I_2 \subset [-1,1]$ of length $2\rho_1$ and $2\rho_2$ respectively, with $\mathrm{dist}(I_1, I_2) \geq \nu$, and $f_{\xi_1} \in \mathcal{S}(\mathbb{R}^2)$ has Fourier support in $P_{\xi_1,\delta}$ for each $\xi_1 \in \Sigma_1$ and $g_{\xi_2} \in \mathcal{S}(\mathbb{R}^2)$ has Fourier support in $P_{\xi_2,\delta}$ for each $\xi_2 \in \Sigma_2$.

Since our ultimate goal is to prove a bound for $D(\delta)$, we would naturally like to estimate $D(\delta)$ by bilinear decoupling constants. This is enabled by the following result:

Lemma 4.15 (Bilinear Reduction)**.** *Let* $0 < \delta \leq \nu \leq 1/2$. *Then,*

$$D(\delta) \lesssim \nu^{-O(1)} M_{2,4}(\delta, 2\nu, \nu, \nu) + D(\delta/\nu).$$

Remark 4.16. When used in this form, $O(1)$ simply refers to a finite constant that is independent of all parameters. Recall that when using this notation, we allow the constant to change each line.

Proof. Let $\Sigma \subset [-1,1]$ be δ-separated, and for each $\xi \in \Sigma$, let $f_\xi \in \mathcal{S}(\mathbb{R}^2)$ have Fourier support in $P_{\xi,\delta}$. We wish to prove that

$$\Big\| \sum_{\xi \in \Sigma} f_\xi \Big\|_{L^6(\mathbb{R}^2)} \lesssim (\nu^{-O(1)} M_{2,4}(\delta, 2\nu, \nu, \nu) + D(\delta/\nu)) \Big(\sum_{\xi \in \Sigma} \|f_\xi\|^2_{L^6(\mathbb{R}^2)} \Big)^{1/2}.$$

By scale invariance, we may normalise so that $\sum_{\xi\in\Sigma}\|f_\xi\|^2_{L^6(\mathbb{R}^2)} = 1$. Raising both sides to the power of 6 and observing that $(\nu^{-O(1)}M_{2,4}(\delta,\nu,\nu,\nu) + D(\delta/\nu))^6 \sim \nu^{-O(1)}M_{2,4}(\delta,\nu,\nu,\nu)^6 + D(\delta/\nu)^6$, we see that it suffices to prove

$$\int_{\mathbb{R}^2}\Big|\sum_{\xi\in\Sigma} f_\xi\Big|^6\,dx \lesssim \nu^{-O(1)}M_{2,4}(\delta,\nu,\nu,\nu)^6 + D(\delta/\nu)^6. \tag{4.8}$$

Let $\mathcal{I}$ denote the collection of $O(\nu^{-1})$ intervals

$$\mathcal{I} = \{[-1+2(k-1)\nu, -1+2k\nu] : k = 1,\cdots,\lfloor \nu^{-1}\rfloor\} \cup \{[1-2\nu, 1]\} \tag{4.9}$$

of length 2ν covering $[-1,1]$, and partition Σ into subsets Σ_I such that $\Sigma_I \subset I$ for each $I\in\mathcal{I}$. Then, by our normalisation and the definition of $M_{2,4}(\delta,2\nu,\nu,\nu)$, we have

$$\sum_{I,J\in\mathcal{I}:\mathrm{dist}(I,J)\geq 2\nu}\int_{\mathbb{R}^2}\Big|\sum_{\xi\in\Sigma_I} f_\xi\Big|^2\Big|\sum_{\xi\in\Sigma_J} f_\xi\Big|^4\,dx \lesssim \nu^{-O(1)}M_{2,4}(\delta,2\nu,\nu,\nu)^6. \tag{4.10}$$

Moreover, by parabolic rescaling, we also have

$$\begin{aligned}\sum_{I\in\mathcal{I}}\int_{\mathbb{R}^2}\Big|\sum_{\xi\in\Sigma_I} f_\xi\Big|^6\,dx &\leq D(\delta/\nu)^6\sum_{I\in\mathcal{I}}\Big(\sum_{\xi\in\Sigma_I}\|f_\xi\|^2_{L^6(\mathbb{R}^2)}\Big)^3\\ &\leq D(\delta/\nu)^6\Big(\sum_{I\in\mathcal{I}}\sum_{\xi\in\Sigma_I}\|f_\xi\|^2_{L^6(\mathbb{R}^2)}\Big)^3 = D(\delta/\nu)^6,\end{aligned} \tag{4.11}$$

where we have again used our normalisation. By (4.10) and (4.11), we see that to prove (4.8), it suffices to prove a pointwise bound of the form

$$\Big|\sum_{\xi\in\Sigma} f_\xi\Big|^6 \lesssim \nu^{-O(1)}\sum_{I,J\in\mathcal{I}\,:\,\mathrm{dist}(I,J)\geq 2\nu}\Big|\sum_{\xi\in\Sigma_I} f_\xi\Big|^2\Big|\sum_{\xi\in\Sigma_J} f_\xi\Big|^4 + \sum_{I\in\mathcal{I}}\Big|\sum_{\xi\in\Sigma_I} f_\xi\Big|^6. \tag{4.12}$$

To do so, let $A = |\sum_{\xi\in\Sigma} f_\xi|$ and $A_I = |\sum_{\xi\in\Sigma_I} f_\xi|$ for each $I\in\mathcal{I}$, and let $N = |\mathcal{I}| = O(\nu^{-1})$. The triangle inequality gives $A \leq \sum_{I\in\mathcal{I}} A_I$, and it is clear that $\sum_{I:A_I\leq A/2N} A_I \leq A/2$. It must therefore be the case that $\sum_{I:A_I>A/2N} A_I \geq A/2$. It follows that if $\#\{I : A_I > A/2N\} \leq 3$, then there must exist some $I\in\mathcal{I}$ such that $A_I \geq A/6$, in which case upon raising to the power of 6 and recalling the definition of A and A_I, we find that

$$\Big|\sum_{\xi\in\Sigma} f_\xi\Big|^6 \lesssim \Big|\sum_{\xi\in\Sigma_I} f_\xi\Big|^6 \tag{4.13}$$

holds for that interval I. If, on the other hand, $\#\{I : A_I > A/2N\} > 3$, then there must be at least 3 intervals I in the collection $\mathcal{I}\setminus\{[1-2\nu,1]\}$ satisfying

$A_I > A/2N$, at least two of which are separated by a distance of at least 2ν. It follows that in this case,

$$\Big|\sum_{\xi\in\Sigma} f_\xi\Big|^6 \lesssim \nu^{-O(1)} \sum_{I,J\in\mathcal{I}\,:\,\mathrm{dist}(I,J)\geq 2\nu} \Big|\sum_{\xi\in\Sigma_I} f_\xi\Big|^2 \Big|\sum_{\xi\in\Sigma_J} f_\xi\Big|^4. \tag{4.14}$$

Since either of the cases leading to (4.13) or (4.14) must hold, we may add the right-hand sides to obtain the pointwise bound (4.12), which completes the proof. □

Given $0 < \delta \leq 2\rho_1', 2\rho_1, 2\rho_2 \leq \nu \leq 1$ with $\rho_1' \leq \rho_1$, it is clear by the definition of the bilinear decoupling constant that $M_{2,4}(\delta,\nu,\rho_1',\rho_2) \leq M_{2,4}(\delta,\nu,\rho_1,\rho_2)$. The next lemma says that if ρ_1' is not too small relative to ρ_2, then $M_{2,4}(\delta,\nu,\rho_1',\rho_2)$ cannot be too much smaller than $M_{2,4}(\delta,\nu,\rho_1,\rho_2)$. This ability to bound bilinear decoupling constants by those with a smaller ρ_1 will be key in our proof of the decoupling theorem, in which we iteratively shrink ρ_1 in this manner.

Lemma 4.17 (Key Estimate)**.** *If* $0 < \delta \leq 2\rho_1', 2\rho_1, 2\rho_2 \leq \nu \leq 1$ *with* $\rho_2^2 \leq \rho_1' \leq \rho_1$*, then*

$$M_{2,4}(\delta,\nu,\rho_1,\rho_2) \lesssim \nu^{-O(1)} M_{2,4}(\delta,\nu,\rho_1',\rho_2).$$

Proof. Let $\Sigma_1, \Sigma_2 \subset [-1,1]$ be δ-separated subsets of intervals $I_1, I_2 \subset [-1,1]$ of length $2\rho_1$ and $2\rho_2$ respectively, with $\mathrm{dist}(I_1,I_2) \geq \nu$. Let $f_{\xi_1} \in \mathcal{S}(\mathbb{R}^2)$ have Fourier support in $P_{\xi_1,\delta}$ for each $\xi_1 \in \Sigma_1$, and let $g_{\xi_2} \in \mathcal{S}(\mathbb{R}^2)$ have Fourier support in $P_{\xi_2,\delta}$ for each $\xi_2 \in \Sigma_2$. By scale invariance, it suffices to prove

$$\int_{\mathbb{R}^2} \Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^2 \Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^4 dx \lesssim \nu^{-O(1)} M_{2,4}(\delta,\nu,\rho_1',\rho_2)^6$$

assuming the normalisation $\sum_{\xi_1\in\Sigma_1} \|f_{\xi_1}\|^2_{L^6(\mathbb{R}^2)} = \sum_{\xi_2\in\Sigma_2} \|g_{\xi_2}\|^2_{L^6(\mathbb{R}^2)} = 1$. To do so, we first cover I_1 by a collection $\mathcal{I}'$ of intervals of length $2\rho_1'$ analogously to the covering of $[-1,1]$ by the collection $\mathcal{I}$ in (4.9). Then, partition Σ_1 into subsets $\Sigma_{1,I'}$ such that $\Sigma_{1,I'} \subset I'$ for each $I' \in \mathcal{I}'$. We have $\mathrm{dist}(I',J) \geq \nu$ since $I' \subset I$ for each $I' \in \mathcal{I}'$, hence

$$\int_{\mathbb{R}^2} \Big|\sum_{\xi_1\in\Sigma_{1,I'}} f_{\xi_1}\Big|^2 \Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^4 dx \leq M_{2,4}(\delta,\nu,\rho_1',\rho_2)^6 \Big(\sum_{\xi_1\in\Sigma_{1,I'}} \|f_{\xi_1}\|^2_{L^6(\mathbb{R}^2)}\Big)$$

by the definition of the bilinear decoupling constant and using the normalisation $\sum_{\xi_2\in\Sigma_2} \|g_{\xi_2}\|^2_{L^6(\mathbb{R}^2)} = 1$. It therefore suffices to prove

$$\int_{\mathbb{R}^2} \Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^2 \Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^4 dx \lesssim \nu^{-O(1)} \sum_{I'\in\mathcal{I}'} \int_{\mathbb{R}^2} \Big|\sum_{\xi_1\in\Sigma_{1,I'}} f_{\xi_1}\Big|^2 \Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^4 dx. \tag{4.15}$$

Letting $F_{I'} = \sum_{\xi_1 \in \Sigma_{1,I'}} f_{\xi_1}$ for each $I' \in \mathcal{I}$ and letting $G = \Big(\sum_{\xi_2 \in \Sigma_2} g_{\xi_2} \Big)^2$, (4.15) is equivalent to the decoupling inequality

$$\Big\| \sum_{I' \in \mathcal{I}'} F_{I'} G \Big\|^2_{L^2(\mathbb{R}^2)} \lesssim \nu^{-O(1)} \sum_{I' \in \mathcal{I}'} \| F_{I'} G \|^2_{L^2(\mathbb{R}^2)}.$$

By Example 4.2, it suffices to show the Fourier supports of each of the $F_{I'}G$ have an overlap of $O(\nu^{-O(1)})$.

Let $I_2 = [\xi_0 - \rho_2, \xi_0 + \rho_2]$. The overlap of the Fourier supports of the $F_{I'}G$ is unchanged by a bijection of $\mathbb{R}^2$, so by applying the invertible affine transformation G_{ξ_0} as defined in the proof of parabolic rescaling, it suffices to consider the case where $I_2 = [-\rho_2, \rho_2]$, and $I_1 \subset [-2, 2]$ satisfies $\operatorname{dist}(I_1, \{0\}) \geq \nu$. Under these simplifications, it is not hard to see from the assumption $\delta \leq 2\rho_2$ that there is a rectangle of the form

$$R = \{(\xi_1, \xi_2) : \xi_1 = O(\rho_2)\,;\, \xi_2 = O(\rho_2^2 + \delta)\}.$$

such that $P_{\xi,\delta}$ is contained in R for all $\xi \in \Sigma_2$. It follows that $\sum_{\xi_2 \in \Sigma_2} g_{\xi_2}$ has Fourier support in R, and recalling how convolution affects supports, it is clear that $G = \Big(\sum_{\xi_2 \in \Sigma_2} g_{\xi_2} \Big)^2$ has Fourier support in a rectangle of the same form (after suitably enlarging the implied constants in the $O(\cdot)$ notation).

Let $\xi_{I'}$ be the centre of I' for each $I' \in \mathcal{I}'$, so $I' = [\xi_{I'} - \rho_1', \xi_{I'} + \rho_1']$. It is not hard to see from the assumption $\delta \leq 2\rho_1'$ that there is a horizontal strip of the form

$$S_{I'} = \{(\xi_1, \xi_2) : \xi_2 = \xi_{I'}^2 + O(\rho_1')\}.$$

such that $P_{\xi,\delta}$ is contained in $S_{I'}$ for all $\xi \in \Sigma_{1,I'}$. It follows that $F_{I'}$ has Fourier support in $S_{I'}$ and hence, using the assumptions $\rho_2^2 \leq \rho_1'$ and $\delta \leq 2\rho_1'$ and comparing the Fourier supports of $F_{I'}$ and G, we see that $F_{I'}G$ has Fourier support in a strip of the same form (after suitably enlarging the implied constants in the $O(\cdot)$ notation). Thus, if $(\xi_1, \xi_2) \in S_{I'}$ for some $I' \in \mathcal{I}'$, then (ξ_1, ξ_2) can only also lie in $S_{J'}$ for those $J' \in \mathcal{I}'$ for which $\xi_{I'}^2 - \xi_{J'}^2 = O(\rho_1')$. But this is only true if $\xi_{I'} - \xi_{J'} = O(\rho_1' / |\xi_{I'} + \xi_{J'}|) = O(\rho_1' \nu^{-1})$, where the last equality holds since $|\xi_{I'} + \xi_{J'}| = |\xi_{I'}| + |\xi_{J'}| \gtrsim \nu$. Since $|\xi_{I'} - \xi_{J'}| = 2\rho_1'$ for adjacent intervals $I', J' \in \mathcal{I}'$ by construction (with one possible exception), there can be at most $O(\nu^{-1})$ such J', and the Fourier supports of the $F_{I'}G$ therefore have an overlap of $O(\nu^{-O(1)})$ as required. □

The final two results we need before proving the decoupling theorem are each simple consequences of Hölder's inequality and parabolic rescaling. The first describes how ρ_1 and ρ_2 can be interchanged in the bilinear decoupling constant, and the second is a counterpart to bilinear reduction, allowing us to estimate bilinear decoupling constants by ordinary linear decoupling constants.

Lemma 4.18. *Let* $0 < \delta \leq 2\rho_1, 2\rho_2 \leq \nu \leq 1$. *Then,*

$$M_{2,4}(\delta, \nu, \rho_1, \rho_2) \leq M_{2,4}(\delta, \nu, \rho_2, \rho_1)^{1/2} D(\delta/\rho_2)^{1/2}.$$

Proof. With the same setup as in the proof of Lemma 4.17, we apply Hölder's inequality with the splitting

$$\Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^2\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^4 = \Big(\Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^2\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|\Big)\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^3$$

to give

$$\begin{aligned}\int_{\mathbb{R}^2}\Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^2\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^4 dx \leq \Big(\int_{\mathbb{R}^2}\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^2\Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^4 dx\Big)^{1/2} \\ \Big(\int_{\mathbb{R}^2}\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^6 dx\Big)^{1/2}.\end{aligned}$$

Estimating the first term using the definition of $M_{2,4}(\delta, \nu, \rho_2, \rho_1)$ and the second term by parabolic rescaling gives the result. □

Lemma 4.19. *Let* $0 < \delta \leq 2\rho_1, 2\rho_2 \leq \nu \leq 1$. *Then,*

$$M_{2,4}(\delta, \nu, \rho_1, \rho_2) \leq D(\delta/\rho_1)^{1/3} D(\delta/\rho_2)^{2/3}.$$

Proof. With the same setup as in the proof of Lemma 4.17, we apply Hölder's inequality to give

$$\int_{\mathbb{R}^2}\Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^2\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^4 dx \leq \Big(\int_{\mathbb{R}^2}\Big|\sum_{\xi_1\in\Sigma_1} f_{\xi_1}\Big|^6 dx\Big)^{1/3}\Big(\int_{\mathbb{R}^2}\Big|\sum_{\xi_2\in\Sigma_2} g_{\xi_2}\Big|^6 dx\Big)^{2/3}.$$

Estimating both terms by parabolic rescaling gives the result. □

We are now ready to prove the decoupling theorem. The idea is that by bilinear reduction, it suffices to estimate $M_{2,4}(\delta, 2\nu, \nu, \nu)$ for an appropriately chosen ν. The key estimate allows us to control this term by $M_{2,4}(\delta, 2\nu, \nu^2, \nu)$ with a loss of $\nu^{-O(1)}$, and Lemma 4.18 allows us to estimate this by $M_{2,4}(\delta, 2\nu, \nu, \nu^2)^{1/2} D(\delta/\nu)^{1/2}$. Iterating the process of lowering ρ_1 to ρ_2^2 and interchanging leads to the result.

Proof (Decoupling for the Parabola): Consider the set

$$\Lambda = \{\lambda \geq 0 : D(\delta) \lesssim_\varepsilon \delta^{-\lambda-\varepsilon} \text{ for all } 0 < \delta \leq 1 \text{ and all } \varepsilon > 0\},$$

and note that Λ is nonempty since any δ-separated subset of $[-1,1]$ has cardinality $O(\delta^{-1})$, from which the trivial bound (4.3) gives $D(\delta) \lesssim \delta^{-1/2} \lesssim_\varepsilon \delta^{-1/2-\varepsilon}$, hence $1/2 \in \Lambda$.

Let $\lambda = \inf \Lambda$. We wish to prove that $\lambda = 0$, so assume for the sake of contradiction that $\lambda > 0$. Choose N large enough that

$$\frac{O(1)}{\lambda} - \frac{5}{3} - N \leq -2$$

(where the $O(1)$ term is a sufficiently large multiple of that appearing in the key estimate), and suppose $0 < \delta \leq \frac{1}{2^{2^{N+1}}}$. Letting $2\nu = \delta^{1/2^{N+1}}$, we have $0 < \delta \leq \nu \leq 1/2$, so bilinear reduction gives

$$D(\delta) \lesssim \nu^{-O(1)} M_{2,4}(\delta, 2\nu, \nu, \nu) + D(\delta/\nu). \tag{4.16}$$

We also have $0 < \delta \leq 2\nu^2 \leq 2\nu \leq 1$, so the key estimate followed by Lemma 4.18 gives

$$M_{2,4}(\delta, 2\nu, \nu, \nu) \lesssim \nu^{-O(1)} M_{2,4}(\delta, 2\nu, \nu, \nu^2)^{1/2} D(\delta/\nu)^{1/2}.$$

In fact, by our choice of δ and ν, we have $0 < \delta \leq 2\nu^{2^j} \leq 2\nu \leq 1$ for all $1 \leq j \leq N$, so we may iterate the key estimate and Lemma 4.18 in this fashion to obtain

$$\begin{aligned} M_{2,4}(\delta, 2\nu, \nu, \nu) &\lesssim_N \nu^{-O(1)\sum_{j=0}^{N-1} 2^{-j}} M_{2,4}(\delta, 2\nu, \nu^{2^{N-1}}, \nu^{2^N})^{\frac{1}{2^N}} \prod_{j=1}^{N} D(\delta/\nu^{2^{j-1}})^{\frac{1}{2^j}} \\ &\lesssim_N \nu^{-O(1)} D(\delta/\nu^{2^{N-1}})^{\frac{1}{3\cdot 2^N}} D(\delta/\nu^{2^N})^{\frac{2}{3\cdot 2^N}} \prod_{j=1}^{N} D(\delta/\nu^{2^{j-1}})^{\frac{1}{2^j}}, \end{aligned} \tag{4.17}$$

where we have used Lemma 4.19 on the last line. Together with equation (4.16), (4.17) gives

$$\begin{aligned} D(\delta) &\lesssim_N \nu^{-O(1)} D(\delta/\nu^{2^{N-1}})^{\frac{1}{3\cdot 2^N}} D(\delta/\nu^{2^N})^{\frac{2}{3\cdot 2^N}} \prod_{j=1}^{N} D(\delta/\nu^{2^{j-1}})^{\frac{1}{2^j}} + D(\delta/\nu) \\ &= I + II. \end{aligned} \tag{4.18}$$

Now, choose $\lambda' \in \Lambda$, and note that $D(\delta) \lesssim_\varepsilon \delta^{-\lambda'-\varepsilon}$ for all $\varepsilon > 0$ by the definition of Λ. It follows from our choice of ν that for any $1 \leq j \leq N$ and all $\varepsilon > 0$, we have

$$D(\delta/\nu^{2^j}) = D(2^{2^j}\delta^{1-\frac{2^j}{2^{N+1}}}) \lesssim_\varepsilon 2^{-2^j(\lambda'+\varepsilon)}\delta^{-\lambda'(1-\frac{2^j}{2^{N+1}})-\varepsilon}$$
$$\lesssim_\varepsilon \delta^{-\lambda'(1-\frac{2^j}{2^{N+1}})-\varepsilon}. \tag{4.19}$$

Letting C_ε be the logarithm of the implied constant associated to ε in (4.19) (which we allow to change each line), taking logarithms gives

$$\log(D(\delta/\nu^{2^j})) \leq -\log(\delta)\Big(\lambda'\Big(1-\frac{2^j}{2^{N+1}}\Big)+\varepsilon\Big)+C_\varepsilon. \tag{4.20}$$

We may therefore compute

$$\begin{aligned}\log(I) &\leq -\log(\delta)\Big(\lambda'\Big(\frac{O(1)}{\lambda' 2^{N+1}}+\frac{1}{2^{N+2}}+\frac{1}{3\cdot 2^N}+\sum_{j=1}^{N}\Big(\frac{1}{2^j}-\frac{1}{2^{N+2}}\Big)\Big)+\varepsilon\Big)+C_\varepsilon\\ &= -\log(\delta)\Big(\lambda'\Big(1+\frac{1}{2^{N+2}}\Big(\frac{O(1)}{\lambda'}-\frac{5}{3}-N\Big)\Big)+\varepsilon\Big)+C_\varepsilon\\ &\leq -\log(\delta)\Big(\lambda'\Big(1-\frac{1}{2^{N+1}}\Big)+\varepsilon\Big)+C_\varepsilon,\end{aligned} \tag{4.21}$$

where the last line follows by our choice of N (noting that $\lambda' \geq \lambda$). Equation (4.20) also gives

$$\log(II) \leq -\log(\delta)\Big(\lambda'\Big(1-\frac{1}{2^{N+1}}\Big)+\varepsilon\Big)+C_\varepsilon, \tag{4.22}$$

so by equations (4.18), (4.21), and (4.22), we have

$$D(\delta) \lesssim_{N,\varepsilon} \delta^{-\lambda'(1-\frac{1}{2^{N+1}})-\varepsilon}$$

for all $0 < \delta \leq \frac{1}{2^{2^{N+1}}}$ and all $\varepsilon > 0$. Since this is also trivially true for $\frac{1}{2^{2^{N+1}}} < \delta \leq 1$, we see that $(1-\frac{1}{2^{N+1}})\lambda' \in \Lambda$. But by choosing $\lambda' \in \Lambda$ small enough initially, we can ensure $(1-\frac{1}{2^{N+1}})\lambda' < \lambda$, contradicting the fact that λ is a lower bound for Λ. It follows that we must have $\lambda = 0$. □

We will not treat any higher-dimensional cases of the decoupling theorem for the paraboloid, though we note that they rely inductively upon the $n = 2$ case.

Chapter 5

A New Proof of the Tomas Restriction Theorem

In this chapter, we unite the fields of Fourier restriction and decoupling by presenting a new proof of the Tomas restriction theorem for the paraboloid based on decoupling. We will see that the local extension estimates $R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1}\,;\,\varepsilon)$ for all $\varepsilon > 0$ follow quite easily from the decoupling theorem for the paraboloid; the main difficulty lies in proving an appropriate ε-removal theorem to convert this family of local extension estimates to a suitable family of global extension estimates. In fact, we will prove a more general result, showing that for any pair of exponents (p,q) with $2 \le p \le q \le \infty$, knowing $R^*_{P^{n-1}}(p \to q\,;\,\varepsilon)$ for all $\varepsilon > 0$ implies the global extension estimates $R^*_{P^{n-1}}(p \to r)$ for all $r > q$.

5.1 The Local Tomas-Stein Estimate

Our first goal is to prove $R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1}\,;\,\varepsilon)$ for all $\varepsilon > 0$, which we will call the *local Tomas-Stein estimate.* In light of Proposition 3.17, it suffices to prove that whenever $R \ge 1$ and $g \in \mathcal{S}(\mathbb{R}^n)$ has Fourier support in $\mathcal{N}_{R^{-1}}(P^{n-1})$, we have $\|g\|_{L^{\frac{2(n+1)}{n-1}}(B(0,R))} \lesssim_\varepsilon R^{\varepsilon-1/2}\|\hat{g}\|_{L^2(\mathbb{R}^n)}$ for all $\varepsilon > 0$. It is evident that there is some similarity between this estimate and the conclusion of the decoupling theorem for the paraboloid; indeed, it is in proving this estimate that the decoupling theorem becomes useful.

There is a manifestation of the uncertainty principle known as *Bernstein's inequality* which states that if $g \in \mathcal{S}(\mathbb{R}^n)$ has Fourier support in the ball $B(0,R)$, then $\|g\|_{L^q(\mathbb{R}^n)} \lesssim_{p,q} |B(0,R)|^{1/p-1/q}\|g\|_{L^p(\mathbb{R}^n)}$ for all $1 \le p \le q \le \infty$. We will require a variant of Bernstein's inequality specific to the regions $P_{\xi,\delta}$ featured in the decoupling theorem for the paraboloid. Our proof of this variant is inspired by that given for discs and ellipsoids in Chapter 5 of [Wol03].

Theorem 5.1. *Let $\xi \in \mathbb{R}^{n-1}$ and $\delta > 0$, and suppose $g \in \mathcal{S}(\mathbb{R}^n)$ has Fourier support in $P_{\xi,\delta}$. Then, for all $1 \le p \le q \le \infty$, we have*

$$\|g\|_{L^q(\mathbb{R}^n)} \lesssim_{p,q} |P_{\xi,\delta}|^{1/p-1/q}\|g\|_{L^p(\mathbb{R}^n)}.$$

Proof. Fix some $\varphi \in \mathcal{S}(\mathbb{R}^n)$ with $\hat{\varphi} \equiv 1$ on $[-1,1]^n$, and define an invertible affine transformation $G_{\xi,\delta} : \mathbb{R}^n \to \mathbb{R}^n$ by

$$G_{\xi,\delta}(\xi', \xi_n) = (\delta^{-1}(\xi' - \xi), \delta^{-2}(\xi_n - |\xi|^2 - 2\xi \cdot (\xi' - \xi))).$$

It is clear by the definitions that $G_{\xi,\delta}$ maps $P_{\xi,\delta}$ to the cube $[-1,1]^n$, so we have $\hat{\varphi} \circ G_{\xi,\delta} \equiv 1$ on $P_{\xi,\delta}$. Since g has Fourier support in $P_{\xi,\delta}$, it follows that $\hat{g} = \hat{g}(\hat{\varphi} \circ G_{\xi,\delta})$, from which we see that $g = g * (\hat{\varphi} \circ G_{\xi,\delta})\check{}$ upon Fourier inversion. Letting r satisfy $1/p + 1/r = 1 + 1/q$ (which is possible since $1 \le p \le q \le \infty$), Young's convolution inequality gives

$$\|g\|_{L^q(\mathbb{R}^n)} = \|g * (\hat{\varphi} \circ G_{\xi,\delta})\check{}\|_{L^q(\mathbb{R}^n)} \le \|g\|_{L^p(\mathbb{R}^n)}\|(\hat{\varphi} \circ G_{\xi,\delta})\check{}\|_{L^r(\mathbb{R}^n)}. \tag{5.1}$$

Now, $G_{\xi,\delta}$ is the composition of a translation and the invertible linear map $T_{\xi,\delta} : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$T_{\xi,\delta}(\xi', \xi_n) = (\delta^{-1}\xi', \delta^{-2}(\xi_n - 2\xi \cdot \xi')).$$

By translation invariance, we have $\|(\hat{\varphi} \circ G_{\xi,\delta})\check{}\|_{L^r(\mathbb{R}^n)} = \|(\hat{\varphi} \circ T_{\xi,\delta})\check{}\|_{L^r(\mathbb{R}^n)}$, and by Proposition 2.2 and Fourier inversion, we have

$$(\hat{\varphi} \circ T_{\xi,\delta})\check{} = |\det T_{\xi,\delta}|^{-1}(\varphi \circ T_{\xi,\delta}^{-t}).$$

It follows by a change of variables that

$$\begin{aligned} \|(\hat{\varphi} \circ G_{\xi,\delta})\check{}\|_{L^r(\mathbb{R}^n)} &= |\det T_{\xi,\delta}|^{-1}\Big(\int_{\mathbb{R}^n} |\varphi \circ T_{\xi,\delta}^{-t}(x)|^r\,dx\Big)^{1/r} \\ &= |\det T_{\xi,\delta}|^{-1/r'}\|\varphi\|_{L^r(\mathbb{R}^n)}. \end{aligned} \tag{5.2}$$

Since φ is a fixed function, $\|\varphi\|_{L^r(\mathbb{R}^n)}$ depends only on r, which depends only on p and q. Moreover, $|\det T_{\xi,\delta}|^{-1} \sim |P_{\xi,\delta}|$. Noting that $1/r' = 1/p - 1/q$, equations (5.1) and (5.2) therefore give the result. □

Having established an appropriate variant of Bernstein's inequality, we are ready to prove $R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1}\,;\varepsilon)$ for all $\varepsilon > 0$.

Theorem 5.2 (The Local Tomas-Stein Estimate). *$R^*_{P^{n-1}}(2 \to \frac{2(n+1)}{n-1}; \varepsilon)$ holds for all $\varepsilon > 0$.*

Proof. By Proposition 3.17, it suffices to prove that whenever $R \geq 1$ and $g \in \mathcal{S}(\mathbb{R}^n)$ has Fourier support in $\mathcal{N}_{R^{-1}}(P^{n-1})$, we have

$$\|g\|_{L^{\frac{2(n+1)}{n-1}}(B(0,R))} \lesssim_\varepsilon R^{\varepsilon-1/2}\|\hat{g}\|_{L^2(\mathbb{R}^n)} \tag{5.3}$$

for all $\varepsilon > 0$. In fact, we will prove the stronger statement in which the inequality (5.3) is replaced by

$$\|g\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \lesssim_\varepsilon R^{\varepsilon-1/2}\|\hat{g}\|_{L^2(\mathbb{R}^n)}.$$

Suppose $g \in \mathcal{S}(\mathbb{R}^n)$ has Fourier support in $\mathcal{N}_{R^{-1}}(P^{n-1})$, and let $\Sigma_R = R^{-1/2}\mathbb{Z}^{n-1} \cap [-1,1]^{n-1}$. Note that $\mathcal{N}_{R^{-1}}$ is covered by the collection $\mathcal{P}_R = \{P_{\xi,R^{-1/2}} : \xi \in \Sigma_R\}$, after possibly making the regions $P_{\xi,R^{-1/2}}$ taller by a fixed factor (recalling that the decoupling theorem remains true after such a scaling). By Theorem 2.39, let $(\eta_P)_{P\in\mathcal{P}_R}$ be a partition of unity subordinate to this cover, and for each $P \in \mathcal{P}_R$, let $g_P = (\hat{g}\eta_P)^{\vee}$. Each g_P is Schwartz as the inverse Fourier transform of a Schwartz function, and by Fourier inversion, we have that $\widehat{g_P} = \hat{g}\eta_P$ is supported in P. It follows from the property $\sum_{P\in\mathcal{P}_R} \eta_P = 1$ that $\sum_{P\in\mathcal{P}_R} \widehat{g_P} = \hat{g}$ and hence, upon Fourier inversion, we have $\sum_{P\in\mathcal{P}_R} g_P = g$. The decoupling theorem for the paraboloid therefore gives

$$\|g\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} = \Big\|\sum_{P\in\mathcal{P}_R} g_P\Big\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \lesssim_\varepsilon R^\varepsilon\Big(\sum_{P\in\mathcal{P}_R} \|g_P\|^2_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)}\Big)^{1/2} \tag{5.4}$$

for all $\varepsilon > 0$. Since $\widehat{g_P}$ is supported in P, Theorem 5.1 and Plancherel's theorem give

$$\begin{aligned}\|g_P\|^2_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} &\lesssim |P|^{1-\frac{n-1}{n+1}}\|g_P\|^2_{L^2(\mathbb{R}^n)} \\ &\lesssim (R^{-(n+1)/2})^{1-\frac{n-1}{n+1}}\|\widehat{g_P}\|^2_{L^2(\mathbb{R}^n)} \\ &= R^{-1}\|\widehat{g_P}\|^2_{L^2(\mathbb{R}^n)},\end{aligned}$$

and substituting this into (5.4) gives

$$\begin{aligned}\|g\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} &\lesssim_\varepsilon R^\varepsilon\Big(\sum_{P\in\mathcal{P}_R} R^{-1}\|\widehat{g_P}\|^2_{L^2(\mathbb{R}^n)}\Big)^{1/2} \\ &= R^{\varepsilon-1/2}\Big(\sum_{P\in\mathcal{P}_R} \|\widehat{g_P}\|^2_{L^2(\mathbb{R}^n)}\Big)^{1/2} \qquad (5.5)\end{aligned}$$

for all $\varepsilon > 0$. Now, each $\xi \in \mathbb{R}^n$ lies in at most 2^{n-1} of the regions P, and it follows that $\sum_{P\in\mathcal{P}_R} |\eta_P|^2 \leq 2^{n-1}$. We therefore have

$$\sum_{P\in\mathcal{P}_R} \|\widehat{g_P}\|_{L^2(\mathbb{R}^n)}^2 = \sum_{P\in\mathcal{P}_R} \int_{\mathbb{R}^n} |\hat{g}\eta_P|^2\, dx = \int_{\mathbb{R}^n} |\hat{g}|^2 \Big(\sum_{P\in\mathcal{P}_R} |\eta_P|^2 \Big)\, dx$$
$$\lesssim \|\hat{g}\|_{L^2(\mathbb{R}^n)}^2,$$

which together with (5.5) gives

$$\|g\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \lesssim_\varepsilon R^{\varepsilon-1/2} \|\hat{g}\|_{L^2(\mathbb{R}^n)}. \qquad \square$$

5.2 ε-Removal

We now treat the problem of bootstrapping the local extension estimate of the previous section to a global extension estimate. We will prove a more general ε-removal theorem which states that for all pairs of exponents (p, q) in a particular range, a local extension estimate $R^*_{P^{n-1}}(p \to q\,;\,\varepsilon)$ with $\varepsilon > 0$ sufficiently small implies a global extension estimate $R^*_{P^{n-1}}(p \to r)$ for all r larger than q by a certain threshold, with the threshold becoming arbitrarily small as $\varepsilon \to 0$.

Theorem 5.3 (ε-removal)**.** *For all $2 \leq p \leq q \leq \infty$ and all $\varepsilon > 0$ sufficiently small, $R^*_{P^{n-1}}(p \to q\,;\,\varepsilon)$ implies $R^*_{P^{n-1}}(p \to r)$ whenever $\frac{1}{r} < \frac{1}{q_0} := \frac{1}{q} - \frac{4\log 2}{q\log(1/\varepsilon q)}$.*

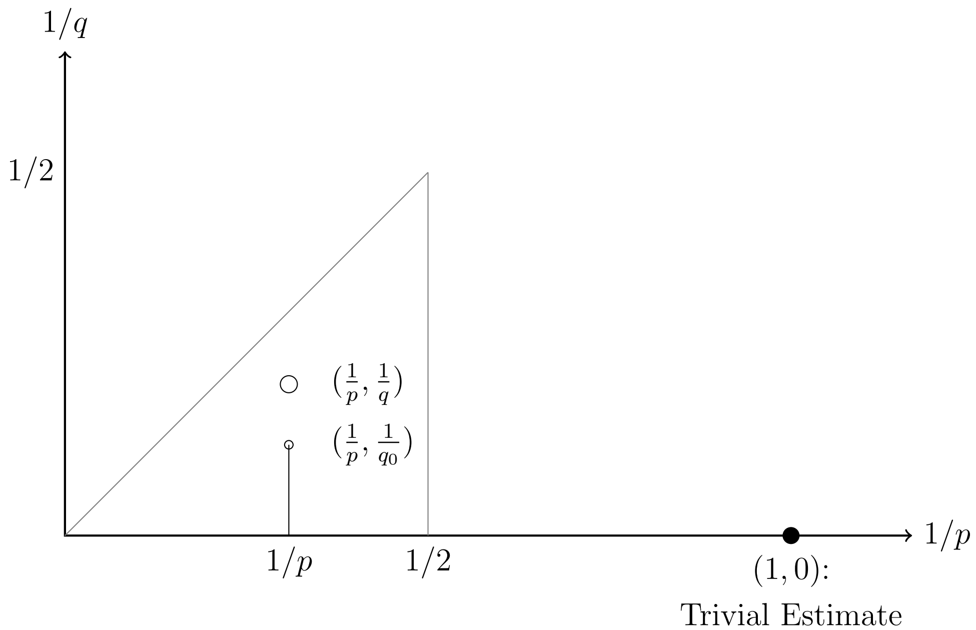

Figure 5.1: Theorem 5.3 visualised.

Theorem 5.3, and our approach to proving it, can be easily visualised using the above strong-type diagram for the extension operator $E_{P^{n-1}}$. Given the local extension estimate $R^*_{P^{n-1}}(p \to q\,;\varepsilon)$ (where $(\frac{1}{p}, \frac{1}{q})$ lies in the closed triangular region), we wish to deduce the global extension estimate $R^*_{P^{n-1}}(p \to r)$ for all r with $\frac{1}{r} < \frac{1}{q_0} = \frac{1}{q} - \frac{4\log 2}{q\log(1/\varepsilon q)}$. That is, we know a local extension estimate corresponding to the open circle at $(\frac{1}{p}, \frac{1}{q})$, and we wish to prove the global extension estimate for all pairs of exponents corresponding to the vertical line between $(\frac{1}{p}, \frac{1}{q_0})$ and $(\frac{1}{p}, 0)$. If we prove that E is of weak-type (p, q_0), then by real interpolation with the known strong-type (p, ∞) estimate (which follows from the trivial estimate and Lemma 3.11), we will obtain a global extension estimate for all pairs of exponents in the required range. It therefore suffices to prove

$$\|Ef\|_{L^{q_0,\infty}(\mathbb{R}^n)} \lesssim \|f\|_{L^p(P^{n-1},d\sigma)},$$

for all $f \in L^p(P^{n-1}, d\sigma)$, and our first step will be to bootstrap the local extension estimate $\|Ef\|_{L^q(B(0,R))} \lesssim R^\varepsilon \|f\|_{L^p(P^{n-1},\,d\sigma)}$ to an analogous estimate for multiple R-balls $\bigcup_i B(x_i, R)$, provided the balls are sufficiently separated, or *sparse*. We will then show that the superlevel sets $\{x \in \mathbb{R}^n : |Ef(x)| \geq t\}$ can be covered by a reasonably small number of sparse collections of balls of sufficiently small radius, from which we will be able to deduce the result. Our approach is based on that taken to prove a slightly different ε-removal result in [Tao99].

Definition 5.4. A collection $\{B(x_i, R)\}_{i=1}^N$ of R-balls is said to be *sparse* if $|x_j - x_i| \gtrsim (RN)^{\frac{2}{n-1}}$ for all $i \neq j$.

Remark 5.5. The implied constant in the definition of sparseness should be considered fixed but yet to be determined.

The following lemma and its proof are inspired by Lemma 3.2 of [Tao99].

Lemma 5.6. *Let $2 \leq p \leq q \leq \infty$, and suppose $R^*_{P^{n-1}}(p \to q\,;\varepsilon)$ holds. Then for all $R \geq 1$ and all $f \in L^p(P^{n-1}, d\sigma)$, we have*

$$\|Ef\|_{L^q(\bigcup_{i=1}^N B(x_i,R))} \lesssim R^\varepsilon \|f\|_{L^p(P^{n-1},\,d\sigma)}$$

whenever $\{B(x_i, R)\}_{i=1}^N$ is a sparse collection of R-balls.

Proof. Let $R \geq 1$. By Proposition 3.18 and tranlation invariance, $R^*_{P^{n-1}}(p \to q\,;\varepsilon)$ implies that for all $x \in \mathbb{R}^n$ and all $g \in \mathcal{S}(\mathbb{R}^n)$ with Fourier support in $\mathcal{N}_{R^{-1}}(P^{n-1})$, we have

$$\|g\|_{L^q(B(x,R))} \lesssim R^{\varepsilon - 1/p'} \|\hat{g}\|_{L^p(\mathcal{N}_{R^{-1}}(P^{n-1}))}. \tag{5.6}$$

By Proposition 2.13, choose some nonnegative $\psi \in \mathcal{S}(\mathbb{R}^n)$ such that $\psi \sim 1$ on $B(0,1)$, with $\hat{\psi}$ also nonnegative and supported in $B(0,1)$. Given a sparse collection of R-balls $\{B(x_i,R)\}_{i=1}^N$, we let $\psi_i(x) = \psi((x-x_i)/R)$. Given $f \in L^p(P^{n-1}, d\sigma)$, the same argument as in the derivation of equation (3.22) gives

$$\|Ef\|_{L^q(\cup_i B(x_i,R))} \sim \Big(\sum_{i=1}^N \|\psi_i Ef\|^q_{L^q(B(x_i,R))}\Big)^{1/q}. \tag{5.7}$$

By similar steps which led to equation (3.23), we find that $\widehat{\psi_i Ef} = \hat{\psi}_i * (f d\sigma)$, which is smooth and supported in $\mathcal{N}_{R^{-1}}(P^{n-1})$. We may therefore apply (5.6), giving

$$\|\psi_i Ef\|_{L^q(B(x_i,R))} \lesssim R^{\varepsilon - 1/p'} \|\hat{\psi}_i * (f d\sigma)\|_{L^p(\mathcal{N}_{R^{-1}}(P^{n-1}))}.$$

Combining with (5.7), we see that

$$\begin{aligned}\|Ef\|_{L^q(\cup_i B(x_i,R))} &\lesssim R^{\varepsilon-1/p'}\Big(\sum_{i=1}^N \|\hat{\psi}_i * (f d\sigma)\|^q_{L^p(\mathcal{N}_{R^{-1}}(P^{n-1}))}\Big)^{1/q} \\ &\le R^{\varepsilon-1/p'}\Big(\sum_{i=1}^N \|\hat{\psi}_i * (f d\sigma)\|^p_{L^p(\mathcal{N}_{R^{-1}}(P^{n-1}))}\Big)^{1/p},\end{aligned}$$

where the last inequality follows since $p \le q$ and the ℓ^p norm is decreasing in p. It therefore suffices to prove

$$\Big(\sum_{i=1}^N \|\hat{\psi}_i * (f d\sigma)\|^p_{L^p(\mathcal{N}_{R^{-1}}(P^{n-1}))}\Big)^{1/p} \lesssim R^{1/p'} \|f\|_{L^p(P^{n-1}, d\sigma)}. \tag{5.8}$$

Now, define a linear operator T mapping functions on P^{n-1} to functions on $\mathbb{R}^n \times \{1, \cdots, N\}$ by $Tf(\xi, i) = \hat{\psi}_i * (f d\sigma)(\xi)$. Then, (5.8) is equivalent to the statement that T is of strong-type (p,p) with $\|T\|_{L^p \to L^p} \lesssim R^{1/p'}$ (where $\mathbb{R}^n \times \{1, \cdots, N\}$ is given the product measure of the Lebesgue measure on $\mathbb{R}^n$ and counting measure on $\{1, \cdots, N\}$). Since p is in the range $2 \le p \le \infty$, it suffices by Riesz-Thorin interpolation to prove the bound (5.8) for $p = 2$ and $p = \infty$.
For $p = \infty$, we note that for any i and any $\xi \in \mathbb{R}^n$, the triangle inequality gives

$$|\hat{\psi}_i * (f d\sigma)(\xi)| \le \int_{P^{n-1}} |\hat{\psi}_i(\xi - \omega)||f(\omega)|\, d\sigma(\omega).$$

But $\hat{\psi}_i(\xi - \omega)$ is supported in $B(\xi, R^{-1})$ (as a function of ω), and it follows that

$$|\hat{\psi}_i * (f d\sigma)(\xi)| \le \|\hat{\psi}_i\|_{L^\infty(\mathbb{R}^n)} \|f\|_{L^\infty(P^{n-1}\, d\sigma)} \int_{P^{n-1}} \chi_{B(\xi, R^{-1})}(\omega)\, d\sigma(\omega). \tag{5.9}$$

Lemma 2.46 gives $\int_{P^{n-1}} \chi_{B(\xi,R^{-1})}(\omega)\,d\sigma(\omega) \lesssim R^{-(n-1)}$; moreover, $\|\hat{\psi}_i\|_{L^\infty(\mathbb{R}^n)} = R^n\|\hat{\psi}\|_{L^\infty(\mathbb{R}^n)} \sim R^n$, from which equation (5.9) gives

$$\|\hat{\psi}_i * (fd\sigma)\|_{L^\infty(\mathbb{R}^n)} \lesssim R\|f\|_{L^\infty(P^{n-1}\,d\sigma)}.$$

Since this is true for all i, we have $\max_i \|\hat{\psi}_i * (fd\sigma)\|_{L^\infty(\mathbb{R}^n)} \lesssim R\|f\|_{L^\infty(P^{n-1}\,d\sigma)}$, which is (5.8) for $p=\infty$.

For $p=2$, we exploit the fact that L^2 is a Hilbert space, so T has an adjoint T^*. In what follows, we will have frequent need to refer to $L^2 \to L^2$ operator norms, so we tidy our notation by using $\|\cdot\|$ to denote $\|\cdot\|_{L^2\to L^2}$.

Using the identity $\|TT^*\| = \|T\|^2$, we see that it suffices to prove $\|TT^*\| \lesssim R$ (this is known as the TT^* *method*). Identifying T^*, we see that it suffices to prove the estimate

$$\Big(\sum_{j=1}^N \Big\|\psi_j EE^*\Big(\sum_{i=1}^N \psi_i g_i\Big)\Big\|^2_{L^2(\mathbb{R}^n)}\Big)^{1/2} \lesssim R\Big(\sum_{i=1}^N \|g_i\|^2_{L^2(\mathbb{R}^n)}\Big)^{1/2}$$

whenever $g_1,\cdots,g_N \in L^2(\mathbb{R}^n)$. To estimate the left-hand side, we first note that by the triangle inequality, we have

$$\Big\|\psi_j EE^*\Big(\sum_{i=1}^N \psi_i g_i\Big)\Big\|_{L^2(\mathbb{R}^n)} \le \sum_{i=1}^N \|\psi_j EE^*\psi_i\|\|g_i\|_{L^2(\mathbb{R}^n)},$$

where we are now viewing each $\psi_j EE^*\psi_i$ as an operator. Hence,

$$\Big(\sum_{j=1}^N \Big\|\psi_j EE^*\Big(\sum_{i=1}^N \psi_i g_i\Big)\Big\|^2_{L^2(\mathbb{R}^n)}\Big)^{1/2} \le \Big(\sum_{j=1}^N \Big(\sum_{i=1}^N \|\psi_j EE^*\psi_i\|\|g_i\|_{L^2(\mathbb{R}^n)}\Big)^2\Big)^{1/2}. \tag{5.10}$$

But for each j, the Cauchy-Schwarz inequality gives

$$\Big(\sum_{i=1}^N \|\psi_j EE^*\psi_i\|\|g_i\|_{L^2(\mathbb{R}^n)}\Big)^2 \le \Big(\sup_j \sum_{i=1}^N \|\psi_j EE^*\psi_i\|\Big)\Big(\sum_{i=1}^N \|\psi_j EE^*\psi_i\|\|g_i\|^2_{L^2(\mathbb{R}^n)}\Big),$$

hence,

$$\sum_{j=1}^N \Big(\sum_{i=1}^N \|\psi_j EE^*\psi_i\|\|g_i\|_{L^2(\mathbb{R}^n)}\Big)^2 \le \Big(\sup_j \sum_{i=1}^N \|\psi_j EE^*\psi_i\|\Big)\Big(\sup_i \sum_{j=1}^N \|\psi_j EE^*\psi_i\|\Big)\Big(\sum_{i=1}^N \|g_i\|^2_{L^2(\mathbb{R}^n)}\Big). \tag{5.11}$$

But for each i and j, we have

$$\|\psi_j EE^*\psi_i\| = \|(\psi_j EE^*\psi_i)^*\| = \|\psi_i EE^*\psi_j\|,$$

where the last equality holds since the ψ_i, ψ_j were chosen to be real-valued. In particular, we have

$$\sup_j \sum_{i=1}^N \|\psi_j EE^*\psi_i\| = \sup_i \sum_{j=1}^N \|\psi_j EE^*\psi_i\|,$$

so equations (5.10) and (5.11) give

$$\Big(\sum_{j=1}^N \Big\|\psi_j EE^*\Big(\sum_{i=1}^N \psi_i g_i\Big)\Big\|_{L^2(\mathbb{R}^n)}^2\Big)^{1/2} \leq \Big(\sup_j \sum_{i=1}^N \|\psi_j EE^*\psi_i\|\Big)\Big(\sum_{i=1}^N \|g_i\|_{L^2(\mathbb{R}^n)}^2\Big)^{1/2};$$

it therefore suffices to prove

$$\sup_j \sum_{i=1}^N \|\psi_j EE^*\psi_i\| \lesssim R. \tag{5.12}$$

Given any $j \in \{1, \cdots, N\}$, the TT^* method gives

$$\|\psi_j EE^*\psi_j\| = \|(\psi_j E)(\psi_j E)^*\| = \|\psi_j E\|^2.$$

But by Plancherel's theorem and the same steps which led to equation (3.25), we have that for all $f \in L^p(P^{n-1}, d\sigma)$,

$$\|\psi_j Ef\|_{L^2(\mathbb{R}^n)} = \|(\psi_j Ef)\widehat{\ }\|_{L^2(\mathbb{R}^n)} \lesssim R^{1/2}\|f\|_{L^2(P^{n-1}, d\sigma)}.$$

Hence, $\|\psi_j E\|^2 \lesssim R$, giving $\|\psi_j EE^*\psi_j\| \lesssim R$.
For $i \neq j$, we will estimate $\|\psi_j EE^*\psi_i\|$ using Schur's test. As such, we must identify the integral kernel for the operator $\psi_j EE^*\psi_i$. Given $g \in \mathcal{S}(\mathbb{R}^n)$, noting that $E^*_{P^{n-1}} = R_{P^{n-1}}$ by equation (3.2), Fubini's theorem gives

$$\begin{aligned}(\psi_j EE^*\psi_i g)(x) &= \psi_j(x)\int_{\mathbb{R}^n} \psi_i(y)g(y)\Big(\int_{P^{n-1}} e^{-2\pi i(y-x)\cdot\xi}\, d\sigma(\xi)\Big)\, dy \\ &= \int_{\mathbb{R}^n} [\psi_j(x)\psi_i(y)\widehat{d\sigma}(y-x)]g(y)\, dy,\end{aligned}$$

so $\psi_j EE^*\psi_i$ has integral kernel $K(x,y) = \psi_j(x)\psi_i(y)\widehat{d\sigma}(y-x)$. By Theorem 2.44, we have

$$|\widehat{d\sigma}(y-x)| \lesssim |y-x|^{-(n-1)/2};$$

it follows that for any $x \in \mathbb{R}^n$,

$$\int_{\mathbb{R}^n} |K(x,y)|\,dy \lesssim \psi_j(x) \int_{\mathbb{R}^n} \psi_i(y)|y-x|^{-(n-1)/2}\,dy. \tag{5.13}$$

By translation invariance, we may assume without loss of generality that $x_i = 0$ and $|x_j| \gtrsim (RN)^{\frac{2}{n-1}}$. Since ψ_i and ψ_j are Schwartz, we then have

$$\psi_i(y) \lesssim (1+|y|)^{-(n+1)}$$

and

$$\psi_j(x) \lesssim (1+|x-x_j|)^{-(n+1)}.$$

Equation (5.13) therefore gives

$$\int_{\mathbb{R}^n} |K(x,y)|\,dy \lesssim (1+|x-x_j|)^{-(n+1)} \int_{\mathbb{R}^n} (1+|y|)^{-(n+1)}|y-x|^{-(n-1)/2}\,dy. \tag{5.14}$$

We estimate this integral by the "divide and conquer" approach, treating separately the cases when x is close to the origin (relative to x_j), and when x is far from the origin.

First, suppose $|x| < |x_j|/2$. Then, $|x-x_j| > |x_j|/2$, so (5.14) gives

$$\int_{\mathbb{R}^n} |K(x,y)|\,dy \lesssim |x_j|^{-(n+1)} \int_{\mathbb{R}^n} (1+|y|)^{-(n+1)}|y-x|^{-(n-1)/2}\,dy.$$

But $(1+|y|)^{-(n+1)}$ is bounded above by 1, so

$$\begin{aligned}\int_{|y-x|<1} (1+|y|)^{-(n+1)}|y-x|^{-(n-1)/2}\,dy &\leq \int_{|y-x|<1} |y-x|^{-(n-1)/2}\,dy \\ &\sim 1.\end{aligned}$$

We also have

$$\begin{aligned}\int_{|y-x|\geq 1} (1+|y|)^{-(n+1)}|y-x|^{-(n-1)/2}\,dy &\leq \int_{\mathbb{R}^n} (1+|y|)^{-(n+1)}\,dy \\ &\sim 1,\end{aligned}$$

and it follows that if $|x| < |x_j|/2$, then

$$\int_{\mathbb{R}^n} |K(x,y)|\,dy \lesssim |x_j|^{-(n+1)}. \tag{5.15}$$

Next, suppose $|x| \geq |x_j|/2$. Then, since $(1+|x-x_j|)^{-(n+1)}$ is bounded above by 1, equation (5.14) gives

$$\int_{\mathbb{R}^n} |K(x,y)|\,dy \lesssim \int_{\mathbb{R}^n} (1+|y|)^{-(n+1)}|y-x|^{-(n-1)/2}\,dy.$$

Clearly, we have

$$\begin{aligned}\int_{|y-x|\geq|x_j|/4}(1+|y|)^{-(n+1)}|y-x|^{-(n-1)/2}\,dy &\lesssim |x_j|^{-(n-1)/2}\int_{\mathbb{R}^n}(1+|y|)^{-(n+1)}\,dy\\ &\sim |x_j|^{-(n-1)/2}.\end{aligned}$$

Moreover, if $|y-x|<|x_j|/4$, then $|y|>|x_j|/4$, hence

$$\begin{aligned}\int_{|y-x|<|x_j|/4}(1+|y|)^{-(n+1)}|y-x|^{-\frac{(n-1)}{2}}\,dy &\lesssim |x_j|^{-(n+1)}\int_{|y-x|<|x_j|/4}|y-x|^{-\frac{(n-1)}{2}}\,dy\\ &\sim |x_j|^{-(n+1)/2}.\end{aligned}$$

It follows that if $|x|\geq|x_j|/2$, then

$$\begin{aligned}\int_{\mathbb{R}^n}|K(x,y)|\,dy &\lesssim |x_j|^{-(n-1)/2}+|x_j|^{-(n+1)/2}\\ &\lesssim |x_j|^{-(n-1)/2}.\end{aligned} \tag{5.16}$$

Comparing equations (5.15) and (5.16), we see that

$$\begin{aligned}\sup_{x\in\mathbb{R}^n}\int_{\mathbb{R}^n}|K(x,y)|\,dy &\lesssim |x_j|^{-(n-1)/2}\\ &\lesssim (RN)^{-1},\end{aligned}$$

where we have used the assumption $|x_j|\gtrsim(RN)^{\frac{2}{n-1}}$. By symmetry, we also have

$$\sup_{y\in\mathbb{R}^n}\int_{\mathbb{R}^n}|K(x,y)|\,dx\lesssim(RN)^{-1},$$

and it follows by Schur's test that

$$\|\psi_j EE^*\psi_i\|\lesssim(RN)^{-1}.$$

Combining this with the estimate $\|\psi_j EE^*\psi_j\|\lesssim R$, we find that

$$\sum_{i=1}^{N}\|\psi_j EE^*\psi_i\|\lesssim R+(N-1)(RN)^{-1}\lesssim R,$$

which proves equation (5.12) since $j\in\{1,\cdots,N\}$ was arbitrary. □

We can therefore bootstrap local extension estimates to sparse collections of balls. Recall that we also require a covering lemma, allowing us to cover a set of the form $\{x\in\mathbb{R}^n:|Ef(x)|\geq t\}$ by an appropriate family of sparse collections of balls. The following lemma provides an appropriate covering for sets which are a union of unit cubes, with the idea being that the uncertainty principle dictates that the set $\{x\in\mathbb{R}^n:|Ef(x)|\geq t\}$ will not be too different to a union of unit cubes (since $\widehat{Ef}$ is supported in a cube).

Lemma 5.7 (Lemma 3.3, [Tao99])**.** *Let A be a union of unit cubes. For any $N \geq 1$, A may be covered by $O(N|A|^{1/N})$ sparse collections of balls of radius $O(|A|^{2^N})$.*

Proof. If A is a single unit cube the result is clear. We may therefore assume without loss of generality that A is a union of at least two unit cubes, in which case $|A| \geq 2$.

Let $N \geq 1$. Define the radii R_k for $0 \leq k \leq N$ inductively by $R_0 = 1$ and $R_{k+1} = |A|^2 R_k^2$, and note that $R_k = O(|A|^{2^k})$. Define

$$A_1 = \{x \in A : |A \cap B(x, R_0)| \leq |A|^{1/N}\},$$

and for $1 < k \leq N$, define

$$A_k = \{x \in A : x \notin A_j \text{ for } j < k, \text{ and } |A \cap B(x, R_k)| \leq |A|^{k/N}\}.$$

Since $A = \bigcup_{k=1}^N A_k$, it suffices to prove that each A_k can be covered by $O(|A|^{1/N})$ sparse collections of balls of radius $O(|A|^{2^N})$. In fact, we will show that each A_k can be covered by $O(|A|^{1/N})$ sparse collections of balls of radius $5R_{k-1}$ (noting that $5R_{k-1} = O(|A|^{2^{k-1}}) = O(|A|^{2^N})$).

Fix $1 \leq k \leq N$, and let $x \in A_k$. We will first cover the set $A_x = A_k \cap B(x, R_k/2)$ by $O(|A|^{1/N})$ balls of radius $5R_{k-1}$. To do so, consider the cover $\{B(y, R_{k-1}) : y \in A_x\}$ of the set A_x by R_{k-1}-balls. By the Vitali covering lemma, we may extract a countable subset $A'_x \subset A_x$ such that the balls $\{B(y, R_{k-1}) : y \in A'_x\}$ are pairwise disjoint, and $\{B(y, 5R_{k-1}) : y \in A'_x\}$ covers A_x. We claim that the set A'_x must have cardinality $O(|A|^{1/N})$; indeed, given any $y \in A'_x$, we have $y \in A_k$, hence $|A \cap B(y, R_{k-1})| \gtrsim |A|^{(k-1)/N}$ by the definition of A_k (using the fact that A is a union of unit cubes in the case $k = 1$). We also have $y \in B(x, R_k/2)$, hence $B(y, R_{k-1}) \subset B(x, R_k)$, since the assumption $|A| \geq 2$ implies $R_{k-1} \leq R_k/2$. It follows that for any distinct $y_1, \cdots, y_m \in A'_x$, we have

$$\begin{aligned} m|A|^{(k-1)/N} &\lesssim \sum_{j=1}^m |A \cap B(y_j, R_{k-1})| = \Big|A \cap \bigcup_{j=1}^m B(y_j, R_{k-1})\Big| \\ &\leq |A \cap B(x, R_k)| \leq |A|^{k/N}, \end{aligned}$$

hence, $m \lesssim |A|^{1/N}$ (where we have used the disjointness of the balls $B(y_j, R_{k-1})$ and the definition of A_k). It follows that A'_x has cardinality $O(|A|^{1/N})$, and since $x \in A_k$ was arbitrary, we see that the set $A_k \cap B(x, R_k/2)$ may be covered by $O(|A|^{1/N})$ balls of radius $5R_{k-1}$ for any $x \in A_k$.

The set A_k may therefore be covered by $O(|A|^{1/N})$ collections of $5R_{k-1}$-balls of separation $\gtrsim R_k$, where each collection has cardinality $O(|A|)$. Since $R_k \gtrsim (5R_{k-1}O(|A|))^{\frac{2}{n-1}}$, it follows that each of these collections is sparse provided the implied constant in the definition of sparseness is small enough. $\square$

We now have all of the tools necessary to prove the ε-removal theorem. In what follows, we will have frequent need to consider various superlevel sets, so we introduce some notation here for convenience:

Notation 5.8. Let $f : \mathbb{R}^n \to \mathbb{C}$ be measurable. Given $t > 0$, we let $\eta_f(t)$ denote the superlevel set

$$\eta_f(t) = \{x \in \mathbb{R}^n : |f(x)| \geq t\}.$$

The distribution function $\lambda_f(t)$ as introduced in Section 2.2 is therefore given by $\lambda_f(t) = |\eta_f(t)|$.
We will also use the following:

Notation 5.9. Given $r > 0$, we let $Q_r = [-r, r]^n$.

Proof (ε-Removal): Let $2 \leq p \leq q \leq \infty$, $\varepsilon > 0$, and suppose that $R^*_{P^{n-1}}(p \to q\,;\,\varepsilon)$ holds. Recall from the discussion following the statement of the theorem that it suffices to prove

$$\|Ef\|_{L^{q_0,\infty}(\mathbb{R}^n)} \lesssim \|f\|_{L^p(P^{n-1},\,d\sigma)} \tag{5.17}$$

for all $f \in L^p(P^{n-1}, d\sigma)$, where $\frac{1}{q_0} = \frac{1}{q} - \frac{4\log 2}{q\log(1/\varepsilon q)}$.
Let $f \in L^p(P^{n-1}, d\sigma)$, and assume without loss of generality that $\|f\|_{L^p(P^{n-1},\,d\sigma)} = 1$. We begin by bounding the left-hand side of (5.17) by the $L^{q_0,\infty}(\mathbb{R}^n)$ norm of the average $|Ef| * \chi_{Q_{1/4}}$, motivated by the heuristic that the superlevel sets $\eta_{|Ef|*\chi_{Q_{1/4}}}(t)$ resemble unions of unit cubes, so Lemma 5.7 will become applicable. By Fubini's theorem, we may compute

$$\widehat{\chi_{Q_{1/4}}}(\xi) = \prod_{i=1}^{n} \frac{\sin(\pi\xi_i/2)}{\pi\xi_i},$$

so $\widehat{\chi_{Q_{1/4}}} \sim 1$ on $Q_{3/2}$. Let $\varphi \in C_c^\infty(Q_{3/2})$ be a bump function with $\varphi \equiv 1$ on Q_1, and define

$$\psi := \begin{cases} \varphi/\widehat{\chi_{Q_{1/4}}} & \text{on } Q_{3/2}\,; \\ 0 & \text{on } \mathbb{R}^n \setminus Q_{3/2}. \end{cases}$$

Since $\varphi \equiv 1$ on $P^{n-1} \subset Q_1$, Fubini's theorem gives $Ef = Ef * \check{\varphi}$. But clearly, $\varphi = \widehat{\chi_{Q_{1/4}}}\psi$, from which we see by Fourier inversion that $\check{\varphi} = \chi_{Q_{1/4}} * \check{\psi}$, hence $Ef = Ef * \chi_{Q_{1/4}} * \check{\psi}$. Young's convolution inequality for weak L^p spaces therefore gives $\|Ef\|_{L^{q_0,\infty}(\mathbb{R}^n)} \lesssim \||Ef| * \chi_{Q_{1/4}}\|_{L^{q_0,\infty}(\mathbb{R}^n)}$, so by our normalisation, it suffices to prove $\||Ef| * \chi_{Q_{1/4}}\|_{L^{q_0,\infty}(\mathbb{R}^n)} \lesssim 1$. That is, it suffices to prove that for all $t > 0$,

$$|\eta_{|Ef|*\chi_{Q_{1/4}}}(t)|^{1/q_0} \lesssim \frac{1}{t}. \tag{5.18}$$

We first note that (5.18) is clear for all t larger than a particular threshold K independent of f, so we need only consider $0 < t \lesssim 1$. Indeed, Young's convolution inequality and the trivial estimate give

$$\||Ef| * \chi_{Q_{1/4}}\|_{L^\infty(\mathbb{R}^n)} \leq \|Ef\|_{L^\infty(\mathbb{R}^n)} \|\chi_{Q_{1/4}}\|_{L^1(\mathbb{R}^n)} \lesssim \|f\|_{L^1(P^{n-1}, d\sigma)},$$

and since P^{n-1} is compact, Hölder's inequality combined with our normalisation $\|f\|_{L^p(P^{n-1}, d\sigma)} = 1$ implies $\|f\|_{L^1(P^{n-1}, d\sigma)} \lesssim 1$. It follows that $\||Ef| * \chi_{Q_{1/4}}\|_{L^\infty(\mathbb{R}^n)} \leq K$ for some constant K independent of f, from which we see that $|\eta_{|Ef|*\chi_{Q_{1/4}}}(t)| = 0$ for $t > K$, in which case (5.18) clearly holds. We may therefore assume without loss of generality that $0 < t \leq K$.

Given $t > 0$, let A_t be the union of all unit cubes of the form $z + Q_{1/2}$ for some $z \in \mathbb{Z}^n$ which have nonempty intersection with the superlevel set $\eta_{|Ef|*\chi_{Q_{1/4}}}(t)$. Clearly, $\eta_{|Ef|*\chi_{Q_{1/4}}}(t) \subset A_t$, hence $|\eta_{|Ef|*\chi_{Q_{1/4}}}(t)|^{1/q_0} \leq |A_t|^{1/q_0}$, so it now suffices to prove

$$|A_t|^{1/q_0} \lesssim \frac{1}{t} \tag{5.19}$$

for all $0 < t \leq K$. Since A_t is a union of unit cubes, for any $N \geq 1$, Lemma 5.7 allows us to cover A_t by sparse collections $\mathcal{C}_1, \cdots, \mathcal{C}_m$ of balls of radius $O(|A_t|^{2^N})$, where $m = O(N|A_t|^{1/N})$. Since these collections cover A_t, we have

$$A_t = \bigcup_{i=1}^m \Big(A_t \cap \bigcup_{B \in \mathcal{C}_i} B\Big),$$

hence,

$$|A_t| \leq \sum_{i=1}^m \Big|A_t \cap \bigcup_{B \in \mathcal{C}_i} B\Big|. \tag{5.20}$$

Now, if $x \in A_t$, then by definition, there exists $z \in \mathbb{Z}^n$ and $y \in \mathbb{R}^n$ such that $x, y \in z + Q_{1/2}$, and $|Ef| * \chi_{Q_{1/4}}(y) \geq t$. It follows that $|x - y| \leq \sqrt{n}$ and hence,

$|Ef|*\chi_{Q_{1/4+\sqrt{n}}}(x) \geq |Ef|*\chi_{Q_{1/4}}(y) \geq t$. We therefore have $A_t \subset \eta_{|Ef|*\chi_{Q_{1/4+\sqrt{n}}}}(t)$, and in particular, equation (5.20) gives

$$\begin{aligned}|A_t| \leq \sum_{i=1}^{m} \Big|A_t \cap \bigcup_{B\in\mathcal{C}_i} B\Big| &\leq \sum_{i=1}^{m} \Big|\eta_{|Ef|*\chi_{Q_{1/4+\sqrt{n}}}}(t) \cap \bigcup_{B\in\mathcal{C}_i} B\Big| \\ &= \sum_{i=1}^{m} |\eta_{(|Ef|*\chi_{Q_{1/4+\sqrt{n}}})\chi_{\bigcup_{B\in\mathcal{C}_i} B}}(t)| \\ &\leq \frac{1}{t^q}\sum_{i=1}^{m} \||Ef| * \chi_{Q_{1/4+\sqrt{n}}}\|^q_{L^q(\bigcup_{B\in\mathcal{C}_i} B)}, \end{aligned} \tag{5.21}$$

where the last line follows by Chebyshev's inequality. But for all $x \in \bigcup_{B\in\mathcal{C}_i} B$, we have

$$|Ef| * \chi_{Q_{1/4+\sqrt{n}}}(x) = (|Ef|\chi_{\bigcup_{B\in\mathcal{C}_i}(B+Q_{1/4+\sqrt{n}})}) * \chi_{Q_{1/4+\sqrt{n}}}(x).$$

It follows that

$$\begin{aligned}\||Ef| * \chi_{Q_{1/4+\sqrt{n}}}\|^q_{L^q(\bigcup_{B\in\mathcal{C}_i} B)} &\leq \|(|Ef|\chi_{\bigcup_{B\in\mathcal{C}_i}(B+Q_{1/4+\sqrt{n}})}) * \chi_{Q_{1/4+\sqrt{n}}}\|^q_{L^q(\mathbb{R}^n)} \\ &\lesssim \|Ef\|^q_{L^q(\bigcup_{B\in\mathcal{C}_i}(B+Q_{1/4+\sqrt{n}}))}, \end{aligned} \tag{5.22}$$

where the last line follows by Young's convolution inequality. But clearly, $B(x,R)+Q_{1/4+\sqrt{n}} \subset B(x, R+2n)$ for any ball $B(x,R) \subset \mathbb{R}^n$. Letting $\tilde{\mathcal{C}}_i$ denote the collection of balls obtained by enlarging the radius of each ball in $\mathcal{C}_i$ by $2n$, it follows that

$$\|Ef\|^q_{L^q(\bigcup_{B\in\mathcal{C}_i}(B+Q_{1/4+\sqrt{n}}))} \leq \|Ef\|^q_{L^q(\bigcup_{B\in\tilde{\mathcal{C}}_i} B)},$$

and combining this with (5.21) and (5.22) gives

$$|A_t| \lesssim \frac{1}{t^q}\sum_{i=1}^{m} \|Ef\|^q_{L^q(\bigcup_{B\in\tilde{\mathcal{C}}_i} B)}. \tag{5.23}$$

Noting that each of the collections $\tilde{\mathcal{C}}_i$ is also sparse and is comprised of balls of radius $\lesssim |A_t|^{2^N} + 2n$, Lemma 5.6 and our normalisation give

$$\|Ef\|^q_{L^q(\bigcup_{B\in\tilde{\mathcal{C}}_i} B)} \lesssim (|A_t|^{2^N} + 2n)^{\varepsilon q}. \tag{5.24}$$

Now, if $|A_t| < 2n$, we get $|A_t|^{1/q_0} \lesssim 1 \leq K/t$ (recalling that $0 < t \leq K$), which gives the conclusion (5.19). We may therefore assume without loss of generality that $2n \leq |A_t|$, in which case (5.24) gives

$$\|Ef\|^q_{L^q(\bigcup_{B\in\tilde{\mathcal{C}}_i} B)} \lesssim |A_t|^{\varepsilon q 2^N}. \tag{5.25}$$

Combining (5.23) and (5.25) and recalling that $m = O(N|A_t|^{1/N})$, we find that

$$|A_t| \lesssim \frac{1}{t^q} N |A_t|^{\varepsilon q 2^N + 1/N}. \tag{5.26}$$

Setting $N = \frac{\log(1/\varepsilon q)}{2\log 2}$, we get $\varepsilon q 2^{2N} = 1$, hence $\varepsilon q 2^N = 2^{-N} \leq 1/N$. Noting that $|A_t| \geq 1$, (5.26) then gives $|A_t| \lesssim \frac{1}{t^q} N |A_t|^{2/N}$ for our particular choice of N. Hence, if $|A_t|$ is finite, we may rearrange to conclude that $|A_t|^{1/q - 2/Nq} \lesssim \frac{N^{1/q}}{t}$. Substituting our expression for N, this is equivalent to

$$|A_t|^{\frac{1}{q} - \frac{4\log 2}{q\log(1/\varepsilon q)}} \lesssim \frac{1}{t}, \tag{5.27}$$

which is the conclusion (5.19). To see that $|A_t|$ is indeed finite, we note that by Remark 2.45 and a density argument, $Ef(x)$ decays to 0 as $|x| \to \infty$. It follows that the superlevel set $\eta_{|Ef| * \chi_{Q_{1/4}}}(t)$ is bounded, from which it is clear that A_t is also bounded, and therefore has finite measure. Our manipulations leading to (5.27) are therefore justified, and we are done. □

The most important feature of Theorem 5.3 is that the threshold of $\frac{1}{r} < \frac{1}{q} - \frac{4\log 2}{q\log(1/\varepsilon q)}$ below which $R^*_{P^{n-1}}(p \to q\,;\varepsilon)$ implies $R^*_{P^{n-1}}(p \to r)$ limits to $1/q$ as $\varepsilon \to 0$. This leads to the following simple corollary:

Corollary 5.10. *If $2 \leq p \leq q \leq \infty$ and $R^*_{P^{n-1}}(p \to q\,;\varepsilon)$ holds for all $\varepsilon > 0$ sufficiently small, then $R^*_{P^{n-1}}(p \to r)$ holds for all $r > q$.* □

In particular, when combined with Theorem 5.2, Corollary 5.10, yields the Tomas restriction theorem:

Corollary 5.11. *$R^*_{P^{n-1}}(2 \to q)$ holds for all $q > \frac{2(n+1)}{n-1}$.* □

5.3 Future Directions

It is natural to ask whether the regime of using decoupling to prove a family of local extension estimates and then applying ε-removal may be used to prove further global extension estimates for other submanifolds (not necessarily hypersurfaces). Indeed, we suspect that whenever one has an ℓ^2 decoupling inequality in L^p for a submanifold S (analogous to Theorem 4.5 for the paraboloid), it may be possible to use decoupling to prove the local extension estimates $R^*_S(2 \to p\,;\varepsilon)$ for all $\varepsilon > 0$ (analogous to Theorem 5.2), following which one may use ε removal to obtain the global extension estimates $R^*_S(2 \to q)$ for all $q > p$.

One pertinent possibility is the question of whether one may use the decoupling theorem for the moment curve (Theorem 4.8) to prove $R^*_{\Gamma^n}(2 \to n(n+1)\,;\,\varepsilon)$ for all $\varepsilon > 0$. Unfortunately, when attempting to emulate our proof of Theorem 5.2, one encounters some geometric difficulties regarding the possibility of covering a R^{-1}-neighbourhood of the moment curve by a family of regions of the form $\theta_{\xi,R^{-k}}$ for a suitable exponent k. Despite this, it is our conviction that with sufficient insight, our technique may be successfully applied.
There is a wide range of decoupling theorems which have been proven for different submanifolds in recent years (many of which are discussed in [Tao20a]), so the possible applications of our techniques are numerous.

Bibliography

[BD15] Jean Bourgain and Ciprian Demeter. The proof of the l^2 decoupling conjecture. *Ann. of Math. (2)*, 182(1):351–389, 2015.

[BDG16] Jean Bourgain, Ciprian Demeter, and Larry Guth. Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three. *Ann. of Math. (2)*, 184(2):633–682, 2016.

[BL76] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976.

[GLY21] Shaoming Guo, Zane Kun Li, and Po-Lam Yung. A bilinear proof of decoupling for the cubic moment curve. *Trans. Amer. Math. Soc.*, 374(8):5405–5432, 2021.

[Gra04] Loukas Grafakos. *Classical and modern Fourier analysis*. Pearson Education, Inc., Upper Saddle River, NJ, 2004.

[HZ20] Jonathan Hickman and Joshua Zahl. A note on fourier restriction and nested Polynomial Wolff axioms. *arXiv:2010.02251*, 2020.

[Li21] Zane Kun Li. An l^2 decoupling interpretation of efficient congruencing: the parabola. *Rev. Mat. Iberoam.*, 37(5):1761–1802, 2021.

[O'N63] Richard O'Neil. Convolution operators and $L(p, q)$ spaces. *Duke Math. J.*, 30:129–142, 1963.

[Pie19] Lillian B. Pierce. The Vinogradov mean value theorem [after Wooley, and Bourgain, Demeter and Guth]. Number 407, pages Exp. No. 1134, 479–564. 2019. Séminaire Bourbaki. Vol. 2016/2017. Exposés 1120–1135.

[Sog91] Christopher D. Sogge. Propagation of singularities and maximal functions in the plane. *Invent. Math.*, 104(2):349–376, 1991.

[SS03] Elias M. Stein and Rami Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003. An introduction.

[Str77] Robert S. Strichartz. Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Math. J.*, 44(3):705–714, 1977.

[SW71] Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.

[Tao99] Terence Tao. The Bochner-Riesz conjecture implies the restriction conjecture. *Duke Math. J.*, 96(2):363–375, 1999.

[Tao04] Terence Tao. Some recent progress on the restriction conjecture. In *Fourier analysis and convexity*, Appl. Numer. Harmon. Anal., pages 217–243. Birkhäuser Boston, Boston, MA, 2004.

[Tao08] Terence Tao. The divisor bound `https://terrytao.wordpress.com/2008/09/23/the-divisor-bound/`, 2008.

[Tao09] Terence Tao. Interpolation of L^p spaces `https://terrytao.wordpress.com/2009/03/30/245c-notes-1-interpolation-of-lp-spaces/`, 2009.

[Tao15] Terence Tao. Decoupling and the bourgain-demeter-guth proof of the vinogradov main conjecture `https://terrytao.wordpress.com/2015/12/10/decoupling-and-the-bourgain-demeter-guth-proof-of-the-vinogradov-main-conjecture/`, 2015.

[Tao20a] Terence Tao. Decoupling theory `https://terrytao.wordpress.com/2020/04/13/247b-notes-2-decoupling-theory/`, 2020.

[Tao20b] Terence Tao. Restriction theory `https://terrytao.wordpress.com/2020/03/29/247b-notes-1-restriction-theory/`, 2020.

[Tom75] Peter A. Tomas. A restriction theorem for the Fourier transform. *Bull. Amer. Math. Soc.*, 81:477–478, 1975.

[Wol00] T. Wolff. Local smoothing type estimates on L^p for large p. *Geom. Funct. Anal.*, 10(5):1237–1288, 2000.

[Wol03] Thomas H. Wolff. *Lectures on harmonic analysis*, volume 29 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2003. With a foreword by Charles Fefferman and a preface by Izabella Łaba, Edited by Łaba and Carol Shubin.

[Zyg74] A. Zygmund. On Fourier coefficients and transforms of functions of two variables. *Studia Math.*, 50:189–201, 1974.