Decoupling and restriction inequalities for smooth compact surfaces in three dimensions

by
Jianhui Li

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)

at the
University of Wisconsin-Madison
2023

Date of Final Oral Exam: 05/01/2023
The dissertation is approved by the following members of the Final Oral Committee:

Betsy Stovall, Professor, Mathematics
Shaoming Guo, Associate Professor, Mathematics
Andreas Seeger, Professor, Mathematics
Hung Vinh Tran, Associate Professor, Mathematics

Decoupling and restriction inequalities for smooth compact surfaces in three dimensions

Jianhui Li

Abstract

This dissertation establishes general decoupling and restriction inequalities for smooth compact surfaces in $\mathbb{R}^3$. First, we prove decoupling inequalities in $\mathbb{R}^3$ for the graphs of all bivariate polynomials of degree at most k with bounded coefficients over a compact set, with the decoupling constant uniform in the coefficients of those polynomials. As a consequence, we prove a decoupling inequality for smooth compact surfaces in $\mathbb{R}^3$. This extends the decoupling theorem of Bourgain and Demeter to all smooth surfaces. Second, we prove sharp L^2 Fourier restriction inequalities for smooth compact surfaces in $\mathbb{R}^3$ equipped with the affine surface measure or a power thereof. The estimates are uniform for all surfaces defined by the graph of polynomials of degrees at most k with bounded coefficients. The primary tool is a variant of the aforementioned decoupling inequalities for these surfaces.

Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my advisor Betsy Stovall for her consistent assistance and encouragement throughout my doctoral study at the University of Wisconsin-Madison. Without her support, this dissertation would be impossible to complete. Second, I would also like to thank my coauthor Tongou Yang for allowing me to adapt our joint work as one of the main chapters in this document. Third, I would like to thank my master advisor Po Lam Yung for his invaluable guidance, which built a solid foundation for my doctoral research in harmonic analysis.

I would also like to acknowledge support from National Science Foundation grants DMS-1653264 and the Wisconsin Alumni Research Foundation.

Lay summary

Harmonic analysis is a quantitative study of functions and their oscillations. The Fourier transform, which decomposes a general function, such as the one that describes a musical signal, into a superposition of waves with varying intensities and frequencies, is a fundamental component of the study. In practice, errors occur when sampling, transforming, and reconstructing the signals. Quantitatively studying these operators controls the errors and thus establishes stability results.

Over the past sixty years, harmonic analysts have spent significant efforts studying functions formed by the superposition of waves with frequencies concentrated on hypersurfaces. One motivation is to study dispersive partial differential equations. Examples include the Schrödinger, wave, and Helmholtz equations, the solutions to which have frequencies confined on the paraboloids, cones, and spheres, respectively. While the above three surfaces are better understood in the literature, this dissertation attempts to extend the study to all smooth surfaces.

Consider a wave with frequencies on a small neighborhood of a smooth surface centered at a point ξ_0. Let $\vec{n}$ be the normal direction of the surface at ξ_0. Then, the wave travels along $\vec{n}$ in the physical space. Now, we describe a dichotomy. If all waves have frequencies on a flat plane, they travel along the same direction in physical space. In this scenario, no interference occurs as the solution evolves. In an opposite scenario, if all waves have frequencies on a curved surface, for example, a sphere, they travel along different directions in physical space. In this case, interference among waves occurs as time evolves. A way to describe interference is to measure how "big" the resulting function can be. In the case

when frequencies lie in a plane, the best we can say is that energy is conserved. In the latter case of curved surfaces, a famous theorem proved by Tomas and Stein (also known as Strichartz's inequality) provides some excellent quantitative decay estimates.

What happens to surfaces that are not flat while also not as curved as a sphere? Given how well we understand both extreme cases, we want to decompose the frequency support of the function into smaller pieces, each of which can be enlarged to one of these two extreme cases. Decoupling theory describes a technique for controlling the constructive interference arising from multiple piece of this ensemble. Chapter 2 formulates and proves decoupling inequalities for general smooth surfaces. Chapter 3 applies these inequalities to prove the best possible extension of the Tomas-Stein theory to general smooth surfaces with a measure that compensates flat regions to allow similar decay estimates.

Contents

Abstract **i**

Acknowledgements **ii**

Lay summary **iii**

1 Introduction **1**

1.1 Decoupling theory 1

1.2 Affine restriction theory 5

1.3 Proof Strategies 10

1.3.1 Passing to polynomials 10

1.3.2 Iterative decoupling 11

1.3.3 Dyadic decomposition 12

1.3.4 Projection and 2D general decoupling 12

1.3.5 Surfaces with small Gaussian curvature on the entire domain 13

1.3.6 Induction on scale 14

1.3.7 Affine restriction estimates 14

1.4 Notations 15

2 Decoupling theory **18**

2.1 Overview 18

2.2 Two reductions 22

2.2.1 Removal of ε-dependence of the covering $\mathcal{P}$ 22

2.2.2 Taylor approximation 23
2.3 General 2D decoupling 24
2.3.1 Uniform decoupling of polynomial-like rational functions 25
2.3.2 The case of constant gradient 27
2.3.3 Induction on degrees of polynomials 34
2.4 Polynomials with small Hessian determinant 37
2.4.1 A reduction 38
2.4.2 The representing line 40
2.5 Proof of Proposition 2.1.2 42
2.5.1 Two simple cases 43
2.5.2 A bootstrapping argument 44
2.5.3 Induction on scale 48
2.6 Variants of decoupling inequalities 50

3 Affine restriction theory **53**
3.1 Overview 53
3.2 An equivalent version 55
3.3 A scaling argument 56
3.4 Proof of Proposition 3.1.1 58
3.4.1 The set with tiny Gaussian curvature 58
3.4.2 Proof of the extra damped estimate 59
3.4.3 Proof of the estimate with affine surface measure 60
3.5 A counterexample 61

References **63**

Chapter 1

Introduction

This dissertation concerns the theory of decoupling and affine restriction for surfaces in $\mathbb{R}^3$. This introductory chapter attempts to survey the recent development of the two theories and the results based on two recent preprints [44, 46]. The exact formulation and proof are presented in Chapters 2 and 3 for decoupling and affine restriction theories, respectively.

1.1 Decoupling theory

Decoupling theory originated from Wolff [74] as a powerful tool to study local smoothing type estimates. We now describe the general setting of the decoupling theory for hypersurfaces. Consider a smooth compact hypersurface $\mathcal{M}$ in $\mathbb{R}^n$. By a partition of unity, we may assume without loss of generality that $\mathcal{M}$ is given by the graph of a smooth function $\phi : [-1,1]^{n-1} \to \mathbb{R}$. Let $d\mu_{\mathcal{M}}$ be a measure defined on $\mathcal{M}$ and $d\mu$ be the push-forward measure of $d\mu_{\mathcal{M}}$ under the projection map $(\xi, \phi(\xi)) \mapsto \xi$. We assume that the measure $d\mu$ is absolutely continuous with respect to the Lebesgue measure $d\xi$. For any $g \in L^1([-1,1]^{n-1}, d\mu)$, define the extension operator $E_A = E_A^{\phi,\mu}$ on $A \subseteq [-1,1]^{n-1}$ with measure $d\mu$:

$$E_A g(x', x_n) = \int_A g(\xi) e^{2\pi i(\xi\cdot x' + \phi(\xi)x_n)} d\mu(\xi), \quad x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}. \tag{1.1}$$

We define the extension operator $E := E_{[-1,1]^{n-1}}$.

In decoupling theory, the Radon–Nikodym derivative of the measure $d\mu$ can be absorbed into g. In this subsection, we will consider the Lebesgue measure $d\xi$ for simplicity. The pullback of $d\xi$ under the projection map agrees with the surface measure up to a constant depending only on the supremum of $|\nabla\phi|$ over $[-1,1]^{n-1}$.

Let $\mathcal{P}$ be a finitely overlapping covering of $[-1,1]^{n-1}$. We want to understand how different pieces $E_\omega g$, for $\omega \in \mathcal{P}$, interact with each other. These can be measured by the L^p norm of Eg over a large ball. More precisely, for $p, q \geq 2$ and $R \gg 1$, our goal is to determine the smallest constant $Dec_{p,q}(\phi, R, \mathcal{P})$ such that the following $\ell^q L^p$ decoupling inequality holds for all $B \subset \mathbb{R}^n$ of radius at least R and $g \in L^1([-1,1]^{n-1})$:

$$\|Eg\|_{L^p(B)} \leq Dec_{p,q}(\phi, R, \mathcal{P}) \left(\sum_{\omega\in\mathcal{P}} \|E_\omega g\|^q_{L^p(w_B)} \right)^{1/q}, \tag{1.2}$$

where w_B is a weight with exponentially decaying tails outside B. See equation (1.25) below for the exact formulation. To avoid technicalities, we here present a morally correct version that treats w_B as the characteristic function on B.

The uncertainty principle tells us that localizing the physical observation of Eg on a ball of radius R blurs its Fourier transform at scale R^{-1}. We therefore shall consider the R^{-1} neighborhood of the graph of ϕ over $[-1,1]^{n-1}$. This principle is made precise in Section 3.2.

With no extra assumption on ϕ and some mild assumptions[1] on $\mathcal{P}$, we have the following general upper bound by interpolating Plancherel's identity and the trivial L^∞ estimate:

$$Dec_{p,q}(\phi, R, \mathcal{P}) \lesssim |\#\mathcal{P}|^{1-1/p-1/q} \tag{1.3}$$

for $2 \leq q \leq p \leq \infty$. Here and throughout this dissertation, we say $A \lesssim_{\alpha,\beta,\ldots} B$ if there exists a constant C depending on parameters $\alpha, \beta, \ldots$, dimension n, Lebesgue exponents p, q and degree k, that will be introduced later, such that $A \leq CB$. This implicit constant

[1] For instance, there are $O(1)$ sub-collections of $\mathcal{P}$ such that each sub-collection contains balls having pairwise disjoint doubles.

C may change from line to line.

Let $\mathcal{M}$ be a flat hypersurface that lies in R^{-1} neighborhood of a hyperplane. Let $\mathcal{P}$ be a partition of $[-1,1]^{n-1}$ into disjoint congruent rectangular slabs having $n-2$ sides of length 1. Then, the general decoupling inequality (1.2) turns out to be the best we can say for $\mathcal{M}$ and $\mathcal{P}$. See for instance Proposition 9.5 and Exercise 12.13 of [18] for cases when $q=2$ and $p=q$ respectively. Other cases follow similarly.

The opposite scenario is when the compact hypersurface $\mathcal{M}$ is far from being a hyperplane. This is measured by the lower bound of the Hessian determinant of ϕ, which is roughly the Gaussian curvature of $\mathcal{M}$ up to a constant depending on the upper bound of $\nabla\phi$. Bourgain and Demeter proved the following sharp decoupling inequalities.

Theorem 1.1.1 (Bourgain-Demeter's decoupling inequalities, [10, 11])**.** *Let* $2 \le p \le \frac{2(n+1)}{n-1}$ *and* $\varepsilon > 0$. *Let* ϕ *be a smooth function on* $[-1,1]^2$. *Suppose that* $|\det D^2\phi| \gtrsim 1$ *and* $\mathcal{P}$ *is a collection of finitely overlapping squares of side length* $R^{-1/2}$ *that covers* $[-1,1]^2$. *Then, there exist* C_1, C_2 *that depend only on* ε, n, $\|\phi\|_{C^3}$ *and the infimum of* $|\det D^2\phi|$ *such that*

$$C_1|\#\mathcal{P}|^{1/2-1/p} \le Dec_{p,p}(\phi,\delta,\mathcal{P}) \le C_2R^{\varepsilon}|\#\mathcal{P}|^{1/2-1/p}. \tag{1.4}$$

Moreover, if $D^2\phi$ *is positive definite, then we have*

$$1 \le Dec_{p,2}(\phi,\delta,\mathcal{P}) \le C_2R^{\varepsilon}. \tag{1.5}$$

It is worth noting that the lifting of each $\omega \in \mathcal{P}$ onto the surface $\mathcal{M}$ lies in a CR^{-1} neighborhood of the tangent plane of the surface at the center of ω, denoted by $c(\omega)$. More precisely, ω satisfies

$$\sup_{\xi\in\omega} |\phi(\xi) - \phi(c(\omega)) - \nabla\phi(c(\omega))\cdot(\xi - c(\omega))| \lesssim R^{-1}. \tag{1.6}$$

We call parallelograms ω satisfying (1.6) (ϕ, R^{-1})-flat. As mentioned earlier, no nontrivial decoupling other than (1.3) can happen on ω. Therefore, this is a canonical covering for hypersurfaces $\mathcal{M}$ with non-vanishing Gaussian curvature.

Decoupling inequalities of the form (1.2) can also be formulated for curves and, more generally, manifolds of higher co-dimensions. Theorem 1.1.1 and its variants have rich applications in harmonic analysis, PDE, and number theory. See, for example, [12, 23, 24, 25, 30, 36, 49] for curves and [3, 10, 19, 26, 27, 28, 29, 31] for more general manifolds.

For hypersurfaces, while the case when ϕ has a non-vanishing Hessian determinant is settled in Theorem 1.1.1, the general decoupling theory for smooth hypersurfaces of vanishing Gaussian curvature is still under-developed. In $\mathbb{R}^2$, a smooth hypersurface is just a smooth curve, and its decoupling theory is relatively well-understood. In Section 12.6 of [18], Demeter proved a slightly more refined decoupling inequality for every compact analytic curve, with the partition chosen to be adapted to the curve as in (1.6). Later, in [75], Yang proved a decoupling inequality for the family of all polynomials of degree at most d and with coefficients bounded by 1, with the decoupling constant depending only on d but not the individual polynomial. This, together with a brute-force Taylor approximation described in Section 2.2.2, implies a decoupling inequality for every smooth compact curve, further generalizing the result in [18]. The first part of this dissertation extends the above results to smooth surfaces in $\mathbb{R}^3$.

In the rest of this subsection, we consider $n = 3$. The range of p in Theorem 1.1.1 is now $2 \leq p \leq 4$. Prior to this dissertation, various partial progress has been made. In [14], Bourgain, Demeter, and Kemp proved decoupling inequalities for all real-analytic surfaces of revolution in $\mathbb{R}^3$. Later in [42, 43], Kemp respectively proved decoupling inequalities for surfaces with constantly zero Gaussian curvature but without umbilical points and for a broad class of C^5 surfaces in $\mathbb{R}^3$ lacking planar points. Recently, Yang and the author proved in [45] a decoupling inequality for surfaces given as graphs of mixed-homogeneous polynomials in $\mathbb{R}^3$. Note that none of the previous partial results implies any of the others.

We now state our main decoupling result.

Theorem 1.1.2 (Yang-L., [46]). *Let $2 \leq p \leq 4$, $n = 3$ and $R \gg 1$. Let ϕ be a smooth function on $[-1, 1]^2$. There exists a family $\mathcal{P}$ of parallelograms ω covering $[-1, 1]^2$ such* *that for every* $\varepsilon > 0$,

1. *each* $\omega \in \mathcal{P}$ *is* (ϕ, R^{-1})*-flat, i.e. it satisfies* (1.6)*;*

2. $\mathcal{P}$ *has* $O_\varepsilon(R^\varepsilon)$*-bounded overlap in the sense that* $\sum_{\omega\in\mathcal{P}} 1_T \lesssim_{\phi,\varepsilon} R^\varepsilon$*;*

3. *We have the following* $\ell^p(L^p)$ *decoupling inequality:*

$$Dec_{p,p}(\phi, R, \mathcal{P}) \lesssim_{\phi,\varepsilon} R^\varepsilon |\#\mathcal{P}|^{1/2-1/p}. \tag{1.7}$$

If, in addition, $D^2\phi$ *is positive semidefinite on* $[-1,1]^2$*, then* (1.7) *can be strengthened to the* $\ell^2(L^p)$ *inequality:*

$$Dec_{p,2}(\phi, R, \mathcal{P}) \lesssim_{\phi,\varepsilon} R^\varepsilon. \tag{1.8}$$

Moreover, the implicit constants in (1.7) *and* (1.8) *can be made uniform over all polynomials* ϕ *of degree up to* k *with bounded coefficients.*

Note that (1.7) and (1.8) agree, up to R^ε loss, with the lower bounds in (1.4) and (1.5), respectively. Thus, up to R^ε loss, we prove the optimal $\ell^p(L^p)$ and $\ell^2(L^p)$ decoupling inequalities for general smooth compact surfaces in $\mathbb{R}^3$ and $2 \le p \le 4$.

1.2 Affine restriction theory

Originated from a discovery of Stein in 1967, the restriction problem [61] asks the following: for which hypersurfaces $\mathcal{M} \subseteq \mathbb{R}^n$ and which $1 \le p' \le 2$, can the Fourier transform of an $L^{p'}(\mathbb{R}^n)$ function can be meaningfully restricted? More precisely, given a hypersurface $\mathcal{M} \subset \mathbb{R}^n$ equipped with a measure $\mu_{\mathcal{M}}$, what are the pairs of exponents $(p', q') \in [1,\infty]^2$ such that the following inequality holds:

$$\|f|_{\mathcal{M}}\|_{L^{q'}(\mathcal{M}, d\mu_{\mathcal{M}})} \lesssim_{\mathcal{M},\mu_{\mathcal{M}},p',q'} \|f\|_{L^{p'}(\mathbb{R}^n)}, \tag{1.9}$$

for all $f \in L^{p'}(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$.

In this dissertation, we consider only compact hypersurfaces. As in the decoupling theory, we assume that $\mathcal{M}$ is given by the graph of a smooth function $\phi : [-1,1]^{n-1} \to \mathbb{R}$ and $d\mu$ is the pushforward measure of $d\mu_{\mathcal{M}}$ under the projection map $(\xi, \phi(\xi)) \mapsto \xi$.

Recall from (1.1) that the extension operator $E_A = E_A^{\phi,\mu}$ on $A \subseteq [-1,1]^{n-1}$, with measure $d\mu$, is defined by

$$E_A g(x', x_n) = \int_A g(\xi) e^{2\pi i (\xi \cdot x' + \phi(\xi) x_n)} d\mu(\xi), \quad x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}, \tag{1.10}$$

for $g \in L^1([-1,1]^{n-1}, d\mu)$. Recall as well that $E = E_{[-1,1]^{n-1}}$.

The extension operator E is dual to the restriction operator $f \mapsto \hat{f}|_{\mathcal{M}}$. In the literature, the restriction problem is usually reduced to studying the equivalent extension estimates of the form:

$$\|Eg\|_{L^p(\mathbb{R}^n)} \lesssim_{\phi,\mu,p,q} \|g\|_{L^q(d\mu)}, \tag{1.11}$$

where $\frac{1}{p} + \frac{1}{p'} = \frac{1}{q} + \frac{1}{q'} = 1$. Here and throughout this dissertation, we adopt the notation that $\frac{1}{\infty} = 0$.

Harmonic analysts have spent enormous efforts solving for the best range of exponents (p, q) for which (1.11) holds. The problem has only been solved for various low-dimensional cases. These include circles $\mathbb{S}^1$ by Fefferman [22] (who credits Stein) and Zygmund [76], three, four and five-dimensional cones by Taberner [66], Wolff [73] and Ou-Wang [50] respectively. For higher dimensions, the problem is still largely open. Various methods have been developed. They include bilinear methods [4, 67, 68, 69, 70, 73], multilinear methods [5, 6, 7, 8, 9, 13, 33, 35], and polynomial methods [32, 34, 37, 41, 50].

Using tools only applicable to Hilbert spaces, the case where $q = 2$ is significantly simpler. Nevertheless, the following theorem is significant to the study of Schrödinger equations.

Theorem 1.2.1 (Tomas-Stein, [60, 71])**.** *Let $p \geq \frac{2(n+1)}{n-1}$. Let ϕ be a smooth function on $[-1,1]^{n-1}$ and $E = E_{[-1,1]^{n-1}}$ be as in* (1.10)*. Suppose that $|\det D^2\phi| \gtrsim 1$ and $d\xi$ is the $n-1$ dimensional Lebesgue measure. There exists C that depends only on p, n, $\|\phi\|_{C^3}$*

and the infimum of $|\det D^2\phi|$ *such that*

$$\|Eg\|_{L^p(\mathbb{R}^n)} \leq C\|g\|_{L^2(d\xi)}, \tag{1.12}$$

for any $g \in L^2([-1,1]^{n-1}, d\xi) \cap L^1([-1,1]^{n-1}, d\xi)$.

Theorem 1.2.1 is a special case of the Strichartz estimates [65], in which mixed norm Lebesgue spaces are considered in place of $L^p(\mathbb{R}^n)$ in (1.12).

The range of p is optimal. This can be shown by Knapp example, which we now describe. Note that rotation and translation of the surface $\mathcal{M}$ correspond to rotation and modulation in physical space. Since these operations do not change the norm we are considering, we may assume without loss of generality that $\phi(0) = 0$ and $\nabla\phi(0) = 0$. Let $\delta \ll 1$ be a small number. Let g be the characteristic function on the square $\omega := [-\delta^{1/2}, \delta^{1/2}]^{n-1}$. This means the convex hull of the lifting of ω onto the surface $\mathcal{M}$ is essentially an axis-parallel box of dimensions roughly $\delta^{1/2} \times ... \times \delta^{1/2} \times \delta$. For $x' \in \mathbb{R}^{n-1}$ with $|x'| < \delta^{-1/2}/(10n)$ and $x_n \in [-\delta^{-1}/(10n), \delta^{-1}/(10n)]$, we have $|\xi \cdot x' + \phi(\xi)x_n| < 1/10$. Thus, for $x = (x', x_n)$

$$|Eg(x)| = |E_\tau g(x)| \geq \left| \int_\tau \cos(1/10) d\mu(\xi) \right| \sim \mu(\tau) = \delta^{(n-1)/2},$$

and

$$\|Eg\|_{L^p(\mathbb{R}^n)} \gtrsim_n \delta^{(n-1)/2}\delta^{-(n+1)/(2p)}.$$

On the other hand,

$$\|g\|_{L^2([-1,1]^2, d\mu)} = \mu(\tau)^{1/2} \sim \delta^{(n-1)/4}.$$

By sending $\delta \to 0$, we see that (1.12) can hold only if $\frac{n-1}{2} - \frac{n+1}{2p} \geq \frac{n-1}{4}$, which is equivalent to $p \geq \frac{2(n+1)}{n-1}$.

It is worth noting that ω satisfies 1.6 and is (ϕ, δ)-flat. Moreover, one can repeat the above calculation for squares ω of side length $\sim \delta^{1/2}$ at any point on ω, since all these ω

are (ϕ, δ)-flat.

We now turn to general smooth hypersurfaces $\mathcal{M}$. Without the Gaussian curvature assumption $|\det D^2\phi| \gtrsim 1$, the size of ω is different at different points of the surface $\mathcal{M}$. Thus, the necessary condition on the Lebesgue exponent $p \geq p_\phi$ differs for different $\mathcal{M}$ if we insist on using the Lebesgue measure. The goal is then to determine the smallest p_ϕ so that (1.12) holds for $p \geq p_\phi$ for some C depending on ϕ. When $n = 2$, this problem is completely settled. It follows from the methods of Stein and Tomas. When $n = 3$, this becomes significantly harder because the set where the Hessian determinant $\det D^2\phi$ vanishes is much more complicated. Nevertheless, for a large class of smooth functions ϕ, including those that are analytic, p_ϕ is fully determined in [39], see also [40, 47, 51, 53, 72] for related results and developments.

In this dissertation, we take up a different question. We introduce the affine surface measure μ^0 defined by

$$d\mu^0(\xi) = |\det D^2\phi(\xi)|^{\frac{1}{n+1}} d\xi, \tag{1.13}$$

such that the extension estimate

$$\|Eg\|_{L^{\frac{2(n+1)}{n-1}}(\mathbb{R}^n)} \leq C\|g\|_{L^2(d\mu)}, \tag{1.14}$$

is invariant under affine transformation that sends the surface $\mathcal{M}$ to $\mathcal{M}'$.

Heuristically, μ^0 puts small weight on the parts of $\mathcal{M}$ with small Gaussian curvature and hence mitigates the obstruction to (1.12) due to the lack of curvature conditions.

In the literature, the restriction estimates with affine arclength measure for curves in $\mathbb{R}^n$ are well understood. See [1, 2, 20, 21, 64]. For uniform results for polynomial curves up to degree k, see [17, 55, 56, 59, 64]. On the other hand, the affine surface measure has also been considered in [15, 57, 58] for convex surfaces of revolutions in $\mathbb{R}^3$, in [16] for surfaces in $\mathbb{R}^3$ given by the graphs of homogeneous polynomials, and in [52] for surfaces in $\mathbb{R}^3$ given by the graphs of mixed homogeneous polynomials. Other notable results include [38, 48, 63].

We successfully solve the L^2 affine restriction problem in $\mathbb{R}^3$ for smooth surfaces. For polynomial surfaces, the result is uniform for polynomials of bounded degree. The primary tool comes from the decoupling theory. It is worth noting that decoupling inequalities are also invariant under affine transformation.

To state our result, we introduce a slightly smaller μ^ε defined by

$$d\mu^\varepsilon(\xi) = |\det D^2\phi(\xi)|^{\frac{1}{n+1}+\varepsilon} d\xi, \quad \text{for any } \varepsilon > 0. \tag{1.15}$$

The measure μ^ε overdamps the variance of the Gaussian curvature of $\mathcal{M}$

Theorem 1.2.2 (L., [44])**.** *Let $n = 3$, $\varepsilon > 0$ and $R \gg 1$. Let ϕ be a smooth function on $[-1,1]^{n-1}$, and $E^{\phi,\mu} = E^{\phi,\mu}_{[-1,1]^{n-1}}$ be as in* (1.10)*. Let $d\mu^0$, $d\mu^\varepsilon$ be defined in* (1.13) *and* (1.15)*, respectively. For any $g \in \mathcal{S}$, we have*

$$\|E^{\phi,\mu^0} g\|_{L^4(B)} \lesssim_{\varepsilon,\phi} R^\varepsilon \|g\|_{L^2(d\mu^0)}, \quad \textit{for any ball } B \textit{ of radius } R, \tag{1.16}$$

$$\|E^{\phi,\mu^\varepsilon} g\|_{L^4} \lesssim_{\varepsilon,\phi} \|g\|_{L^2(d\mu^\varepsilon)}, \tag{1.17}$$

and

$$\|E^{\phi,\mu^0} g\|_{L^p} \lesssim_{\varepsilon,\phi} \|g\|_{L^2(d\mu^0)}, \quad \textit{for any } p > 4. \tag{1.18}$$

Moreover, the implicit constants in (1.16)*,* (1.17)*, and* (1.18) *are uniform over all polynomials ϕ of degree at most k with bounded coefficients.*

By abuse of notation, we define the measure $d\mu^{-1/4}$ on S to be the pullback of the two-dimensional Lebesgue measure $d\xi$. Then by trivial estimates, we have for any $g \in \mathcal{S}$,

$$\|E^{\phi,\mu^{-1/4}} g\|_{L^\infty(d\mu_{-1/4})} \lesssim \|g\|_{L^2(d\mu^{-1/4})} \tag{1.19}$$

and

$$\|E^{\phi,\mu} g\|_{L^\infty(d\mu^0)} \lesssim \|g\|_{L^1(d\mu^0)}. \tag{1.20}$$

By complex interpolation, see for instance [62], among (1.17), (1.19) and (1.20), we get the following affine restriction estimates off the scaling line for affine surface measure μ.

Corollary 1.2.3. *Let $1 < q \le 2$ and $\frac{2}{p} < 1 - \frac{1}{q}$. Let $E^{\phi,\mu}$ and μ^0 be as in Theorem 1.2.2. For any $g \in \mathcal{S}$, we have*

$$\|E^{\phi,\mu} g\|_{L^p(d\mu^0)} \lesssim_\phi \|g\|_{L^q(d\mu^0)}. \tag{1.21}$$

Moreover, the implicit constant in (1.21) *can be made uniform over all polynomials ϕ of degree up to k with bounded coefficients.*

We remark that the ε losses in (1.16) and (1.17), and the off-scaling line conditions $p > 4$ in (1.18) and $2/p < 1 - 1/q$ in (1.21) are necessary for the case of general smooth surfaces. This may be seen by considering the highly oscillatory function

$$\phi(\xi) = \begin{cases} e^{-1/|\xi|} \sin(|\xi|^{-3}) \quad \text{if } \xi \neq 0; \\ 0 \quad \text{if } \xi = 0, \end{cases} \tag{1.22}$$

modified from Sjölin's two-dimensional example in [59]. See Section 3.5.

1.3 Proof Strategies

In this section, we briefly describe the strategies to prove Theorem 1.1.2 and 1.2.2.

1.3.1 Passing to polynomials

Let $n = 3$. In all of our results, we can afford to lose R^ε. We shall see later that such a loss can be transferred to the measure μ^ε in (1.17) and the non-endpoint exponent $p > 4$ in (1.18). In what follows, we write $A \lessapprox B$ to mean $A \lesssim_\varepsilon R^\varepsilon B$ for any $\varepsilon > 0$.

The first reduction is to approximate a smooth function ϕ by its Taylor polynomial of degree $2/\varepsilon$ on each square of side length $R^{-\varepsilon/2}$. We trivially decompose, via the trivial decoupling inequality (1.2), our domain $[-1, 1]^2$ into a finitely overlapping cover of these squares of side length $R^{-\varepsilon/2}$. It suffices to prove Theorem 1.1.2 for all polynomials ϕ of degrees up to k with bounded coefficients, with implicit constants depending on k but not the specific polynomial ϕ. The details are in Section 2.2.2. In the rest of this subsection, we assume that ϕ is a polynomial of degree at most k with bounded coefficients.

1.3.2 Iterative decoupling

Note that decoupling inequalities of the form (1.2) can be iterated. More precisely, suppose that ω_0 can be decoupled into $\omega_1 \in \mathcal{P}_1$ at a cost of D_1, i.e.

$$\|E_{\omega_0} g\|_{L^p(w_B)} \leq D_1 \left(\sum_{\omega_1 \in \mathcal{P}_1} \|E_{\omega_1} g\|^q_{L^p(w_B)} \right)^{1/q},$$

where w_B is defined in (1.25), and each $\omega_1 \in \mathcal{P}_1$ can be decoupled into $\omega_2 \in \mathcal{P}_2(\omega_1)$ at an uniform cost of D_2, i.e.

$$\|E_{\omega_1} g\|_{L^p(w_B)} \leq D_2 \left(\sum_{\omega_2 \in \mathcal{P}_2(\omega_1)} \|E_{\omega_2} g\|^q_{L^p(w_B)} \right)^{1/q}.$$

Then, ω_0 can be decoupled into $\omega_2 \in \mathcal{P}_2 := \cup_{\omega_1} \mathcal{P}_2(\omega_1)$ at a cost of $D_1 D_2$, i.e.

$$\|E_{\omega_0} g\|_{L^p(w_B)} \leq D_1 D_2 \left(\sum_{\omega_2 \in \mathcal{P}_2} \|E_{\omega_2} g\|^q_{L^p(w_B)} \right)^{1/q}.$$

See Proposition 9.17 of [18] for the proof.

For this reason, we use the phrase "ω can be decoupled into $\omega' \in \mathcal{P}$ at the cost of D" to mean inequalities of the form above. We usually keep track of all the constant losses after describing the iteration.

Moreover, if the number of elements in $\mathcal{P}$ is bounded by insignificant constants, including those of order $O_\varepsilon(\delta^{-\varepsilon})$ for any $\varepsilon > 0$, we use triangle and Hölder inequalities to get:

$$Dec_{p,q}(\phi, R, \mathcal{P}) \lesssim |\#\mathcal{P}|^{1-1/q}. \tag{1.23}$$

The loss from the above trivial estimate is also of order $O_\varepsilon(\delta^{-\varepsilon})$ for any $\varepsilon > 0$, which is tolerable. An important situation in which we can apply this procedure is the dyadic decomposition by Gaussian curvature which we describe below.

1.3.3 Dyadic decomposition

Since Bourgain-Demeter's decoupling theorem (Theorem 1.1.1) and Tomas-Stein theorem (Theorem 1.2.1) already cover the case where $|\det D^2\phi| \gtrsim 1$, our enemy is the set where $\det D^2\phi$ is small. Therefore, the first step is to dyadically decompose our domain $[-1,1]^{n-1}$ into sets of the form $\{|\det D^2\phi| \sim \sigma\}$ for some dyadic number $R^{-C} < \sigma \leq 1$ and absolute constant C, depending on degree k, to be determined.

We take a detour to look at the case when $n = 2$. In this case, the sets $\{|\det D^2\phi| \sim \sigma\} = \{|\phi''| \sim \sigma\}$ are significantly simpler: they are unions of at most $2k$ intervals. By the trivial decoupling inequality, it suffices to consider one of these at a tolerable loss depending only on k and the exponents p, q. The part of the curve $\mathcal{M}$ over these intervals can be directly rescaled to part of a parabola. Since our problems are affine invariant and they are solved in parabolic cases, this closes the proof.

1.3.4 Projection and 2D general decoupling

Our strategy for $n = 3$ is more complex than the $n = 2$ case. The first major obstacle is that the sub-level sets $\{|\det D^2\phi| \sim \sigma\}$ are in generally curved regions that cannot be rescaled to the unit square, as opposed to the rescaling of intervals to the unit interval when $n = 2$. We follow similar ideas in [42, 45] to project the part of $\mathcal{M}$ lying over $\{|\det D^2\phi| \sim \sigma\}$ down to a curved region in $\mathbb{R}^2$. Although a projection forgets certain geometry of the surface, the decoupling of the projected sets, as a subset of $\mathbb{R}^2$, is easier to study.

Write $P = \det D^2\phi$. Note that P is a polynomial of degree at most $2k$ and with bounded coefficients. We will decouple the projected set $\{|P| < \delta\} \subset [-1,1]^2$ into parallelograms. We call this generalized 2D uniform decoupling. The special case where $|P_y| \sim 1$ is more manageable because it is essentially a δ-neighborhood of an algebraic curve. We apply the Pramanik-Seeger iteration [54] to approximate the curve by graphs of some carefully chosen polynomial-like functions in different scales. We then prove this special case by applying the uniform decoupling on polynomial-like functions in $\mathbb{R}^2$ [75]. For the general case, we use induction on the degree of the polynomial. Morally, the set $\{|\nabla P| \sim \sigma\}$, for some dyadic number σ, can be treated as sub-level sets of polynomials of one degree less and thus can be decoupled into parallelograms by the induction hypothesis. Correct applications of rescaling arguments led us to the known case where $|\nabla P|$ is bounded below. The rigorous argument requires an in-depth analysis of the geometric properties of the sets between consecutive steps.

1.3.5 Surfaces with small Gaussian curvature on the entire domain

By 2D uniform decoupling, we are able to localize ϕ to parallelograms with small Hessian determinants. After rescaling these parallelograms to the unit square, the problem is reduced to studying the polynomial surfaces with essentially constant Gaussian curvatures. By the iterative structure of the decoupling inequality (1.2), it suffices to decouple each of these rescaled pieces further. We denote the rescaled function $\tilde{\phi}$.

For $n = 2$, one can always divide $\tilde{\phi}$, which corresponds to rescaling the last variable in the frequency space so that the new function has Hessian determinant bounded below. This does not work for $n = 3$. The best we can say when the Gaussian curvature is small on the entire unit square is that there exists a direction along which the surface can be projected into a tiny neighborhood of the graph of a polynomial in one variable. Quantitatively, if $\tilde{\phi}$ has Hessian determinant $\sim \nu \ll 1$ on $[-1, 1]^2$, it admits the form

$$\tilde{\phi} \circ \rho = A(\xi_1) + \nu^{\alpha} B(\xi_1, \xi_2) \tag{1.24}$$

for some rotation ρ, $\alpha = \alpha(k) \in (0, 1)$ and polynomials A, B with bounded coefficients.

It is remarked that (1.24) is no longer true for non-polynomial ϕ. For example, the cone has zero Gaussian curvature and hence function ϕ representing the truncated cone $\{(\xi, |\xi|) : |\xi| \in [1, 2]\}$ has zero Hessian determinant. Nevertheless, ϕ does not admit the form (1.24).

1.3.6 Induction on scale

Now, we may approximate $\tilde{\phi} \circ \rho$ by the graph of a cylinder $(\xi, A(\xi_1))$. This allows us to apply a cylindrical decoupling to $\tilde{\phi} \circ \rho$ at the scale ν^{α}. Each decoupled piece can be enlarged to a polynomial surface living in the unit ball and having Gaussian curvature $\sim \nu^{1-\alpha/2}$. Note that we start with a surface having Gaussian curvature $\sim \nu \ll \nu^{1-\alpha/2}$. Iteratively, we arrive at the situation that all the resulting surfaces have Gaussian curvature bounded away from 0. We then conclude Theorem 1.1.2 by applying Bourgain-Demeter's decoupling result (Theorem 1.1.1) to each resulting surface.

1.3.7 Affine restriction estimates

To prove affine restriction results for all smooth surfaces, including those having principle curvatures of opposite signs, we should not apply Bourgain-Demeter's decoupling result (Theorem 1.1.1) to surfaces having Gaussian curvature bounded away from 0. The loss $|\#\mathcal{P}\|^{1/2-1/4}$ in (1.4) cannot be tolerated in the sharp restriction result. We instead apply the Tomas-Stein inequality (Theorem 1.2.1) directly to these surfaces. It is worth noting that except for the last step where we applied decoupling to surfaces having Gaussian curvature bounded away from 0, all decouplings we used are cylindrical and is $\ell^2(L^4)$. By affine invariance, we rescale back all these pieces and add them up in ℓ^2 norms to obtain the desired estimate (1.16).

For the over-damped situation, the measure compensates an extra factor of σ^{ε} for the region $\{|\det D^2\phi| \sim \sigma\}$. Thus, we shall keep track of all the cylindrical decouplings in the iteration, ensuring the loss is at most $O_{\varepsilon}(\sigma^{-\varepsilon/2})$. This allows us to sum all dyadic pieces and obtain the $L^2 - L^4$ endpoint estimate (1.17).

The non-endpoint result $p > 4$ for the affine surface measure μ in (1.18) is a consequence of interpolating the over-damped estimate (1.17) and the trivial $L^1 - L^{\infty}$ estimate.

1.4 Notations

In this section, we summarize the notations used throughout the dissertation.

1. For integer $n \geq 2$, $\mathcal{M}$ is a smooth compact hypersurface in $\mathbb{R}^n$ defined by the graph of a smooth function $\phi : [-1,1]^{n-1} \to \mathbb{R}$.

2. $x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$ denotes variables in physical space and $(\xi, \eta) \in \mathbb{R}^{n-1} \times \mathbb{R}$ denotes variables in frequency space.

3. Dyadic numbers are numbers of the form $(1+c)^{-k}, k \in \mathbb{Z}_{\geq 0}$, for some positive constant c that will not change throughout the dissertation.

4. $d\xi$ is the $n-1$ dimensional Lebesgue measure. $d\mu^0_{\mathcal{M}}$ and $d\mu^\varepsilon_{\mathcal{M}}$, for $\varepsilon > 0$, are respectively affine surface and over-damped affine surface measures on $\mathcal{M}$ defined by
$$d\mu^0_{\mathcal{M}}(\xi, \phi(\xi)) = |\det D^2\phi(\xi)|^{\frac{1}{n+1}} d\xi,$$
and
$$d\mu^\varepsilon_{\mathcal{M}}(\xi, \phi(\xi)) = |\det D^2\phi(\xi)|^{\frac{1}{n+1}+\varepsilon} d\xi.$$

5. Measures $d\mu$, respectively $d\mu^0$ and $d\mu^\varepsilon$, defined on $[-1,1]^{n-1}$ are the pushforward measure of $d\mu_{\mathcal{M}}$, respectively $d\mu^0_{\mathcal{M}}$ and $d\mu^\varepsilon_{\mathcal{M}}$, under the projection map $(\xi, \phi(\xi)) \mapsto \xi$.

6. For any L^1 function g, define the extension operator $E_A = E_A^{\phi,\mu}$ on $A \subseteq [-1,1]^{n-1}$ with measure $d\mu$:
$$E_A g(x', x_n) = \int_A g(\xi) e^{2\pi i(\xi\cdot x' + \phi(\xi)x_n)} d\mu(\xi), \quad x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}.$$

7. We use the standard notation $A = O_{\alpha,\beta}(B)$, or $A \lesssim_{\alpha,\beta} B$ to mean there is a constant C depending on α, β, dimension n, Lebesgue exponents p, q and degree k that may

change from line to line such that $A \leq CB$. We also use $a \sim_A b$ to mean $a \lesssim_A b$ and $b \lesssim_A a$.

8. For parallelograms $\omega \subseteq [-1,1]^{n-1}$, we denote $c(\omega)$ to be the center of ω. ω is said to be (ϕ, R^{-1})-flat if

$$\sup_{\xi \in \omega} |\phi(\xi) - \phi(c(\omega)) - \nabla\phi(c(\omega)) \cdot (\xi - c(\omega))| \leq R^{-1}.$$

9. Given a ball $B \subset \mathbb{R}^n$ of radius R centered at c, define the weight w_B to be

$$w_B(x) := (1 + \frac{|x-c|}{R})^{-100n}. \tag{1.25}$$

10. For any set $A \subseteq [-1,1]^{n-1}$ and any $\delta > 0$, denote the (vertical) δ-neighbourhood of the graph of ϕ above A by

$$\mathcal{N}_\delta^\phi(A) = \{(\xi, \eta) : \xi \in A, |\eta - \phi(\xi)| < \delta\}. \tag{1.26}$$

11. For $F : \mathbb{R}^n \to \mathbb{C}$ and $A \subseteq [-1,1]^{n-1}$, we denote by F_A the Fourier restriction of F to the strip $A \times \mathbb{R}$, namely, F_A is defined by the relation

$$\widehat{F_A}(\xi, \eta) = \hat{F}(\xi, \eta) 1_A(\xi). \tag{1.27}$$

12. For any smooth function ϕ, we define the norm of ϕ to be

$$\|\phi\| := \sup_{[-1,1]^2} |\phi|. \tag{1.28}$$

13. For any polynomial $P : \mathbb{R}^2 \to \mathbb{R}$ defined by $P = \sum_\alpha c_\alpha \xi^\alpha$ of degree at most k, the norm $\|P\|$ defined above can defined equivalently as below:

$$\|P\| := \sup_{[-1,1]^2} |P| \sim_k \max_\alpha |c_\alpha| \sim_k \sum_\alpha |c_\alpha|. \tag{1.29}$$

We say a polynomial P is bounded if $\|P\| \lesssim_k 1$.

14. For parallelogram $\omega \subseteq [-1,1]^2$, define T_ω be the invertible affine map that maps $[-1,1]^2$ to ω. T_ω sends the smooth function ϕ to ϕ_{T_ω} as follows:

$$\phi_{T_\omega}(\xi) := \phi \circ T_\omega(\xi) - \nabla(\phi \circ T_\omega)(0,0) \cdot \xi - \phi \circ T_\omega(0,0). \tag{1.30}$$

We normalise T_ω by

$$\bar{\phi}_{T_\omega} = \frac{\phi_{T_\omega}}{\|\phi_{T_\omega}\|}. \tag{1.31}$$

Note that ω is (ϕ, δ)-flat if and only if $\|\phi_{T_\omega}\| \lesssim R^{-1}$.

15. For parallelogram $\omega \subseteq [-1,1]^2$, the width of ω is the diameter of the largest ball contained in ω.

16. We use the phrase "ω can be $\ell^2(L^p)$ decoupled into $\omega' \in \mathcal{P}$ at a cost of D" to mean: for each $g \in L^1([-1,1]^{n-1}$, we have

$$\|E_\omega g\|_{L^p(w_B)} \leq D \left(\sum_{\omega' \in \mathcal{P}} \|E_{\omega'} g\|^q_{L^p(w_B)} \right)^{1/q}.$$

$\ell^2(L^p)$ is often omitted if it is clear in the context what p is. If $D \lessapprox 1$ and there is finitely many iterations, we sometimes also omit D.

Chapter 2

Decoupling theory

In this chapter, we prove the decoupling inequalities for compact smooth surfaces in $\mathbb{R}^3$, stated in Theorem 1.1.2.

2.1 Overview

In (1.2), we formulated the decoupling problem in terms of the extension operator. The extension formulation involves more technical details to work with, mainly due to the necessity of weight w_B. We will instead use an alternative formulation in which functions have frequency support on a neighborhood of the surface. Let $D_{p,q}(\phi, \delta, \mathcal{P})$ be the smallest constant such that the following $\ell^q(L^p)$ decoupling inequality holds for all F Fourier supported on $\mathcal{N}_\delta^\phi(\cup\mathcal{P})$:

$$\|F\|_{L^p(\mathbb{R}^n)} \leq D_{p,q}(\phi, \delta, \mathcal{P}) \left(\sum_{\omega \in \mathcal{P}} \|F_\omega\|_{L^p(\mathbb{R}^n)}^q \right)^{1/q}. \tag{2.1}$$

By Proposition 9.15 of [18], Theorem 1.1.2 is implied by the following equivalent version under an extra assumption on the size of ω:

Proposition 2.1.1 (Yang-L., [44, 46])**.** *Let $2 \leq p \leq 4$, $n = 3$, and $0 < \delta \ll 1$. Let $\phi : [-1, 1]^2 \to \mathbb{R}$ be a smooth function. There exists a family $\mathcal{P} = \mathcal{P}(\delta, \phi)$ of parallelograms ω covering $[-1, 1]^2$ such that, for any $\varepsilon > 0$*

1. *each $\omega \in \mathcal{P}$ is (ϕ, δ)-flat;*

2. *$\mathcal{P}$ has $O_\varepsilon(\delta^{-\varepsilon})$-bounded overlap in the sense that $\sum_{\omega\in\mathcal{P}} 1_\omega \lesssim_{\phi,\varepsilon} \delta^{-\varepsilon}$;*

3. *The width of each $\omega \in \mathcal{P}$ is at least δ;*

4. *We have the following $\ell^p(L^p)$ decoupling inequality:*

$$D_{p,p}(\phi, \delta, \mathcal{P}) \lesssim_{\phi,\varepsilon} \delta^{-\varepsilon} |\#\mathcal{P}|^{1/2-1/p}. \tag{2.2}$$

If, in addition, $D^2\phi$ is positive semidefinite on $[-1, 1]^2$, then (2.2) *can be strengthened to the $\ell^2(L^p)$ inequality:*

$$D_{p,2}(\phi, \delta, \mathcal{P}) \lesssim_{\phi,\varepsilon} \delta^{-\varepsilon}. \tag{2.3}$$

Moreover, the implicit constants in (2.2) *and* (2.3) *can be made uniform over all polynomials ϕ of degree up to k with bounded coefficients.*

Conditions 1 and 3 above ensure that the following assumption in Proposition 9.15 of [18] is met: for each $\omega \in \mathcal{P}$, there is a rectangular box R_ω such that

$$R_\omega \subset \omega \quad \text{and} \quad \omega + B(0, \delta) \subset R_\omega + T, \tag{2.4}$$

for $O(1)$ many points T independent of ω and δ. We briefly explain why this assumption is important. In a standard localization argument, we multiply a smooth cutoff function ψ_B, whose Fourier support is in $B(0, \delta)$, to Eg in (1.2), where $R = \delta^{-1}$. Let $F = Eg\psi_B$. Then $\hat{F} = \widehat{Eg} * \hat{\psi}_B$ in distributional sense. Thus, to relate F_ω and Eg_ω, ω needs to satisfy (2.4).

It is remarked that Proposition 2.1.1 is first proven by Yang and the author in [46], without condition 3. In this dissertation, we describe the approach taken by the author in [44]. As a side product of the approach, we prove some variants, Propositions 2.6.1 and 2.6.2, to Proposition 2.1.1, which will be used to prove the affine restriction results in the next chapter.

In Proposition 2.1.1, $\mathcal{P}$ is not allowed to depend on ε. It turns out that it suffices to prove the superficially weaker version in which $\mathcal{P}$ is allowed to depend on ε.

Furthermore, the general smooth case is a corollary of the uniform polynomial case.

The following summarizes what we need to prove.

Proposition 2.1.2. *Let $2 \le p \le 4$, $n = 3$, $\varepsilon > 0$ and $0 < \delta \ll 1$. Let ϕ be a bivariate polynomial of degree at most k with bounded coefficients. There exists a family $\mathcal{P} = \mathcal{P}(\delta, \phi, \varepsilon)$ of parallelograms ω covering $[-1,1]^2$ such that the following hold:*

1. *each $\omega \in \mathcal{P}$ is (ϕ, δ)-flat;*
2. *$\mathcal{P}$ has $O_{k,\varepsilon}(\delta^{-\varepsilon})$-bounded overlap in the sense that $\sum_{\omega\in\mathcal{P}} 1_\omega \lesssim_{k,\varepsilon} \delta^{-\varepsilon}$;*
3. *The width of each $\omega \in \mathcal{P}$ is at least δ;*
4. *We have the following $\ell^p(L^p)$ decoupling inequality:*

$$D_{p,p}(\phi, \delta, \mathcal{P}) \lesssim_{k,\varepsilon} \delta^{-\varepsilon} |\#\mathcal{P}|^{1/2-1/p}. \tag{2.5}$$

If, in addition, $D^2\phi$ is positive semidefinite on $[-1,1]^2$, then (2.5) can be strengthened to the $\ell^2(L^p)$ inequality:

$$D_{p,2}(\phi, \delta, \mathcal{P}) \lesssim_{k,\varepsilon} \delta^{-\varepsilon}. \tag{2.6}$$

The proof of Proposition 2.1.1 from Proposition 2.1.2 is contained in Section 2.2.

The following two propositions are the key ingredients of Proposition 2.1.2.

Proposition 2.1.3 (Generalised 2D uniform decoupling inequality with size estimates)**.** *Let $\varepsilon > 0$, $2 \le p \le 6$, $0 < \delta, \lambda \le 1$ be dyadic numbers and $P : \mathbb{R}^2 \to \mathbb{R}$ be a bounded polynomial of degree at most k. For each dyadic number $\sigma \in [\lambda, 1]$, there exists a family $\mathcal{P}^0_{\sigma,\lambda} = \mathcal{P}^0_{\sigma,\lambda}(\delta, P)$ of parallelograms such that the following statements hold:*

1. *for each λ, $\mathcal{P}^0_\lambda := \cup_\sigma \mathcal{P}^0_{\sigma,\lambda}$, where the union is over dyadic numbers $\sigma \in [\lambda, 1]$, covers*

$[-1,1]^2$ in the sense that

$$\sum_{\omega\in\mathcal{P}_\lambda} 1_\omega(\xi) \geq 1 \quad \textit{for } \xi \in [-1,1]^2;$$

2. *$\mathcal{P}_\lambda^0$ has bounded overlap in the sense that $\sum_{\omega\in\mathcal{P}_\lambda^0} 1_{100\omega} \lesssim_k 1$;*

3. *for each $\omega \in \mathcal{P}_{\sigma,\lambda}^0$, at least one of the following holds:*

   (a) $\lambda < \sigma \leq 1$, $|P| \sim \sigma$ over 10ω, and the width of ω is $\gtrsim \max\{\sigma,\delta\}$;

   (b) $\lambda < \sigma \leq \delta$, $|P| \lesssim \sigma$ over 10ω, and the width of ω is $\sim \delta$.

   (c) $\sigma = \lambda$, $|P| \lesssim \lambda$ over 10ω, and the width of ω is $\gtrsim \max\{\sigma,\delta\}$.

4. *For $\lambda' < \lambda < \sigma$, $\mathcal{P}_{\sigma,\lambda} = \mathcal{P}_{\sigma,\lambda'}$.*

5. *the following 2D $\ell^2(L^p)$ decoupling inequality holds: for any $f : \mathbb{R}^2 \to \mathbb{C}$ whose Fourier transform is supported on $\bigcup_{\omega\in\mathcal{P}_{\sigma,\lambda}^0} \omega$, we have*

$$\|f\|_{L^p(\mathbb{R}^2)} \lesssim_{\varepsilon,k} \sigma^{-\varepsilon} \left(\sum_{\omega\in\mathcal{P}_{\sigma,\lambda}^0} \|f_\omega\|_{L^p(\mathbb{R}^2)}^2 \right)^{1/2}, \tag{2.7}$$

   where f_ω is the frequency projection of f onto ω, defined by $\hat{f_\omega} = \hat{f} 1_\omega$.

Before proceeding, we make several remarks regarding Proposition 2.1.3. First, the introduction of the parameter λ is primarily for induction purposes. When applying Proposition 2.1.3, we will take $\lambda \sim \delta^C$ for some C depending on k. Second, the width estimates on ω in statement 3 are crucial in proving Proposition 2.1.1, which requires a lower bound on the width of ω. Nevertheless, the other lower bound, σ, in items (3a) and (3c), as well as the decoupling constant $\sigma^{-\varepsilon}$, instead of the usual $\delta^{-\varepsilon}$ in (2.7), are for the purpose of the next chapter. (Compare Propositions 2.6.1 and 2.6.2 at the end of the chapter with Propositions 2.1.1 and 2.1.2 above.) Heuristically, we shall not "over" decouple sets in $\mathbb{R}^2$ by stopping at steps where the sub-level set $\{|P| \lesssim \sigma\}$ already has width δ. On the other hand, since P has bounded derivatives, the set $\{|P| \sim \sigma\}$ contains the σ neighborhood of the algebraic varieties $\{|P| = \sigma\}$. Thus, it is expected that the decoupled set has a width of at least σ.

Proposition 2.1.3 will be proved in Section 2.3.

Proposition 2.1.4. *For each $2 \leq k \in \mathbb{N}$, there is a constant $\alpha = \alpha(k) \in (0, 1]$ such that the following holds. Let $\phi : \mathbb{R}^2 \to \mathbb{R}$ be a bounded polynomial of degree at most k, without linear terms. If $\|\det D^2\phi\| \leq \nu \in (0, 1)$, then there exist a rotation $\rho : \mathbb{R}^2 \to \mathbb{R}^2$, and bounded polynomials A, B such that*

$$\phi(\xi) = A \circ \rho(\xi) + \nu^\alpha B \circ \rho(\xi),$$

and A is one-dimensional, i.e. for any $\xi \in \mathbb{R}^2$, $A(\xi_1, \xi_2) = A(\xi_1, 0)$.

Proposition 2.1.4 will be proved in Section 2.4.

In Section 2.5, we prove Proposition 2.1.2 using Propositions 2.1.3 and 2.1.4. This will finish the proof of Proposition 2.1.1 and hence Theorem 1.1.2.

Finally, in section 2.6, we present variants of Propositions 2.1.1 and 2.1.2 for the application of the next chapter.

2.2 Two reductions

In this section, we prove that Proposition 2.1.2 implies Proposition 2.1.1.

2.2.1 Removal of ε-dependence of the covering $\mathcal{P}$

In this subsection, we remove ε-dependence of the covering $\mathcal{P}$ in Proposition 2.1.2. The proof presented here was suggested by Joshua Zahl.

Let $\varepsilon = \varepsilon(\delta)$ be a function such that $\varepsilon \searrow 0$ as $\delta \searrow 0$, with the rate of decay to be determined.

Now with each $\delta > 0$, there exists a family $\mathcal{P}_\delta = \mathcal{P}(\delta, \phi, \varepsilon(\delta))$ of parallelograms ω satisfying the prescribed properties. Note that by definition, $\mathcal{P}(\delta)$ depends on δ only.

Let $\varepsilon > 0$ be arbitrary.

If $\varepsilon < 2\varepsilon(\delta)$, then using the monotonicity of the fixed function $\varepsilon(\delta)$, δ is bounded below by a positive number depending on ε only. In this case, we have a trivial decoupling inequality, with the constant depending on ϕ, ε.

If $\varepsilon \geq 2\varepsilon(\delta)$, then it suffices to find a suitable $C'_{\phi,\varepsilon}$ such that

$$C_{\phi,\varepsilon(\delta)}\delta^{-\varepsilon(\delta)} \leq C'_{\phi,\varepsilon}\delta^{-\varepsilon}, \text{ for all } \varepsilon > 0,\, \delta > 0,$$

where $C_{\phi,\varepsilon}$ is the implicit constant in the ϵ depend version of decoupling. Since $\varepsilon/2 \geq \varepsilon(\delta)$, we have $\delta^{-\varepsilon(\delta)} \leq \delta^{-\varepsilon/2}$, and thus it suffices to show that

$$C_{\phi,\varepsilon(\delta)}\delta^{\varepsilon/2} \leq C'_{\phi,\varepsilon}, \text{ for all } \varepsilon > 0,\, \delta > 0. \tag{2.8}$$

We may assume $C_{\phi,\varepsilon} \nearrow \infty$ as $\varepsilon \searrow 0$. If we choose $\varepsilon(\delta)$ to decrease slowly enough as $\delta \searrow 0$, then we may have $C_{\phi,\varepsilon(\delta)} \leq \log \delta^{-1}$ for all δ small enough. Then for all $\varepsilon > 0$, we have

$$\lim_{\delta \to 0^+} C_{\phi,\varepsilon(\delta)}\delta^{\varepsilon/2} \leq \lim_{\delta \to 0^+} \delta^{\varepsilon/2} \log \delta^{-1} = 0.$$

Thus, by choosing a suitable constant $C'_{\phi,\varepsilon}$, we have (2.8). The ε-bounded overlap follows from the same proof. Thus we have removed the ε dependence of the family $\mathcal{P}$.

2.2.2 Taylor approximation

In this subsection, we prove the decoupling of general smooth functions by using the uniform decoupling inequalities for polynomials of degree k with bounded coefficients. The idea is a standard Taylor polynomial approximation.

Fix $\phi \in C^\infty([-1,1]^2)$. We will first prove the general case without assuming that $D^2\phi$ is positive semidefinite.

Given $\varepsilon > 0$ and $\delta > 0$, we first partition $[-1,1]^2$ into squares Q of side length δ^ε. Since we allow a loss of $\delta^{-\varepsilon}$ in the decoupling inequality, by the triangle and Hölder's inequalities, it suffices to decouple each Q into (ϕ, δ)-flat rectangles.

Let k be the smallest integer greater than or equal to ε^{-1}. For each Q, we may approximate ϕ by its k-th degree Taylor polynomial P_Q, which depends on ε, δ. The error is at most

$$C_k \sup_{[-1,1]^2} |D^{k+1}\phi(x,y)|(\delta^\varepsilon)^{k+1} \leq C_k \sup_{[-1,1]^2} |D^{k+1}\phi(x,y)|\delta^{1+\varepsilon}.$$

For δ small enough (depending on ε), the above error will be less than $\delta/2$. Hence, to find a cover of Q by (ϕ, δ)-flat rectangles, it suffices to find a cover of Q by $(P_Q, \delta/2)$-flat rectangles.

Since all coefficients of P_Q are bounded by a large constant depending on ϕ, ε only, we may normalize P_Q to $\tilde{P}_Q$ with maximum coefficient having magnitude 1, and apply Proposition 2.1.2 to $\tilde{P}_Q$ find a cover $\mathcal{P}_Q = \mathcal{P}_Q(\delta, \tilde{P}_Q)$ of Q that satisfies the prescribed properties. The final collection $\mathcal{P}_\delta$ is simply defined to be the union of the collections $\cup_Q \mathcal{P}_Q$. The overlap between rectangles covering different Q's is trivially bounded since each covering rectangle of Q is a subset of $2Q$, and $2Q$'s have bounded overlap.

Although P_Q depends on δ, our main uniform decoupling Proposition 2.1.2 ensures a uniform bound of all decoupling constants as Q varies, and this uniform bound is independent of δ. Thus, the final decoupling inequality (2.2) follows immediately by the triangle and Hölder's inequalities.

Lastly, in the special case when $D^2\phi$ is positive semidefinite, by choosing the degree of P_Q to be greater than ε^{-3} if necessary, we may approximate ϕ by P_Q such that $\det D^2 P_Q \geq -\delta^3$. By normalising P_Q to $\tilde{P}_Q$ with maximum coefficient having magnitude 1 and choosing δ small enough, we also have $1 \gtrsim_\varepsilon \det D^2 \tilde{P}_Q \geq -\delta^2$. In this case, we get uniform ℓ^2 decoupling inequalities from Proposition 2.1.2. This finishes the proof.

2.3 General 2D decoupling

In this section, we prove Proposition 2.1.3. The proof of the proposition involves three steps. First, we show Proposition 2.1.3 under the assumptions that $P = A(\xi_1) - \xi_2 B(\xi_1)$ and that A/B is polynomial-like, the decoupling inequalities of which have been solved by Yang in [75]. Second, we prove the case when $|\nabla P| \sim \kappa$. In this case, $|P| \sim \sigma$ is roughly the σ/κ neighborhood of the zero set of P. This can be solved by a Pramanik-Seeger iteration by approximating P at each scale by polynomials obeying the condition of the first case. Third, we fully prove Proposition 2.1.3 by induction on the degrees of polynomials P. Note that high enough derivatives of P are constant. Applying the second step to derivatives of P iteratively, we conclude the proof of Proposition 2.1.3.

2.3.1 Uniform decoupling of polynomial-like rational functions

We first introduce the following definition of polynomial-like functions:

Definition 2.3.1. *ψ is said to be a polynomial-like function of degree k over I_0 if, for each $0 < \sigma \ll 1$ and each interval $J \subseteq I_0$, the set*

$$B(\psi, \sigma, J) := \{\xi_1 \in J : |\psi''(\xi_1)| < \sigma(\sup_{\xi_1 \in J} |\psi''(\xi_1)| + |J| \sup_{\xi_1 \in J} |\psi'''(\xi_1)|)\}$$

is a disjoint union of at most $O_k(1)$ subintervals of J, and satisfies

$$|B(\psi, \sigma, J)| \lesssim_k \sigma^{\frac{1}{3k}} |J|.$$

Note that polynomials of degree at most k with bounded coefficients are polynomial-like functions of degree k; see Lemma 1.7 of [75].

We need the following rescaled version of the uniform decoupling for polynomial-like functions from [75].

Proposition 2.3.2 (Uniform decoupling for polynomial-like functions)**.** *Let $k \geq 1$, $\varepsilon > 0$, $0 < \delta \ll 1$, and let ψ be a polynomial-like function of degree k over $I_0 \subseteq [-1, 1]$. There exists a partition I_δ of I_0 such that*

1. *each $I \in I_\delta$ is (ψ, δ)-flat.*
2. *each $I \in I_\delta$ has length at least $\delta^{1/2}$.*
3. *for any $C > 1$, there exists at most $O_C(1)$ many neighboring intervals whose union is $(\psi, C(\delta))$-flat.*

4. *the following 2D $\ell^2(L^p)$ decoupling inequality holds: for any $f : \mathbb{R}^2 \to \mathbb{C}$ Fourier supported in $\mathcal{N}^{\psi}_{O(\delta)}(I_0)$, we have*

$$\|f\|_{L^p(\mathbb{R}^2)} \lesssim_{\varepsilon,k} \delta^{-\varepsilon} \left(\sum_{I \in \mathcal{I}_\delta} \|f_I\|^2_{L^p(\mathbb{R}^2)} \right)^{\frac{1}{2}} .$$

We remark that the above theorem will also be used repeatedly in the bootstrapping argument, Section 2.5.2 below, for the special case when ψ is a polynomial.

Combined with the following lemma, we prove the uniform decoupling inequalities for rational functions.

Lemma 2.3.3. *Let $0 < \kappa \le 1$. Suppose that A, B are univariate polynomials of degree at most k such that $|A| \lesssim \kappa$ and $|B| \sim \kappa$ over an interval $I_0 \subseteq [-1,1]$. Then $\psi := A/B$ is a polynomial-like function of degree $3k$.*

Proof. By rescaling, it suffices to check these conditions for $J = I_0 = [-1,1]$. Then, the coefficients of A and B are bounded by κ. By dividing A and B by κ, we may assume without loss of generality that $\kappa = 1$. By direct computation, $\psi'' = A_1/B^3$ for some polynomial A_1 of degree at most $3k$ such that $|A_1| \lesssim \kappa^3$ over $[-1,1]$. Thus, by the fundamental theorem of algebra, $B(\psi, \sigma, [-1,1])$ is a union of $O_k(1)$ many subintervals of $[-1,1]$.

On the other hand,

$$\sup_{[-1,1]} |\psi'''| = \sup_{[-1,1]} \left| \frac{A_1'}{B^3} - \frac{3A_1 B^2 B'}{B^6} \right| \lesssim \sup_{[-1,1]} |A_1'| + |A_1| \lesssim \sup_{[-1,1]} |A_1|,$$

where the last inequality follows from (1.29) because A_1 is a polynomial.

Thus, $B(\psi, \sigma, J)$ is contained in $\{|A_1| \lesssim \sigma \sup_{[-1,1]} |A_1|\}$. Since degree of A_1 is bounded by $3k$, the size of this set is $\lesssim_k \sigma^{\frac{1}{3k}}$ as desired. $\square$

2.3.2 The case of constant gradient

In this subsection, we consider the case when $|\nabla P| \sim \kappa$ for some $\kappa \ll 1$ over a parallelogram ω_0 of width at least $\kappa^{-1}\sigma$. Our goal is to decouple $\{|P| \leq \sigma\} \cap \omega_0$ into bounded overlap parallelograms of width $\sim \min\{\kappa^{-1}\sigma, 1\}$.

We first need the following structure lemma for the set $\{|P| \leq \sigma\} \cap \omega_0$.

Lemma 2.3.4. *Let $0 < \sigma, \kappa \ll 1$. Let $\omega_0 \subseteq [-1,1]^2$ be an axis-parallel rectangle. Suppose that P is a polynomial of degree at most k such that $|\nabla P| \sim \kappa$ and $|\partial_{\xi_2} P| \sim \kappa$ over ω_0. There exist $O_k(1)$ many disjoint intervals I and smooth functions ψ_I on I, depending on σ and κ, such that*

$$\{|P| \leq \sigma\} \cap \omega_0 \subseteq \bigsqcup_I \mathcal{N}^{\psi_I}_{O(\kappa^{-1}\sigma)}(I), \tag{2.9}$$

and, for each I,

$$\mathcal{N}^{\psi_I}_{O(\kappa^{-1}\sigma)}(I) \cap \omega_0 \subseteq \{|P| \lesssim \sigma\} \tag{2.10}$$

where $B(0,r)$ is the ball of radius r centered at 0.

See Figure 2.1 below for an illustration of this lemma.

Proof. By the fundamental theorem of algebra, there are $O_k(1)$ many ξ_1 on $[-1,1]$ such that $P(\xi_1,\xi_2) = \pm\sigma$ and $(\xi_1,\xi_2) \in \partial\omega_0$ for some ξ_2. These ξ_1 form intervals I on which there exist ξ_2 such that $|P(\xi_1,\xi_2)| \leq \sigma$. Define ψ_I implicitly by $P(\xi_1, \psi_I(\xi_1)) = \pm\sigma$ using implicit function theorem, where the sign is chosen as follows. We use $-\sigma$ (resp. σ) if the graph of ψ_I intersects the top (resp. bottom) of ω_0. If it intersects both the top and the bottom, we cut the interval into halves and choose the sign separated as described above. The function ψ_I is smooth with bounded derivatives because all derivatives of $\partial_{\xi_1} P$ are of magnitude $\lesssim \kappa$.

To see (2.9), let $(\xi_1,\xi_2) \in \{|P| \leq \sigma\} \cap \omega_0$. By construction, there exists I as above such that $\xi_1 \in I$. By the mean value theorem, there exists ξ_2' between ξ_2 and $\psi_I(\xi_1)$ such that

$$2\sigma \geq |P(\xi_1,\xi_2) - P(\xi_1,\psi_I(\xi_1))| = |\partial_{\xi_2} P(\xi_2')||\xi_2 - \psi_I(\xi_1)|.$$

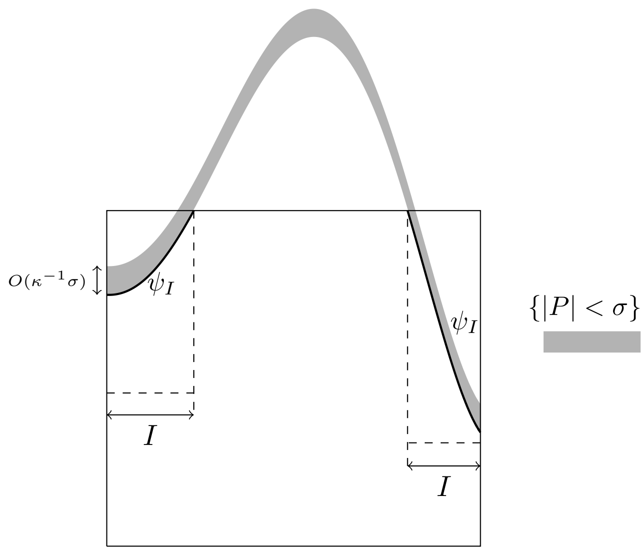

Figure 2.1: Illustration of the structure lemma

The way we construct ψ_I ensures that $(\xi_1, \xi_2') \in \omega_0$, and hence $|\partial_{\xi_2} P(\xi_2')| \sim \kappa$. Therefore, $|\xi_2 - \psi_I(\xi_1)| \lesssim \kappa^{-1}\sigma$.

It remains to show (2.10). Let $(\xi_1, \xi_2) \in \omega_0$ be such that $|\xi_2 - \psi_I(\xi_1)| \lesssim \kappa^{-1}\sigma$. Then, by mean value theorem again, there exists ξ_2' between $\xi_2, \psi_I(\xi_1)$ such that

$$|P(\xi_1, \xi_2)| = |P(\xi_1, \xi_2) - P(\xi_1, \psi_I(\xi_1))| = |\partial_{\xi_2} P(\xi_2')||\xi_2 - \psi_I(\xi_1)| \lesssim \kappa(\kappa^{-1}\sigma) = \sigma$$

as desired. □

The above lemma reduces the decoupling of $\{|P| \leq \sigma\} \cap \omega_0$ to the decoupling of the neighborhood of level sets of polynomials P. Since there are $O_k(1)$ many intervals, it suffices to consider each interval separately.

Proposition 2.3.5. *Let $0 < \sigma \ll 1$ and $I_0 \subseteq [-1, 1]$. Suppose that $\psi : I_0 \to \mathbb{R}$ is a smooth function and $P(\xi_1, \psi(\xi_1)) = 0$ for some polynomial P of degree at most k with bounded coefficients. Moreover, $|\nabla P| \lesssim |\partial_{\xi_2} P| \sim \kappa$ on ω_0 for some axis parallel rectangle ω_0 containing the graph of ψ above I_0. Then, there exists a collection of (ψ, σ)-flat disjoint* *intervals I such that*

1. *the following 2D $\ell^2(L^p)$ decoupling inequality holds: for any $f : \mathbb{R}^2 \to \mathbb{C}$ Fourier supported in $\mathcal{N}_\sigma^\psi(I_0)$, we have*

$$\|f\|_{L^p(\mathbb{R}^2)} \lesssim_{\varepsilon,k} \sigma^{-\varepsilon} \left(\sum_{I \in \mathcal{I}} \|f_I\|^2_{L^p(\mathbb{R}^2)} \right)^{\frac{1}{2}};$$

2. *each I has length at least $\sigma^{1/2}$;*

3. *for any absolute constant $C > 1$, any $\xi_1 \in \mathbb{R}$ stays in $O(1)$ many $I \in \mathcal{I}$.*

Proof. Let $\mathcal{I}_0$ contain the singleton I_0. Let $\sigma_0 \lesssim 1$ be the smallest number such that I_0 is (ψ, σ)-flat. Also, the side of ω_0 that is parallel to the to ξ_2 axis has a length of at least σ_0. Therefore, for any $(\xi_1, \xi_2) \in \omega_0$,

$$|P(\xi_1, \xi_2)| = |P(\xi_1, \xi_2) - P(\xi_1, \psi(\xi_1))| \lesssim \sigma_0 \sup_{\omega_0} |\partial_{\xi_2} P| \lesssim \sigma_0 \kappa. \tag{2.11}$$

We now describe the Pramanik-Seeger iteration. At each step, we approximate ψ locally by rational functions. For each $i = 1, ...$, our goal is to decouple each $I \in \mathcal{I}_i$, which is (ψ, σ_i)-flat, for some $\sigma_i < \sigma_0$.

We consider the following one-dimensional variant of the rescaling (1.30). Let T_I be the translation and rescaling that maps $[-1, 1]$ to I. Consider

$$\psi_{T_I}(\xi_1) := \psi \circ T_I(\xi_1) - \nabla(\psi \circ T_I)(0)\xi_1 - \psi \circ T_I(0),$$

and

$$P_{T_I}(\xi_1, \xi_2) = P(T_I(\xi_1), \xi_2 + \nabla(\psi \circ T_I)(0)\xi_1 + \psi \circ T_I(0))$$

so that $P_{T_I}(\xi_1, \psi_{T_1}(\xi_1)) = 0$.

To decouple ψ over I, it suffices to decouple ψ_{T_I} over $[-1, 1]$.

Write

$$P_{T_I}(\xi_1, \xi_2) = A(\xi_1) - \xi_2 B(\xi_1) + \xi_2^2 E(\xi_1, \xi_2)$$

for some polynomials A, B, E of degree at most k with bounded coefficients.

Note that $\psi_{T_I}(0) = 0$ and $P(0,0) = 0$. Recall from (1.29) that the norm of a polynomial $\|\cdot\|$ is equivalent to the supremum of the polynomial over the unit interval. We may bound A and B by:

$$\|A\| = \sup_{\xi_1 \in [-1,1]} |P_{T_I}(\xi_1, 0)| \lesssim \sup_{\omega_0} |P| \lesssim \sigma_0 \kappa$$

and

$$\|B\| = \sup_{\xi_1 \in [-1,1]} |\partial_{\xi_2} P_{T_I}(\xi_1, 0)| \sim \kappa.$$

Thus, A/B satisfies the assumptions in Lemma 2.3.3 and is polynomial-like of degree $3k$. We now approximate ψ_{T_I} by A/B.

Since I is (ψ, σ_i)-flat, we have $|\psi_{T_I}(\xi_1)| \lesssim \sigma_i$. Moreover, for any $|\xi_2 - \psi_T(\xi_1)| \leq \sigma_0$, we have

$$|\xi_2^2 E(\xi_1, \xi_2)| \leq \sup_{\xi \in \omega_0} |P(\xi)| + \|A\| + \sigma_0 \|B\| \lesssim \sigma_0 \kappa,$$

and hence for any fixed $\xi_1 \in [-1, 1]$,

$$(\sigma_0 \xi_2)^2 E(\xi_1, \sigma_0 \xi_2) \lesssim \sigma_0 \kappa.$$

This means that all coefficients of $(\xi_1, \xi_2) \mapsto E(\xi_1, \sigma_0 \xi_2)$ are bounded by κ/σ_0, and therefore $E(\xi_1, \xi_2) \lesssim \kappa/\sigma_0$ for any $|\xi_2| \leq \sigma_0$.

On the other hand, we have

$$0 = P_{T_I}(\xi_1, \psi_{T_I}(\xi_1)) = A(\xi_1) - \psi_{T_1}(\xi_1) B(\xi_1) + \psi_{T_1}^2(\xi_1) E(\xi_1, \psi_{T_1}(\xi_1)).$$

Using $|\psi_{T_I}(\xi_1)| \lesssim \sigma_i$, we have

$$|\psi_{T_1}(\xi_1) - A/B(\xi_1)| \leq C_0 \sigma_i^2 / \sigma_0.$$

This allows us to approximate $\mathcal{N}_\sigma^{\psi_{T_I}}([-1,1])$ by a $\sigma_{i+1} := \max\{\sigma, C_0 \sigma_i^2/\sigma_0\}$ neighborhood of the graph of A/B, a polynomial-like function of degree $3k$. By Proposition 2.3.2,

$[-1,1]$ can be decoupled into $(A/B, \sigma_{i+1})$-flat disjoint intervals I'. Let $\mathcal{I}_{i+1}(I)$ denote the collection of all $T_I(I')$. Rescaling back, we successfully decouple I into the partition $\mathcal{I}_{i+1}(I)$, each interval of which is (ψ, σ_{i+1})-flat, at a cost of $O_\varepsilon(\sigma_{i+1}^{-\varepsilon})$.

Let $\mathcal{I}_{i+1}$ be the union of $\mathcal{I}_{i+1}(I)$ for all $I \in \mathcal{I}_i$. Since intervals I may be decoupled into incomplete sub-intervals near both ends. In this case, we concatenate the incomplete sub-intervals with its neighbor to ensure that the size of each interval in $\mathcal{I}_{i+1}$ is at least $\sigma_{i+1}^{1/2}$. This is possible at $O(1)$ cost.

We are ready to conclude Proposition 2.3.5 using the above induction steps.

We start with $\sigma_1 = \sigma_0/(10C_0)$. The initial collection $\{I_0\} = \mathcal{I}_0$ can be decoupled into $O(1)$ many (ψ, σ_1)-flat intervals by the trivial decoupling inequality (1.3), at $O(1)$ cost. We then iteratively decouple each $I \in \mathcal{I}_i$, for $i = 1, ..., N$ until $\sigma_N = \sigma$. We then take $\mathcal{I} = \mathcal{I}_N$. Note that $\sigma_{i+1}/\sigma_0 = C_0(\sigma_i/\sigma_0)^2$ for all $1 \leq i \leq N-1$. Thus, $\sigma/\sigma_0 \sim (C_0\sigma_1/\sigma_0)^{(3/2)^{N-1}} = (1/2)^{2^{N-1}}$. Therefore, $N \sim \log\log(\sigma/\sigma_0)^{-1}$. The total cost is bounded above by

$$C_\varepsilon^N \prod_{i=1}^N \sigma_i^{-\varepsilon} \leq (\log(\sigma/\sigma_0)^{-1})^{\log C_\varepsilon} \sigma^{-\varepsilon(1+2/3+(2/3)^2+\dots} \leq \sigma^{-O(\varepsilon)}.$$

Since $\varepsilon > 0$ is arbitrary, we have proved the decoupling inequality. Also, the size estimate is evident in the construction. It remains to show the bounded overlapping condition.

Fix $\xi_1 \in \mathbb{R}$ and assume there is a collection $\mathcal{I}'$ of intervals $I \in \mathcal{I}$ such that $\xi \in CI$.

Suppose there are N intervals $\{I_i\}_{i=1}^N \subseteq \mathcal{I}$ to the right of ξ_1 that lie in $\mathcal{I}'$. Then

$$c(I_i) - \xi_1 \leq C|I_i|/2,$$

where $c(I)$ denotes the center of the interval I.

On the other hand, by disjointness of the intervals I_i, we have

$$c(I_i) - \xi_1 \geq \sum_{j=1}^{i-1} |I_j|.$$

Thus we have for every $1 \leq i \leq N$,

$$|I_i| \geq 2C^{-1} \sum_{j=1}^{i-1} |I_j|.$$

Therefore

$$|I_N| \geq 2C^{-1} \sum_{j=1}^{N-1} |I_j| \geq C^{-1}(1+C^{-1}) \sum_{j=1}^{N-2} |I_j| \geq ... \geq 2C^{-1}(1+C^{-1})^{N-2}|I_1|.$$

Since I_N is (ψ, σ)-flat, we have $CI_N \cap [-1,1]$ is $(\psi, O(\sigma))$-flat, i.e.

$$\sup_{\xi_1 \in CI_N \cap [-1,1]} |\psi''(\xi_1)||I_N|^2 \lesssim \sigma.$$

Combining the above two display equations, we have

$$\sup_{\xi_1 \in CI_N} |\psi''(\xi_1)||I_I|^2 \lesssim \sigma(1+C^{-1})^{-N}.$$

But since $CI_N \cap [-1,1]$ contains I_1, we have

$$\sup_{x \in I_1} |g''(x)||I_1|^2 \lesssim \sigma(1+C^{-1})^{-N}.$$

Since $\mathcal{I}'$ satisfies (3), we have $N = O_C(1)$. The argument for intervals lying to the left of ξ_1 is the same. This concludes the proof of Proposition 2.3.5. $\square$

Note that Lemmas 2.3.4 and 2.3.5 require ω_0 to be an axis-parallel rectangle. This condition can be removed under some slightly stronger assumptions on P. The following proposition concludes this subsection by establishing the decoupling inequality in the case of a constant gradient.

Proposition 2.3.6. *Let $0 < \kappa' \leq \kappa \ll 1$, $0 < \sigma \ll 1$ and $2 \leq p \leq 6$. Let $\omega_0 \subseteq [-1,1]^2$ be a convex set and P be a polynomial of degree at most k with bounded coefficients. Suppose* *that $\partial_{\xi_1} P \sim \kappa'$ and $\partial_{\xi_2} P \sim \kappa$ over $3\omega_0$, defined by*

$$3\omega_0 := \{\xi \in \mathbb{R}^2 : \exists \xi_0, \xi_1 \in \omega_0 \ s.t. \ |\xi - \xi_0| \leq 3|\xi - \xi_1|\}.$$

There exists a family $\mathcal{P}$ of disjoint parallelograms such that the following hold:

1. *$\mathcal{P}$ covers $\{|P| \leq \sigma\} \cap \omega_0$;*
2. *for any absolute constant $C > 1$, $C\mathcal{P}$ has bounded overlapping in the sense that $\sum_{\omega \in \mathcal{P}} 1_{C\omega} \lesssim 1$;*
3. *for each $\omega \in \mathcal{P}$, $|P| \lesssim \sigma$ over 10ω;*
4. *the width of each $\omega \in \mathcal{P}$ is comparable to the minimum of the width of ω_0 and $\kappa^{-1}\sigma$;*
5. *the following 2D $\ell^2(L^p)$ decoupling inequality holds: for any $f : \mathbb{R}^2 \to \mathbb{C}$ whose Fourier transform is supported on $\bigcup_{\omega \in \mathcal{P}} \omega$, we have*

$$\|f\|_{L^p(\mathbb{R}^2)} \lesssim_{\varepsilon,k} \sigma^{-\varepsilon} \left(\sum_{\omega \in \mathcal{P}} \|f_\omega\|_{L^p(\mathbb{R}^2)}^2 \right)^{1/2}, \tag{2.12}$$

   where f_ω is defined by $\hat{f}_\omega = (\hat{f} 1_\omega)$.

Proof. First, we reduce ω_0 to a rectangle by the following. By John's ellipsoid theorem, there exists an ellipse containing ω_0 that is contained in $2\omega_0$. This ellipse is contained in a rectangle, and that rectangle is contained in $\sqrt{2}$ enlargement of the ellipse. Therefore, there exists a rectangle containing ω_0 that is contained in $2\sqrt{2}\omega_0 \subset 3\omega_0$. Thus, it suffices to consider rectangles ω_0 on which $\nabla P \lesssim \partial_{\xi_2} P \sim \kappa$ over ω_0.

Second, we rotate ω_0 to an axis-parallel rectangle $\tilde{\omega}_0$. Let ρ be the counterclockwise rotation by angle $\theta \in (-\pi/4, \pi/4]$. Let $\tilde{P} = P \circ \rho$. By direct computation,

$$\partial_{\xi_1} \tilde{P} = (\partial_{\xi_1} P \circ \rho) \cos\theta + (\partial_{\xi_2} P \circ \rho) \sin\theta, \quad \partial_{\xi_2} \tilde{P} = -(\partial_{\xi_1} P \circ \rho) \sin\theta + (\partial_{\xi_2} P \circ \rho) \cos\theta.$$

If $\kappa' < \kappa/2$, then $\nabla P \lesssim \partial_{\xi_2} \tilde{P} \sim \kappa$ over an axis-parallel rectangle. Otherwise, either

$\nabla P \lesssim \partial_{\xi_2}\tilde{P} \sim \kappa$ or $\nabla P \lesssim \partial_{\xi_1}\tilde{P} \sim \kappa$ over an axis-parallel rectangle. By re-labeling the axis if necessary, $\tilde{P}$ and $\tilde{\omega}_0$ meet the requirements of Lemma 2.3.4 and Proposition 2.3.5. For notational simplicity, we write $\tilde{P}$ and $\tilde{\omega}_0$ as P and ω_0, respectively, in the rest of the proof.

By Lemma 2.3.4, decoupling of the set $\{|P| \leq \sigma\} \cap \omega_0$ reduces to the decoupling of $\mathcal{N}^{\psi_I}_{O(\kappa^{-1}\sigma)}(I)$ for $O_k(1)$ many I. By abuse of notation, we proceed by writing P as $P \pm \sigma$, so that $P(\xi_1, \psi_I(\xi_1)) = 0$. Doing so does not impact ∇P and the set $\{|P| \lesssim \sigma\}$, up to a possible absolute constant loss.

We now apply Proposition 2.3.5 to decouple I into $(\psi_I, \kappa^{-1}\sigma)$-flat disjoint intervals $I \in \mathcal{I}$. Equivalently, we have decoupled $\mathcal{N}^{\psi_I}_{O(\kappa^{-1}\sigma)}(I)$ into a family $\mathcal{P}$ of parallelograms ω given by

$$\omega = \{(\xi_1, \xi_2) : \xi_1 \in I, |\xi_2 - \psi_I(\xi_1)| \lesssim \kappa^{-1}\sigma\} \cap \omega_0.$$

A minor issue concerns the intersection of small parallelograms near the boundary of ω_0. In this case, the intersection may be a convex set that is not a parallelogram. We then find a comparable parallelogram of this set to replace it. The properties still hold in a slight enlargement of ω_0.

Items 1, 2, 4 and 5 are immediate from Proposition 2.3.5. Item 3 follows from (2.10) in Lemma 2.3.4. $\square$

2.3.3 Induction on degrees of polynomials

We are now ready to prove Proposition 2.1.3 by induction on the polynomial degree. Heuristically, we seek to decouple $[-1, 1]^2$ into sets having essentially constant $|\nabla P|$. This can be achieved by our induction hypothesis. The key observation is that if $|\nabla P| \lesssim \sigma$ over $\omega_0 \subseteq [-1, 1]^2$, either $|P| \lesssim \sigma$ or $|P| \sim \sigma'$, for some $\sigma' > \sigma$, on ω_0. Thus, to decouple the sub-level set $\{|P| \lesssim \sigma\}$, we need to consider only sets on which $|\nabla P| \lesssim \sigma$ or $|\nabla P| \sim \sigma'$ for some $\sigma' > \sigma$.

Proof of Proposition 2.1.3. Proposition 2.1.3 is clear when P is a constant function. By induction on degree of the polynomial, we may assume that there exist families $\mathcal{P}^{\xi_1}_{\sigma_1,\lambda}, \mathcal{P}^{\xi_2}_{\sigma_2,\lambda}$, for functions $\partial_{\xi_1} P, \partial_{\xi_2} P$ replacing P in Proposition 2.1.3 respectively.

For each $\lambda \leq \sigma$, we now construct the family $\mathcal{P}_{\sigma,\lambda}$. Note that the parallelograms in $\mathcal{P}_{\sigma,\lambda}$ are subsets of $\{|P| \lesssim \sigma\}$.

Since the union of the $\mathcal{P}^{\xi_i}_{\sigma_i,\sigma/C}$ over all dyadic $\sigma_i \in [\sigma, 1]$ covers $[-1,1]^2$, for $i = 1, 2$, and some absolute constant C, we may restrict our attentions to $\omega_1 \cap \omega_2$, for some $\omega_i \in \mathcal{P}^{\xi_i}_{\sigma_i,\sigma/C}$, $i = 1, 2$. Note that

$$\text{width}(\omega_1 \cap \omega_2) \gtrsim \max\{\min\{\sigma_1, \sigma_2\}, \delta\} \gtrsim \max\{\sigma, \delta\}.$$

Without loss of generality, we may assume that $\sigma_1 \leq \sigma_2$. We now consider the following 3 cases corresponding to the items (3a), (3b), and (3c), one of which ω_2 satisfies by our induction hypothesis.

Case 1: $|\partial_{\xi_2} P| \sim \sigma_2$ over $10\omega_2$ for some $\sigma_2 > \sigma$.

If $\sigma \in (\max\{\lambda, \sigma_2\delta\}, \sigma_2]$, the set $\{|P - \sigma| < \sigma/2\} \cap \omega_1 \cap \omega_2$ can be decoupled into parallelograms ω of width $\gtrsim \sigma_2^{-1}\sigma \geq \max\{\sigma, \delta\}$ by Proposition 2.3.6. We put these ω in $\mathcal{P}^0_{\sigma,\lambda}$. They satisfy item (3a).

If $\sigma = \max\{\lambda, \sigma_2\delta\}$ by Proposition 2.3.6 again, the set $\{|P| < \sigma\} \cap \omega_1 \cap \omega_2$ can be decoupled into parallelograms ω. We put these ω in $\mathcal{P}^0_{\sigma,\lambda}$. If $\sigma = \sigma_2\delta$, ω has width $\sim \delta$ and satisfies item (3b). If $\sigma = \lambda$, ω satisfies item (3c).

For σ in either of these regime, the ω that we just put in $\mathcal{P}^0_{\sigma,\lambda}$ already cover $\{|P| \lesssim \sigma_2\} \cap \omega_1 \cap \omega_2$.

For other values of σ, $\mathcal{P}^0_{\sigma,\lambda}$ has no element intersecting $\omega_1 \cap \omega_2$.

Case 2: $|\partial_{\xi_2} P| \lesssim \sigma_2$ over $10\omega_2$ for some $\sigma < \sigma_2 \leq \delta$, and ω_2 has width δ.

If $\sup_{10\omega_1 \cap \omega_2} |P| < 4\sigma \lesssim \delta$, we put the set $\omega_1 \cap \omega_2$ in $\mathcal{P}^0_{\sigma,\lambda}$. Recall that ω_1 has width at least $\max\{\sigma, \delta\}$. Therefore, $\omega_1 \cap \omega_2$ has width $\sim \delta$ and hence satisfies item (3b).

If $\sup_{10\omega_1 \cap \omega_2} |P| > 4\sigma_2$, $|P| \sim \sigma'$ for some dyadic number $\sigma' > \sigma_2 > \sigma$ over $\omega_1 \cap \omega_2$. This is because $|\nabla P| \sim \sigma_2$ and $\omega_1 \cap \omega_2 \subseteq [-1, 1]^2$. Again, we put no element that has an intersection with $\omega_1 \cap \omega_2$ inside $\mathcal{P}^0_{\sigma,\lambda}$. We shall see later that this kind of set has already been considered in some $\mathcal{P}^0_{\sigma',\lambda}$.

Case 3: $|\partial_{\xi_2} P| < \sigma/10$ over $10\omega_2$.

This assumption is guaranteed by choosing the absolute constant C in $\mathcal{P}_{\sigma_i,\sigma/C}$ large enough.

This case is similar to Case 2. If $\sup_{10\omega_1\cap\omega_2} |P| < \sigma/10$, $\omega_1 \cap \omega_2$ is put in $\mathcal{P}^0_{\sigma,\sigma}$ and satisfies item (3c). If $|P| \sim \sigma$ on $10\omega_1 \cap \omega_2$, $\omega_1 \cap \omega_2$ is put in $\mathcal{P}^0_{\sigma,\lambda}$ and it satisfies item (3a). Otherwise, $|P| \sim \sigma' > \sigma$ and we put no element that has intersection with $\omega_1 \cap \omega_2$ inside $\mathcal{P}^0_{\sigma,\lambda}$.

Verification of the statements in Proposition 2.1.3

$\boxed{\textit{Statement 1:}}$ *for each λ, $\mathcal{P}^0_\lambda := \cup_\sigma \mathcal{P}^0_{\sigma,\lambda}$, where the union is over dyadic numbers $\sigma \in [\lambda, 1]$, covers $[-1,1]^2$.*

Let $\xi \in [-1,1]^2$. Then either $|P(\xi)| \sim \sigma$ for some dyadic number $\sigma > \lambda$ or $|P(\xi)| < \lambda$.

If $|P(\xi)| \sim \sigma$ for some dyadic number $\sigma > \lambda$, then $\xi \in \omega_1 \cap \omega_2$ for some $\omega_i \in \mathcal{P}^{\xi_i}_{\sigma_i,\sigma/C}$, since $\mathcal{P}^{\xi_i}_{\sigma_i,\sigma/C}$ covers $[-1,1]^2$, $i = 1,2$. Note that $\sigma_i \geq \sigma$.

Now, suppose that $\omega_1 \cap \omega_2$ is as in case 1. Since the $\mathcal{P}^0_{\sigma,\lambda}$ in case 1 cover $\omega_1 \cap \omega_2$, $\xi \in \omega$ for some $\omega \in \mathcal{P}^0_{\sigma,\lambda} \subset \mathcal{P}^0_\lambda$.

Second, we suppose that $\omega_1 \cap \omega_2$ is as in case 2 or case 3. Since $\xi \in \omega_1 \cap \omega_2$, $|P| \sim \sigma'$ for some dyadic number $\sigma' > 4\sigma$ cannot holds. Thus, $\omega_1 \cap \omega_2 \in \mathcal{P}^0_{\sigma,\lambda}$.

If $|P(\xi)| < \lambda$, $\xi \in \omega_1 \cap \omega_2$ for some $\omega_i \in \mathcal{P}^{\xi_i}_{\sigma_i,\lambda}$, $i = 1,2$. We may argue similarly if $\omega_1 \cap \omega_2$ is as in case 1 or case 2. For $\omega_1 \cap \omega_2$ as in case 3, $\omega_1 \cap \omega_2 \in \mathcal{P}^0_{\lambda,\lambda}$.

$\boxed{\textit{Statement 2:}}$ *$\mathcal{P}^0_\lambda$ has bounded overlap in the sense that $\sum_{\omega\in\mathcal{P}^0_\lambda} 1_{100\omega} \lesssim_k 1$.*

Since there are only $O(k)$ many times we apply Proposition 2.3.6, the bounded overlapping condition follows from that in Proposition 2.3.6.

$\boxed{\textit{Statement 3:}}$ *for each $\omega \in \mathcal{P}^0_{\sigma,\lambda}$, at least one of the following holds:*

(a) $\lambda < \sigma \leq 1$, $|P| \sim \sigma$ over 10ω, and the width of ω is $\gtrsim \max\{\sigma, \delta\}$;

(b) $\lambda < \sigma \leq \delta$, $|P| \lesssim \sigma$ *over* 10ω, *and the width of* ω *is* $\sim \delta$.

(c) $\sigma = \lambda$, $|P| \lesssim \lambda$ *over* 10ω, *and the width of* ω *is* $\gtrsim \max\{\sigma, \delta\}$.

This follows directly from the construction.

$\boxed{\textit{Statement 4:}}$ *For* $\lambda' < \lambda < \sigma$, $\mathcal{P}_{\sigma,\lambda} = \mathcal{P}_{\sigma,\lambda'}$.

Suppose that $\lambda' < \lambda < \sigma$ and $\omega \in \mathcal{P}_{\sigma,\lambda}$. Note that we apply the induction hypothesis to get the same families $\mathcal{P}^{\xi_i}_{\sigma_i,\sigma/C}$, $i = 1, 2$. If ω comes from Case 1, it is decoupled from either the set $\{|P \pm \sigma| < \sigma/2\} \cap \omega_1 \cap \omega_2$, or the set $\{|P| < \sigma\} \cap \omega_1 \cap \omega_2$. Thus, we get the except same families of ω from these $\omega_1 \cap \omega_2$. Otherwise, ω comes from Case 2 or 3, where $\omega = \omega_1 \cap \omega_2$. Thus, we see that in any case, $\omega \in \mathcal{P}_{\sigma,\lambda'}$. The other direction is similar.

$\boxed{\textit{Statement 5:}}$ *the following 2D* $\ell^2(L^p)$ *decoupling inequality holds: for any* $f : \mathbb{R}^2 \to \mathbb{C}$ *whose Fourier transform is supported on* $\bigcup_{\omega \in \mathcal{P}^0_{\sigma,\lambda}} \omega$, *we have*

$$\|f\|_{L^p(\mathbb{R}^2)} \lesssim_{\varepsilon,k} \sigma^{-\varepsilon} \left(\sum_{\omega \in \mathcal{P}^0_{\sigma,\lambda}} \|f_\omega\|^2_{L^p(\mathbb{R}^2)} \right)^{1/2}, \tag{2.13}$$

where f_ω *is the frequency projection of* f *onto* ω, *defined by* $\hat{f_\omega} = \hat{f} 1_\omega$.

It suffices to consider the decoupling constant. Note that the number of iterations is $O(k)$. Also, all ω put in $\mathcal{P}_{\sigma,\lambda}$ are from $\mathcal{P}^{\xi_i}_{\sigma_i,\sigma/C}$, $i = 1, 2$, and $\sigma_i \geq \sigma$. The decoupling loss from the inductive step is therefore $\lesssim_\varepsilon (\sigma_1\sigma_2)^{-O_k(\varepsilon)} \lesssim \sigma^{-O_k(\varepsilon)}$. The decoupling loss from Proposition 2.3.6 is $\lesssim_\varepsilon \sigma^{-\varepsilon}$. The total loss is $O_\varepsilon(\sigma^{-O_k(\varepsilon)})$. Since ε is arbitrary, we have shown that the decoupling constant is $O_\varepsilon(\sigma^{-\varepsilon})$ as desired.

This finishes the proof of Proposition 2.1.3. □

2.4 Polynomials with small Hessian determinant

In this section, we prove Proposition 2.1.4.

2.4.1 A reduction

In this subsection, we reduce Proposition 2.1.4 to studying a special type of polynomial. We need the following lemma about the eigenvalues of the Hessian matrix of a polynomial.

Lemma 2.4.1. *Let $\phi : \mathbb{R}^2 \to \mathbb{R}$ be a polynomial of degree at most k, without linear terms. Suppose that ϕ has $O(1)$ coefficients, and at least one of them is ~ 1. Then there exists $\xi' \in B(0,1)$ such that one of the eigenvalues of $D^2\phi(\xi')$ is of magnitude ~ 1.*

Proof. By a rotation and replacing ϕ by $-\phi$ if necessary, we may assume that $\partial^2_{\xi_1\xi_1}\phi$ has at least one coefficient ~ 1. There exists $\xi' \in B(0,1)$ such that $\partial^2_{\xi_1\xi_1}\phi(\xi') \sim 1$.

Note that $\partial^2_{\xi_1\xi_1}\phi + \partial^2_{\xi_2\xi_2}\phi$ is the trace, and hence the sum of eigenvalues, of $D^2\phi$. Thus, if $|\partial^2_{\xi_1\xi_1}\phi(\xi') + \partial^2_{\xi_2\xi_2}\phi(\xi')| > \partial^2_{\xi_1\xi_1}\phi(\xi')/2 \sim 1$, we are done. Otherwise, we have $\partial^2_{\xi_2\xi_2}\phi(\xi') < -\partial^2_{\xi_1\xi_1}\phi(\xi')/2 < 0$. However,

$$\det D^2\phi(\xi') < -(\partial^2_{\xi_1\xi_1}\phi)^2(\xi')/2 - (\partial^2_{\xi_1\xi_2}\phi)^2(\xi').$$

Therefore

$$|\det D^2\phi(\xi')| \geq (\partial^2_{\xi_1\xi_1}\phi)^2(\xi')/2 \sim 1.$$

This implies that the product of the eigenvalues of $D^2\phi(\xi')$ is bounded below. Since both eigenvalues are $O(1)$, they have magnitude ~ 1, as desired. $\square$

Lemma 2.4.1 helps us to reduce Proposition 2.1.4 to the following case.

Proposition 2.4.2. *There is a constant $\alpha = \alpha(k) \in (0,1]$ such that any polynomial $Q(\xi)$ satisfying*

1. $\deg Q \leq k$;
2. *Q has no linear term;*
3. *the only second order term of $Q(\xi_1)$ is ξ_1^2;*
4. *all coefficients of Q are $O(1)$;*

5. *all coefficients of* $\det D^2 Q$ *are bounded by some* $\nu \in (0,1)$*;*

is of the form

$$Q(\xi) = \tilde{A}(\xi_1) + \nu^{\alpha} \tilde{B}(\xi), \tag{2.14}$$

where $\tilde{A}, \tilde{B}$ *are polynomials with* $O(1)$ *coefficients.*

We will prove Proposition 2.4.2 in the following subsection.

Proof of Proposition 2.1.4 assuming Proposition 2.4.2. Let ϕ be a polynomial satisfying the assumption of Proposition 2.1.4. By dividing ϕ by its largest coefficient, we may assume without loss of generality that ϕ has some coefficient equal to 1.

Now, we apply Lemma 2.4.1 and get $\xi' \in B(0,1)$ such that one of the eigenvalues λ_1 of $D^2\phi(\xi')$ is of magnitude ~ 1. By dividing ϕ by λ_1 and replacing ϕ by $-\phi$ if necessary, we assume without loss of generality that $\lambda_1 = 1$.

Since $\sup_{B(0,1)} |\det D^2\phi| \lesssim \nu$, the other eigenvalue λ_2 of $D^2\phi(\xi')$ is of magnitude $O(\nu)$. Let ρ be a rotation that sends $\{e_1, e_2\}$ to the unit eigenvectors of $D^2\phi(\xi')$ and τ be a translation that sends the origin to $\rho^{-1}(\xi')$. By the eigenvalue analysis, the second order terms of $\tilde{\phi} = \phi \circ \rho \circ \tau$ are given by $\xi_1^2 + O(\nu)\xi_2^2$. Define Q by removing all the terms of degree at most two except ξ_1^2 in $\tilde{\phi}$. Then we see that all assumptions in Proposition 2.4.2 are satisfied. Hence, Q is of the form (2.14).

Now, we analyze the coefficients of $\phi \circ \rho = \tilde{\phi} \circ \tau^{-1}$. First, $Q \circ \tau^{-1}$ is also of the form (2.14). Moreover, the linear terms in $\tilde{\phi}$ have no impact on the higher order terms in $\tilde{\phi} \circ \tau^{-1}$ and $O(\nu)\xi_2^2$ can be absorbed to $\tilde{B}$. On the other hand, ϕ has no linear term, nor does $\phi \circ \rho$. In conclusion, we see that

$$\phi \circ \rho(\xi) = A_1(\xi_1) + \nu^{\alpha} B(\xi),$$

for some polynomials A, B with $O(1)$ coefficients. This implies Proposition 2.1.4 because

$$\phi(\xi) = A \circ \rho^{-1}(\xi) + \nu^{\alpha} B \circ \rho^{-1}(\xi),$$

where $A(\xi) = A_1(\xi_1)$ is one-dimensional. $\square$

2.4.2 The representing line

In this subsection, we prove Proposition 2.4.2. We first introduce the following terminology.

Definition 2.4.3. *Let $\phi : \mathbb{R}^2 \to \mathbb{R}$ be a polynomial. We consider the set of multi-indices appearing in ϕ, namely,*

$$N(\phi) := \{(m, n) \in \mathbb{Z}^2_{\geq 0} : \textit{the coefficient of } \xi_1^m \xi_2^n \textit{ in } \phi \textit{ is nonzero}\}.$$

Given a straight line $\ell \subset \mathbb{R}^2$, we denote by $\phi|_\ell$ the sum of monomials in ϕ with multi-indices lying on ℓ.

We have the following simple lemma.

Lemma 2.4.4. *Let ℓ be such that all points of $N(\phi)$ lie on one side of ℓ, inclusive. Then all points of $N(\det D^2\phi)$ lie on the same side of $\ell' := \ell + \ell - (2, 2)$. Moreover,*

$$(\det D^2\phi)|_{\ell'} = \det D^2(\phi|_\ell).$$

Proof. Write $\phi = \phi|_\ell + R$. Then

$$\det D^2\phi = \det D^2(\phi|_\ell) + \partial^2_{\xi_1\xi_1}(\phi|_\ell)\partial^2_{\xi_2\xi_2}R + \partial^2_{\xi_2\xi_2}(\phi|_\ell)\partial^2_{\xi_1\xi_1}R - \partial^2_{\xi_1\xi_2}(\phi|_\ell)\partial^2_{\xi_1\xi_2}R + \det D^2R.$$

All the terms except $\det D^2(\phi|_\ell)$ are strictly on one side of $\ell + \ell - (2, 2)$. $\square$

By Lemma 2.4.4, the boundaries of the convex hull of $N(\phi)$ are important in the analysis. Since we are interested in polynomials whose only second-order term is ξ_1^2, the boundaries containing the term ξ_1^2 are particularly interesting.

Definition 2.4.5. *Let Q be a polynomial that satisfies the five assumptions in Proposition 2.4.2. $\ell \subset \mathbb{R}^2$ is called a representing line of Q if*

1. *ℓ contains $(2,0)$ and at least one other point in $N(Q)$;*

2. *ℓ is not horizontal;*

3. *all points of $N(Q)$ lie on one side of ℓ, inclusive.*

Proposition 2.4.6. *Let Q be a polynomial that satisfies the five assumptions in Proposition 2.4.2 and let ℓ be a representing line of Q. Then*

$$Q|_\ell = \xi_1^2 + \nu^\beta B(\xi),$$

for some $\beta \in (0,1]$ depending only on d and some polynomial B with $O(1)$ coefficients.

Proof. Express ℓ in the (m,n)-plane by the equation $m = tn + 2$ where $t \in \mathbb{R}$. By Lemma 2.4.4, we see that $\det D^2(Q|_\ell)$ has coefficients bounded by ν.

Case $t < 0$.

Since we are in $\mathbb{Z}^2_{\varepsilon 0}$ we see $Q|_\ell$ is either of the form $Q|_\ell(\xi) = \xi_1^2 + a\xi_2^k$ where $k \geq 3$ or

$$Q|_\ell(\xi) = \xi_1^2 + a_1\xi_1\xi_2^k + a_2\xi_2^{2k}$$

where $k \geq 2$.

In the former case, $\det D^2(Q|_\ell) = 2ak(k-1)\xi_2^{k-2}$. Then $a = O(\nu)$ and we get the form we want.

In the latter case, a direct computation shows that

$$\det D^2(Q|_\ell) = 2a_1k(k-1)\xi_1\xi_2^{k-2} + [4a_2k(2k-1) - a_1^2k^2]\xi_2^{2k-2}.$$

Since $\det D^2(Q|_\ell)$ has coefficients bounded by ν, we have $a_1 = O(\nu)$, $a_2 = O(\nu)$ as desired.

Case $t \geq 0$.

In this case, ℓ is either vertical or has a positive slope. Let $c\xi_1^{k_1}\xi_2^{k_2}$ be the highest order term of $Q|_\ell$. If $k_2 \geq 1$, then by direct computation,

$$\det D^2(Q|_\ell) = -c^2 k_1 k_2 (k_1 + k_2 - 1)\xi_1^{2k_1-2}\xi_2^{2k_2-2} + \text{lower order terms}.$$

In particular, we see that $c = O(\nu^{1/2})$. Thus we can approximate $Q|_\ell$ by $\tilde{Q} = Q|_\ell - c\xi_1^{k_1}\xi_2^{k_2}$ and see that all coefficients of $\det D^2\tilde{Q}$ are of the order $O(\nu^{1/2})$. Note that $\tilde{Q}$ satisfies the same five assumptions with ν replaced by $\nu^{1/2}$. Each such approximation reduces the degree by at least 1. This process can be repeated at most k times until we arrive at ξ_1^2. In conclusion, we see that $Q|_\ell$ is of the form $\xi_1^2 + \nu^\beta B(\xi)$ where B has bounded coefficients, and $\beta = 2^{2-d}$.

□

We are now ready to prove Proposition 2.4.2, thus completing the proof of Proposition 2.1.4.

Proof of Proposition 2.4.2. Let Q be a polynomial that satisfies the five assumptions. If there is no representing line of Q, we are done because, in this case, Q is a function of ξ_1. Otherwise, by Proposition 2.4.6, all coefficients on representing lines ℓ of Q are $O(\nu^\beta)$. We can then approximate Q by $Q_1 = Q - Q|_\ell + \xi_1^2$ and $\det D^2 Q_1$ has $O(\nu^\beta)$ coefficients. For $j \geq 1$, repeat the process with ν replaced ν^{β^j} and let $Q_{j+1} = Q_j - (Q_j)|_{\ell_j}$, for some representing lines ℓ_j of Q_j. The process will be terminated in less than $(d+1)^2$ steps when no representing lines are available. In summary, we see that all coefficients of Q, except those terms containing ξ_1 only, are $O(\nu^{\beta^{(d+1)^2}})$. Thus, we obtain (2.14) with $\alpha = \beta^{(d+1)^2}$. □

2.5 Proof of Proposition 2.1.2

In this section, we prove Proposition 2.1.2.

We start by applying Proposition 2.1.3 with $P = \det D^2\phi$ to decouple $[-1,1]^2$ into level sets of $\det D^2\phi$. By the projection property, also known as cylindrical decoupling

(See Exercise 9.22 in [18] for details), we can lift these sets in $[-1,1]^2$ onto the surfaces $\mathcal{M}$, given by the graph of ϕ over these sets.

More precisely, let $\lambda = \delta^{1/\alpha}$, where $\alpha = \alpha(k)$ in Proposition 2.1.4. Then $[-1,1]^2$ can be decoupled into parallelograms in $\mathcal{P}^0_\lambda$ by Proposition 2.1.3. To decouple $\mathcal{N}^\phi_\delta([-1,1]^2)$, it suffices to decouple $\mathcal{N}^\phi_\delta(\omega)$ for each $\omega_0 \in \mathcal{P}^0_\lambda = \cup_\sigma \mathcal{P}^0_{\sigma,\lambda}$. There exists some dyadic number $\sigma \in [\lambda, 1]$ such that $\omega_0 \in \mathcal{P}^0_{\sigma,\lambda}$.

We recall the following notations in (1.30) and (1.31). Let T_ω be the invertible affine map such that $T_\omega([-1,1]^2) = \omega$. Let

$$\phi_{T_\omega} := \phi \circ T_\omega - \nabla(\phi \circ T_\omega)(0,0) \cdot \xi - \phi \circ T_\omega(0,0) \tag{2.15}$$

and its normalisation

$$\bar{\phi}_{T_\omega} = \frac{\phi_{T_\omega}}{\|\phi_{T_\omega}\|}. \tag{2.16}$$

Note that the graph of $\bar{\phi}_{T_\omega}$ is a translated, rotated, and enlarged copy of the graph of ϕ over ω. Thus, the Hessian determinant of $\bar{\phi}$ is also dyadically a constant over $[-2,2]^2$.

2.5.1 Two simple cases

We claim that if ω_0 satisfies either item (3b) or item (3c) of Proposition 2.1.3, then

$$\phi_{T_{\omega_0}} = A \circ \rho(\xi) + O(\delta), \tag{2.17}$$

for some one dimensional function $A(\xi_1, \xi_2) = A(\xi_1, 0)$ for all ξ_2.

We first prove the claim. Suppose that ω_0 satisfies item (3b). Let T be the composition of rotation and translation that ω_0 to $\tilde{\omega}_0$ so that $\tilde{\omega}_0$ is centered at 0 and has the shorter side of length $\sim \delta$ perpendicular to ξ_1-axis. Therefore,

$$\phi \circ T^{-1}(\xi_1, \xi_2) = \phi \circ T^{-1}(\xi_1, 0) + O(\delta).$$

This proves (2.17).

Suppose that ω_0 satisfies item (3c). Note that $\det D^2\phi \lesssim \delta^{1/\alpha}$ over ω_0. Thus, $\det D^2\phi_{T_{\omega_0}} \lesssim \delta^{1/\alpha}$ over $[-1,1]^2$. By Proposition 2.1.4, there exists a rotation ρ such that (2.17) holds because $(\delta^{1/\alpha})^\alpha B(\xi) = O(\delta)$ for some polynomial B with bounded coefficients.

The claim is proven. We approximate $\mathcal{N}_\delta^{\phi_{T_\omega}}([-1,1]^2)$ by $\mathcal{N}_\delta^{A}([-1,1]^2)$. Since A is one-dimensional, by cylindrical decoupling, it suffices to decouple a δ-neighborhood of the curve $\xi_1 \to A(\xi_1, 0)$ in $\mathbb{R}^2$ into $(A(\cdot,0),\delta)$-flat intervals. This can be done by applying Proposition 2.3.2.

We have $\ell^2(L^6)$ decoupled $\omega_0 \in \mathcal{P}^0_{\sigma,\lambda}$ into (ϕ,δ)-flat rectangles.

2.5.2 A bootstrapping argument

It remains to consider ω_0 satisfying item (3a) of Proposition 2.1.3. We will deal with this case by induction on scale argument.

We induct the Hessian determinant of the normalized function:

$$H(\omega) := \inf_{\xi\in[-1,1]^2} |\det D^2\bar{\phi}_{T_\omega}(\xi)| \sim \|\phi_{T_\omega}\|^{-2}|\det T_\omega|^2\sigma = \|\phi_{T_\omega}\|^{-2}|\omega|^2\sigma \lesssim 1. \tag{2.18}$$

By the size estimate (3a), $|\omega| \gtrsim \sigma^2$ and hence $H(\omega) \gtrsim \sigma^5$.

By Theorem 1.1.1, if $H(\omega) \sim 1$, $[-1,1]^2$ can be decoupled into $(\bar{\phi}_{T_\omega}, \delta/\|T_\omega\|)$-flat parallelograms. Rescaling back, ω is decoupled into (ϕ,δ)-flat parallelograms as desired. On the other hand, if $\|\phi_{T_\omega}\| \lesssim \delta$, ω is (ϕ,δ)-flat and we are done. So our goal is to decouple inductively until we achieve either of the conditions for all decoupled pieces.

Now, we are ready to state and prove the key induction step:

Proposition 2.5.1. *Let $\varepsilon > 0$, $\delta \ll 1$, $2 \le d \in \mathbb{N}$, $2 \le p \le 6$, $\lambda := \delta^\alpha < \sigma \le 1$, α as in Proposition 2.1.4 and ϕ be a bounded polynomial of degree at most k. Let $\omega_0 \subseteq [-1,1]^2$ be a parallelogram. Let T_{ω_0}, $\bar{\phi}_{T_{\omega_0}}$ and $H(\omega_0)$ be as above. Suppose that $H(\omega_0) \gtrsim \sigma^5$. Then, there exists a covering $\mathcal{P}^{\omega_0} = \mathcal{P}^{\omega_0}_{iter} \sqcup \mathcal{P}^{\omega_0}_{stop}$ of parallelograms ω such that the following holds:*

1. $100\mathcal{P}^{\omega_0}$ *has* $O(1)$*-bounded overlaps in the sense that*

$$\sum_{\omega\in\mathcal{P}_{\omega_0}} 1_{100\omega_0} \lesssim_k 1; \tag{2.19}$$

2. *for each* $\omega \in \mathcal{P}^{\omega_0}_{iter}$, $H(\omega) \gtrsim_k H(\omega_0)^{1-\alpha/2}$;

3. $\mathcal{P}^{\omega_0}_{stop} = \emptyset$ *if* $\sigma \geq \delta^{1/2}$;

4. *for each* $\omega \in \mathcal{P}^{\omega_0}_{stop}$, *the width of* ω *is* $\sim \delta$;

5. *each* $\omega \in \mathcal{P}^{\omega_0}$ *is contained inside* $\left(1 + CH(\omega)^{\alpha/d}\right)\omega$ *for some absolute constant* C;

6. *the following* $\ell^2(L^p)$ *decoupling inequality holds: for any* F *Fourier supported on* $\mathcal{N}^{\phi}_{\sigma}(\omega_0)$,

$$\|F\|_{L^p(\mathbb{R}^3)} \lesssim_{\varepsilon,k} H(\omega_0)^{-\varepsilon}\left(\sum_{\omega\in\mathcal{P}^{\omega_0}} \|F_\omega\|^2_{L^p(\mathbb{R}^3)}\right)^{1/2}. \tag{2.20}$$

Proof of Proposition 2.5.1. The smaller α is, the weaker Proposition 2.1.4 is, so we may, without loss of generality, assume that $\alpha < 1/5$.

We apply Proposition 2.1.4 to $\bar{\phi}_{T_{\omega_0}}$ to obtain a rotation $\rho : \mathbb{R}^2 \to \mathbb{R}^2$ and an one-dimensional bounded polynomial A and bounded polynomial B such that

$$\bar{\phi}_{T_{\omega_0}}(\xi) - A\circ\rho(\xi) = H(\omega_0)^{\alpha} B\circ\rho(\xi) = O(H(\omega_0)^{\alpha}).$$

Take $\delta' = H(\omega_0)^{\alpha} + \sigma \sim H(\omega_0)^{\alpha}$. We see that support of $\hat{F}$ lies in $\mathcal{N}^{A\circ\rho}_{\delta'}(\omega_0)$. Hence, by cylindrical decoupling, along the direction of $\rho^{-1}(e_2)$, the decoupling of the set $\mathcal{N}^{A\circ\rho}_{\delta'}(\omega_0)$ is reduced to the decoupling of the two-dimensional set

$$\{(\xi_1,\xi_2) \in \rho^{-1}([-1,1]^2) \subset [-2,2]^2 : |\xi_2 - A(\xi_1, 0)| \leq \delta'\}.$$

To decouple this set, we apply Proposition 2.3.2 to obtain a partition of $[-2,2]$ into intervals $I \in \mathcal{I}$. For each I, we define ω_I to be a parallelogram such that $(T_\omega \circ \rho)^{-1}(\omega_I)$

is the smallest rectangle of the form $I \times [a, b]$ that contains $((T_\omega \circ \rho)^{-1}(\omega)) \cap (I \times \mathbb{R})$. See Figure 2.2 below.

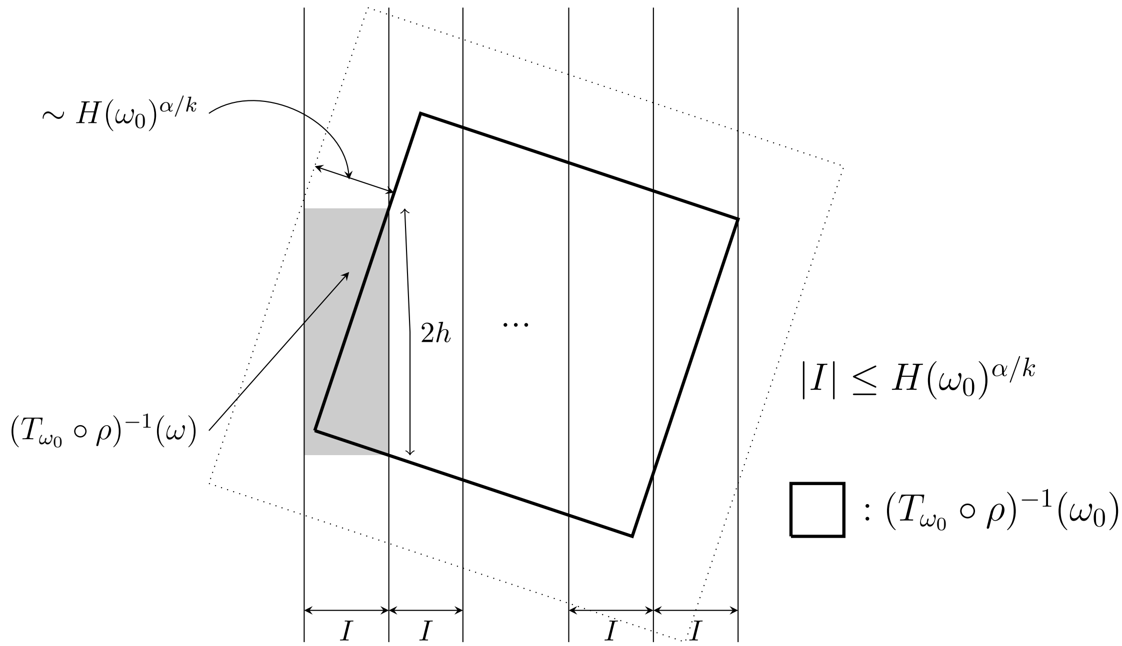

Figure 2.2: The cylindrical decoupling

We merge intervals in $\mathcal{I}$ so that the width of ω_I is at least δ. If merges are involved, ω_I has width δ and we put ω_I into $\mathcal{P}^{\omega_0}_{stop}$. In this case, we also write $I = I'$. We do the following for each remaining (A, δ')-flat intervals I.

Recall that $(T_{\omega_0} \circ \rho)^{-1}(\omega_I) = I \times [a, b]$. Since we allow finite overlapping, we may assume without loss of generality that for ξ_1 in half of I, $\{\xi_1\} \times \mathbb{R}$ has a nonempty intersection with $(T_{\omega_0} \circ \rho)^{-1}(\omega_0)$. Therefore, we have $2h := b - a \geq |I|/2$. Let T_0 be the translation that maps $I \times [-h, h]$ to $I \times [a, b]$. By putting all terms involving ξ_1 only to $A \circ T_0$ if necessary, we may assume that

$$\bar{\phi}_{T_\omega} \circ \rho^{-1} \circ T_0 - A \circ T_0 = H(\omega)^\alpha \xi_2 \tilde{B},$$

for some bounded polynomial $\tilde{B}$. Since $|\xi_2| \leq h$ and $\delta' \sim H(\omega_0)^\alpha$, we have

$$\bar{\phi}_{T_{\omega_0}} \circ \rho^{-1} \circ T_0 - A \circ T_0 = O(\delta' h).$$

Thus, we may apply Proposition 2.3.2 again to decouple I into $(A\circ T_0, \delta' h)$-flat intervals $I' \in \mathcal{I}_I$. We define $\omega_{I'}$ to be a parallelogram such that $(T_{\omega_0} \circ \rho)^{-1}(\omega_{I'}) = (I' \times \mathbb{R}) \cap \big((T_\omega \circ \rho)^{-1}(\omega_I)\big)$. Similarly, we merge the intervals I' if necessary so that the width of $\omega_{I'}$ is at least δ and put these $\omega_{I'}$ into $\mathcal{P}^{\omega_0}_{stop}$. We put all remaining parallelograms $\omega = \omega_{I'}$ into $\mathcal{P}^{\omega_0}_{iter}$.

We now check Properties (1), (5) and (6) for this family $\mathcal{P}^{\omega_0} = \mathcal{P}^{\omega_0}_{iter} \sqcup \mathcal{P}^{\omega_0}_{stop}$. Property (1) is inherited from that of $\mathcal{P}$ and $\mathcal{P}_I$. Note that $(T_\omega \circ \rho)^{-1}(\omega_I)$ is contained in the dotted rectangle, $(1 + CH(\omega_0)^{\alpha/k})(T_\omega \circ \rho)^{-1}(\omega)$. Thus, rescaling back, we obtain property (5) for $\omega \subset \omega_I$. The decoupling inequality (2.20) in the property (6) follows from Proposition 2.3.2, invariance of decoupling inequalities under affine transformations, cylindrical decoupling, and the fact that $\delta' \gtrsim H(\omega_0)$.

Now, we check property (2). For δh-flat intervals I', $\|\phi_{T_\omega}\| \leq \delta h \|\phi_{T_\omega}\|$. On the other hand, in the size estimate in Proposition 2.3.2, we have

$$|\omega| = |I'|(2h)|\omega_0| \gtrsim h^{3/2}\delta'^{1/2}|\omega_0|,$$

and $h \gtrsim |I| \gtrsim \delta^{1/2}$.

Therefore,

$$\begin{aligned} \|\phi_{T_\omega}\|^{-2}|\omega|^2\sigma &\gtrsim (h\delta)^{-2}(h^{3/2}\delta^{1/2})^2\|\phi_{T_\omega}\|^{-2}|\omega_0|^2\sigma \\ &\sim H(\omega_0)h\delta \gtrsim \delta^{1/2}H(\omega_0) \sim H(\omega_0)^{1-\alpha/2}, \end{aligned}$$

as desired.

Property (4) is immediate by the definition of the collection $\mathcal{P}^{\omega_0}_{stop}$. For intervals I' such that $\omega_{I'}$ has width δ, we have, $|I'|\sigma \lesssim \delta$ because ω has width at least σ. By the size estimate, $|I'| \gtrsim h^{1/2}\delta'^{1/2} \gtrsim \delta'^{3/4} \sim H(\omega)^{3\alpha/4} \geq \sigma^{15\alpha/4}$. Thus, $\sigma^{15\alpha/4+1} \lesssim \delta$. Since $0 \leq \alpha < 1/5$, we have $\sigma \lesssim \delta^{\frac{1}{15\alpha/4+1}} \ll \delta^{1/2}$. Thus, if $\sigma \geq \delta^{1/2}$, such $\omega_{I'}$ cannot exist. This proves property (3).

We have obtained all properties for the family $\mathcal{P}^{\omega_0}$, and hence Proposition 2.5.1 is proved. □

2.5.3 Induction on scale

In this section, we continue the proof of Proposition 2.1.2. Recall that we want to decouple $\omega_0 \in \mathcal{P}^0_{\sigma,\lambda}$ satisfying item (3a) of Proposition 2.1.3 for some $\sigma \in [\lambda, 1]$, $\lambda = \delta^{1/\alpha}$ and $\alpha = \alpha(k)$ in Proposition 2.1.4.

Let $K = K(\varepsilon) \gg 1$ to be determined. Note that $H(\omega_0) \geq |\omega_0|^2\sigma \gtrsim \sigma^5$. We apply Proposition 2.5.1 to decouple each ω_0 into $\omega_1 \in \mathcal{P}^{\omega_0}$. We apply Proposition 2.5.1 again to decouple each $\omega_1 \in \mathcal{P}^{\omega_0}_{iter}$ into $\omega_2 \in \mathcal{P}^{\omega_1}$, and so on. The process stops if $\omega_{\bullet-1} \in \mathcal{P}^{\omega_{\bullet-2}}_{stop}$, or $H(\omega_{\bullet-1}) \geq 1/K$. If $\omega_{\bullet-1} \in \mathcal{P}^{\omega_{\bullet-2}}_{stop}$, then $\omega_{\bullet-1}$ has width δ and $\sigma \leq \delta^{1/2}$. We repeat subsection 2.5.1 to get a decoupling of $\omega_{\bullet-1}$ into (ϕ, δ)-flat parallelograms $\omega_\bullet$, with decoupling constant $\delta^{-\varepsilon} = \sigma^{-O(\varepsilon)}$.

The tree diagram (Figure 2.3) below describes the process where every branch leads to some set $\omega_\bullet$ satisfying either $\omega_\bullet$ is (ϕ, δ)-flat or $H(\omega_\bullet) \geq 1/K$. The value next to each edge in the tree diagram represents an upper bound of the decoupling constant in that step.

Let $\{\omega_i\}_{i=0}^N$ be a sequence of sets that forms a branch of the tree diagram. Since $H(\omega_{i+1}) \gtrsim_k H(\omega_i)^{1-\alpha}$, we have $H(\omega_{i+1}) > H(\omega_i)^{1-\alpha/2}$ by picking K large enough. We see that the maximum number of steps N we need is $\leq c_\alpha \log(\sigma^{-1})/\log(K)$.

Moreover, the cost to decouple in each iteration in this sequence is $C_\varepsilon H(\omega_i)^{-\varepsilon}$, except possibly the last one, and so the total cost is at most

$$C_\varepsilon \sigma^{-O(\varepsilon)} C_\varepsilon^N \prod_{i=0}^N H(\omega_i)^{-\varepsilon} \leq C_\varepsilon^{N+2} \sigma^{-O(\varepsilon)} \left(\sigma^3 \cdot \sigma^{3(1-\alpha/2)} \cdot \ldots \cdot \sigma^{3(1-\alpha/2)^N}\right)^{-\varepsilon} \lesssim_\varepsilon \sigma^{-O_d(\varepsilon)},$$

if we pick

$$\log K \sim c_\alpha \frac{\log C_\varepsilon}{\varepsilon} \implies \log C_\varepsilon \cdot \frac{c_\alpha \log \sigma^{-1}}{\log K} \sim \varepsilon \log \sigma^{-1} \implies C_\varepsilon^{N+2} \leq \sigma^{-\varepsilon}.$$

Now, we apply Bourgain-Demeter's decoupling inequalities, Theorem 1.1.1, on $\bar{\phi}_{T_{\omega_N}}$.

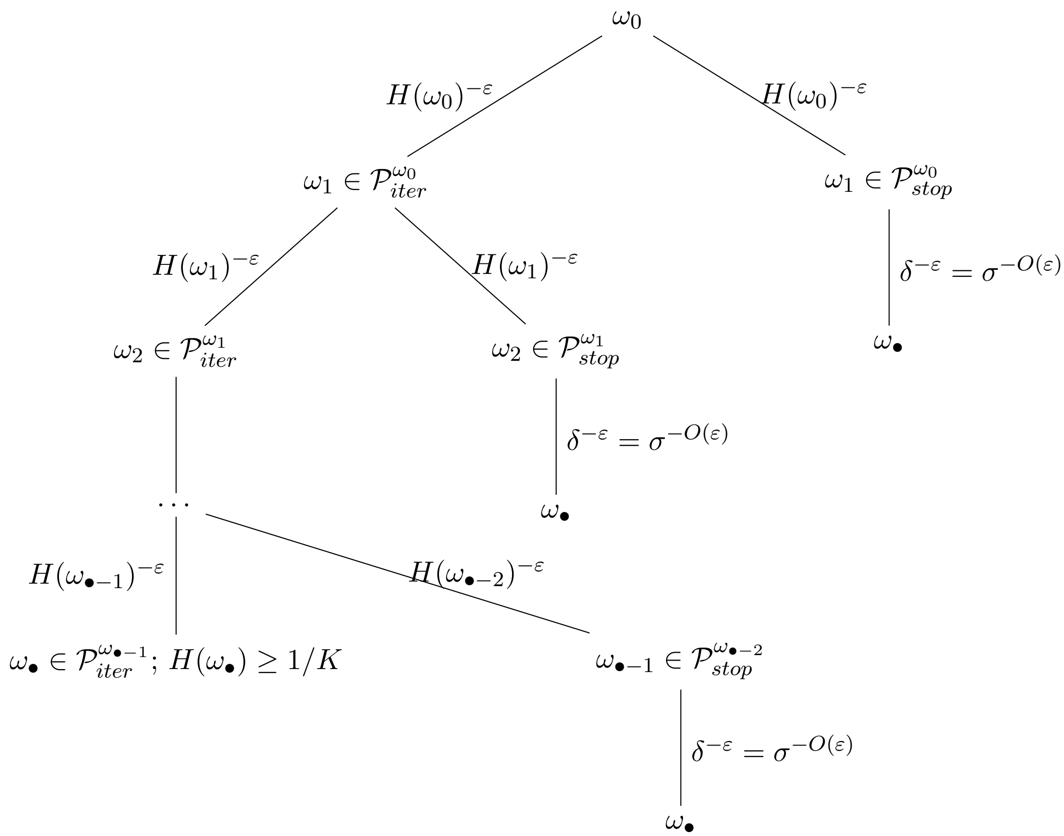

Figure 2.3: Tree diagram for the induction

Since $H(\omega_N) \sim 1$, $[-1,1]^2$ can be $\ell^4(L^4)$ (or $\ell^2(L^4)$ if $\det D^2\bar{\phi}_{T_\omega} > 0$) decoupled into $(\bar{\phi}_{T_\omega}, \delta/\|T_\omega\|)$-flat parallelograms. Rescaling back, ω is decoupled into (ϕ,δ)-flat parallelograms as desired. On the other hand, if $\|\phi_{T_\omega}\| \lesssim \delta$, ω is (ϕ,δ)-flat and we are done. This proves items 1 and 4 in Proposition 2.1.2. Item 3 is evident from our construction. See the corresponding size estimate for Propositions 2.1.3 and 2.5.1.

We now prove the overlap bound among $\mathcal{P}$, item 2. We recall that $H(\omega_i) < H(\omega_{i+1})^{1-\alpha/2}$ and hence $H(\omega_i) < H(\omega_N)^{(1-\alpha/2)^{N-i}} < K^{-(1-\alpha/2)^{N-i}}$. Therefore,

$$\sum_i \log(1+CH(\omega_i)^{\alpha/k}) \lesssim \sum_i H(\omega_i)^{\alpha/k} \lesssim ...+K^{-(1-\alpha/2)^2\alpha/k}+K^{-(1-\alpha/2)\alpha/k}+K^{-\alpha/k} \leq \log 2,$$

by choosing K large enough. The last inequality in the above display equation follows from estimating the sum by a geometric series with common ratio $K^{-(1-\alpha/2)}$. Thus, we have

$$\prod_i (1 + CH(\omega_i)^{\alpha/k}) \le 2 \tag{2.21}$$

and hence $50\omega_N \subset 100\omega_0$. Since $100\mathcal{P}_0$ has $O(1)$-bounded overlaps and in each iteration, the overlaps are also $O(1)$, $50\mathcal{P}$ has $O_\varepsilon(\sigma^{-\varepsilon})$ overlaps.

We have finished the proof of Proposition 2.1.2.

2.6 Variants of decoupling inequalities

We end the chapter by recording two versions of decoupling inequalities. The theorem below can be seen directly by not applying Bourgain-Demeter's decoupling inequalities on ω_N, but instead stopping when we arrive at $\omega_\bullet$ in Figure 2.3. These inequalities are essential tools in Chapter 3.

Proposition 2.6.1. *Let $2 \le p \le 6$, $n = 3$, $\varepsilon > 0$, $0 < \delta \ll 1$ and $\lambda = \delta^{1/\alpha}$ for $\alpha = \alpha(k)$ as in Proposition 2.1.4. Let ϕ be a bivariate polynomial of degree at most k with bounded coefficients. For each dyadic number $\sigma \in [\lambda, 1]$, there exist families $\mathcal{P}_{\sigma,\lambda} = \mathcal{P}_{\sigma,\lambda}(\delta, \phi, \varepsilon)$ of parallelograms such that the following statements hold:*

1. *$\mathcal{P}_\lambda := \cup_\sigma \mathcal{P}_{\sigma,\lambda}$ (the union is over dyadic numbers $\sigma \in [\lambda, 1]$) covers $[-1, 1]^2$;*

2. *each $\omega \in \mathcal{P}_\lambda$ satisfies either of the followings*

   (a) ω is (ϕ, δ)-flat;

   (b) $\bar{\phi}_{T_\omega}$, defined as in (1.31), has Hessian determinant $\sim_\varepsilon 1$;

3. *$\mathcal{P}_\lambda$ has $O_{k,\varepsilon}(\delta^{-\varepsilon})$-bounded overlap in the sense that $\sum_{\omega \in \mathcal{P}_\lambda} 1_\omega \lesssim_{k,\varepsilon} \delta^{-\varepsilon}$;*

4. *the width of each $\omega \in \mathcal{P}_\lambda$ is at least δ;*

5. *for any $\lambda < \sigma$, $\mathcal{P}_{\ge\sigma,\lambda} := \cup_{\sigma \le \sigma' \le 1} \mathcal{P}_{\sigma',\lambda}$ covers $\{|\det D^2\phi| \sim \sigma\}$;*

6. *the following $\ell^2(L^p)$ decoupling inequality holds: for any $F : \mathbb{R}^3 \to \mathbb{C}$ whose Fourier*

transform is supported on $\mathcal{N}_\delta^\phi(\cup_{\omega\in\mathcal{P}_{\geq\sigma,\lambda}}\omega)$*,*

$$\|F\|_{L^p} \lesssim_{\varepsilon,k} \sigma^{-\varepsilon}\left(\sum_{\omega\in\mathcal{P}_{\geq\sigma,\lambda}}\|F_\omega\|_{L^p}^2\right)^{1/2}. \tag{2.22}$$

We tile $[-1,1]^2$ by squares Q of side length σ^ε, followed by an approximation of each square Q of side length σ^ε by polynomials of degree $\sim \frac{1}{\varepsilon}$ as in Subsection 2.2.2. We arrive at the following theorem.

Proposition 2.6.2. *Let* $2 \leq p \leq 6$*,* $n = 3$*,* $\varepsilon > 0$*,* $0 < \delta \ll 1$ *and* $\lambda = \delta^{1/\alpha}$ *for some* $\alpha = \alpha(\varepsilon) \in (0, 1/4)$*. Let* $\phi : [-1,1]^2 \to \mathbb{R}$ *be a smooth function. Let* Q *be a tiling of* $[-1,1]^2$ *by squares of side length* σ^ε*. For each dyadic number* $\sigma \in [\lambda, 1]$*, there exist families* $\mathcal{P}_{\sigma,\lambda}^Q = \mathcal{P}_{\sigma,\lambda}^Q(\delta,\phi,\varepsilon)$ *of parallelograms such that the following statements hold:*

1. $\mathcal{P}_\lambda^Q := \cup_\sigma \mathcal{P}_{\sigma,\lambda}^Q$ *(the union is over dyadic numbers* $\sigma \in [\lambda, 1]$*) covers* Q*;*
2. *each* $\omega \in \mathcal{P}_\lambda^Q$ *satisfies one of the following*
   - (a) ω *is* (ϕ,δ)*-flat;*
   - (b) $\bar{\phi}_{T_\omega}$*, defined as in* (1.31)*, has Hessian determinant* $\sim_\varepsilon 1$*;*
3. $\mathcal{P}_\lambda^Q$ *has* $O_{\phi,\varepsilon}(\delta^{-\varepsilon})$*-bounded overlap in the sense that* $\sum_{\omega\in\mathcal{P}_\lambda^Q} 1_\omega \lesssim_{\phi,\varepsilon} \delta^{-\varepsilon}$*;*
4. *the width of each* $\omega \in \mathcal{P}_\lambda^Q$ *is at least* δ*;*
5. *for any* $\lambda < \sigma$*,* $\mathcal{P}_{\geq\sigma,\lambda}^Q := \cup_{\sigma\leq\sigma'\leq 1}\mathcal{P}_{\sigma',\lambda}^Q$ *covers* $\{|\det D^2\phi| \sim \sigma\} \cap Q$*;*
6. *the following* $\ell^2(L^p)$ *decoupling inequality holds: for any* $F : \mathbb{R}^3 \to \mathbb{C}$ *whose Fourier transform is supported on* $\mathcal{N}_\delta^\phi(\cup_{\omega\in\mathcal{P}_{\geq\sigma,\lambda}^Q}\omega)$*,*

$$\|F\|_{L^p} \lesssim_{\varepsilon,\phi} \sigma^{-\varepsilon}\left(\sum_{\omega\in\mathcal{P}_\sigma}\|F_\omega\|_{L^p}^2\right)^{1/2}, \tag{2.23}$$

where the implicit constant is independent of the choice of Q*.*

Note that we restrict the range of $\alpha(\varepsilon) \in (0, 1/4)$ for purely technical reasons in the next chapter. This can be achieved by picking $\alpha = \alpha(k)$ in Proposition 2.1.4 smaller than 1/4.

Chapter 3

Affine restriction theory

In this chapter, we prove Theorem 1.2.2, the L^2 affine restriction theorem for compact smooth surfaces in $\mathbb{R}^3$. We also demonstrate the optimality of Theorem 1.2.2 by establishing counterexamples in Section 3.5. We works in the ambient dimension $n = 3$ throughout the chapter.

3.1 Overview

The neighborhood formulation for the decoupling theorem serves as an essential ingredient in the proof of our restriction theorem. As in our proof of the decoupling result, we will prove Theorem 1.2.2 by proving the following equivalent version.

Proposition 3.1.1. *Let $\varepsilon > 0$ and $R \ll 1$. Let $\phi : [-1,1]^2 \to \mathbb{R}$ be a smooth function. Define measures M, M_ε on $(\xi, \eta) \in [-1,1]^2 \times \mathbb{R}$ by*

$$dM(\xi,\eta) = dM^\phi(\xi,\eta) = |\det D^2\phi(\xi)|^{-1/4} d\xi d\eta \tag{3.1}$$

and

$$dM_\varepsilon(\xi,\eta) = dM_\varepsilon^\phi(\xi,\eta) = |\det D^2\phi(\xi)|^{-1/4-\varepsilon} d\xi d\eta. \tag{3.2}$$

Then for any ball, B_R of radius R and any function F such that $\hat{F}$ is supported on the

R^{-1} vertical neighborhood of the graph of ϕ above $[-1,1]^2$, we have

$$\|F\|_{L^4(B_R)} \lesssim_{\varepsilon,\phi} R^{\varepsilon-1/2}\|\hat{F}\|_{L^2(dM)}, \quad \text{if } \hat{F} \in L^2(dM), \tag{3.3}$$

and

$$\|F\|_{L^4(B_R)} \lesssim_{\varepsilon,\phi} R^{-1/2}\|\hat{F}\|_{L^2(dM_\varepsilon)}, \quad \text{if } \hat{F} \in L^2(dM_\varepsilon). \tag{3.4}$$

Moreover, the implicit constants in (3.3) *and* (3.4) *can be made uniform over all polynomials ϕ of degree up to k with bounded coefficients.*

The deduction of Theorem 1.2.2 from Proposition 3.1.1 is given in Section 3.2.

We recall the following notation from (1.30) and (1.31). For parallelogram $\omega \subseteq [-1,1]^2$, T_ω is an invertible affine map that maps $[-1,1]^2$ to ω, and T_ω sends the smooth function ϕ to ϕ_{T_ω} as follows:

$$\phi_{T_\omega}(\xi) := \phi \circ T_\omega(\xi) - \nabla(\phi \circ T_\omega)(0,0) \cdot \xi - \phi \circ T_\omega(0,0).$$

We normalise ϕ_{T_ω} by

$$\bar{\phi}_{T_\omega} = \frac{\phi_{T_\omega}}{\|\phi_{T_\omega}\|},$$

where $\|\phi_{T_\omega}\| := \sup_{[-1,1]^2} |\phi_{T_\omega}|$ is as defined in (1.28).

From the definition of (ϕ,δ)-flatness in (1.6), we see that ω is (ϕ,δ)-flat if and only if $\|\phi_{T_\omega}\| \lesssim \delta$.

We need the following definition of admissible sets

Definition 3.1.2. *For $\varepsilon > 0$, $0 < \sigma \leq 1$, $R \gg 1$, and ϕ a smooth function over $[-1,1]^2$, a parallelogram $\omega \subseteq [-1,1]^2$ is said to be $(\phi,\sigma,R,\varepsilon)$-admissible if either of the following holds:*

1. *$|\det D^2\phi| \lesssim \sigma$ over 2ω and $\|\phi_{T_\omega}\| \lesssim R^{-1}$;*

2. $|\det D^2\phi| \sim \sigma$ *over* 2ω *and for any* $\xi \in [-1,1]^2$,

$$|\det D^2\bar{\phi}_{T_\omega}(\xi)| \sim_\varepsilon 1 \quad \textit{and} \quad \sum_{|\alpha|=2,3} |D^\alpha \bar{\phi}_{T_\omega}(\xi)| \lesssim_\varepsilon 1. \tag{3.5}$$

The motivation for this definition is the following proposition.

Proposition 3.1.3. *Let* $\varepsilon > 0$, $R \geq 1$, $0 < \sigma \leq 1$. *Assume the parallelogram* $\omega \subseteq [-1,1]^2$ *is* $(\phi, \sigma', R, \varepsilon)$*-admissible for some* $\sigma' \gtrsim \sigma$. *For any function* F *such that* $\mathrm{supp}\hat{F} \subseteq \mathcal{N}^\phi_{R^{-1}}(\omega \cap \{|\det D^2\phi| \sim \sigma\})$, *we have*

$$\|F\|_{L^4(\mathbb{R}^3)} \lesssim_\varepsilon R^{-1/2}\sigma^{\varepsilon/2}\|\hat{F}\|_{L^2(dM_\varepsilon)}, \tag{3.6}$$

where the implicit constant is independent of ω *and* ϕ.

This allows us to estimate the left-hand side of (3.4) if $\hat{F}$ is further restricted to an admissible set ω. The proof of Proposition 3.1.3 is given in Section 3.3.

We recall from Propositions 2.6.1 and 2.6.2 in Chapter 2 that we can $\ell^2(L^4)$ decouple $\mathcal{N}^\phi_{R^{-1}}([-1,1]^2)$ into admissible rectangles, regardless of the sign of the Hessian determinant of ϕ. This is the primary tool used in Section 3.4 to prove Proposition 3.1.1.

Finally, in Section 3.5, we compute a counterexample to show that the ε losses in (1.16), (1.17) and non-endpoint exponent $p > 4$ in (1.18) are necessary for the case of general smooth surfaces.

3.2 An equivalent version

In this section, we deduce Theorem 1.2.2 from Proposition 3.1.1. The proof follows from Proposition 1.27 of [18] with adaption to the new measures. We only prove the implication from (3.4) to (1.17). The other cases are similar.

For $R \geq 1$, let $\psi_R \in \mathcal{S}(\mathbb{R}^3)$ be such that

1. $\hat{\psi}_1$ is non negative and supported on $B(0,1)$;

2. $1_{B(0,1)} \le \psi_1$;

3. $\psi_R = \psi_1(R^{-1}\cdot)$.

Let $F = Eg\psi_R$. Then we have

$$\hat{F}(\xi,\eta) = g(\xi)|\det D^2\phi(\xi)|^{1/4+\varepsilon}\hat{\psi}_R(\eta - \phi(\xi)).$$

We also define dm_ε such that

$$dM_\varepsilon(\xi,\eta) = |\det D^2\phi(\xi)|^{-1/4-\varepsilon} d\xi d\eta = dm_\varepsilon(\xi)d\eta.$$

Now, we have

$$\begin{aligned}
\|Eg\|_{L^4(B_R)} &\le \|F\|_{L^4(B_R)} \\
&\lesssim_{\phi,\varepsilon} R^{-1/2}\|\hat{F}\|_{L^2(dM_\varepsilon)} \\
&\lesssim_{\psi} R^{-1/2}\left\|g\,|\det D^2\phi|^{1/4+\varepsilon}\right\|_{L^2(dm_\varepsilon)} R^{1/2} \\
&= \|g\|_{L^2(d\mu_\varepsilon)},
\end{aligned}$$

where the second-to-last inequality follows from (3.4) and Fubini's theorem. By letting $R \to \infty$, we have (1.17).

3.3 A scaling argument

In this section, we prove Proposition 3.1.3 by a scaling argument.

Since we have $|\det D^2\phi| \sim \sigma$ on the support of $\hat{F}$, we have $dM_\varepsilon \sim \sigma^{-\varepsilon} dM$. Hence, it suffices to show that

$$\|F\|_{L^4(\mathbb{R}^3)} \lesssim_\varepsilon R^{-1/2}\|\hat{F}\|_{L^2(dM)}. \tag{3.7}$$

We need the following scaling lemma:

Lemma 3.3.1 (Affine invariance of measure M)**.** *Let $R^{-1} \le s \le 1$, $\omega \subseteq [-1,1]^2$ be a* *parallelogram, and* ϕ *be a smooth function over* $[-1,1]^2$*. Let* ϕ_{T_ω} *be defined as in* (1.30)*. Let* $\bar{\phi} = s^{-1}\phi_{T_\omega}$*. Let* F *be such that* $\mathrm{supp}\hat{F} \subseteq \mathcal{N}^{\phi}_{R^{-1}}(\omega)$*. Define* $\hat{G}(\xi,\eta) = \hat{F}(T(\xi), s\eta)$ *such that* $\hat{G}$ *is supported on* $\mathcal{N}^{\bar{\phi}}_{(sR)^{-1}}([-1,1]^2)$*. Then we have*

$$\frac{\|F\|_{L^4(\mathbb{R}^3)}}{R^{-1/2}\|\hat{F}\|_{L^2(dM^{\phi})}} = \frac{\|G\|_{L^4(\mathbb{R}^3)}}{(sR)^{-1/2}\|\hat{G}\|_{L^2(dM^{\bar{\phi}})}}. \tag{3.8}$$

Proof. Rotation and translation have no impact on the quantity of the left-hand side of (3.8). Thus, we may assume without loss of generality that the center of ω is the origin and that $\phi(0,0) = 0$, $\nabla\phi(0,0) = (0,0)$.

Now, we keep track of the scaling, and we have $\hat{G}(\xi,\eta) = \hat{F}(T_\omega(\xi), s\eta)$. Direct computation shows that

$$G(x) = \frac{1}{s|\det T_\omega|} F(T_\omega^{-t}(x_1,x_2), s^{-1}x_3),$$

where T_ω^{-t} is the inverse transpose of T_ω. Therefore, we have

$$\|F\|_{L^4(\mathbb{R}^3)} = (s|\det T_\omega|)^{1-1/4}\, \|G\|_{L^4(\mathbb{R}^3)}, \tag{3.9}$$

On the other hand, $|(\det D^2\phi)(T_\omega\cdot)||\det T_\omega|^2 = s^2|\det D^2\bar{\phi}(\cdot)|$. Hence,

$$\|\hat{F}\|_{L^2(dM^{\phi})} = (s^{-2}|\det T_\omega|^2)^{(1/4)(1/2)}(s|\det T_\omega|)^{1/2}\|\hat{G}\|_{L^2(dM^{\bar{\phi}})}. \tag{3.10}$$

Combining (3.9) and (3.10), we obtain (3.8) as desired. □

Now, after rescaling, $(\phi, \sigma', R, \varepsilon)$-admissible sets are in one of the following situations:

1. $\|\phi_{T_\omega}\| \lesssim R^{-1}$ and $\mathrm{supp}\,\hat{G} \subset \mathcal{N}_1^{R\phi_{T_\omega}}([-1,1]^2)$, and $s = R^{-1}$;
2. $\|\bar{\phi}_{T_\omega}\| \sim_\varepsilon 1$ and $\mathrm{supp}\,\hat{G} \subset \mathcal{N}^{\bar{\phi}_{T_\omega}}_{(sR)^{-1}}([-1,1]^2)$, where $s = \|\phi_{T_\omega}\|$ and $\det D^2\bar{\phi}_{T_\Omega} \sim_\varepsilon 1$ over $[-1,1]^2$.

Now, we consider the first case. Thus, we estimate by Hausdorff-Young and Hölder's inequalities:

$$\|G\|_{L^4(\mathbb{R}^3)} \lesssim \|\hat{G}\|_{L^{4/3}(\mathbb{R}^3)} \lesssim \|\hat{G}\|_{L^2(dM^{R\phi_{T_\omega}})}.$$

The last inequality follows from the facts that the support of $\hat{G}$ is of size $\lesssim 1$, and that $R\phi_{T_\omega}$ has bounded coefficients. Recall that $s = R^{-1}$, the right-hand side of (3.8) is bounded, and thus we obtain (3.7) as desired.

We consider the second case. By Theorem 1.2.1 in an equivalent formulation, we have

$$\|G\|_{L^4(\mathbb{R}^3)} \lesssim_\varepsilon (sR)^{-1/2}\|\hat{G}\|_{L^2(d\xi d\eta)} \lesssim_\varepsilon (sR)^{-1/2}\|\hat{G}\|_{L^2(dM^{\bar{\phi}_{T_\omega}})}.$$

See Proposition 1.27 and Exercise 1.34 of [18]. Again, the right-hand side of (3.8) is bounded. We obtain (3.7). This finishes the proof of Proposition 3.1.3.

3.4 Proof of Proposition 3.1.1

In this section, we prove Proposition 3.1.1 by using Propositions 2.6.1 and 2.6.2.

3.4.1 The set with tiny Gaussian curvature

In this subsection, we additionally assume F is Fourier supported on $\mathcal{N}^{\phi}_{R^{-1}}(\{\xi \in [-1,1]^2 : |\det D^2\phi(\xi)| \lesssim R^{-4}\})$. In this case, we have

$$\|F\|_{L^4(\mathbb{R}^3)} \lesssim \|\hat{F}\|_{L^{4/3}(d\xi d\eta)}, \tag{3.11}$$

by the Hausdorff-Young inequality. Since the support of $\hat{F}$ is of size $O_\phi(R^{-1})$ (or $O_k(R^{-1})$ in the polynomial case), (3.11) can be further bounded by

$$\lesssim_\phi \text{ (or } \lesssim_k) R^{-1/4}\|\hat{F}\|_{L^2(d\xi d\eta)} \lesssim R^{-1/2-\varepsilon}\|\hat{F}\|_{L^2(dM_\varepsilon)}, \tag{3.12}$$

where we used

$$dM_\varepsilon \gtrsim (R^{-4})^{-1/4-\varepsilon} d\xi d\eta.$$

This completes the proof of Proposition 3.1.1 for the set with tiny Gaussian curvature.

3.4.2 Proof of the extra damped estimate

In this subsection, we prove the extra damped estimate (3.4). First, we decompose

$$F = F_0 + \sum_{\substack{\sigma \text{ dyadic} \\ R^{-4} < \sigma \le 1}} F_\sigma \tag{3.13}$$

where $F_0, F_\sigma \in \mathcal{S}$ are such that F_0 is Fourier supported on $\mathcal{N}^{\phi}_{R^{-1}}(\{\xi \in [-1,1]^2 : |\det D^2\phi(\xi)| \lesssim R^{-4}\})$ and F_σ is Fourier supported on $\mathcal{N}^{\phi}_{R^{-1}}(\{\xi \in [-1,1]^2 : |\det D^2\phi(\xi)| \sim \sigma\})$.

We obtained the required estimate for F_0 in section 3.4.1. For F_σ, let Q be a tiling of $[-1,1]^2$ by squares of side length $\sigma^{\varepsilon/8}$. We apply Proposition 2.6.2 with ε replaced by $\varepsilon/8$ and $\lambda = R^{-1/\alpha}$ to get a partition $\mathcal{P}^{Q}_{\ge\sigma,\lambda}$ such that

$$\|(F_\sigma)_Q\|_{L^4} \lesssim_{\varepsilon,\phi} \sigma^{-\varepsilon/8} \left(\sum_{\omega \in \mathcal{P}^{Q}_{\ge\sigma,\lambda}} \|(F_\sigma)_{Q\cap\omega}\|_{L^4}^2 \right)^{1/2}, \tag{3.14}$$

where the decoupling constant is independent of Q, and each $\omega \in \mathcal{P}^{Q}_{\ge\sigma,\lambda}$ is $(\phi, \sigma', R, \varepsilon)$-admissible for some $\sigma' \ge \sigma$.

By Proposition 3.1.3 and the fact that $(F_\sigma)_\omega$ is Fourier supported on a $(\phi, \sigma', R, \varepsilon)$-admissible set for some $\sigma' \ge \sigma$, we have

$$\|(F_\sigma)_{Q\cap\omega}\|_{L^4} \lesssim_{\varepsilon} R^{-1/2}\sigma^{\varepsilon/2} \|\widehat{(F_\sigma)_{Q\cap\omega}}\|_{L^2(dM_\varepsilon)}. \tag{3.15}$$

Putting (3.15) to (3.14), we obtain

$$\begin{aligned} \|(F_\sigma)_Q\|_{L^4} &\lesssim_{\varepsilon,\phi} \sigma^{-\varepsilon/8} \left(\sum_{\omega \in \mathcal{P}_{\ge\sigma,\lambda}} R^{-1/2}\sigma^{\varepsilon/2} \|\widehat{(F_\sigma)_{Q\cap\omega}}\|_{L^2(dM_\varepsilon)}^2 \right)^{1/2} \\ &\lesssim_{\varepsilon,\phi} R^{-1/2}\sigma^{\varepsilon/4} \|\widehat{(F_\sigma)_Q}\|_{L^2(dM_\varepsilon)}. \end{aligned}$$

By trivial decoupling (1.3) and the fact that there are $\sigma^{-\varepsilon/4}$ many squares Q, we have

$$\|F_\sigma\|_{L^4} \lesssim_{\varepsilon,\phi} (\sigma^{-\varepsilon/4})^{1-1/2-1/4} \left(\sum_Q \|(F_\sigma)_Q\|_{L^4}^2\right)^{1/2} \lesssim_{\varepsilon,\phi} R^{-1/2}\sigma^{3\varepsilon/16}\|\hat{F}_\sigma\|_{L^2(dM_\varepsilon)}.$$

Finally, we sum up the dyadic pieces:

$$\begin{aligned}
\|F\|_{L^4(B_R)} &\leq \|F_0\|_{L^4(B_R)} + \sum_{\substack{\sigma \text{ dyadic} \\ R^{-4}<\sigma\leq 1}} \|F_\sigma\|_{L^4(B_R)} \\
&\lesssim_{\varepsilon,\phi} R^{-1/2-\varepsilon}\|\hat{F}_0\|_{L^2(dM_\varepsilon)} + \sum_{\substack{\sigma \text{ dyadic} \\ R^{-4}<\sigma\leq 1}} R^{-1/2}\sigma^{3\varepsilon/16}\|\hat{F}_\sigma\|_{L^2(dM_\varepsilon)} \\
&\leq R^{-1/2}\left(R^{-\varepsilon} + \sum_{\substack{\sigma \text{ dyadic} \\ R^{-4}<\sigma\leq 1}} \sigma^{3\varepsilon/8}\right)^{1/2} \left(\|\hat{F}_0\|_{L^2(dM_\varepsilon)}^2 + \sum_{\substack{\sigma \text{ dyadic} \\ R^{-4}<\sigma\leq 1}} \|\hat{F}_\sigma\|_{L^2(dM_\varepsilon)}^2\right)^{1/2} \\
&\lesssim R^{-1/2}\|\hat{F}\|_{L^2(dM_\varepsilon)},
\end{aligned}$$

as desired.

For the case where ϕ is a bounded polynomial of degree at most k, the same proof applies except we apply the uniform estimates from Proposition 2.6.1 so that every $\lesssim_{\varepsilon,\phi}$ is replaced by $\lesssim_{\varepsilon,k}$.

3.4.3 Proof of the estimate with affine surface measure

In this subsection, we prove the estimate with the affine surface measure (3.3). This is implied by (3.4) and the estimates in Section 3.4.1.

We decompose $F = F_0 + F_1$ so that F_0 is Fourier supported on $\mathcal{N}^\phi_{R^{-1}}(\{\xi \in [-1,1]^2 : |\det D^2\phi(\xi)| \lesssim R^{-4}\})$ and F_1 is Fourier supported on $\mathcal{N}^\phi_{R^{-1}}(\{\xi \in [-1,1]^2 : |\det D^2\phi(\xi)| \geq R^{-4}\})$. The estimate for F_0 was already obtained in Section 3.4.1. It suffices to estimate F_1.

Note that on the set where $|\det D^2\phi(\xi)| \geq R^{-4}$, we have

$$dM_\varepsilon \leq R^{4\varepsilon} dM.$$

Therefore, by using (3.4), we have

$$\|F\|_{L^4} \lesssim_{\varepsilon,\phi} R^{-1/2}\|\hat{F}\|_{L^2(dM_\varepsilon)} \leq R^{4\varepsilon-1/2}\|\hat{F}\|_{L^2(dM)}. \tag{3.16}$$

Since (3.16) is true for arbitrary $\varepsilon > 0$, we have obtained (3.3).

The case where ϕ is a bounded polynomial of degree at most k is similar. We have proved Proposition 3.1.1.

3.5 A counterexample

In this section, we prove that the estimates in Theorem 1.2.2, Corollary 1.2.3, and Proposition 3.1.1 are optimal in the following sense. The following counterexample is modified from Sjölin's two dimensional example in [59].

Proposition 3.5.1. *Let $1 < q \leq \infty$, $\frac{2}{p} = 1 - \frac{1}{q}$ and $3 \leq k \in \mathbb{N}$. Let $\phi(\xi) = \psi(|\xi|)$, where*

$$\psi(r) = \begin{cases} e^{-1/r}\sin(r^{-k}) & \text{if } r > 0; \\ 0 & \text{if } r = 0. \end{cases}$$

Let μ^0 be defined as in (1.13) and $E = E^{\phi,\mu^0}$ be as in (1.1). Then, there is no constant C such that the following holds for all $g \in L^q(d\mu^0)$:

$$\|Eg\|_{L^p} \leq C\|g\|_{L^q(d\mu^0)}. \tag{3.17}$$

Proof. Let p, q be on the scaling line: $2 - \frac{2}{p} = \frac{1}{q}$ and $1 \leq q < \infty$. Suppose that there exists constant C such that the following holds for all $g \in L^q(d\mu^0)$:

$$\|Eg\|_{L^p} \leq C\|g\|_{L^q(d\mu^0)}. \tag{3.18}$$

Let $g = \chi_{B(0,\frac{1}{n})}$ for some $n \gg 1$.

Note that if $|(x_1, x_2)| \leq \frac{n}{10}$ and $|x_3| \leq e^n/10$, we have

$$|Eg(x)| \gtrsim \int_{B(0,\frac{1}{n})} |\det D^2\phi(\xi)|^{\frac{1}{4}} d\xi.$$

Thus,

$$\|Eg\|_{L^p} \gtrsim (ne^n)^{1/p} \int_{B(0,\frac{1}{n})} |\det D^2\phi(\xi)|^{\frac{1}{4}} d\xi.$$

On the other hand,

$$\|g\|_{L^q(d\mu)} = \left(\int_{B(0,\frac{1}{n})} |\det D^2\phi(\xi)|^{\frac{1}{4}} d\xi\right)^{1/q}.$$

Rearranging, (3.18) can be rewritten as

$$\left(\int_{B(0,\frac{1}{n})} |\det D^2\phi(\xi)|^{\frac{1}{4}} d\xi\right)^{1/q'} \leq C(n^{-1}e^{-n})^{1/p}. \tag{3.19}$$

We now compute the integral. Using $\det D^2\phi(\xi) = \frac{\psi'(|\xi|)\psi''(|\xi|)}{|\xi|}$, we have $\det D^2\phi(\xi)$ is a finite sum of terms of the forms

$$c_1 e^{-2/r} r^{-c_2} \sin r^{-k} \cos r^{-k}, c_1 e^{-2/r} r^{-c_2} \sin r^{-k} \sin r^{-k}, c_1 e^{-2/r} r^{-c_2} \cos r^{-k} \cos r^{-k}$$

where $|\xi| = r$. The term involving the largest power (c_2) of r typically dominates when r is small. It is given by

$$c_k e^{-2/r} r^{-(3k+4)} \sin r^{-k} \cos r^{-k},$$

for some $c_k \neq 0$.

Therefore, the integral on the left-hand side (3.19) is bounded below by

$$\begin{aligned}\int_{B(0,\frac{1}{n})} |\det D^2\phi(\xi)|^{\frac{1}{4}} d\xi &\gtrsim \int_0^{2\pi}\int_0^{\frac{1}{n}} e^{-1/2r} r^{-(3k+4)/4} |\sin(2r^{-k})|^{1/4} r dr d\theta \\ &\sim \int_n^{\infty} e^{-t/2} t^{3k/4-2} |\sin(2t^k)|^{1/4} dt \\ &\geq e^{-n/2} n^{3k/4-2}.\end{aligned}$$

Now, (3.19) and above implies that

$$(e^{-n/2} n^{3k/4-2})^{1/q'} \leq C'(n^{-1}e^{-n})^{1/p},$$

for some finite C'.

Using $1/q' = 2/p$, we have

$$n^{3k/2-3} \leq C'^p n^{-1}.$$

This is impossible for large enough n unless $p = \infty$. The scaling condition then implies $q' = \infty$, contradicting the assumption. We have finished proving Proposition 3.5.1. $\square$

Bibliography

[1] J.-G. Bak, D. M. Oberlin, and A. Seeger. "Restriction of Fourier transforms to curves and related oscillatory integrals". *Amer. J. Math.* 131.2 (2009), 277–311.

[2] J.-G. Bak, D. M. Oberlin, and A. Seeger. "Restriction of Fourier transforms to curves: an endpoint estimate with affine arclength measure". *J. Reine Angew. Math.* 682 (2013), 167–205.

[3] S. Basu, S. Guo, Z. K. Li, P. L. Yung, and P. Zorin-Kranich. "A stationary set method for estimating oscillatory integrals". *preprint* (2021). arXiv:2103.08844.

[4] I. Bejenaru. "Optimal bilinear restriction estimates for general hypersurfaces and the role of the shape operator". *Int. Math. Res. Not. IMRN* 23 (2017), 7109–7147.

[5] I. Bejenaru. "Optimal multilinear restriction estimates for a class of hypersurfaces with curvature". *Anal. PDE* 12.4 (2019), 1115–1148.

[6] I. Bejenaru. "The multilinear restriction estimate: a short proof and a refinement". *Math. Res. Lett.* 24.6 (2017), 1585–1603.

[7] J. Bennett. "Aspects of multilinear harmonic analysis related to transversality". *Harmonic analysis and partial differential equations.* Vol. 612. Contemp. Math. Amer. Math. Soc., Providence, RI, 2014, 1–28.

[8] J. Bennett, A. Carbery, and T. Tao. "On the multilinear restriction and Kakeya conjectures". *Acta Math.* 196.2 (2006), 261–302.

[9] J. Bennett, An. Carbery, M. Christ, and T. Tao. "Finite bounds for Hölder-Brascamp-Lieb multilinear inequalities". *Math. Res. Lett.* 17.4 (2010), 647–666.

[10] J. Bourgain and C. Demeter. "Decouplings for surfaces in $\mathbb{R}^4$". *J. Funct. Anal.* 270.4 (2016), 1299–1318.

[11] J. Bourgain and C. Demeter. "The proof of the l^2 decoupling conjecture". *Ann. of Math. (2)* 182.1 (2015), 351–389.

[12] J. Bourgain, C. Demeter, and L. Guth. "Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three". *Ann. of Math. (2)* 184.2 (2016), 633–682.

[13] J. Bourgain and L. Guth. "Bounds on oscillatory integral operators based on multilinear estimates". *Geom. Funct. Anal.* 21.6 (2011), 1239–1295.

[14] Jean Bourgain, Ciprian Demeter, and Dominique Kemp. "Decouplings for real analytic surfaces of revolution". *Geometric aspects of functional analysis. Vol. I.* Vol. 2256. Lecture Notes in Math. Springer, Cham, 2020, 113–125.

[15] A. Carbery, C. Kenig, and S. Ziesler. "Restriction for flat surfaces of revolution in $\mathbb{R}^3$". *Proc. Amer. Math. Soc.* 135.6 (2007), 1905–1914.

[16] A. Carbery, C. E. Kenig, and S. N. Ziesler. "Restriction for homogeneous polynomial surfaces in $\mathbb{R}^3$". *Trans. Amer. Math. Soc.* 365.5 (2013), 2367–2407.

[17] M. Christ. "On the restriction of the Fourier transform to curves: endpoint results and the degenerate case". *Trans. Amer. Math. Soc.* 287.1 (1985), 223–238.

[18] C. Demeter. *Fourier restriction, decoupling, and applications.* Vol. 184. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2020, xvi+331.

[19] C. Demeter, S. Guo, and F. Shi. "Sharp decouplings for three dimensional manifolds in $\mathbb{R}^5$". *Rev. Mat. Iberoam.* 35.2 (2019), 423–460.

[20] S. W. Drury and B. P. Marshall. "Fourier restriction theorems for curves with affine and Euclidean arclengths". *Math. Proc. Cambridge Philos. Soc.* 97.1 (1985), 111–125.

[21] S. W. Drury and B. P. Marshall. "Fourier restriction theorems for degenerate curves". *Math. Proc. Cambridge Philos. Soc.* 101.3 (1987), 541–553.

[22] C. Fefferman. "Inequalities for strongly singular convolution operators". *Acta Math.* 124 (1970), 9–36.

[23] P. T. Gressman, S. Guo, L. B. Pierce, J. Roos, and P.-L. Yung. "Reversing a philosophy: from counting to square functions and decoupling". *J. Geom. Anal.* 31.7 (2021), 7075–7095.

[24] S. Guo, Z. K. Li, P. L. Yung, and P. Zorin-Kranich. "Improved discrete restriction for the parabola". *preprint* (2021). arXiv:2103.09795.

[25] S. Guo, Z. K. Li, and P.-L. Yung. "A bilinear proof of decoupling for the cubic moment curve". *Trans. Amer. Math. Soc.* 374.8 (2021), 5405–5432.

[26] S. Guo, C. Oh, J. Roos, Yung, and P. Zorin-Kranich. "Decoupling for two quadratic forms in three variables: a complete characterization". *Rev. Mat. Iberoam.* (2022).

[27] S. Guo and R. Zhang. "On integer solutions of Parsell-Vinogradov systems". *Invent. Math.* 218.1 (2019), 1–81.

[28] S. Guo and P. Zorin-Kranich. "Decoupling for certain quadratic surfaces of low codimensions". *J. Lond. Math. Soc. (2)* 102.1 (2020), 319–344.

[29] S. Guo and P. Zorin-Kranich. "Decoupling for moment manifolds associated to Arkhipov-Chubarikov-Karatsuba systems". *Adv. Math.* 360 (2020), 106889, 56.

[30] Shaoming Guo, Zane Kun Li, Po-Lam Yung, and Pavel Zorin-Kranich. "A short proof of ℓ^2 decoupling for the moment curve". *Amer. J. Math.* 143.6 (2021), 1983–1998.

[31] Shaoming Guo, Changkeun Oh, Ruixiang Zhang, and Pavel Zorin-Kranich. "Decoupling inequalities for quadratic forms". *Duke Math. J.* 172.2 (2023), 387–445.

[32] L. Guth. "A restriction estimate using polynomial partitioning". *J. Amer. Math. Soc.* 29.2 (2016), 371–413.

[33] L. Guth. “A short proof of the multilinear Kakeya inequality”. *Math. Proc. Cambridge Philos. Soc.* 158.1 (2015), 147–153.

[34] L. Guth. “Restriction estimates using polynomial partitioning II”. *Acta Math.* 221.1 (2018), 81–142.

[35] L. Guth. “The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture”. *Acta Math.* 205.2 (2010), 263–286.

[36] L. Guth and D. Maldague. “Small cap decoupling for the moment curve in $\mathbb{R}^3$”. *preprint* (2022). arXiv:2206.01574.

[37] L. Guth and J. Zahl. “Polynomial Wolff axioms and Kakeya-type estimates in $\mathbb{R}^4$”. *Proc. Lond. Math. Soc. (3)* 117.1 (2018), 192–220.

[38] J. Hickman. “An affine Fourier restriction theorem for conical surfaces”. *Mathematika* 60.2 (2014), 374–390.

[39] I. A. Ikromov and D. Müller. *Fourier restriction for hypersurfaces in three dimensions and Newton polyhedra*. Vol. 194. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 2016, viii+258.

[40] I. A. Ikromov and D. Müller. “On adapted coordinate systems”. *Trans. Amer. Math. Soc.* 363.6 (2011), 2821–2848.

[41] N. H. Katz and K. M. Rogers. “On the polynomial Wolff axioms”. *Geom. Funct. Anal.* 28.6 (2018), 1706–1716.

[42] D. Kemp. “Decoupling via Scale-based Approximations of Limited Efficacy”. *preprint* (2021). arXiv:2104.04115.

[43] Dóminique Kemp. “Decouplings for surfaces of zero curvature”. *J. Funct. Anal.* 283.12 (2022), Paper No. 109663, 23.

[44] J. Li. “L^2 affine Fourier restriction theorems for smooth surfaces in $\mathbb{R}^3$”. *preprint* (2022). arXiv:2210.15015.

[45] J. Li and T. Yang. "Decoupling for mixed-homogeneous polynomials in $\mathbb{R}^3$". *Math. Ann.* 383.3-4 (2022), 1319–1351.

[46] J. Li and T. Yang. "Decoupling for smooth surfaces in $\mathbb{R}^3$". *preprint* (2021). arXiv:2110.08441.

[47] D. Müller. "Problems of harmonic analysis related to finite-type hypersurfaces in $\mathbb{R}^3$, and Newton polyhedra". *Advances in analysis: the legacy of Elias M. Stein.* Vol. 50. Princeton Math. Ser. Princeton Univ. Press, Princeton, NJ, 2014, 303–345.

[48] D. M. Oberlin. "A uniform Fourier restriction theorem for surfaces in $\mathbb{R}^d$". *Proc. Amer. Math. Soc.* 140.1 (2012), 263–265.

[49] C. Oh. "Small cap decoupling inequalities: bilinear methods". *Rev. Mat. Iberoam.* 38.1 (2022), 33–52.

[50] Y. Ou and H. Wang. "A cone restriction estimate using polynomial partitioning". *J. Eur. Math. Soc. (JEMS)* 24.10 (2022), 3557–3595.

[51] L. Palle. "Mixed norm Strichartz-type estimates for hypersurfaces in three dimensions". *Math. Z.* 297.3-4 (2021), 1529–1599.

[52] L. Palle. "Strichartz estimates for mixed homogeneous surfaces in three dimensions". *preprint* (2020). arXiv:2004.07751.

[53] D. H. Phong and E. M. Stein. "The Newton polyhedron and oscillatory integral operators". *Acta Math.* 179.1 (1997), 105–152.

[54] M. Pramanik and A. Seeger. "L^p regularity of averages over curves and bounds for associated maximal operators". *Amer. J. Math.* 129.1 (2007), 61–103.

[55] E. Prestini. "A restriction theorem for space curves". *Proc. Amer. Math. Soc.* 70.1 (1978), 8–10.

[56] E. Prestini. "Restriction theorems for the Fourier transform to some manifolds in $\mathbb{R}^n$". *Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1.* Proc. Sympos. Pure Math., XXXV, Part. Amer. Math. Soc., Providence, R.I., 1979, 101–109.

[57] B. Shayya. “Affine restriction for radial surfaces”. *Math. Z.* 262.1 (2009), 41–55.

[58] B. Shayya. “An affine restriction estimate in $\mathbb{R}^3$”. *Proc. Amer. Math. Soc.* 135.4 (2007), 1107–1113.

[59] P. Sjölin. “Fourier multipliers and estimates of the Fourier transform of measures carried by smooth curves in R^2”. *Studia Math.* 51 (1974), 169–182.

[60] E. M. Stein. “Oscillatory integrals in Fourier analysis”. *Beijing lectures in harmonic analysis (Beijing, 1984)*. Vol. 112. Ann. of Math. Stud. Princeton Univ. Press, Princeton, NJ, 1986, 307–355.

[61] E. M. Stein. “Some problems in harmonic analysis”. *Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1*. Proc. Sympos. Pure Math., XXXV, Part. Amer. Math. Soc., Providence, R.I., 1979, 3–20.

[62] E. M. Stein and G. Weiss. “Interpolation of operators with change of measures”. *Trans. Amer. Math. Soc.* 87 (1958), 159–172.

[63] B. Stovall. “Linear and bilinear restriction to certain rotationally symmetric hypersurfaces”. *Trans. Amer. Math. Soc.* 369.6 (2017), 4093–4117.

[64] B. Stovall. “Uniform estimates for Fourier restriction to polynomial curves in $\mathbb{R}^d$”. *Amer. J. Math.* 138.2 (2016), 449–471.

[65] R. S. Strichartz. “Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations”. *Duke Math. J.* 44.3 (1977), 705–714.

[66] B. B. Taberner. “On the restriction of the Fourier transform to a conical surface”. *Trans. Amer. Math. Soc.* 292.1 (1985), 321–333.

[67] T. Tao. “A sharp bilinear restrictions estimate for paraboloids”. *Geom. Funct. Anal.* 13.6 (2003), 1359–1384.

[68] T. Tao. “Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates”. *Math. Z.* 238.2 (2001), 215–268.

[69] T. Tao and A. Vargas. "A bilinear approach to cone multipliers. I. Restriction estimates". *Geom. Funct. Anal.* 10.1 (2000), 185–215.

[70] T. Tao, A. Vargas, and L. Vega. "A bilinear approach to the restriction and Kakeya conjectures". *J. Amer. Math. Soc.* 11.4 (1998), 967–1000.

[71] P. A. Tomas. "A restriction theorem for the Fourier transform". *Bull. Amer. Math. Soc.* 81 (1975), 477–478.

[72] A. N. Varchenko. "Newton polyhedra and estimates of oscillating integrals". *Funct. Anal. Its Appl.* 10.10 (1976), 175–196.

[73] T. Wolff. "A sharp bilinear cone restriction estimate". *Ann. of Math. (2)* 153.3 (2001), 661–698.

[74] T. Wolff. "Local smoothing type estimates on L^p for large p". *Geom. Funct. Anal.* 10.5 (2000), 1237–1288.

[75] T. Yang. "Uniform l^2-decoupling in $\mathbb{R}^2$ for polynomials". *J. Geom. Anal.* (2021).

[76] A. Zygmund. "On Fourier coefficients and transforms of functions of two variables". *Studia Math.* 50 (1974), 189–201.