Acta Math., 221 (2018), 81–142
DOI: 10.4310/ACTA.2018.v221.n1.a3

Restriction estimates using polynomial partitioning II

by

Larry Guth

Massachusetts Institute of Technology
Cambridge, MA, U.S.A.

1. Introduction

This paper gives improved restriction estimates for the paraboloid in high dimensions. Recall that the extension operator for the paraboloid can be written in the form

$$Ef(x) := \int_{B^{n-1}} e^{i(x_1\omega_1+\ldots+x_{n-1}\omega_{n-1}+x_n|\omega|^2)} f(\omega)\, d\omega, \tag{1.1}$$

where B^{n-1} denotes the unit ball in $\mathbb{R}^{n-1}$ and $x\in\mathbb{R}^n$. Stein [S1] conjectured that the extension operator should obey the inequality

$$\|Ef\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^p(B^{n-1})} \tag{1.2}$$

for all $p>2n/(n-1)$. We prove new partial results towards this conjecture in dimension $n\geqslant 4$.

Theorem 1.1. *For $n\geqslant 2$, the operator E obeys the estimate (1.2) if*

$$p > 2\frac{3n+1}{3n-3} \quad \textit{for } n \textit{ odd}, \tag{1.3}$$

$$p > 2\frac{3n+2}{3n-2} \quad \textit{for } n \textit{ even}. \tag{1.4}$$

The best previous estimates for the problem were proven by Tao [T2] for $n=4$ and by Bourgain and the author [BG] for $n\geqslant 5$. For $n=4$, the conjecture is that (1.2) holds for $p>2\frac{2}{3}$. Theorem 1.1 gives the range $p>2.8$, and the best previous estimate was $p>3$. Asymptotically, for large n, the conjecture is that (1.2) holds for p bigger than the lower bound

$$\frac{2n}{n-1} = 2+2n^{-1}+O(n^{-2}).$$

The lower bound for p in Theorem 1.1 is

$$2+\tfrac{8}{3}n^{-1}+O(n^{-2}),$$

and the lower bound in the best previous estimate was $2+3n^{-1}+O(n^{-2})$.

The new ingredient of our argument has to do with algebraic structure. Roughly speaking, the argument shows that, if $\|Ef\|_{L^p}$ is large, then the region where $|Ef|$ is large must be organized into thin neighborhoods of low-degree algebraic varieties. Exploiting this structure leads to improved bounds on $\|Ef\|_{L^p}$. We find this algebraic structure using the tool of polynomial partitioning, which was introduced by Katz and the author in [GK].

Polynomial partitioning was first applied to the restriction problem in [G], which gave the best current restriction estimate in dimension 3. In this paper we combine that approach with ideas from [BG]. Besides making incremental progress on the restriction conjecture, the methods in this paper are related to sharp results for some other problems in the field, which we describe in the next two subsections.

1.1. Related work

In this subsection, we describe two papers which build on this one, and adapt the methods to other problems.

In [GHI], Hickman, Iliopoulou and the author generalize Theorem 1.1 to the setting of Hörmander-type operators with positive-definite phase. For this more general class of operators, the estimates are sharp up to the endpoint. Hörmander-type operators with positive-definite phase can be thought of as small perturbations of the extension operator E. To formulate this precisely, we first write E in a slightly different form. We define the phase function

$$\Psi_{par}(y,\omega) := y_1\omega_1+\ldots+y_{n-1}\omega_{n-1}+y_n|\omega|^2. \tag{1.5}$$

We restrict x to a ball B^n_R of radius R. For $x\in B^n_R$, we can write $Ef(x)$ in the form

$$Ef(x) = \int_{B^{n-1}} e^{iR\Psi_{par}(x/R,\omega)} f(\omega)\, d\omega.$$

We think of Ψ_{par} as a function from $B^n\times B^{n-1}$ to $\mathbb{R}$. We now consider other phase functions $\Psi(y,\omega)$, which are small C^∞ perturbations of $\Psi_{par}(y,\omega)$ on $B^n\times B^{n-1}$. For each such phase function Ψ, and each scale R, we define an operator

$$Tf(x) = \int_{B^{n-1}} e^{iR\Psi(x/R,\omega)} f(\omega)\, d\omega. \tag{1.6}$$

Hörmander introduced this type of operator in [H]. As we said above, operators of the form (1.6) can be thought of as small perturbations of the extension operator E. Hörmander raised the question whether all such operators obey the L^p bounds conjectured to hold for E, and he proved that this is the case when $n=2$. But it turns out to be false for all $n\geqslant 3$. A counterexample was found by Wisewell [Wi] (cf. also [BG]). These counterexamples build on a well-known counterexample of Bourgain from [Bo1] for a related but slightly different problem. These counterexamples are surprising, because they show that a C^∞ small perturbation of the phase function can cause a major change in the behavior of the operator. In this context, it is reasonable to ask about the best L^p estimates that hold for all operators of the form (1.6)—the best estimates that are robust to such small perturbations. Hörmander [H] answered this question in dimension $n=2$, and Lee [L] did so in dimension $n=3$. The paper [GHI] does so for all n. It shows that

$$\|Tf\|_{L^p(B_R^n)} \lesssim \|f\|_{L^p(B^{n-1})} \quad \text{for the range of } p \text{ in Theorem 1.1.}$$

The counterexamples from [Wi] and [BG] show that, up to the endpoint, this is the sharp range of p in every dimension.

In another direction, in [OW], Ou and Wang adapt the methods here to the case of the cone. They prove the sharp range of restriction estimates for the cone in dimension $n\leqslant 5$. Previously, Wolff [Wol3] proved the sharp range of restriction estimates for the cone in dimension $n\leqslant 4$.

1.2. k-linear estimates and k-broad estimates

Multilinear estimates have played a key role in the recent developments in restriction theory. Our main new result, which leads to Theorem 1.1, is a weaker version of a k-linear restriction estimate, which we call a k-broad estimate. The exponents in our k-broad estimate are sharp for all k. We recall some background on multilinear estimates and then formulate this new result.

We begin by recalling the wave packet decomposition. Suppose we want to study Ef on a large ball $B_R\subset\mathbb{R}^n$. We decompose the domain B^{n-1} into balls θ of radius $R^{-1/2}$. Then, we decompose f in the form

$$f=\sum_{\theta,v} f_{\theta,v},$$

where $f_{\theta,v}$ is supported in θ and has Fourier transform essentially supported in a ball around v of radius $R^{1/2}$. In the sum, θ ranges over our set of finitely overlapping balls covering B^{n-1}, and v ranges over $R^{1/2}\mathbb{Z}^{n-1}$. For each pair (θ,v), the restriction of

$Ef_{\theta,v}$ to B_R is essentially supported on a tube $T_{\theta,v}$ with radius $R^{1/2}$ and length R. The direction of this tube depends only on θ, and we denote it by $G(\theta)\in S^{n-1}$. We call $Ef_{\theta,v}$ a wave packet.

We can now describe multilinear restriction estimates. Given subsets $U_1,\dots,U_k$ of B^{n-1}, we say that they are *transverse* if, for any choice of $\theta_j\subset U_j$, the directions $G(\theta_1),...,G(\theta_k)$ are quantitatively transverse in the sense that

$$|G(\theta_1)\wedge...\wedge G(\theta_k)|\gtrsim 1. \tag{1.7}$$

Building on important work of Wolff [Wol3], Tao [T2] proved a sharp bilinear estimate for the extension operator E.

THEOREM 1.2. (2-linear restriction, [T2]) *If $U_1,U_2\subset B^{n-1}$ are transverse, and f_j is supported in U_j, then*

$$\Big\|\prod_{j=1}^{2}|Ef_j|^{1/2}\Big\|_{L^p(B_R)}\lesssim R^{\varepsilon}\prod_{j=1}^{2}\|f_j\|^{1/2}_{L^2(B^{n-1})} \tag{1.8}$$

for $p\geqslant 2(n+2)/n$.

By an argument of Tao, Vargas, and Vega, [TVV], this bilinear estimate implies that $\|Ef\|_{L^p(B_R)}\lesssim R^{\varepsilon}\|f\|_{L^p(B^{n-1})}$ in the same range $p\geqslant 2(n+2)/n$. The ε-removal theorem ([T1]) then implies $\|Ef\|_{L^p(B_R)}\lesssim\|f\|_{L^p(B^{n-1})}$ for all $p>2(n+2)/n$.

A few years after the bilinear results, Bennett, Carbery, and Tao [BCT] proved a sharp n-linear estimate for E.

THEOREM 1.3. (n-linear restriction, [BCT]) *If $U_1,...,U_n\subset B^{n-1}$ are transverse, and f_j is supported in U_j, then*

$$\Big\|\prod_{j=1}^{n}|Ef_j|^{1/n}\Big\|_{L^p(B_R)}\lesssim R^{\varepsilon}\prod_{j=1}^{n}\|f_j\|^{1/n}_{L^2(B^{n-1})} \tag{1.9}$$

for $p\geqslant 2n/(n-1)$.

This theorem is important and remarkable in part because it involves the sharp exponent for the restriction problem: $p>2n/(n-1)$. The paper [BG] gives a technique to exploit multilinear restriction estimates in order to get improved estimates on the original restriction problem. Since then, multilinear restriction has had many applications, including the striking recent work of Bourgain and Demeter on decoupling (see [BD] and many followup papers).

Given this 2-linear estimate and this n-linear estimate, it is natural to try to prove a k-linear estimate for all $2\leqslant k\leqslant n$ which would include these two estimates as special cases. Here is what looks to me like the natural conjecture, which I first learned from Jonathan Bennett.

Conjecture 1.4. (k-linear restriction) If $U_1,...,U_k \subset B^{n-1}$ are transverse, and f_j is supported in U_j, then

$$\Big\| \prod_{j=1}^{k} |Ef_j|^{1/k} \Big\|_{L^p(B_R)} \lesssim R^{\varepsilon} \prod_{j=1}^{k} \|f_j\|^{1/k}_{L^2(B^{n-1})} \tag{1.10}$$

for $p \geqslant \bar{p}(k,n) := 2(n+k)/(n+k-2)$.

Having the full range of k-linear estimates available would improve the results from [BG]. Combining Conjecture 1.4 with the method from [BG] would give $\|Ef\|_{L^p} \lesssim \|f\|_{L^p}$ for exactly the range of p in Theorem 1.1.

For $3 \leqslant k \leqslant n-1$, Conjecture 1.4 is open. In [Be1] and [Be2], Bejenaru proves multilinear estimates for certain curved hypersurfaces, but not including the paraboloid. The surfaces he considers are foliated by $(k-1)$-planes and they are curved in the transverse directions, in an appropriate sense. The main new result of this paper is a weak version of Conjecture 1.4, which we call a k-broad restriction inequality. To motivate this inequality, let us recall the approach from [BG] for deducing linear estimates from multilinear ones.

We decompose B^{n-1} into balls τ of radius K^{-1}, where K is a large constant. This decomposition is much coarser than the decomposition into balls θ of radius $R^{-1/2}$. We write $f = \sum_\tau f_\tau$, where f_τ is supported in τ. Next we subdivide B^n_R into much smaller balls. In [BG], we used balls of radius K, but it will be slightly more convenient here to use balls of radius K^2. For each $B_{K^2} \subset B_R$, we consider $\int_{B_{K^2}} |Ef_\tau|^p$ for each τ. We say that τ contributes significantly to B_{K^2} if

$$\int_{B_{K^2}} |Ef_\tau|^p \gtrsim K^{-10n} \int_{B_{K^2}} |Ef|^p.$$

We let $S(B_{K^2})$ denote the set of τ which contribute significantly to B_{K^2}. We now break the balls B_{K^2} into two classes. We label a ball B_{K^2} as k-transverse if there are k significant τ's which are k-transverse in the sense above. We label a ball B_{K^2} as k-non-transverse otherwise. A k-linear restriction estimate gives a good bound for the integral of $|Ef|^p$ over the union of all of the k-transverse balls. The paper [BG] then gives an inductive argument to control the contribution from the k-non-transverse balls.

This inductive argument can be described most cleanly using the language of decoupling. Building on [BG], Bourgain proved a decoupling theorem in [Bo4] which implies that, for each k-non-transverse ball B_{K^2}, one has

$$\|Ef\|^2_{L^p(B_{K^2})} \lesssim \sum_{\tau \in S(B_{K^2})} \|Ef_\tau\|^2_{L^p(B_{K^2})}$$

for a certain range of p which covers the exponents we study. For a k-non-transverse ball B_{K^2}, all the significant τ have direction $G(\tau)$ within K^{-1} of some $(k-1)$-plane. In particular, $|S(B_{K^2})| \lesssim K^{k-2}$. Using this bound and Hölder's inequality, we get

$$\int_{B_{K^2}} |Ef|^p \lesssim K^{\alpha} \sum_{\tau \in S(B_{K^2})} \int_{B_{K^2}} |Ef_\tau|^p, \tag{1.11}$$

with $\alpha = (K^{k-2})^{(p-2)/2}$, which is optimal. Summing this inequality over all the k-non-transverse balls gives

$$\int_{\bigcup k\text{-non-transverse balls}} |Ef|^p \lesssim K^{\alpha} \sum_{\tau} \int_{B_R} |Ef_\tau|^p.$$

The right-hand side may then be controlled by induction on scales.

To make this strategy work, we do not need a full k-linear bound. The decoupling estimate that we used to control the k-non-transverse balls applies whenever all the significant τ for a ball B_{K^2} have directions $G(\tau)$ lying within the $O(K^{-1})$-neighborhood of $O(1)$ $(k-1)$-planes. We call such a ball *k-narrow*. To get the argument to work, we only need to bound $\int |Ef|^p$ over the remaining balls—the k-broad balls.

Here is a little notation using which we can state our k-broad bound precisely. We let $G(\tau) = \bigcup_{\theta \subset \tau} G(\theta)$. The set $G(\tau) \subset S^{n-1}$ is a spherical cap with radius $\sim K^{-1}$, representing the possible directions of wave packets in Ef_τ. If $V \subset \mathbb{R}^n$ is a subspace, then we write $\mathrm{Angle}(G(\tau), V)$ for the smallest angle between any non-zero vectors $v \in V$ and $v' \in G(\tau)$. For each ball $B_{K^2} \subset B_R$, we consider $\int_{B_{K^2}} |Ef_\tau|^p$ for every τ. To define the k-broad norm, we discount the contributions of f_τ with $G(\tau)$ lying near a few $(k-1)$-planes, and we record the largest remaining contribution. More formally, for a parameter A, we define

$$\mu_{Ef}(B_{K^2}) := \min_{V_1, \ldots, V_A} \left(\max_{\tau} \int_{B_{K^2}} |Ef_\tau|^p \right), \tag{1.12}$$

where the minimum is over $(k-1)$-subspaces of $\mathbb{R}^n$ and the maximum is over all τ such that $\mathrm{Angle}(G(\tau), V_a) > K^{-1}$ for all a.

We can now define the k-broad part of $\|Ef\|_{L^p(B_R)}$ by

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} := \sum_{B_{K^2} \subset B_R} \mu_{Ef}(B_{K^2}). \tag{1.13}$$

Our main new result is an estimate for this k-broad norm.

THEOREM 1.5. *For any $2 \leqslant k \leqslant n$ and any $\varepsilon > 0$ there is a large constant A such that the following holds (for any value of K):*

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim_{K,\varepsilon} R^{\varepsilon} \|f\|_{L^2(B^{n-1})}, \tag{1.14}$$

for $p \geqslant \bar{p}(k,n) = 2(n+k)/(n+k-2)$.

The range of p in Theorem 1.5 is sharp for all k and n. Using the method from [BG] outlined above, Theorem 1.5 implies the restriction estimate Theorem 1.1.

We should mention that $\mathrm{BL}^p_{k,A}$ is not literally a norm, but it has some similar properties, which is why we use the norm notation. In particular, $\mathrm{BL}^p_{k,A}$ obeys the following weak version of the triangle inequality: if $f=g+h$, then

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim \|Eg\|_{\mathrm{BL}^p_{k,A/2}(B_R)} + \|Eh\|_{\mathrm{BL}^p_{k,A/2}(B_R)}. \tag{1.15}$$

The reason for introducing the parameter A is to use this version of the triangle inequality. If we choose $A(\varepsilon)$ very large, then we can effectively use the triangle inequality $O_\varepsilon(1)$ times during our proof, and so $\mathrm{BL}^p_{k,A}$ behaves almost like a norm. We were not able to prove Conjecture 1.4, and the main issue is that in the true k-linear setting, we do not have a substitute for this triangle inequality.

1.3. Examples

To help digest Theorem 1.5, we describe a couple of examples. These examples show that the range of exponents p in Theorem 1.5 is sharp.

In one example, the wave packets $Ef_{\theta,v}$ concentrate in the $R^{1/2}$-neighborhood of a k-plane. We denote this neighborhood by W. Each wave packet $Ef_{\theta,v}$ has $|Ef_{\theta,v}(x)|\sim 1$ on the tube $T_{\theta,v}$, and rapidly decaying outside $T_{\theta,v}$. It is not hard to arrange that each point in the slab W lies in many wave packets $T_{\theta,v}$, pointing in many directions within the k-plane. For each ball B_{K^2} in the slab, only a tiny fraction of the wave packets through this ball lie near any $(k-1)$-plane. In this scenario, $\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \sim \|Ef\|_{L^p(B_R)}$. We can also arrange that $|Ef(x)|^2 \sim \sum_{\theta,v} |Ef_{\theta,v}(x)|^2$ at most of the points x, by replacing $f_{\theta,v}$ by $\pm f_{\theta,v}$ with independent random sings. We can distribute the wave packets evenly, so that $|Ef(x)|$ is roughly constant on the slab. Moreover, by a standard orthogonality argument $\|Ef\|_{L^2(B_R)} \sim R^{1/2}\|f\|_{L^2}$. Therefore, we get

$$\frac{\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}}{\|f\|_{L^2}} \sim R^{1/2} \frac{\|Ef\|_{L^p(B_R)}}{\|Ef\|_{L^2(B_R)}} \sim R^{1/2}|W|^{1/p-1/2}.$$

Since $|W| \sim R^k R^{(1/2)(n-k)}$, a short calculation shows that the ratio $\|Ef\|_{L^p(B_R)}/\|f\|_{L^2}$ is bounded for $p \geqslant \bar{p}(k,n)$ and blows up for $p < \bar{p}(k,n)$.

But there are also more complicated sharp examples, coming from low-degree algebraic varieties. This type of example was first pointed out to me by Josh Zahl. For instance, consider the quadric hypersurface $Z \subset \mathbb{R}^4$ defined by

$$\left(\frac{x_1}{R}\right)^2 + \left(\frac{x_2}{R}\right)^2 - \left(\frac{x_3}{R}\right)^2 - \left(\frac{x_4}{R}\right)^2 = 1.$$

Each point of Z lies in a 1-parameter family of lines in Z. The union of the lines through a given point form a 2-dimensional cone. For example, the point $(R,0,0,0)$ lies in the cone defined by

$$x_1 = R \quad \text{and} \quad \left(\frac{x_2}{R}\right)^2 - \left(\frac{x_3}{R}\right)^2 - \left(\frac{x_4}{R}\right)^2 = 0.$$

If we take the $R^{1/2}$-neighborhoods of lines in Z, we can find many tubes in the $R^{1/2}$-neighborhood of Z. We can now build an example like the one above using wave packets concentrated in the $R^{1/2}$-neighborhood of Z. For each ball B_{K^2} in this neighborhood, the wave packets through B_{K^2} fill out a 2-dimensional cone, and very few of them lie near any 2-dimensional plane. Therefore, $\|Ef\|_{\mathrm{BL}^p_{3,A}(B_R)} \sim \|Ef\|_{L^p(B_R)}$. The rest of the discussion in the planar slab example applies also here, and so we see that this example is sharp for Theorem 1.5 in dimension 4 with $k=3$.

For larger n, there are more variations on this example. The dimension of the variety Z in these examples is k. The degree of Z may be larger than 2, although in the known examples it is always bounded by $C(n)$. Similar examples apply to Conjecture 1.4.

These examples help to suggest that algebraic varieties could be relevant to Theorem 1.5. Polynomial partitioning is a tool that helps us to find and exploit the type of algebraic structure in these examples. If we run through the proof of Theorem 1.5 on this type of examples, then the argument will find the variety Z.

The new difficulty in this paper, compared with [G], is that it is harder to find a k-dimensional variety for small k than it is to find a hypersurface. In the next section, we will describe the polynomial partitioning process and give a sense of the issues involved.

1.4. A direction for further improvement

The paper [G] applies polynomial partitioning to the restriction problem in three dimensions. It proves an estimate which is stronger than Theorem 1.1, namely $\|Ef\|_{L^p(\mathbb{R}^3)} \lesssim \|f\|_{L^\infty}$ for $p > 3.25$. This estimate relies on one additional ingredient: an estimate for how many different θ can be represented by wave packets $T_{\theta,v}$ in the $R^{1/2}$-neighborhood of a low-degree variety Z. In the 3-dimensional case, recall that there are $\sim R$ balls $\theta \subset B^2$, each with radius $R^{-1/2}$. Let $\Theta(Z)$ denote the set of θ such that at least one wave packet $T_{\theta,v}$ is contained in the $R^{1/2}$-neighborhood of Z. Lemma 3.6 of [G] proves that, if Z is a 2-dimensional variety in $\mathbb{R}^3$ of degree $\lesssim 1$, then $|\Theta(Z)| \lesssim_\varepsilon R^{1/2+\varepsilon}$. (If Z is a 2-plane, then $|\Theta(Z)| \sim R^{1/2}$, and so the result says that the example of a plane is nearly the worst possible.)

When $\|Ef\|_{L^p}$ is large, the polynomial partitioning method locates algebraic pieces that contribute most of $\|Ef\|_{L^p}$. The bound for $|\Theta(Z)|$ gives a stronger estimate for the

contribution of each such piece in terms of $\|f\|_{L^\infty}$ or $\|f\|_{L^p}$. Notice that, for Hörmander-type operators of positive-definite phase in three dimensions, the estimate $\|Tf\|_{L^p(B^3_R)} \lesssim R^\varepsilon \|f\|_{L^\infty(B^2)}$ is false for all $p < \frac{10}{3}$. To prove the bound $\|Ef\|_{L^p} \lesssim \|f\|_{L^\infty}$ for all $p > 3.25$, the argument from [G] has to distinguish E from more general Hörmander-type operators of positive-definite phase. The bound on $|\Theta(Z)|$ is the step that does this. In the Hörmander case, $Tf_{\theta,v}$ is concentrated on a curved tube. And in the counterexample from [Wi] or [BG], there is a low-degree variety Z whose $R^{1/2}$-neighborhood contains one such curved tube for every θ.

I have not been able to prove a good bound for $|\Theta(Z)|$ in higher dimensions. Such a bound would lead to further improvements in the restriction exponents in high dimensions. We will discuss this issue more in the final section of the paper.

Acknowledgements. I was supported by a Simons Investigator Award during this work. I would also like to thank Marina Iliopoulou, Jongchon Kim, and the referee for helpful comments on a draft of the paper.

2. Sketch of the proof

In this section, we sketch the proof of the k-broad estimate, Theorem 1.5. We actually give two sketches. The first sketch aims to show the main ideas of the argument. The second sketch brings into play more of the technical issues, and it provides a detailed outline of the argument in the paper.

The proof begins with a wave packet decomposition. We decompose the domain B^{n-1} into balls θ of radius $R^{1/2}$. We then decompose the function $f: B^{n-1} \to \mathbb{C}$ as

$$f = \sum_{\theta,v} f_{\theta,v},$$

where $f_{\theta,v}$ is supported on θ and the Fourier transform of $f_{\theta,v}$ is essentially supported on a ball of radius $R^{1/2}$ around v. In the sum, v ranges over $R^{1/2}\mathbb{Z}^{n-1}$. On B_R, $Ef_{\theta,v}$ is essentially supported on a tube $T_{\theta,v}$ of radius $R^{1/2}$ and length R. In addition, the functions $f_{\theta,v}$ are essentially orthogonal. In particular, we have

$$c_{L^2}^2 \sim \sum_{\theta,v} \|f_{\theta,v}\|_{L^2}^2. \tag{2.1}$$

Our goal is to prove that

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant C(\varepsilon) R^\varepsilon \|f\|_{L^2} \quad \text{for } p = \bar{p}(k,n) := \frac{2(n+k)}{n+k-2}. \tag{2.2}$$

The proof will be by induction. So we assume that (2.2) holds for balls of smaller radii, and in lower dimension.

Recall that

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} = \sum_{B_{K^2}\subset B_R} \mu_{Ef}(B_{K^2}),$$

where $\mu_{Ef}(B_{K^2})$ was defined in (1.12). We can extend μ_{Ef} to be a measure on B_R, making it a constant multiple of the Lebesgue measure on each B_{K^2}. In particular $\mu_{Ef}(B_R) = \|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)}$.

We now introduce polynomial partitioning. We let D be a large constant that we can choose later. For a polynomial P on $\mathbb{R}^n$, we write $Z(P)$ for the zero set of P. By [G, Theorem 1.4], there is a (non-zero) polynomial P of degree at most D on $\mathbb{R}^n$ such that $\mathbb{R}^n \backslash Z(P)$ is a disjoint union of $\sim D^n$ open cells O_j, and the measures $\mu_{Ef}(O_j)$ are all equal.

Next, we consider how the wave packets $Ef_{\theta,v}$ interact with this partition. We note that a line can cross $Z(P)$ at most D times, and so a line can enter at most $D+1$ of the $\sim D^n$ cells O_j. The tube $T_{\theta,v}$ can still enter many or all cells O_j, but it can only penetrate deeply into $D+1$ cells. To make this precise, we define W to be the $R^{1/2}$-neighborhood of $Z(P)$, and we define O'_j to be $O_j \backslash W$. If a tube $T_{\theta,v}$ enters O'_j, then the axis of $T_{\theta,v}$ must enter O_j, and so we get that

$$\text{each tube } T_{\theta,v} \text{ enters at most } D+1 \text{ cells } O'_j. \tag{2.3}$$

We now have

$$\mu_{Ef}(B_R) = \sum_j \mu_{Ef}(O'_j) + \mu_{Ef}(W).$$

We say that we are in the *cellular case* when the contribution of the cells dominates and in the *algebraic case* when the contribution of W dominates. If we are in the cellular case, then there must be $\sim D^n$ cells O'_j such that

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim D^n \mu_{Ef}(O'_j). \tag{2.4}$$

Next, we study Ef on each of these cells O'_j. We define f_j by

$$f_j := \sum_{\substack{\theta,v \\ T_{\theta,v}\cap O'_j \neq \varnothing}} f_{\theta,v}. \tag{2.5}$$

Since $Ef_{\theta,v}$ is essentially supported on $T_{\theta,v}$, we see that Ef_j is almost equal to Ef on O'_j. Therefore, $\mu_{Ef_j}(O'_j) \sim \mu_{Ef}(O'_j)$. Now we study $\mu_{Ef_j}(O'_j)$ using induction. Since

f_j involves fewer wave packets than f, it is a simpler object, and so it makes sense to assume by induction that our theorem holds for f_j. This leads to the following bound:

$$\mu_{Ef}(O'_j) \sim \mu_{Ef_j}(O'_j) \leqslant \|Ef_j\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim (C(\varepsilon)R^{\varepsilon})^p \|f_j\|^p_{L^2}. \tag{2.6}$$

Next, we analyze $\|f_j\|_{L^2}$. By the orthogonality of the $f_{\theta,v}$, we have

$$\sum_j \|f_j\|^2_{L^2} \sim \sum_j \sum_{\substack{\theta,v \\ T_{\theta,v}\cap O'_j \neq \varnothing}} \|f_{\theta,v}\|^2_{L^2} \sim \sum_{\theta,v} \#\{O'_j : T_{\theta,v} \cap O'_j \neq \varnothing\} \|f_{\theta,v}\|^2_{L^2}.$$

By (2.3), each tube $T_{\theta,v}$ enters $\lesssim D$ cells O'_j, and so we get

$$\sum_j \|f_j\|^2_{L^2} \lesssim D \sum_{\theta,v} \|f_{\theta,v}\|^2_{L^2} \sim D\|f\|^2_{L^2}.$$

Since there are $\sim D^n$ cells O'_j that obey (2.4), we see that most of them must also obey

$$\|f_j\|^2_{L^2} \lesssim D^{1-n}\|f\|^2_{L^2}. \tag{2.7}$$

Combining this bound with (2.4) and our inductive assumption (2.6), we get

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant C D^{n+(1-n)p/2} (C(\varepsilon)R^{\varepsilon})^p \|f\|^p_{L^2}.$$

In this equation, the constant C is the implicit constant from the various $\lesssim$'s. It does not depend on D. The induction closes as long as the term in brackets is $\leqslant 1$. Since we can choose the constant D, we can arrange that the induction closes as long as the exponent of D is negative. Given our value of p, we can check that the exponent of D is $\leqslant 0$, and the induction closes.

Now we turn to the algebraic case. In this case, the measure μ_{Ef} is concentrated in W—the $R^{1/2}$-neighborhood of $Z(P)$—a degree-D algebraic variety of dimension $n-1$. There are two types of wave packets that contribute to μ_{Ef} on W, which we describe roughly as follows:

- *Tangential wave packets*: wave packets that are essentially contained in W. For these wave packets, the direction of the tube $T_{\theta,v}$ is (nearly) tangent to $Z = Z(P)$.
- *Transverse wave packets*: wave packets that cut across W.

We first discuss the case that the tangential wave packets dominate. To simplify the exposition, let us imagine for the moment that the variety Z is a hyperplane, so W is a planar slab of dimensions $R^{1/2} \times R \times \ldots \times R$. We can also imagine that f has the form $f = \sum_{\theta,v} f_{\theta,v}$, where all the tubes $T_{\theta,v}$ are contained in W.

We study this case using induction on the dimension. In the tangential algebraic case, the behavior of Ef on the hyperplane Z can be controlled by applying Theorem 1.5 in dimension $n-1$. Here is one way to set this up. There is a standard L^2 estimate giving

$$\|Ef\|_{L^2(B_R)} \lesssim R^{1/2}\|f\|_{L^2}. \tag{2.8}$$

It is not hard to reduce Theorem 1.5 to the case that $\|Ef\|_{L^2(B_R)} \sim R^{1/2}\|f\|_{L^2}$, and so we can think of Theorem 1.5 in the equivalent form

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim R^{-1/2+\varepsilon}\|Ef\|_{L^2(B_R)}.$$

We can apply Theorem 1.5 in dimension $n-1$ to study Ef on Z, and we get the estimate

$$\|Ef\|_{\mathrm{BL}^{\bar{p}(k,n-1)}_{k,A}(Z\cap B_R)} \lesssim R^{-1/2+\varepsilon}\|Ef\|_{L^2(Z\cap B_R)}.$$

Now $\bar{p}(k,n-1) > \bar{p}(k,n) = p$. Interpolating between this estimate and the L^2 estimate (2.8), we get a bound of the form

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(Z\cap B_R)} \lesssim R^{-1/2+e+\varepsilon}\|Ef\|_{L^2(Z\cap B_R)},$$

for $p=\bar{p}(k,n)$, where e is an exponent depending on k and n. The same estimate holds not just for Z, but for any translate of Z in the slab W. These estimates control the behavior of Ef in the directions tangent to Z.

To get a good estimate in the tangential algebraic case, we also have to control the behavior of Ef in the direction transverse to Z. We will show that there is a direction transverse to the hyperplane Z such that $|Ef(x)|$ is approximately constant as we move x in this direction for distance $\lesssim R^{1/2}$. We call this behavior a *transverse equidistribution estimate*. It implies that

$$\|Ef\|^2_{L^2(B_R)} \sim \|Ef\|^2_{L^2(W)} \sim R^{1/2}\|Ef\|^2_{L^2(Z\cap B_R)}$$

and

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \sim \|Ef\|^p_{\mathrm{BL}^p_{k,A}(W)} \sim R^{1/2}\|Ef\|^p_{\mathrm{BL}^p_{k,A}(Z\cap B_R)}.$$

Using these estimates, we can control $\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}$ by induction on the dimension:

$$\begin{aligned}\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} &\sim \|Ef\|^p_{\mathrm{BL}^p_{k,A}(W)} \sim R^{1/2}\|Ef\|^p_{\mathrm{BL}^p_{k,A}(Z\cap B_R)} \\ &\lesssim R^{1/2}R^{(-1/2+e+\varepsilon)p}\|Ef\|^p_{L^2(Z\cap B_R)} \\ &\sim R^{1/2}R^{(-1/2+e+\varepsilon)p}(R^{-1/2}\|Ef\|^2_{L^2(B_R)})^{p/2}.\end{aligned}$$

There are a lot of messy powers of R in this computation. But plugging in $p=\bar{p}(k,n)$ and working out the exponent, one gets $\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim R^{-1/2+\varepsilon}\|Ef\|_{L^2(B_R)} \lesssim R^{\varepsilon}\|f\|_{L^2}$, which closes the induction in the tangential algebraic case. The exponent

$$p=\bar{p}(k,n)=2(n+k)/(n+k-2)$$

is exactly the exponent needed to make the powers of R work out in this computation.

Next, we sketch the reason for this transverse equidistribution. From the definition of E, we see that $Ef_{\theta,v}$ has Fourier transform supported in a ball of radius $R^{-1/2}$ around

$$\xi(\theta)=(\omega_{\theta,1},...,\omega_{\theta,n-1},|\omega_\theta|^2),$$

where ω_θ denotes the center of θ. We know that all the wave packets $Ef_{\theta,v}$ in Ef are supported in tubes $T_{\theta,v}\subset W$. This situation restricts the directions $G(\theta)$, which in turn restricts the frequencies $\xi(\theta)$. The connection between $G(\theta)$ and $\xi(\theta)$ is simplest for a slightly different operator—the extension operator for the sphere. In that case, we have $G(\theta)=\xi(\theta)$. Since the tubes $T_{\theta,v}$ all lie in W, the directions $G(\theta)$ all lie in the $R^{-1/2}$-neighborhood of the subspace $V\subset\mathbb{R}^n$ parallel to Z. Therefore, the frequencies $\xi(\theta)$ all lie in the $R^{-1/2}$-neighborhood of V as well. So, the Fourier transform of Ef is supported in the $R^{-1/2}$-neighborhood of V. If ℓ denotes a line perpendicular to V (or to Z), then the restriction of Ef to ℓ has Fourier transform supported in a ball of radius $R^{-1/2}$. Therefore, $|Ef|$ is approximately constant as we move along the line ℓ for distances $\lesssim R^{1/2}$.

For the extension operator in the case of the paraboloid, the situation is similar but a touch messier. We know that the directions $G(\theta)$ all lie in the $R^{-1/2}$-neighborhood of the plane V. A short calculation shows that the frequencies $\xi(\theta)$ all lie in the $R^{-1/2}$-neighborhood of an affine hyperplane V'. The hyperplane V' is not equal to V, but the angle between V and V' is fairly small. If ℓ is perpendicular to the plane V', then it still follows that $|Ef|$ is approximately constant as we move along ℓ for distances $\lesssim R^{-1/2}$. The line ℓ is no longer exactly perpendicular to the original plane Z, but it is still quantitatively transverse to Z, and this is good enough for our application.

In this sketch, we assumed that Z is a hyperplane. But in the real proof we cannot assume this. We have to set up the induction on the dimension in a different way, taking into account the possibility that Z is curved. We explain this in the next subsection.

Finally, it can happen that the transverse wave packets dominate. In this last case, $\mu_{Ef}(W)$ dominates $\mu_{Ef}(B_R)$, but the wave packets transverse to W make the main contribution to $\mu_{Ef}(W)$. In this case, we can imagine that $f=\sum f_{\theta,v}$, where each tube $T_{\theta,v}$ is transverse to W. Recall that the number of times a line can cross the hypersurface

Z is $\leqslant D \lesssim 1$. Similarly, we will prove that the number of times a tube $T_{\theta,v}$ can cross the surface W is $\lesssim 1$.

In this case, we subdivide the ball B_R into smaller balls B_j with radius $\varrho \ll R$. A tube $T_{\theta,v}$ enters $\sim R/\varrho \gg 1$ of these balls. But, because of the discussion above, a tube $T_{\theta,v}$ can cross W transversely in $\lesssim 1$ balls B_j. We define

$$\mathbb{T}_{j,\text{trans}} := \{(\theta,v) : T_{\theta,v} \text{ crosses } W \text{ transversely in } B_j\},$$
$$f_{j,\text{trans}} := \sum_{(\theta,v)\in\mathbb{T}_{j,\text{trans}}} f_{\theta,v}.$$

Since a tube $T_{\theta,v}$ can cross W transversely in $\lesssim 1$ balls B_j, each (θ,v) belongs to $\lesssim 1$ sets $\mathbb{T}_{j,\text{trans}}$. Therefore,

$$\sum_j \|f_{j,\text{trans}}\|_{L^2}^2 \lesssim \|f\|_{L^2}^2. \tag{2.9}$$

As we assumed that all the wave packets in f intersect W transversely, $Ef_{j,\text{trans}}$ is essentially equal to Ef on $W \cap B_j$. Therefore,

$$\mu_{Ef}(W\cap B_j) \leqslant \mu_{Ef_{j,\text{trans}}}(B_j).$$

Since we are in the algebraic case, and since we assumed that all the tubes $T_{\theta,v}$ intersect W transversely, we have

$$\mu_{Ef}(B_R) \lesssim \mu_{Ef}(W) \sim \sum_j \mu_{Ef}(W\cap B_j) \lesssim \sum_j \mu_{Ef_{j,trans}}(B_j). \tag{2.10}$$

By induction on the radius, we may assume that

$$\mu_{Ef_j}(B_j) \leqslant (C(\varepsilon)\varrho^{\varepsilon})^p \|f_{j,\text{trans}}\|_{L^2}^p.$$

Plugging this bound into (2.10) and then applying (2.9), we get

$$\|Ef\|_{\text{BL}^p_{k,A}(B_R)}^p \lesssim (C(\varepsilon)\varrho^{\varepsilon})^p \sum_j \|f_{j,\text{trans}}\|_{L^2}^p.$$

As $p>2$, we get $\sum_j \|f_{j,\text{trans}}\|_{L^2}^p \leqslant (\sum_j \|f_{j,\text{trans}}\|_{L^2}^2)^{p/2} \lesssim \|f\|_{L^2}^p$. Therefore, we have

$$\|Ef\|_{\text{BL}^p_{k,A}(B_R)}^p \lesssim (C(\varepsilon)\varrho^{\varepsilon})^p \|f\|_{L^2}^p.$$

Since R/ϱ is large, $C\varrho^{\varepsilon} < R^{\varepsilon}$, and so this closes the induction in the transverse algebraic case.

2.1. Studying wave packets tangent to a variety

In the preceding sketch, we considered the special case that the variety Z is a hyperplane. In this special situation, the tangential algebraic case reduces to the original theorem in dimension $n-1$. In the full proof, we need to consider curved varieties Z, and so we have to do the induction on the dimension in a different way.

If Z is an m-dimensional variety in $\mathbb{R}^n$, then we say that the tube $T_{\theta,v}$ is *α-tangent* to Z in B_R if the following two conditions hold:

- *Distance condition*:
$$T_{\theta,v} \subset N_{\alpha R}(Z) \cap B_R.$$
- *Angle condition*: If $z \in Z \cap N_{\alpha R}(T_{\theta,v})$, then
$$\mathrm{Angle}(T_z Z, G(\theta)) \leqslant \alpha.$$

For each dimension m, we will choose an angle α_m slightly larger than $R^{-1/2}$, and then we define
$$\mathbb{T}_Z := \{(\theta, v) : T_{\theta,v} \text{ is } \alpha_m\text{-tangent to } Z \text{ in } B_R\}.$$

Our main technical result is Proposition 8.1. It says that, if Z is an m-dimensional variety of controlled degree and if $f = \sum_{(\theta,v)\in\mathbb{T}_Z} f_{\theta,v}$, then (for a range of exponents $p \geqslant 2$)
$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant M \|f\|_{L^2},$$
where M is a fairly complicated expression which depends on the parameters of the setup, including the exponent p, the radius R, the dimension m, the value of A, etc.

The proof of Proposition 8.1 is by induction on the dimension m. The base of the induction is the case $m = k-1$. In this case, $\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}$ is negligibly small. To see this, consider a small ball B_{K^2}. The function Ef is essentially supported in $N_{R^{1/2}}(Z) \cap B_R$, so we may assume that $B_{K^2} \subset N_{R^{1/2}}(Z) \cap B_R$. Because of the angle condition, all the wave packets $T_{\theta,v} \in \mathbb{T}_Z$ that pass through B_{K^2} have direction $G(\theta)$ within a small angle of a $(k-1)$-plane—the plane $T_z Z$ for a point $z \in Z$ near B_{K^2}—and so the ball B_{K^2} makes a negligible contribution to $\|Ef\|_{\mathrm{BL}^p_{k,A}}$. This provides the base of the induction.

The proof of Proposition 8.1 follows the rough outline of our first sketch. We again use polynomial partitioning. Under the hypotheses of the proposition, we know that Ef is essentially supported in $N_{R^{1/2}}(Z)$ for a given m-dimensional variety Z. We want to find a polynomial P such that $Z(P)$ cuts $N_{R^{1/2}}(Z)$ into smaller cells. To do this, we choose new orthogonal coordinates $y_1, ..., y_n$ such that the projection of Z to the $(y_1, ..., y_m)$-plane is non-degenerate on a significant portion of Z. Then, we let P be a polynomial in $y_1, ..., y_m$. In this case, $Z(P)$ intersects Z transversely (at least on a significant portion

of Z), and this makes the polynomial partitioning work. Essentially everything works as in the first sketch, except that, because P only depends on m variables, the number of cells O_j is only $\sim D^m$. This will affect the final exponents, but the method of the argument is the same. If μ_{Ef} is concentrated on the cells O_j', then we can prove the desired bounds by induction. Otherwise, μ_{Ef} is concentrated in the $R^{1/2}$-neighborhood of a lower-dimensional variety $Y = Z \cap Z(P)$.

Now we turn to this algebraic case. In the algebraic case, there is a lower-dimensional variety Y, with $\operatorname{Deg} Y \lesssim \operatorname{Deg} Z$, so that

$$\mu_{Ef}(N_{R^{1/2}}(Y) \cap B_R) \gtrsim \mu_{Ef}(B_R) \sim \mu_{Ef}(N_{R^{1/2}}(Z) \cap B_R).$$

There are two types of wave packets that contribute to μ_{Ef} on $N_{R^{1/2}}(Y) \cap B_R$: tangential wave packets, which lie in $N_{R^{1/2}}(Y)$ and run tangent to Y, and transverse wave packets, which cut across $N_{R^{1/2}}(Y)$. (Recall that all the wave packets are tangent to Z, by hypothesis.)

In our current setup, if all the wave packets are tangent to Y, then we get the desired estimate just by induction on the dimension m.

However, there may be a mix of transverse and tangential wave packets. If we let f_{tang} be the sum of the tangential wave packets and f_{trans} be the sum of the transverse wave packets, then the quasi-triangle inequality for $\text{BL}^p_{k,A}$ (1.15) gives

$$\|Ef\|_{\text{BL}^p_{k,A}(B_R)} \lesssim \|Ef_{\text{tang}}\|_{\text{BL}^p_{k,A/2}(B_R)} + \|Ef_{\text{trans}}\|_{\text{BL}^p_{k,A/2}(B_R)}.$$

This is the step in the proof where we need to use the quasi-triangle inequality.

We can handle the tangential terms by induction on the dimension, and now we turn to the transverse terms. As in the last sketch, we decompose $B_R = \bigcup_j B_j$, where each ball B_j has radius $\varrho \ll R$. We define $f_{j,\text{trans}}$ as above to be the sum of wave packets that intersect $N_{R^{1/2}}(Y)$ transversely in B_j. As in the previous sketch, geometric arguments show that a tube $T_{\theta,v}$ can intersect $N_{R^{1/2}}(Y)$ transversely in $\lesssim 1$ balls B_j, and so

$$\sum_j \|f_{j,\text{trans}}\|_{L^2}^2 \lesssim \|f\|_{L^2}^2.$$

Next, we want to study $\|Ef_{j,\text{trans}}\|_{\text{BL}^p_{k,A/2}(B_j)}$ by using induction on the radius. This step is more complicated than in the previous sketch. We know that f is concentrated on wave packets that are tangent to Z on B_R, and we need to use that information. We expand $Ef_{j,\text{trans}}$ into wave packets on the ball B_j. Since B_j has radius ϱ, each wave packet is essentially supported on a tube of radius $\varrho^{1/2}$ and length ϱ in B_j. When we examine this wave packet decomposition, it is not exactly true that all the wave packets

are tangent to Z in B_j—in fact, something better is true. All the wave packets of $f_{j,\mathrm{trans}}$ on B_j lie in $N_{R^{1/2}}(Z)\cap B_j$, but they do not necessarily lie in $N_{\varrho^{1/2}}(Z)\cap B_j$. We cover $N_{R^{1/2}}(Z)\cap B_j$ with disjoint translates of $N_{\varrho^{1/2}}(Z)\cap B_j$:

$$N_{R^{1/2}}(Z)\cap B_j = \bigcup_b N_{\varrho^{1/2}}(Z+b)\cap B_j.$$

Now it turns out that each wave packet of $f_{j,\mathrm{trans}}$ lies in one of these translates. We let $f_{j,\mathrm{trans},b}$ be the sum of the wave packets that lie in $N_{\varrho^{1/2}(Z+b)}\cap B_j$. Now we have $f_{j,\mathrm{trans}}=\sum_b f_{j,\mathrm{trans},b}$ and

$$\begin{aligned}\|Ef_{j,\mathrm{trans}}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)} &\sim \sum_b \|Ef_{j,\mathrm{trans}}\|^p_{\mathrm{BL}^p_{k,A/2}(N_{\varrho^{1/2}}(Z+b)\cap B_j)}\\ &\sim \sum_b \|Ef_{j,\mathrm{trans},b}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)}.\end{aligned}$$

The wave packets of $Ef_{j,\mathrm{trans},b}$ are tangent to the m-dimensional variety $Z+b$ on B_j. Therefore, we can study $\|Ef_{j,\mathrm{trans},b}\|_{\mathrm{BL}^p_{k,A/2}(B_j)}$ by induction on the radius: we may assume that

$$\|Ef_{j,\mathrm{trans},b}\|_{\mathrm{BL}^p_{k,A/2}(B_j)} \leqslant M\big(\varrho, \tfrac{1}{2}A\big)\|f_{j,\mathrm{trans},b}\|_{L^2}.$$

(Here we write $M\big(\varrho, \frac{1}{2}A\big)$ because M depends on the radius (which is ϱ) and because we have $\frac{1}{2}A$ in place of A.) Putting together what we have learned so far in the transverse algebraic case, we have

$$\begin{aligned}\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} &\lesssim \sum_j \|Ef_{j,\mathrm{trans}}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)} \sim \sum_{j,b} \|Ef_{j,\mathrm{trans},b}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)}\\ &\leqslant M\big(\varrho, \tfrac{1}{2}A\big)^p \sum_{j,b} \|f_{j,\mathrm{trans},b}\|^p_{L^2}.\end{aligned}$$

To get our final bound, it remains to control $\sum_{j,b}\|f_{j,\mathrm{trans},b}\|^p_{L^2}$. Since each wave packet of $f_{j,\mathrm{trans}}$ lies in exactly one $f_{j,\mathrm{trans},b}$, it is easy to check that

$$\|f_{j,\mathrm{trans}}\|^2_{L^2} \sim \sum_b \|f_{j,\mathrm{trans},b}\|^2_{L^2}.$$

We also already know that $\sum_j \|f_{j,\mathrm{trans}}\|^2_{L^2}\lesssim\|f\|^2_{L^2}$. And so we see that

$$\sum_{j,b}\|f_{j,\mathrm{trans},b}\|^p_{L^2} \lesssim \Big(\max_{j,b}\|f_{j,\mathrm{trans},b}\|^{p-2}_{L^2}\Big)\|f\|^2_{L^2}.$$

The last ingredient of the proof is an estimate for $\max_{j,b}\|f_{j,\mathrm{trans},b}\|_{L^2}$, which has to do with transverse equidistribution. Recall that, since f is concentrated on wave packets

in $\mathbb{T}_Z$, Ef is equidistributed in directions transverse to Z. Recall that $N_{R^{1/2}}(Z)\cap B_j$ is covered by thinner neighborhoods $N_{\varrho^{1/2}}(Z+b)\cap B_j$. The number of these thinner neighborhoods is $\sim(R^{1/2}/\varrho^{1/2})^{n-m}$. Because of transverse equidistribution, each of these neighborhoods receives an even share of the L^2 norm of $Ef_{j,\mathrm{trans}}$. In other words, for each b,

$$\|Ef_{j,\mathrm{trans}}\|^2_{L^2(N_{\varrho^{1/2}}(Z+b)\cap B_j)} \lesssim \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|Ef_{j,\mathrm{trans}}\|^2_{L^2(N_{R^{1/2}}(Z)\cap B_j)}.$$

Using this inequality, we prove the desired estimate for $\|f_{j,\mathrm{trans},b}\|_{L^2}$:

$$\|f_{j,\mathrm{trans},b}\|^2_{L^2} \lesssim \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|f_{j,\mathrm{trans}}\|^2_{L^2} \lesssim \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|f\|^2_{L^2}. \tag{2.11}$$

2.2. Outline of the paper

We carry out the proof of Theorem 1.5 over §§3–8 of the paper. In §3 we review some standard facts about wave packets. In §4 we prove some basic properties of the broad "norms" $\mathrm{BL}^p_{k,A}$. §5 contains tools from algebraic geometry (and differential geometry) that we will use to study the geometry of the algebraic varieties that appear in the proofs. In §6 we prove the first transverse equidistribution estimate. For any ball $B(y,\varrho)\subset\mathbb{R}^n$, there is a wave packet decomposition for Ef on the ball $B(y,\varrho)$. §7 is about the relationship between the original wave packet decomposition on the ball B_R and the wave packet decomposition adapted to a smaller ball $B(y,\varrho)\subset B_R$. This lets us state and prove a second version of the transverse equidistribution estimate, corresponding to (2.11) above. With the background and tools from these sections, we prove Theorem 1.5 in §8.

Following [BG], in §9 we explain how k-broad estimates imply regular L^p estimates of the form $\|Ef\|_{L^p}\lesssim\|f\|_{L^p}$. This argument finishes the proof of Theorem 1.1. §10 is an appendix which helps to keep track of the parameters. Unfortunately, there are quite a few parameters in the paper—various δ's, R, K, A, etc. The appendix lists all the parameters and how they relate to each other. In §11 we discuss further directions and open problems.

3. Basic setup with wave packets

Let f be a function on B^{n-1}. We first break up f into pieces $f_{\theta,v}$ that are localized in both position and frequency.

Cover B^{n-1} by finitely overlapping balls θ of radius $R^{-1/2}$. Let ψ_θ be a smooth partition of unity adapted to this cover, and write $f = \sum_\theta \psi_\theta f$.

Next, we break up $\psi_\theta f$ according to the frequency. Cover $\mathbb{R}^{n-1}$ by finitely overlapping balls of radius $\sim R^{(1+\delta)/2}$, centered at vectors $v \in R^{(1+\delta)/2}\mathbb{Z}^{n-1}$. Let η_v be a smooth partition of unity adapted to this cover. We can now write

$$f = \sum_{\theta,v} (\eta_v (\psi_\theta f)^\wedge)^\vee = \sum_{\theta,v} \eta_v^\vee * (\psi_\theta f).$$

Note that $\eta_v^\vee(x)$ decays rapidly when $|x| \gtrsim R^{-(1+\delta)/2}$. We write $A(R) \leqslant \mathrm{RapDec}(R) B$ to mean that for any power β, there is a constant C_β such that

$$A(R) \leqslant C_\beta R^{-\beta} B \text{ for all } R \geqslant 1.$$

If $|x| \geqslant \frac{1}{1000} R^{-1/2}$, then

$$|\eta_v^\vee(x)| \leqslant \mathrm{RapDec}(R).$$

Now choose smooth functions $\tilde{\psi}_\theta$ such that $\tilde{\psi}_\theta$ is supported on θ, but $\tilde{\psi}_\theta$ is 1 on a small neighborhood of the support of ψ_θ. A bit more precisely, we would like $\tilde{\psi}_\theta = 1$ on a $\left(\frac{1}{1000} R^{-1/2}\right)$-neighborhood of the support of ψ_θ. We now define

$$f_{\theta,v} := \tilde{\psi}_\theta (\eta_v^\vee * (\psi_\theta f)).$$

Because of the rapid decay of $\eta_v^\vee$, one has

$$\| f_{\theta,v} - \eta_v^\vee * (\psi_\theta f) \|_{L^\infty} \leqslant \mathrm{RapDec}(R) \| f \|_{L^2}.$$

Therefore, we see that

$$f = \sum_{\theta,v} f_{\theta,v} + \mathrm{Err}, \quad \text{where } \|\mathrm{Err}\|_{L^\infty} \leqslant \mathrm{RapDec}(R) \|f\|_{L^2}.$$

Terms of the form $\mathrm{RapDec}(R)\|f\|_{L^2}$ are negligibly small in terms of all of our estimates. These rapidly decaying errors will occur from time to time during our arguments.

The functions $f_{\theta,v}$ are approximately orthogonal. For any set $\mathbb{T}$ of pairs (θ, v), we have

$$\Big\| \sum_{(\theta,v)\in\mathbb{T}} f_{\theta,v} \Big\|_{L^2}^2 \sim \sum_{(\theta,v)\in\mathbb{T}} \|f_{\theta,v}\|_{L^2}^2. \tag{3.1}$$

The decomposition $f = \sum_{\theta,v} f_{\theta,v}$ is useful in this problem, because $E f_{\theta,v}$ is localized in space. For each (θ, v), there is a corresponding tube $T_{\theta,v}$, where $E f_{\theta,v}$ is essentially supported. Let ω_θ denote the center of θ. We define $T_{\theta,v}$ by

$$T_{\theta,v} := \{(x', x_n) \in B_R : |x' + 2x_n \omega_\theta + v| \leqslant R^{1/2+\delta}\}. \tag{3.2}$$

LEMMA 3.1. *If* $x \in B_R \backslash T_{\theta,v}$, *then*

$$|Ef_{\theta,v}(x)| \leqslant \mathrm{RapDec}(R)\|f\|_{L^2}.$$

We sketch the proof by stationary phase. We note that $(e^{-iv\omega} f_{\theta,v})$ has Fourier transform essentially supported in $B_{R^{1/2+\delta/2}}$. Therefore, we have

$$\max_{\theta} |e^{-iv\omega} f_{\theta,v}| \lesssim R^{(n-1)\frac{\delta}{2}} \operatorname{Avg}_{\theta} |e^{-iv\omega} f_{\theta,v}|.$$

Moreover, taking derivatives, we see that

$$\max_{\theta} |\partial_{\omega}^{k}(e^{-iv\omega} f_{\theta,v})| \lesssim (R^{1/2+\delta/2})^{k} R^{(n-1)\delta/2} \operatorname{Avg}_{\theta} |e^{-iv\omega} f_{\theta,v}|.$$

Let η_θ be a smooth bump which is equal to 1 on θ. Since θ has diameter $R^{-1/2}$, we can estimate the derivatives $|\partial^k \eta_\theta| \lesssim R^{k/2}$. Now we write $Ef_{\theta,v}(x)$ as

$$Ef_{\theta,v}(x) = \int \eta_\theta e^{i(x'\omega + x_n|\omega|^2 + v\omega)} \cdot e^{-iv\omega} f_{\theta,v}. \tag{3.3}$$

We let $\Psi(\omega) := x'\omega + x_n|\omega|^2 - v\omega$. We note that

$$\partial_\omega \Psi = x' + 2x_n\omega + v.$$

If $x \in B_R \backslash T_{\theta,v}$, then $|x' + 2x_n\omega_\theta + v| = |\partial_\omega \Psi(\omega_\theta)| > R^{1/2+\delta}$. We know that $|x_n| \leqslant R$, and so, for any $\omega \in \theta$, we have $|2x_n\omega - 2x_n\omega_\theta| \lesssim R^{1/2}$. Therefore, for any $\omega \in \theta$,

$$|\partial_\omega \Psi(\omega)| \gtrsim R^{1/2+\delta}.$$

By applying integration by parts to (3.3) many times, we see that

$$|Ef_{\theta,v}(x)| \leqslant \mathrm{RapDec}(R)\|f_{\theta,v}\|_{L^2},$$

as desired.

The tube $T_{\theta,v}$ is a cylinder of length R and radius $\sim R^{1/2+\delta}$. It points in the direction $G(\omega_\theta)$, where $G(\omega)$ is the unit vector given by

$$G(\omega) = \frac{(-2\omega_1, ..., -2\omega_{n-1}, 1)}{|(-2\omega_1, ..., -2\omega_{n-1}, 1)|}.$$

For each $\omega \in B^{n-1}$, we also define a frequency $\xi(\omega)$. Based on the formula for Ef, the frequency $\xi(\omega)$ is given by

$$\xi(\omega) := (\omega_1, ..., \omega_{n-1}, |\omega|^2).$$

We let $\xi(\theta)$ denote the image of θ under ξ:

$$\xi(\theta) := \{(\omega_1, ..., \omega_{n-1}, |\omega|^2) : \omega \in \theta\}.$$

In a distributional sense, the Fourier transform of $Ef_{\theta,v}$ is supported in $\xi(\theta)$. Also, if η_R denotes a smooth bump on B_R (of height 1), then the Fourier transform of $\eta_R Ef_{\theta,v}$ is essentially supported in $N_{R^{-1}}(\xi(\theta))$.

We introduce a little notation. If $\mathbb{T}_\alpha$ is any set of pairs (θ, v), then we say that f is *concentrated on wave packets from* $\mathbb{T}_\alpha$ if

$$f = \sum_{(\theta,v)\in\mathbb{T}_\alpha} f_{\theta,v} + \mathrm{RapDec}(R)\|f\|_{L^2}.$$

Also, for any f, and for any set $\mathbb{T}_\alpha$, we define

$$f_\alpha = \sum_{(\theta,v)\in\mathbb{T}_\alpha} f_{\theta,v}.$$

3.1. Orthogonality

For any fixed x_n, Ef restricted to $\mathbb{R}^{n-1}\times\{x_n\}$ can be described as an inverse Fourier transform:

$$Ef(x_1, ..., x_{n-1}, x_n) = (e^{ix_n|\omega|^2} f(\omega))^\vee(x_1, ..., x_{n-1}).$$

Applying Plancherel's theorem, we get

$$\|Ef\|_{L^2(\mathbb{R}^{n-1}\times\{x_n\})} = \|f\|_{L^2}. \tag{3.4}$$

We record a couple of simple corollaries of this statement.

LEMMA 3.2. *We have*

$$\|Ef\|_{L^2(B_R)} \lesssim R^{1/2}\|f\|_{L^2}.$$

Proof.

$$\int_{B_R} |Ef|^2 \leqslant \int_{-R}^{R} \left(\int_{B^{n-1}(R)} |Ef|^2 dx' \right) dx_n \leqslant 2R\|f\|_{L^2}^2. \qquad \square$$

LEMMA 3.3. *Suppose that f is concentrated on a set of wave packets $\mathbb{T}$ and that, for every $(\theta, v)\in\mathbb{T}$, $T_{\theta,v}\cap(\mathbb{R}^{n-1}\times\{x_n\})\subset B^{n-1}(z_0, r)\times\{x_n\}$. Then,*

$$\|Ef\|_{L^2(B^{n-1}(z_0,r)\times\{x_n\})} = \|f\|_{L^2} + \mathrm{RapDec}(R)\|f\|_{L^2}. \tag{3.5}$$

LEMMA 3.4. *Suppose that f is concentrated on a set of wave packets $\mathbb{T}$ and that, for every $(\theta,v)\in\mathbb{T}$, we have $T_{\theta,v}\cap B(z,r)\neq\varnothing$ for some radius $r\geqslant R^{1/2+\delta}$. Then,*

$$\|Ef\|^2_{L^2(B(z,10r))}\sim r\|f\|^2_{L^2}. \tag{3.6}$$

Proof. For each x_n in the range $z_n-r\leqslant x_n\leqslant z_n+r$, and for each $(\theta,v)\in\mathbb{T}$, the intersection $T_{\theta,v}\cap\mathbb{R}^{n-1}\times\{x_n\}$ is contained in $B(z,5r)$. By the last lemma, we see that

$$\|Ef\|_{L^2(B(z,5r)\cap\mathbb{R}^{n-1}\times\{x_n\})}=\|f\|_{L^2}+\mathrm{RapDec}(R)\|f\|_{L^2}.$$

Applying Fubini's theorem, we get the desired bound. □

4. Properties of the broad "norms" $\mathrm{BL}^p_{k,A}$

We recall the definition of the k-broad "norm" $\mathrm{BL}^p_{k,A}$. Although $\mathrm{BL}^p_{k,A}$ is not literally a norm, it obeys a version of the triangle inequality and a version of Hölder's inequality. These nice algebraic features helped to motivate this particular definition.

Let B^{n-1} be a disjoint union of (approximate) balls τ of radius K^{-1} . For each τ, we define $G(\tau)$ to be the image of τ under the direction map G. If ω_τ is the center of τ, then $G(\tau)$ is essentially a ball of radius K^{-1} around $G(\omega_\tau)$ in S^{n-1}. If $V\subset\mathbb{R}^n$ is a subspace, then we write $\mathrm{Angle}(G(\tau),V)$ for the smallest angle between any non-zero vectors $v\in V$ and $v'\in G(\tau)$.

For any ball B_{K^2} of radius K^2 in B_R, we define μ_{Ef} as in (1.12):

$$\mu_{Ef}(B_{K^2}):=\min_{V_1,\ldots,V_A}\left(\max_\tau\int_{B_{K^2}}|Ef_\tau|^p\right).$$

where the minimum is over $(k-1)$-subspaces of $\mathbb{R}^n$ and the maximum is over all τ such that $\mathrm{Angle}(G(\tau),V_a)>K^{-1}$ for all a.

We abbreviate this expression as

$$\mu_{Ef}(B_{K^2}):=\min_{V_1,\ldots,V_A}\left(\max_{\tau\not\in V_a}\int_{B_{K^2}}|Ef_\tau|^p\right).$$

We remark that it is convenient to allow $A=0$. If $A=0$, then we simply have

$$\mu_{Ef}(B_{K^2})=\max_\tau\int_{B_{K^2}}|Ef_\tau|^p.$$

If $U\subset B_R$ is a finite union of balls B_{K^2}, then we define $\|Ef\|_{\mathrm{BL}^p_{k,A}(U)}$ by

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(U)}:=\sum_{B_{K^2}\subset U}\mu_{Ef}(B_{K^2}). \tag{4.1}$$

The k-broad "norm" obeys a weak version of the triangle inequality.

LEMMA 4.1. *Suppose that $f=g+h$ and suppose that $A=A_1+A_2$, where A and A_j are non-negative integers. Then,*

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(U)} \lesssim \|Eg\|_{\mathrm{BL}^p_{k,A_1}(U)} + \|Eh\|_{\mathrm{BL}^p_{k,A_2}(U)}.$$

Proof. We expand

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(U)} = \sum_{B_{K^2}\subset U} \min_{V_1,\dots,V_A} \left(\max_{\tau\notin V_a} \int_{B_{K^2}} |Ef_\tau|^p \right).$$

Now, for each ball $B_{K^2}\subset U$, we have

$$\begin{aligned}\min_{V_1,\dots,V_A} \left(\max_{\tau\notin V_a} \int_{B_{K^2}} |Ef_\tau|^p \right) &\lesssim \min_{V_1,\dots,V_A} \left(\max_{\tau\notin V_a} \left(\int_{B_{K^2}} |Eg_\tau|^p + \int_{B_{K^2}} |Eh_\tau|^p \right) \right) \\ &\leqslant \min_{V_1,\dots,V_{A_1}} \left(\max_{\substack{\tau\notin V_a \\ 1\leqslant a\leqslant A_1}} \int_{B_{K^2}} |Eg_\tau|^p \right) \\ &\quad + \min_{V_{A_1+1},\dots,V_A} \left(\max_{\substack{\tau\notin V_a \\ A_1+1\leqslant a\leqslant A}} \int_{B_{K^2}} |Eh_\tau|^p \right).\end{aligned}$$

Summing over all $B_{K^2}\subset U$, we get

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(U)} \lesssim \|Eg\|^p_{\mathrm{BL}^p_{k,A_1}(U)} + \|Eh\|^p_{\mathrm{BL}^p_{k,A_2}(U)}.$$ □

The reason for which we need a large value of A in Theorem 1.5 is that we will need to use this triangle inequality many times. If $A=1$, $\mathrm{BL}^p_{k,1}$ does not obey a good triangle inequality. But if we start with A being a large constant, we may use Lemma 4.1 many times. In effect, $\mathrm{BL}^p_{k,A}$ behaves like a norm, as long as we only use the triangle inequality $O_\varepsilon(1)$ times in our argument, and as long as we choose $A=A(\varepsilon)$ large enough.

$\mathrm{BL}^p_{k,A}$ also obeys a version of (a corollary of) Hölder's inequality.

LEMMA 4.2. *Suppose that $1\leqslant p,p_1,p_2<\infty$ and $0\leqslant\alpha_1,\alpha_2\leqslant 1$ obey $\alpha_1+\alpha_2=1$ and*

$$\frac{1}{p} = \alpha_1\frac{1}{p_1} + \alpha_2\frac{1}{p_2}.$$

Also, suppose that $A=A_1+A_2$. Then,

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(U)} \leqslant \|Ef\|^{\alpha_1}_{\mathrm{BL}^{p_1}_{k,A_1}(U)} \|Ef\|^{\alpha_2}_{\mathrm{BL}^{p_2}_{k,A_2}(U)}.$$

Proof. The left-hand side is

$$\left(\sum_{B_{K^2}\subset U} \min_{V_1,\dots V_A} \max_{\tau\notin V_a} \int_{B_{K^2}} |Ef_\tau|^p \right)^{1/p}.$$

Applying the regular Hölder inequality to the inner integral, this expression is

$$\leqslant \Big(\sum_{B_{K^2} \subset U} \min_{V_1, \ldots V_A} \max_{\tau \notin V_a} \Big(\int_{B_{K^2}} |Ef_\tau|^{p_1} \Big)^{\alpha_1 p/p_1} \Big(\int_{B_{K^2}} |Ef_\tau|^{p_2} \Big)^{\alpha_2 p/p_2} \Big)^{1/p}.$$

We can bring the maximum over τ inside, so the last expression is

$$\leqslant \Big(\sum_{B_{K^2} \subset U} \min_{V_1, \ldots V_A} \Big(\max_{\tau \notin V_a} \int_{B_{K^2}} |Ef_\tau|^{p_1} \Big)^{\alpha_1 p/p_1} \Big(\max_{\tau \notin V_a} \int_{B_{K^2}} |Ef_\tau|^{p_2} \Big)^{\alpha_2 p/p_2} \Big)^{1/p}.$$

Now we cannot bring the minimum inside the inner parentheses. But we can split $V_1, ..., V_A$ into $V_1, ..., V_{A_1}$ and $V_{A_1+1}, ..., V_A$. If we weaken the first condition $\tau \notin V_1 \ldots V_A$ to $\tau \notin V_1, ..., V_{A_1}$, and if we weaken the second condition $\tau \notin V_1, ..., V_A$ to $\tau \notin V_{A_1+1}, ..., V_A$, then we see that the last expression is bounded by

$$\leqslant \Big(\sum_{B_{K^2} \subset U} \Big(\min_{V_1, \ldots V_{A_1}} \max_{\substack{\tau \notin V_a \\ 1 \leqslant a \leqslant A_1}} \int_{B_{K^2}} |Ef_\tau|^{p_1} \Big)^{\alpha_1 p/p_1} \times \Big(\min_{V_{A_1+1}, \ldots V_A} \max_{\substack{\tau \notin V_a \\ A_1+1 \leqslant a \leqslant A}} \int_{B_{K^2}} |Ef_\tau|^{p_2} \Big)^{\alpha_2 p/(p_2)} \Big)^{1/p}.$$

We now apply Hölder's inequality to the initial sum over $B_{K^2} \subset U$, and we get

$$\leqslant \Big(\sum_{B_{K^2} \subset U} \min_{V_1, \ldots V_{A_1}} \max_{\substack{\tau \notin V_a \\ 1 \leqslant a \leqslant A_1}} \int_{B_{K^2}} |Ef_\tau|^{p_1} \Big)^{\alpha_1/p_1} \times \Big(\sum_{B_{K^2} \subset U} \min_{V_{A_1+1}, \ldots V_A} \max_{\substack{\tau \notin V_a \\ A_1+1 \leqslant a \leqslant A}} \int_{B_{K^2}} |Ef_\tau|^{p_2} \Big)^{\alpha_2/p_2}$$

$$= \|Ef\|^{\alpha_1}_{\mathrm{BL}^{p_1}_{k,A_1}(U)} \|Ef\|^{\alpha_2}_{\mathrm{BL}^{p_2}_{k,A_2}(U)}. \qquad \square$$

5. Tools from algebraic geometry

5.1. Transverse complete intersections

Over the course of our argument, we will work not just with algebraic hypersurfaces, but with algebraic varieties of all dimensions. We write $Z(P_1, ..., P_{n-m})$ for the set of common zeros of the polynomials $P_1, ..., P_{n-m}$. Throughout the paper, we will work with a nice class of varieties called transverse complete intersections. The variety $Z(P_1, ..., P_{n-m})$ is a transverse complete intersection if

$$\nabla P_1(x) \wedge \ldots \wedge \nabla P_{n-m}(x) \neq 0 \quad \text{for all } x \in Z(P_1, ..., P_{n-m}). \tag{5.1}$$

By the implicit function theorem, a transverse complete intersection $Z(P_1,...,P_{n-m})$ is a smooth m-dimensional manifold. Because of Sard's theorem, there are lots of transverse complete intersections. Here is a lemma making this precise.

LEMMA 5.1. *If P is a polynomial on $\mathbb{R}^n$, then, for almost every $c_0\in\mathbb{R}$, $Z(P+c_0)$ is a transverse complete intersection.*

More generally, suppose that $Z(P_1,...,P_{n-m})$ is a transverse complete intersection and that P is another polynomial. Then, for almost every $c_0\in\mathbb{R}$, $Z(P_1,...,P_{n-m},P+c_0)$ is a transverse complete intersection.

Proof. We begin with the first case. We know that $P:\mathbb{R}^n\to\mathbb{R}$ is a smooth function, and so by Sard's theorem, almost every $y\in\mathbb{R}$ is a regular value for P. But if $-c_0$ is a regular value for P, then $\nabla P(x)\neq 0$ whenever $P(x)+c_0=0$.

The general case is similar. We know that $Z=Z(P_1,...,P_{n-m})$ is a smooth m-dimensional manifold, and $P:Z\to\mathbb{R}$ is a smooth function. By Sard's theorem, almost every $y\in\mathbb{R}$ is a regular value of the map $P:Z\to\mathbb{R}$. If $x\in Z$ and $P(x)$ is a regular value, then $dP_x\neq 0$, where $dP:T_xZ\to T_{P(x)}\mathbb{R}$. In terms of $\nabla P(x)$, this means that

$$\nabla P_1(x)\wedge...\wedge\nabla P_{n-m}(x)\wedge\nabla P(x)\neq 0.$$

So, if $-c_0$ is a regular value for $P:Z\to\mathbb{R}$, then $Z(P_1,...,P_{n-m},P+c_0)$ is a transverse complete intersection. □

5.2. Polynomial partitioning

Polynomial partitioning is a key tool in our arguments. Our presentation here is a minor variation on the polynomial partitioning result from [GK]. We begin by stating a partitioning result from [G].

THEOREM 5.2. ([G, Theorem 1.4]) *Suppose that $W\geqslant 0$ is a (non-zero) L^1 function on $\mathbb{R}^n$. Then, for each D, there a non-zero polynomial P of degree at most D such that $\mathbb{R}^n\backslash Z(P)$ is a union of $\sim D^n$ disjoint open sets O_j, and the integrals $\int_{O_j}W$ are all equal.*

We want to use this result, but we need to upgrade it in a minor way. Because we want all the varieties that appear in our argument to be transverse complete intersections, we need to be able to perturb P a little bit. In order to understand this issue, we need to review some of the proof of Theorem 5.2. The proof is based on the polynomial ham-sandwich theorem, which is due to Stone and Tukey [ST]. Here is a version of the theorem which is convenient for our purposes:

THEOREM 5.3. (Polynomial ham-sandwich theorem; cf. [G, Corollary 1.2]) *Suppose that $W_1, ..., W_N$ are L^1-functions on $\mathbb{R}^n$. Then, there exists a non-zero polynomial P of degree $\leqslant C_n N^{1/n}$ such that, for each W_j,*

$$\int_{\{P>0\}} W_j = \int_{\{P<0\}} W_j.$$

Using the polynomial ham-sandwich theorem iteratively, we get the following partitioning result.

COROLLARY 5.4. *If $W \geqslant 0$ is a (non-zero) L^1-function on $\mathbb{R}^n$, then there is a sequence of polynomials $Q_1, Q_2, \dots$, with $\operatorname{Deg} Q_j \lesssim 2^{j/n}$, satisfying the following equidistribution property.*

If $S \geqslant 1$, and if $\sigma_1, ..., \sigma_S \in \{-1, 1\}$ are any sign conditions, then

$$\int_{\operatorname{Sign}(Q_s) = \sigma_s \text{ for } 1 \leqslant s \leqslant S} W = 2^{-S} \int_{\mathbb{R}^n} W.$$

We can slightly perturb each Q_s by adding a small generic constant: $\widetilde{Q}_s = Q_s + c_s$, where $c_s \in \mathbb{R}$. Using this small perturbation, we will be able to arrange that all the varieties that appear in our arguments are transverse complete intersections. As long as the constants c_s are sufficiently small, we still have the following slightly weaker version of the equidistribution result: if $S \geqslant 1$, and if $\sigma_1, ..., \sigma_S \in \{-1, 1\}$ are any sign conditions, then

$$2^{-S-1} \int_{\mathbb{R}^n} W \leqslant \int_{\operatorname{Sign}(\widetilde{Q}_s) = \sigma_s \text{ for } 1 \leqslant s \leqslant S} W \leqslant 2^{-S+1} \int_{\mathbb{R}^n} W.$$

This gives the following polynomial partitioning result, which is designed to allow small perturbations.

THEOREM 5.5. *Suppose that $W \geqslant 0$ is a (non-zero) L^1 function on $\mathbb{R}^n$. Then, for any degree D, the following holds.*

There is a sequence of polynomials $Q_1, ..., Q_S$ with the following properties. We have $\sum_s \operatorname{Deg} Q_s \lesssim D$ and $2^S \sim D^n$. Let

$$P = \prod_{s=1}^{S} \widetilde{Q}_s = \prod_{s=1}^{S} (Q_s + c_s),$$

where $c_s \in \mathbb{R}$. Let O_j be the open sets given by the sign conditions of $\widetilde{Q}_s$. There are $2^S \sim D^n$ cells O_j and $\mathbb{R}^n \setminus Z(P) = \bigcup_j O_j$.

If the constants c_s are sufficiently small, then, for every O_j,

$$\int_{O_j} W \sim D^{-n} \int_{\mathbb{R}^n} W.$$

For a generic choice of the constants c_s, Lemma 5.1 guarantees that $Z(\widetilde{Q}_s)$ is a transverse complete intersection for each s. This implies that $Z(P)$ is a finite union of transverse complete intersections. Similarly, if $Z(P_1,...,P_{n-m})$ is a transverse complete intersection, then, for a generic choice of the constants c_s, $Z(P_1,...,P_{n-m},\widetilde{Q}_s)$ will also be a transverse complete intersection for each s.

5.3. Controlling the tangent plane of a variety

Suppose that Z is an m-dimensional transverse complete intersection. We know that Z is a smooth m-dimensional manifold. We will consider some subsets of Z, where the tangent plane obeys certain conditions. We will see that these subsets are in fact subvarieties of Z, and that, in generic cases, they are transverse complete intersections.

Let $Z=Z(P_1,...,P_{n-m})$ be a transverse complete intersection. Let $w\in\Lambda^m\mathbb{R}^n$. Define Z_w by

$$Z_w := \{x\in Z : \nabla P_1(x)\wedge...\wedge\nabla P_{n-m}(x)\wedge w=0\}. \tag{5.2}$$

We note that, since w is an m-vector, $\nabla P_1(x)\wedge...\wedge\nabla P_{n-m}(x)\wedge w\in\Lambda^n\mathbb{R}^n$, which we identify with $\mathbb{R}$. Let $g_w:=\nabla P_1(x)\wedge...\wedge\nabla P_{n-m}(x)\wedge w$, which is a polynomial with degree at most $\operatorname{Deg}P_1+...+\operatorname{Deg}P_{n-m}$. The set Z_w is the algebraic variety $Z(P_1,...,P_{n-m},g_w)$.

LEMMA 5.6. *For almost every $w\in\Lambda^m\mathbb{R}^n$, $Z_w=Z(P_1,...,P_{n-m},g_w)$ is a smooth complete intersection.*

The proof uses some ideas from differential topology. The book [GP] is a good reference. In particular, the proof here is closely based on the proof of the transversality theorem in [GP, Chapter 2.3].

Proof. Define a smooth function $g\colon Z\times\Lambda^m\mathbb{R}^n\to\mathbb{R}$ by

$$g(x,w) := \nabla P_1(x)\wedge...\wedge\nabla P_{n-m}(x)\wedge w.$$

The function g is smooth, and it has no critical points, because for any $x\in Z$, $\nabla P_1(x)\wedge...\wedge\nabla P_{n-m}(x)\neq 0$, and the restriction of g to $\{x\}\times\Lambda^m\mathbb{R}^n$ is a non-zero linear function with no critical points. Therefore, $g^{-1}(0)$ is a smooth submanifold M in $Z\times\Lambda^m\mathbb{R}^n$ (of codimension 1).

Consider the smooth map $\pi: M \to \Lambda^m \mathbb{R}^n$ given by $\pi(x,w)=w$. We first note that $\pi^{-1}(w)=Z_w \times \{w\}$. We will use π in order to study Z_w. We claim that Z_w is a transverse complete intersection whenever w is a regular value of π. By Sard's theorem, almost every $w \in \Lambda^m \mathbb{R}^n$ is a regular value of π, and so this claim implies our conclusion. From now on, suppose that w is a regular value of π.

Now $Z_w = Z(P_1, ..., P_{n-m}, g_w)$. To see that it is a complete intersection, we have to check that, for each $x \in Z_w$,

$$\nabla P_1(x) \wedge ... \wedge \nabla P_{n-m}(x) \wedge \nabla g_w(x) \neq 0.$$

This is equivalent to saying that $\nabla g_w(x) \cdot v \neq 0$ for some vector $v \in T_x Z$. This, in turn, is equivalent to saying that x is a regular point for the map $g_w: Z \to \mathbb{R}$. (We remark that saying that x is a regular point for the map $g_w: Z \to \mathbb{R}$ is a stronger condition than just saying that x is a regular point for the map $g_w: \mathbb{R}^n \to \mathbb{R}$.)

So, we have to check that, if $(x,w) \in M$ and w is a regular value for π, then x is a regular point for $g_w: Z \to \mathbb{R}$. Since w is a regular value of π, we know that (x,w) is a regular point for π. So, it suffices to check that whenever (x,w) is a regular point for π, x is a regular point for $g_w: Z \to \mathbb{R}$.

Recall that $(x,w) \in M$ is a regular point for $\pi: M \to \Lambda^m \mathbb{R}^n$ if and only if

$$d\pi: T_{(x,w)} M \to \Lambda^m \mathbb{R}^n$$

is surjective. To understand this condition better, we compute the tangent space $T_{(x,w)} M$. We know that $T_{(x,w)} M \subset T_{(x,w)}(Z \times \Lambda^m \mathbb{R}^n) = T_x Z \times \Lambda^m \mathbb{R}^n$, and more precisely

$$T_{(x,w)} M = \{(v, w') \in T_x Z \times \Lambda^m \mathbb{R}^n : dg_{(x,w)}(v, w') = 0\}.$$

But $dg_{(x,w)}(v,w') = (dg_w)_x(v) + \nabla P_1(x) \wedge ... \wedge \nabla P_{n-m}(x) \wedge w'$. Therefore,

$$T_{(x,w)} M = \{(v, w') \in T_x Z \times \Lambda^m \mathbb{R}^n : (dg_w)_x(v) + \nabla P_1(x) \wedge ... \wedge \nabla P_{n-m}(x) \wedge w' = 0\}.$$

If x is not a regular point of dg_w, then $(dg_w)_x(v)=0$, and so

$$T_{(x,w)} M = \{(v, w') \in T_x Z \times \Lambda^m \mathbb{R}^n : \nabla P_1(x) \wedge ... \wedge \nabla P_{n-m}(x) \wedge w' = 0\}.$$

But, in this case, the projection $d\pi: T_{(x,w)} M \to \Lambda^m \mathbb{R}^n$ is not surjective. (The projection $d\pi$ is just $d\pi(v,w') = w'$.) So, if (x,w) is a regular point of π, then x is a regular point of g_w as desired. □

If $W \subset \Lambda^m \mathbb{R}^n$ is a large finite set, then on each connected component of $Z \setminus \bigcup_{w \in W} Z_w$, the tangent plane TZ is constrained in a small region of the Grassmannian. More precisely, for any small parameter $\beta > 0$, we can choose a finite set $W \subset \Lambda^m \mathbb{R}^n$ so that, for any two points x_1 and x_2 in the same component of $Z \setminus \bigcup_{w \in W} Z_w$, $\mathrm{Angle}(T_{x_1} Z, T_{x_2} Z) < \beta$. We can also choose W generically so that each Z_w is a transverse complete intersection of dimension $m-1$.

5.4. Controlling transverse intersections between a tube and a variety

Suppose that T is a cylinder of radius r with central axis ℓ. Suppose that $Z^m \subset \mathbb{R}^n$ is a transverse complete intersection. Define $Z_{>\alpha}$ by

$$Z_{>\alpha} := \{z \in Z : \text{Angle}(T_z Z, \ell) > \alpha\}.$$

LEMMA 5.7. *Suppose that $Z = Z(P_1, ..., P_{n-m})$ is a transverse complete intersection and that the polynomials P_j have degree at most D. Let T be a tube of radius r as above. Then, for any $\alpha > 0$, $Z_{>\alpha} \cap T$ is contained in a union of $\lesssim D^n$ balls of radius $\lesssim r\alpha^{-1}$.*

The main tool in the proof is the following version of the Bezout theorem.

THEOREM 5.8. (Cf. [CKW, Theorem 5.2] for a short proof) *Let $Z = Z(Q_1, ..., Q_n)$ be a transverse complete intersection in $\mathbb{R}^n$. Then, Z is finite and the cardinality of Z is at most $\prod_{j=1}^n \text{Deg}\, Q_j$.*

Using this Bezout theorem, we now prove Lemma 5.7.

Proof. The proof is by induction on m. When $m = 0$, Theorem 5.8 guarantees that Z consists of at most D^n points, and the conclusion follows.

Now, we turn to the inductive step. Without loss of generality, we may assume that ℓ is the x_n-axis. We let T_r denote the r-neighborhood of the x_n-axis.

Next, we do some scaling to reduce to a special case. By rescaling, we may reduce to the case $r = 1$. Next, by scaling in the x_n-coordinate only, we may reduce to the case $\alpha = 1$. So, we have to show that $Z_{>1} \cap T_1$ is contained in $\lesssim D^n$ balls of radius $\lesssim 1$.

Let Z_w be defined as in the last subsection. We choose $\lesssim 1$ values of w in general position, so that on each connected component of $Z \setminus \bigcup_w Z_w$, the tangent plane of Z varies by an angle at most $\frac{1}{100}$. Since w is generic, $Z_w = Z(P_1, ..., P_{n-m}, g_w)$ is a transverse complete intersection of dimension $m-1$. Also, $\text{Deg}\, g_w \lesssim D$. We may apply our inductive assumption to Z_w, using radius $r = 20$ and $\alpha = \frac{1}{2}$. We see that $Z_{w,>1/2} \cap T_{20}$ is covered by $\lesssim D^n$ balls of radius $\lesssim 1$.

Next, we claim that $Z_{>1/2} \cap Z_w \subset Z_{w,>1/2}$. If $x \in Z_{>1/2} \cap Z_w$, then

$$\inf_{v \in T_x Z} \text{Angle}(v, \ell) > \tfrac{1}{2}.$$

But $T_x Z_w \subset T_x Z$. Therefore, $\inf_{v \in T_x Z_w} \text{Angle}(v, \ell) > \frac{1}{2}$, and so $x \in Z_{w,>1/2}$ as claimed.

Since the total number of w is $\lesssim 1$, we have that

$$\bigcup_w (Z_{>1/2} \cap Z_w \cap T_{20}) \text{ is contained in } \lesssim D^n \text{ balls } B_j \text{ of radius } \lesssim 1.$$

If $B_j = B(x_j, r_j)$, then we write $10B_j$ for $B(x_j, 10r_j)$. The union $\bigcup_j 10B_j$ is a set of $\lesssim D^n$ balls of radius $\lesssim 1$ which covers part of $Z_{>1} \cap T_1$. We still have to cover the remaining part

$$Z_{>1} \cap T_1 \setminus \bigcup_j 10B_j.$$

Consider a point z in this remaining part. We may assume that the radius of each B_j is at least 2, and so we know that the distance from z to $\bigcup_j B_j$ is at least 10. Let A be the connected component of $Z \cap B(z, 10)$ containing z. We claim that A is disjoint from all Z_w. Indeed, suppose that γ was a curve in A starting at z and intersecting $\bigcup_w Z_w$ for the first time at $z' \in A$. Along the curve γ, the tangent plane of TZ is constant up to angle $\frac{1}{100}$. Since γ starts at $z \in Z_{>1}$, we have $\gamma \subset Z_{>1/2}$. Also, $\gamma \subset B(z, 10) \subset T_{20}$. We conclude that $z' \in Z_{>1/2} \cap Z_w \cap T_{20}$, and so $z' \in \bigcup_j B_j$. But $B(z, 10)$ is disjoint from $\bigcup_j B_j$. This contradiction proves the claim. Since A is connected and disjoint from all Z_w, the tangent plane of Z is constant on A up to angle $\frac{1}{100}$.

Therefore, A is a small perturbation of an m-plane that cuts across T_1 in a quantitatively transverse way. Let Π be a random $(n-m)$-plane containing the x_n-axis. With probability $\gtrsim 1$, $\Pi \cap A \cap T_1$ is non-empty. By the Bezout theorem (Theorem 5.8), $|\Pi \cap Z| \leqslant D^{n-m}$ for generic Π. Therefore, there can be at most $\lesssim D^{n-m}$ disjoint sets A of this type. So, we see that the remaining part of $Z_{>1} \cap T_1$ is contained in $\lesssim D^{n-m}$ additional balls of radius $\lesssim 1$. □

6. Transverse equidistribution estimates

In this section we prove a transverse equidistribution estimate. To set up the statement, we first define what it means for a wave packet to be tangent to a transverse complete intersection Z.

Definition 6.1. Suppose that $Z = Z(P_1, ..., P_{n-m})$ is a transverse complete intersection. We say that $T_{\theta,v}$ is $R^{-1/2+\delta_m}$-*tangent* to Z in B_R if

$$T_{\theta,v} \subset N_{R^{1/2+\delta_m}}(Z) \cap B_R \tag{6.1}$$

and, for any $x \in T_{\theta,v}$ and $z \in Z \cap B_R$ with $|x - z| \lesssim R^{1/2+\delta_m}$, we have

$$\text{Angle}(G(\theta), T_z Z) \lesssim R^{-1/2+\delta_m}. \tag{6.2}$$

We define

$$\mathbb{T}_Z := \{(\theta, v) : T_{\theta,v} \text{ is } R^{-1/2+\delta_m}\text{-tangent to } Z \text{ in } B_R\}.$$

We say that f is concentrated in wave packets from $\mathbb{T}_Z$ if

$$\sum_{(\theta,v)\notin\mathbb{T}_Z} \|f_{\theta,v}\|_{L^2} \leqslant \mathrm{RapDec}(R)\|f\|_{L^2}.$$

(In this definition $\delta_m>0$ is a small constant. The estimates we prove in this section hold for any $\delta_m\geqslant 0$. We will choose δ_m in §8.)

Suppose that B is a ball of radius $R^{1/2+\delta_m}$ in $\mathbb{R}^n$. Define

$$\mathbb{T}_{B,Z} := \{(\theta,v)\in\mathbb{T}_Z : T_{\theta,v}\cap B\neq\varnothing\}. \tag{6.3}$$

The main result of this section is the following transverse equidistribution estimate.

LEMMA 6.2. *Let B be a ball of radius $R^{1/2+\delta_m}$ in $B_R\subset\mathbb{R}^n$, and let $\varrho\leqslant R$. Suppose that $g=\sum_{(\theta,v)\in\mathbb{T}_{B,Z}} g_{\theta,v}$. Then,*

$$\int_{B\cap N_{\varrho^{1/2+\delta_m}}(Z)} |Eg|^2 \lesssim R^{O(\delta_m)}\left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \int_{2B} |Eg|^2 + \mathrm{RapDec}(R)\|g\|_{L^2}^2.$$

We build up to the proof via several smaller lemmas. We begin with a version of the Heisenberg uncertainty principle, saying that a function which is concentrated in a small ball in frequency space cannot concentrate too much in physical space.

LEMMA 6.3. *Suppose that $G\colon\mathbb{R}^n\to\mathbb{C}$ is a function, and that $\widehat{G}$ is supported in a ball $B(\xi_0,r)$ of radius r. Then, for any ball $B(x_0,\varrho)$ of radius $\varrho\leqslant r^{-1}$,*

$$\int_{B(x_0,\varrho)} |G|^2 \lesssim \frac{|B_\varrho|}{|B_{r^{-1}}|}\int |G|^2.$$

Proof. Let η be a smooth bump function with $|\eta|\sim 1$ on $B(x_0,\varrho)$ and rapidly decaying outside of it. Then, $|\hat{\eta}(\xi):\sim|B_\varrho|$ on $B_{\varrho^{-1}}$ and rapidly decaying outside of it. Therefore,

$$\int_{B_\varrho} |G|^2 \lesssim \int |\eta G|^2 = \int |\hat{\eta}*\widehat{G}|^2.$$

For $\xi\lesssim\varrho^{-1}$, we bound

$$|\hat{\eta}*\widehat{G}(\xi)| \leqslant \|\hat{\eta}\|_{L^\infty}\,\|\widehat{G}\|_{L^1} \sim |B_\varrho|\int_{B_r}|\widehat{G}|.$$

For $|\xi|$ far from $B_{\varrho^{-1}}$, the rapid decay of η takes over and gives a stronger bound. All together, we have

$$\int |\hat{\eta}*\widehat{G}|^2 \lesssim |B_\varrho^{-1}|\left(|B_\varrho|\int_{B_r}|\widehat{G}|\right)^2 \leqslant |B_\varrho|\,|B_r|\int|\widehat{G}|^2 = \frac{|B_\varrho|}{|B_{r^{-1}}|}\int|G|^2. \qquad \square$$

Next, we need a more local version of this lemma.

LEMMA 6.4. *Suppose that $G: \mathbb{R}^n \to \mathbb{C}$ is a function, and that $\widehat{G}$ is supported in a ball $B(\xi_0, r)$ of radius r. Then, for any ball $B(x_0, \varrho)$ with $\varrho \leqslant r^{-1}$, we have the inequality*

$$\int_{B(x_0,\varrho)} |G|^2 \lesssim \frac{|B_\varrho|}{|B_{r^{-1}}|} \int W_{B(x_0,r^{-1})} |G|^2,$$

where $W_{B(x_0,r^{-1})}$ is a weight function which is equal to 1 on $B(x_0, r^{-1})$ and rapidly decaying outside of it.

Proof. Let ψ be a function with the support of $\hat{\psi} \subset B_r$ and with $|\psi| \sim 1$ on $B(x_0, r^{-1})$ and rapidly decaying outside $B(x_0, r^{-1})$. Our weight function will be $W = |\psi|^2$.

Let $H = \psi \cdot G$. Note that $\widehat{H}$ is supported in $B(\xi_0, 2r)$. Applying Lemma 6.3 to H, we see that

$$\int_{B(x_0,\varrho)} |G|^2 \lesssim \int_{B(x_0,\varrho)} |H|^2 \lesssim \frac{|B_\varrho|}{|B_{r^{-1}}|} \int |H|^2 = \frac{|B_\varrho|}{|B_{r^{-1}}|} \int W |G|^2. \qquad \square$$

Suppose that B is a ball of radius $R^{1/2+\delta_m}$ in $\mathbb{R}^n$, and V is a subspace of $\mathbb{R}^n$. Define

$$\mathbb{T}_{B,V} := \{(\theta, v) : T_{\theta,v} \cap B \neq \varnothing \text{ and } \operatorname{Angle}(G(\theta), V) \lesssim R^{-1/2+\delta_m}\}. \tag{6.4}$$

Let $2B$ denote the ball with the same center as B and twice the radius.

If $g = \sum_{(\theta,v)\in\mathbb{T}_{B,V}} g_{\theta,v}$, then we will show that Eg is equidistributed in B along directions transverse to V. More precisely, we have the following lemma.

LEMMA 6.5. *If $V \subset \mathbb{R}^n$ is a subspace, then there is a subspace $V' \subset \mathbb{R}^n$ with the following properties*:

(1) $\operatorname{Dim} V + \operatorname{Dim} V' = n$.

(2) *V' is transverse to V in the sense that, for any unit vectors $v \in V$ and $v' \in V'$,*

$$\operatorname{Angle}(v, v') \gtrsim 1.$$

(3) *Suppose that*

$$g = \sum_{(\theta,v)\in\mathbb{T}_{B,V}} g_{\theta,v}.$$

If Π is a plane parallel to V' and $x_0 \in \Pi \cap B$, then, for any $\varrho \leqslant R$,

$$\int_{\Pi\cap B(x_0,\varrho^{1/2+\delta_m})} |Eg|^2 \lesssim R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-\operatorname{Dim} V'} \int_{\Pi\cap 2B} |Eg|^2 + \operatorname{RapDec}(R) \|g\|_{L^2}^2. \tag{6.5}$$

Proof. To prove the lemma, we locate the appropriate space V', and then we appeal to Lemma 6.4. Finding the subspace V' and proving transversality involves relating the direction $G(\theta)$ and the frequency $\xi(\theta)$.

Recall that $G(\omega)=G_0(\omega)/|G_0(\omega)|$, where

$$G_0(\omega)=(-2\omega_1,...,-2\omega_{n-1},1).$$

Let $\mathbb{R}^{n-1}\subset\mathbb{R}^n$ be the $(x_1,...,x_{n-1})$-plane. Examining the formula for G_0 and G, we see that, for every $\omega\in B^{n-1}$, $\mathrm{Angle}(G(\omega),\mathbb{R}^{n-1})\geqslant c_{\mathrm{angle}}>0$.

We define

$$\mathrm{Angle}(V,\mathbb{R}^{n-1}):=\max_{v\in V}\mathrm{Angle}(v,\mathbb{R}^{n-1}).$$

If $\mathrm{Angle}(V,\mathbb{R}^{n-1})$ is smaller than $\frac{1}{2}c_{\mathrm{angle}}$, then $G(\omega)$ is never close to V for any $\omega\in B^{n-1}$. In this case, $\mathbb{T}_{B,V}$ is empty and there is nothing to prove. Therefore, we may assume from now on that

$$\mathrm{Angle}(V,\mathbb{R}^{n-1})\gtrsim 1. \tag{6.6}$$

We let $\Omega(V)$ be the set of $\omega\in\mathbb{R}^{n-1}$ such that $G(\omega)\in V$. We note that $G(\omega)\in V$ if and only if $G_0(\omega)\in V$, and so $\Omega(V)=\{\omega\in\mathbb{R}^{n-1}:G_0(\omega)\in V\}$. As G_0 is an affine map, $\Omega(V)$ is an affine subspace of $\mathbb{R}^{n-1}$. Since V is transverse to $\mathbb{R}^{n-1}$, and since $G_0(\omega)=(-2\omega_1,...,-2\omega_{n-1},1)$, we see that $\mathrm{Dim}\,\Omega(V)=\dim V-1$. Moreover, as V is quantitatively transverse to $\mathbb{R}^{n-1}$, we also see that for $\omega\in B^{n-1}$,

$$\mathrm{Dist}(\omega,\Omega(V))\lesssim\mathrm{Angle}(G(\omega),V).$$

Next we define $\Xi(V)$ to be $\Omega(V)\times\mathbb{R}\subset\mathbb{R}^n$, an affine subspace of $\mathbb{R}^n$. Since $\xi(\omega)=(\omega_1,...,\omega_{n-1},|\omega|^2)$, we see that

$$\mathrm{Dist}(\xi(\omega),\Xi(V))=\mathrm{Dist}(\omega,\Omega(V)).$$

The spaces $\Omega(V)\subset\mathbb{R}^{n-1}$ and $\Xi(V)\subset\mathbb{R}^n$ are affine subspaces. We let $V'=\Xi(V)^{\perp}$, the subspace perpendicular to $\Xi(V)$. The dimension of $\Xi(V)=\mathrm{Dim}\,\Omega(V)+1=\mathrm{Dim}\,V$, and so $\mathrm{Dim}\,V+\mathrm{Dim}\,V'=n$ as desired.

We let $\pi_{V'}$ be the orthogonal projection from $\mathbb{R}^n$ to $V'=\Xi(V)^{\perp}$. Combining our estimates above, we see that

$$\pi_{V'}(\{\omega\in B^{n-1}:\mathrm{Angle}(G(\omega),V)\leqslant\alpha\})\subset\text{a ball of radius}\lesssim\alpha. \tag{6.7}$$

Next we turn to the third claim in the lemma. Let Π be an $(n-m)$-plane parallel to V' passing through B. We know that $(Eg_{\theta,v})\widehat{\ }$ is supported in $\xi(\theta)$. The restriction

to Π of $Eg_{\theta,v}$ has Fourier transform supported in $\pi_{V'}(\xi(\theta))$. Now, for all $(\theta,v)\in\mathbb{T}_{B,V}$, $\mathrm{Angle}(G(\theta),V)\lesssim R^{-1/2+\delta_m}$. By equation (6.7), all the $\pi_{V'}(\xi(\theta))$ lie in a single ball of radius $\lesssim R^{-1/2+\delta_m}$. Therefore, if we view Eg as a function $G:\Pi\to\mathbb{C}$, its Fourier transform is supported in a ball of radius $\lesssim R^{-1/2+\delta_m}$. We apply Lemma 6.4, giving

$$\int_{\Pi\cap B(x_0,\varrho^{1/2+\delta_m})}|Eg|^2\lesssim\left(\frac{R^{1/2-\delta_m}}{\varrho^{1/2+\delta_m}}\right)^{-\dim V'}\int_\Pi W_{B(x_0,R^{1/2-\delta_m})}|Eg|^2 \tag{6.8}$$

$$\lesssim R^{O(\delta_m)}\left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-\dim V'}\int_\Pi W_B|Eg|^2. \tag{6.9}$$

Finally, $Eg=\sum_{(\theta,v)\in\mathbb{T}_{B,V}}Eg_{\theta,v}$. Each $Eg_{\theta,v}$ is essentially supported on $T_{\theta,v}$. Since $T_{\theta,v}$ is transverse to Π and intersects B, we see that if $x\in\Pi\setminus 2B$, then

$$|Eg(x)|\leqslant\mathrm{RapDec}(R)\|g\|_{L^2}.$$

So,

$$\int_\Pi W_B|Eg|^2\leqslant\int_{\Pi\cap 2B}|Eg|^2+\mathrm{RapDec}(R)\|g\|_{L^2}^2.$$

Finally, we want to see that V and V' are quantitatively transverse. We define

$$\mathrm{Angle}(V,V'):=\min_{\substack{v\in V\\ w\in V'}}\mathrm{Angle}(v,w).$$

Sublemma 6.6. *We have*

$$\mathrm{Angle}(V,\mathbb{R}^{n-1})=\mathrm{Angle}(V,V').$$

In particular, in the non-vacuous case that $\mathrm{Angle}(V,\mathbb{R}^{n-1})\gtrsim 1$, we see that V and V' are quantitatively transverse.

Proof. The intersection $V\cap G_0(\mathbb{R}^{n-1})$ is an affine space parallel to $\Omega(V)$. Let $v\in V$ be a unit vector perpendicular to $\Omega(V)$. Let $v_1,...,v_{m-1}$ be an orthonormal basis of $\Omega(V)$. Then $v_1,...,v_{m-1},v$ is an orthonormal basis of V.

Let e_n be the nth coordinate unit vector. We see that $v_1,...,v_{m-1},e_n$ is an orthonormal basis for $\Xi(V)$. We also see that $V'=\Xi(V)^\perp\subset\mathbb{R}^{n-1}$ and $V'=\Omega(V)^\perp\subset\mathbb{R}^{n-1}$.

Since $v_1,...,v_{m-1}\subset\Omega(V)\subset\mathbb{R}^{n-1}$,

$$\mathrm{Angle}(V,\mathbb{R}^{n-1})=\mathrm{Angle}(v,\mathbb{R}^{n-1}).$$

As v is perpendicular to $\Omega(V)$, we see that the projection of v to $\mathbb{R}^{n-1}$ actually lies in $\Xi(V)^\perp=V'$, and so

$$\mathrm{Angle}(v,\mathbb{R}^{n-1})=\mathrm{Angle}(v,V').$$

But since $v_1,...,v_{m-1}$ are in $\Xi(V)$, we see that v is the vector in V which makes the smallest angle with $\Xi(V)^\perp=V'$, and so

$$\mathrm{Angle}(v,V')=\mathrm{Angle}(V,V').$$

□

This completes the proof of Lemma 6.5. □

Now we are ready to prove Lemma 6.2.

Proof. Since $(\theta,v)\in\mathbb{T}_Z$, and $T_{\theta,v}\cap B$ is non-empty, we know that, for any $z\in Z\cap 2B$,

$$\mathrm{Angle}(T_zZ,G(\theta))\lesssim R^{-1/2+\delta_m}.$$

Let V be a subspace of lowest possible dimension so that, for all $(\theta,v)\in\mathbb{T}_{B,Z}$,

$$\mathrm{Angle}(V,G(\theta))\lesssim R^{-1/2+\delta_m}.$$

Let V' be the subspace given by Lemma 6.5. We know that $\operatorname{Dim}V+\operatorname{Dim}V'=n$, and we know that V' is quantitatively transverse to V. By (6.8), we also know that, for any plane Π parallel to V',

$$\int_{\Pi\cap B(x_0,\varrho^{1/2+\delta_m})}|Eg|^2\lesssim R^{O(\delta_m)}\left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-\operatorname{Dim}V'}\int_{\Pi\cap 2B}|Eg|^2+\mathrm{RapDec}(R)\|g\|_{L^2}^2. \quad (6.10)$$

We claim that for each $z\in Z\cap B$, T_zZ is quantitatively transverse to V'. If this is not the case, it means that there exists a point $z\in Z$ and a subspace $W\subset T_zZ$ with

$$\operatorname{Dim}W>\operatorname{Dim}Z-\operatorname{Dim}V,$$

such that, for each non-zero $w\in W$,

$$\mathrm{Angle}(w,V')\leqslant o(1).$$

Since V and V' are transverse, this angle condition guarantees that

$$\mathrm{Angle}(w,V)\gtrsim 1.$$

Because of this angle condition, we can construct a linear map $L\colon\mathbb{R}^n\to V$ such that L restricted to V is the identity, L restricted to W is zero, and $|L|\lesssim 1$. Recall that for each $(\theta,v)\in\mathbb{T}_{B,Z}$, $\mathrm{Angle}(G(\theta),V)\lesssim R^{-1/2+\delta_m}$, and so

$$|L(G(\theta))-G(\theta)|\lesssim R^{-1/2+\delta_m}.$$

On the other hand, we know that $G(\theta)\subset N_{R^{-1/2+\delta_m}}(T_zZ)\cap B(1)$, and so $L(G(\theta))$ lies in

$$L(N_{R^{-1/2+\delta_m}}(T_zZ)\cap B(1))\subset N_{CR^{-1/2+\delta_m}}(L(T_zZ)).$$

This shows that, for all $(\theta, v) \in \mathbb{T}_{B,Z}$,

$$\text{Angle}(G(\theta), L(T_z Z)) \lesssim R^{-1/2+\delta_m}.$$

But, since L vanishes on W, $L(T_z Z)$ is a subspace of dimension at most

$$\text{Dim}\, Z - \text{Dim}\, W < \text{Dim}\, V.$$

This contradicts our hypothesis that V has minimal dimension. This finishes the proof of our claim that for each $z \in Z \cap 2B$, $T_z Z$ is quantitatively transverse to V'.

Suppose that Π is a plane parallel to V' and intersecting B. Given the transversality we just proved, it follows that

$$\Pi \cap N_{\varrho^{1/2+\delta_m}}(Z) \cap B \subset N_{C\varrho^{1/2+\delta_m}}(\Pi \cap Z) \cap \Pi \cap 2B.$$

Note that $\Pi \cap Z$ is itself a transverse complete intersection of dimension

$$\text{Dim}\, V' + \text{Dim}\, Z - n.$$

Now, the set $N_{C\varrho^{1/2+\delta_m}}(\Pi \cap Z) \cap \Pi \cap 2B$ can be covered by

$$R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}} \right)^{\text{Dim}\, V' + \text{Dim}\, Z - n}$$

balls in Π of radius $\varrho^{1/2+\delta_m}$ (cf. [Won]). Applying (6.10) on each of these balls and summing, we get the bound

$$\int_{\Pi \cap N_{\varrho^{1/2+\delta_m}}(Z) \cap B} |Eg|^2 \lesssim R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}} \right)^{-(n-m)} \int_{\Pi \cap 2B} |Eg|^2 + \text{RapDec}(R) \|g\|_{L^2}^2.$$

Finally, integrating over planes Π parallel to V' (using Fubini's theorem), we get the desired bound:

$$\int_{B \cap N_{\varrho^{1/2+\delta_m}}(Z)} |Eg|^2 \lesssim R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}} \right)^{-(n-m)} \int_{2B} |Eg|^2 + \text{RapDec}(R) \|g\|_{L^2}^2. \qquad \square$$

7. Adjusting a wave packet decomposition to a smaller ball

Suppose that $B(y, \varrho) \subset B_R$ for some radius ϱ in the range $R^{1/2} < \varrho < R$, and we want to decompose f into wave packets associated with the ball $B(y, \varrho)$. How does the new wave packet decomposition relate to the old wave packet decomposition?

If the center y is not at the origin, then we introduce new coordinates

$$\tilde{x} = x - y.$$

We define

$$\psi_y(\omega) := y_1\omega_1 + ... + y_{n-1}\omega_{n-1} + y_n|\omega|^2.$$

Then, we write

$$Ef(x) = \int e^{i(x_1\omega_1 + ... + x_{n-1}\omega_{n-1} + x_n|\omega|^2)} f(\omega) = \int e^{i(\tilde{x}_1\omega_1 + ... + \tilde{x}_{n-1}\omega_{n-1} + \tilde{x}_n|\omega|^2)} e^{i\psi_y(\omega)} f(\omega).$$

For any function f, we use the notation

$$\tilde{f}(\omega) = e^{i\psi_y(\omega)} f(\omega).$$

In this notation, we now have

$$Ef(x) = E\tilde{f}(\tilde{x}).$$

Next, we decompose $\tilde{f}$ into wave packets adapted to the ball B_ϱ. We follow the construction of wave packets in §3, except with the radius R replaced by ϱ. We cover B^{n-1} with caps $\tilde{\theta}$ of radius $\varrho^{-1/2}$. We cover $\mathbb{R}^{n-1}$ by finitely overlapping balls of radius $\sim \varrho^{(1+\delta)/2}$, centered at vectors $\tilde{v} \in \varrho^{(1+\delta)/2}\mathbb{Z}^{n-1}$. And we decompose $\tilde{f}$ as

$$\tilde{f} = \sum_{\tilde{\theta},\tilde{v}} \tilde{f}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(R)\|f\|_{L^2},$$

where $\tilde{f}_{\tilde{\theta},\tilde{v}}$ is supported in $\tilde{\theta}$ and its Fourier transform is essentially supported in

$$B(\tilde{v}, \varrho^{(1+\delta)/2}).$$

For each $(\tilde{\theta}, \tilde{v})$, $E\tilde{f}_{\tilde{\theta},\tilde{v}}$ is essentially supported on a tube $T_{\tilde{\theta},\tilde{v}}$ of radius $\varrho^{1/2+\delta}$ and length ϱ. In the $\tilde{x}$ coordinates, this tube is contained in B_ϱ, while in the original x coordinates, this tube is contained in $B(y, \varrho)$.

How does the original wave packet decomposition $f = \sum_{\theta,v} f_{\theta,v}$ relate to the new one? The first question we study is, if we expand $f_{\theta,v}$ in wave packets at scale B_ϱ, $(f_{\theta,v})^\sim = \sum_{\tilde{\theta},\tilde{v}} (f_{\theta,v})^\sim_{\tilde{\theta},\tilde{v}}$, then which $(\tilde{\theta}, \tilde{v})$ can have a significant contribution? We answer this question in Lemma 7.1. Before stating the lemma, we need a couple definitions.

For a given y and ω we define

$$\bar{v}(\omega, y) := \partial_\omega \psi_y(\omega),$$

and we compute

$$\begin{aligned}\bar{v}(\omega, y) &= \partial_\omega \psi_y(\omega) = \partial_\omega (y_1\omega_1 + ... + y_{n-1}\omega_{n-1} + y_n|\omega|^2) \\ &= (y_1 + 2\omega_1 y_n, ..., y_{n-1} + 2\omega_{n-1} y_n) = y' + 2y_n\omega.\end{aligned}$$

(Here we use the notation $y' = (y_1, ..., y_{n-1})$.) If ω_θ denotes the center of a cap θ, we also write $\bar{v}(\theta, y)$ for $\bar{v}(\omega_\theta, y)$.

Define

$$\widetilde{\mathbb{T}}_{\theta,v} := \{(\tilde{\theta}, \tilde{v}) : \mathrm{Dist}(\theta, \tilde{\theta}) \lesssim \varrho^{-1/2} \text{ and } |v + \bar{v}(\theta, y) - \tilde{v}| \lesssim R^{1/2+\delta/2}\}. \tag{7.1}$$

Lemma 7.1. *The function $(f_{\theta,v})^\sim$ is concentrated in wave packets from $\widetilde{\mathbb{T}}_{\theta,v}$. In other words,*

$$(f_{\theta,v})^\sim = \sum_{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}} (f_{\theta,v})^\sim_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(R)\|f\|_{L^2}.$$

Proof. Since $f_{\theta,v}$ is supported in θ, the support of $(f_{\theta,v})^\sim$ is clearly contained in

$$\bigcup \{\tilde{\theta} : \mathrm{Dist}(\tilde{\theta}, \theta) \lesssim \varrho^{-1/2}\}.$$

The main point is to check that the Fourier transform of $(f_{\theta,v})^\sim$ is essentially supported in a ball around $v + \bar{v}(\theta, y)$ of radius $\lesssim R^{1/2+\delta/2}$. Let η_θ be a bump function which is 1 on θ and decays to 0 outside of 2θ. Then,

$$(e^{i\psi_y(\omega)} f_{\theta,v})^\wedge = (\eta_\theta e^{i\psi_y(\omega)} \cdot f_{\theta,v})^\wedge = (\eta_\theta e^{i\psi_y(\omega)})^\wedge * (f_{\theta,v})^\wedge.$$

Now, $(f_{\theta,v})^\wedge$ is rapidly decaying outside of $B(v, R^{1/2+\delta/2})$. On the other hand, a stationary phase argument shows that $(\eta_\theta e^{i\psi_y(\omega)})^\wedge$ is rapidly decaying outside of

$$B(\bar{v}(\theta, y), R^{1/2}).$$

(To see this, it helps to note that on the support of η_θ, $\partial_\omega \psi_y$ lies in a ball around $\bar{v}(\theta, y)$ of radius $\lesssim R^{1/2}$.) □

Next, we explore the geometric features of a tube $T_{\tilde{\theta},\tilde{v}}$ with $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$.

Lemma 7.2. *If $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$, then the tube $T_{\tilde{\theta},\tilde{v}}$ obeys the following geometric estimates*:

$$\mathrm{HausDist}(T_{\tilde{\theta},\tilde{v}}, T_{\theta,v} \cap B(y, \varrho)) \lesssim R^{1/2+\delta} \tag{7.2}$$

and

$$\mathrm{Angle}(G(\theta), G(\tilde{\theta})) \lesssim \varrho^{-1/2}. \tag{7.3}$$

Proof. We recall the definition of $T_{\theta,v}$ from (3.2):

$$T_{\theta,v} := \{(x', x_n) \in B_R : |x' + 2x_n\omega_\theta + v| \leqslant R^{1/2+\delta}\}.$$

In the coordinates $\tilde{x}$, since $x = \tilde{x} + y$ and $y' + 2y_n\omega_\theta = \bar{v}(\theta, y)$,

$$T_{\theta,v} \cap B(y, \varrho) = \{(\tilde{x}', \tilde{x}_n) \in B_\varrho : |\tilde{x}' + 2\tilde{x}_n\omega_\theta + \bar{v}(\theta, y) + v| \leqslant R^{1/2+\delta}\}. \tag{7.4}$$

On the other hand,

$$T_{\tilde{\theta},\tilde{v}} := \{(\tilde{x}', \tilde{x}_n) \in B_\varrho : |\tilde{x}' + 2\tilde{x}_n\omega_{\tilde{\theta}} + \tilde{v}| \leqslant \varrho^{1/2+\delta}\}. \tag{7.5}$$

By the definiton of $\widetilde{\mathbb{T}}_{\theta,v}$, $\mathrm{Dist}(\tilde{\theta}, \theta) \leqslant \varrho^{-1/2}$ and so $|\omega_\theta - \omega_{\tilde{\theta}}| \lesssim \varrho^{-1/2}$. Since $|\tilde{x}_n| \leqslant \varrho$, $|2\tilde{x}_n\omega_\theta - 2\tilde{x}_n\omega_{\tilde{\theta}}| \lesssim \varrho^{1/2}$. By the definition of $\widetilde{\mathbb{T}}_{\theta,v}$, $|v + \bar{v}(\theta, y) - \tilde{v}| \lesssim R^{1/2+\delta/2}$. Comparing (7.4) and (7.5), we see that $\mathrm{HausDist}(T_{\theta,v} \cap B(y, \varrho), T_{\tilde{\theta},\tilde{v}}) \lesssim R^{1/2+\delta}$, as desired.

Since $\mathrm{Dist}(\tilde{\theta}, \theta) \leqslant \varrho^{-1/2}$, it follows that $\mathrm{Angle}(G(\theta), G(\tilde{\theta})) \lesssim \varrho^{-1/2}$. □

Many different (θ, v) lead to essentially the same set $\widetilde{\mathbb{T}}_{\theta,v}$. If $\mathrm{Dist}(\theta_1, \theta_2) \leqslant \varrho^{-1/2}$, and $|v_1 + \bar{v}(\theta_1, y) - v_2 - \bar{v}(\theta_2, y)| \leqslant R^{1/2+\delta/2}$, then $\widetilde{\mathbb{T}}_{\theta_1,v_1}$ and $\widetilde{\mathbb{T}}_{\theta_2,v_2}$ are essentially the same. We can organize the possible pairs (θ, v) into equivalence classes in the following way. If $\tilde{\theta}$ is one of our caps of radius $\varrho^{-1/2}$, and $w \in R^{1/2+\delta/2}\mathbb{Z}^{n-1}$, then we define

$$\mathbb{T}_{\tilde{\theta},w} := \{(\theta, v) : \mathrm{Dist}(\theta, \tilde{\theta}) \lesssim \varrho^{-1/2} \text{ and } |v + \bar{v}(\theta, y) - w| \lesssim R^{1/2+\delta/2}\}. \tag{7.6}$$

If (θ_1, v_1) and (θ_2, v_2) lie in the same set $\mathbb{T}_{\tilde{\theta},w}$, then $\widetilde{\mathbb{T}}_{\theta_1,v_1}$ and $\widetilde{\mathbb{T}}_{\theta_2,v_2}$ are essentially the same. They are both contained in (and essentially equal to)

$$\widetilde{\mathbb{T}}_{\tilde{\theta},w} := \{(\tilde{\theta}_1, \tilde{v}) : \mathrm{Dist}(\tilde{\theta}_1, \tilde{\theta}) \lesssim \varrho^{-1/2} \text{ and } |w - \tilde{v}| \lesssim R^{1/2+\delta/2}\}. \tag{7.7}$$

Now Lemma 7.1 gives the following corollary.

LEMMA 7.3. *If g is concentrated in wave packets in $\mathbb{T}_{\tilde{\theta},w}$, then $\tilde{g}$ is concentrated in wave packets in $\widetilde{\mathbb{T}}_{\tilde{\theta},w}$. In other words, if*

$$g = \sum_{(\theta,v) \in \mathbb{T}_{\tilde{\theta},w}} g_{\theta,v} + \mathrm{RapDec}(R)\|g\|_{L^2},$$

then

$$\tilde{g} = \sum_{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_{\tilde{\theta},w}} \tilde{g}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(R)\|g\|_{L^2}.$$

We also note that the sets $\mathbb{T}_{\tilde{\theta},w}$ are essentially disjoint, and their union contains all the possible pairs (θ, v). Similarly, the sets $\widetilde{\mathbb{T}}_{\tilde{\theta},w}$ are essentially disjoint, and their union contains all possible pairs $(\tilde{\theta}, \tilde{v})$. With this in mind, we define

$$g_{\tilde{\theta},w} := \sum_{(\theta,v)\in\mathbb{T}_{\tilde{\theta},w}} g_{\theta,v}. \tag{7.8}$$

$$(\tilde{g})_{\tilde{\theta},w} := \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}_{\tilde{\theta},w}} (\tilde{g})_{\tilde{\theta},\tilde{v}}. \tag{7.9}$$

For any g, we get a decomposition $g=\sum_{\tilde{\theta},w} g_{\tilde{\theta},w}$ obeying with

$$\|g\|_{L^2}^2 \sim \sum_{(\tilde{\theta},w)} \|g_{\tilde{\theta},w}\|_{L^2}^2, \tag{7.10}$$

Similarly, for any $\tilde{g}$, we get a decomposition $\tilde{g}=\sum_{\tilde{\theta},w} \tilde{g}_{\tilde{\theta},w}$ with

$$\|\tilde{g}\|_{L^2}^2 \sim \sum_{(\tilde{\theta},w)} \|\tilde{g}_{\tilde{\theta},w}\|_{L^2}^2. \tag{7.11}$$

By Lemma 7.2, for all the pairs $(\theta, v)\in\mathbb{T}_{\tilde{\theta},w}$, the sets $T_{\theta,v}\cap B(y,\varrho)$ are essentially the same. We denote this intersection by $T_{\tilde{\theta},w}\subset B(y,\varrho)$. It is a tube of radius $R^{1/2+\delta}$ and length ϱ. The set $\mathbb{T}_{\tilde{\theta},w}$ can be described geometrically as the set of pairs (θ, v) such that $T_{\theta,v}\cap B(y,\varrho)$ is essentially $T_{\tilde{\theta},w}$, and such that the direction of $T_{\theta,v}$ obeys the inequality $\mathrm{Angle}(G(\theta), G(\tilde{\theta}))\lesssim\varrho^{-1/2}$.

We will need to study the following situation. We have a function g which is concentrated on wave packets in $\mathbb{T}_Z$, and we want to study Eg on a smaller ball $B(y,\varrho)\subset B_R$. If we decompose g into wave packets associated with the ball $B(y,\varrho)$, what can we say about the new wave packet decomposition?

First of all, we point out the wave decompositon of $\tilde{g}$ at scale ϱ is not necessarily concentrated on wave packets that are tangent to Z on $B(y,\varrho)$. By Lemma 7.1, we do know that $\tilde{g}$ is concentrated on wave packets in $\bigcup_{(\theta,v)\in\mathbb{T}_Z}\widetilde{\mathbb{T}}_{\theta,v}$. If $(\theta, v)\in\mathbb{T}_Z$, then we know that $T_{\theta,v}$ is tangent to Z on B_R, which implies that

$$T_{\theta,v}\cap B(y,\varrho)\subset N_{R^{1/2+\delta_m}}(Z)\cap B(y,\varrho),$$

and that, for any $x\in T_{\theta,v}$ and $z\in Z\cap B(y,\varrho)$ with $|x-z|\lesssim R^{1/2+\delta_m}$,

$$\mathrm{Angle}(G(\theta), T_z Z)\lesssim R^{-1/2+\delta_m}.$$

If now $(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}_{\theta,v}$, then (7.2) and (7.3) imply that

$$T_{\tilde{\theta},\tilde{v}}\subset N_{R^{1/2+\delta_m}}(Z)\cap B(y,\varrho),$$

and that, for any $x \in T_{\tilde{\theta},\tilde{v}}$ and $z \in Z \cap B(y,\varrho)$ with $|x-z| \lesssim R^{1/2+\delta_m}$,

$$\mathrm{Angle}(G(\theta), T_z Z) \lesssim R^{-1/2+\delta_m} + \varrho^{-1/2} \leqslant \varrho^{-1/2+\delta_m}.$$

The angle condition is more than strong enough for $T_{\tilde{\theta},\tilde{v}}$ to be tangent to Z in $B(y,\varrho)$, but it is not true that $T_{\tilde{\theta},\tilde{v}} \subset N_{\varrho^{1/2+\delta_m}}(Z) \cap B(y,\varrho)$. If $T_{\tilde{\theta},\tilde{v}}$ intersects

$$N_{\varrho^{1/2+\delta_m}}(Z) \cap B(y,\varrho),$$

then the angle condition guarantees that $T_{\tilde{\theta},\tilde{v}}$ is contained in

$$N_{2\varrho^{1/2+\delta_m}}(Z) \cap B(y,\varrho).$$

A bit more generally, if b is a vector with $|b| \leqslant R^{1/2+\delta_m}$, and if $T_{\tilde{\theta},\tilde{v}}$ intersects

$$N_{\varrho^{1/2+\delta_m}}(Z+b) \cap B(y,\varrho),$$

then $T_{\tilde{\theta},\tilde{v}}$ is contained in

$$N_{2\varrho^{1/2+\delta_m}}(Z+b) \cap B(y,\varrho),$$

and $T_{\tilde{\theta},\tilde{v}}$ is tangent to $Z+b$ in B_j.

For any $b \in B_{R^{1/2+\delta_m}}$, we define

$$\widetilde{\mathbb{T}}_{Z+b} := \{(\tilde{\theta},\tilde{v}) : T_{\tilde{\theta},\tilde{v}} \text{ is tangent to } Z+b \text{ in } B_j\} \quad \text{and} \quad \tilde{g}_b := \sum_{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_{Z+b}} \tilde{g}_{\tilde{\theta},\tilde{v}}.$$

So, we see that, if g is concentrated on wave packets in $\mathbb{T}_Z$, then $\tilde{g}$ is concentrated on wave packets in $\bigcup_{|b| \lesssim R^{1/2+\delta_m}} \widetilde{\mathbb{T}}_{Z+b}$. For any $(\tilde{\theta},\tilde{v}) \in \bigcup_{(\theta,v)\in\mathbb{T}_Z} \widetilde{\mathbb{T}}_{\theta,v}$, we saw above that either $(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_{Z+b}$, or else $T_{\tilde{\theta},\tilde{v}}$ is disjoint from $N_{\varrho^{1/2+\delta_m}}(Z) \cap B(y,\varrho)$. Therefore, for $x = y + \tilde{x} \in B(y,\varrho)$,

$$|E\tilde{g}_b(\tilde{x})| \sim \chi_{N_{\varrho^{1/2+\delta_m}}(Z+b)}(x) Eg(x). \tag{7.12}$$

To get finer information, it is helpful to decompose g as above as $g = \sum_{\tilde{\theta},w} g_{\tilde{\theta},w}$, and to think about the wave packet decomposition of each piece $(g_{\tilde{\theta},w})^\sim$ on $B(y,\varrho)$. For brevity, we let $h = g_{\tilde{\theta},w}$. We choose a ball $B(x_0, R^{1/2+\delta_m})$ with $x_0 \in T_{\tilde{\theta},w} \subset B(y,\varrho)$. For any $(\theta,v) \in \mathbb{T}_{\tilde{\theta},w}$, $T_{\theta,v} \cap B(y,\varrho) \subset T_{\tilde{\theta},w}$, and so $T_{\theta,v}$ intersects $B(x_0, R^{1/2+\delta_m})$ in a tube segment of length $R^{1/2+\delta_m}$.

By Lemma 3.4, we have

$$\|h\|_{L^2}^2 \sim R^{-1/2-\delta_m} \|Eh\|_{L^2(B(x_0,R^{1/2+\delta_m}))}^2. \tag{7.13}$$

Now we know that $E\tilde{h}(\tilde{x}) = Eh(\tilde{x}+y)$. Also, for $x = y+\tilde{x} \in B(y,\varrho)$, we know by (7.12) that

$$|E\tilde{h}_b(\tilde{x})| \sim \chi_{N_{\varrho^{1/2+\delta_m}}(Z+b)}(x) Eh(x).$$

Using Lemma 3.4 again, we have

$$\|\tilde{h}_b\|_{L^2}^2 \sim R^{-1/2-\delta_m} \|Eh\|^2_{L^2(B(x_0,R^{1/2+\delta_m}) \cap N_{\varrho^{1/2+\delta_m}}(Z+b))}. \tag{7.14}$$

These observations lead to a couple estimates about how $\|\tilde{h}_b\|_{L^2}$ relates to $\|h\|_{L^2}$.

LEMMA 7.4. *Suppose that h is concentrated on wave packets in $\mathbb{T}_{\tilde{\theta},w}$ for some $(\tilde{\theta},w)$, and x_0 is in the tube $T_{\tilde{\theta},w}$. If we choose a set of vectors $b \in B_{R^{1/2+\delta_m}}$ so that the sets*

$$B(x_0, R^{1/2+\delta_m}) \cap N_{\varrho^{1/2+\delta_m}}(Z+b)$$

are disjoint, then

$$\sum_b \|\tilde{h}_b\|_{L^2}^2 \lesssim \|h\|_{L^2}^2.$$

Finally, we come to transverse equidistribution estimates. Combining the transverse equidistribution estimate in Lemma 6.2 with the considerations in this section, we get the following estimates.

LEMMA 7.5. *Suppose that h is concentrated on wave packets in $\mathbb{T}_Z$ and also on wave packets in $\mathbb{T}_{\tilde{\theta},w}$ for some $(\tilde{\theta},w)$. Then, for any $b \in B_{R^{1/2+\delta_m}}$, we have*

$$\|\tilde{h}_b\|_{L^2}^2 \leqslant R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|h\|_{L^2}^2.$$

Proof. We combine (7.14), Lemma 6.2, and (7.13) to get

$$\begin{aligned}
\|\tilde{h}_b\|_{L^2}^2 &\sim R^{-1/2-\delta_m} \|Eh\|^2_{L^2(B(x_0,R^{1/2+\delta_m}) \cap N_{\varrho^{1/2+\delta_m}}(Z+b))} \\
&\lesssim R^{-1/2-\delta_m} R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|Eh\|^2_{L^2(B(x_0,R^{1/2+\delta_m}))} \\
&\sim R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|h\|_{L^2}^2.
\end{aligned}$$

□

We can now combine the different $(\tilde{\theta},w)$ in order to get estimates for g.

LEMMA 7.6. *If g is concentrated in wave packets in $\mathbb{T}_Z$, then, for any b,*

$$\|\tilde{g}_b\|_{L^2}^2 \leqslant R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|g\|_{L^2}^2.$$

Proof. We first expand $g=\sum_{\tilde\theta,w} g_{\tilde\theta,w}$. The wave packets contributing significantly to each $g_{\tilde\theta,w}$ are a subset of those contributing to g, and so each $g_{\tilde\theta,w}$ is concentrated on wave packets in $\mathbb{T}_Z$. For each $\tilde\theta$ and w, Lemma 7.5 tells us that

$$\|(g_{\tilde\theta,w})_b^\sim\|_{L^2}^2 \leqslant R^{O(\delta_m)}\left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)}\|g_{\tilde\theta,w}\|_{L^2}^2.$$

We know that the $g_{\tilde\theta,w}$ are orthogonal, and so

$$\|g\|_{L^2}^2 \sim \sum_{\tilde\theta,w}\|g_{\tilde\theta,w}\|_{L^2}^2.$$

The operation $f\mapsto\tilde f_b$ is a linear map, and so

$$\tilde g_b = \sum_{\tilde\theta,w}(g_{\tilde\theta,w})_b^\sim.$$

We claim that this is also an orthogonal decomposition. By Lemma 7.3, $(g_{\tilde\theta,w})^\sim$ is concentrated on wave packets in $\widetilde{\mathbb{T}}_{\tilde\theta,w}$. But then $(g_{\tilde\theta,w})_b^\sim$ is also concentrated on wave packets in $\widetilde{\mathbb{T}}_{\tilde\theta,w}$. The different sets $\widetilde{\mathbb{T}}_{\tilde\theta,w}$ are disjoint, and so the functions $(g_{\tilde\theta,w})_b^\sim$ are orthogonal, as claimed. Therefore

$$\|\tilde g_b\|_{L^2}^2 \sim \sum_{\tilde\theta,w}\|(g_{\tilde\theta,w})_b^\sim\|_{L^2}^2.$$

Combining these estimates gives the desired conclusion. □

8. Proof of Theorem 1.5.

We now formulate the inductive estimate that proves Theorem 1.5.

PROPOSITION 8.1. *For $\varepsilon>0$, there are small constants*

$$0<\delta\ll\delta_{n-1}\ll\ldots\ll\delta_1\ll\delta_0\ll\varepsilon,$$

and a large constant $\bar A$ such that the following holds.

Let m be a dimension in the range $1\leqslant m\leqslant n$ and let $R\geqslant 1$. Let $Z=Z(P_1,\ldots,P_{n-m})$ be a transverse complete intersection, with $\operatorname{Deg} P_j\leqslant D_Z$. Suppose that f is concentrated on wave packets from $\mathbb{T}_Z$. Then, for any $1\leqslant A\leqslant\bar A$,

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant C(K,\varepsilon,m,D_Z)R^{m\varepsilon}R^{\delta(\log\bar A-\log A)}R^{-e+1/2}\|f\|_{L^2} \tag{8.1}$$

for all

$$2\leqslant p\leqslant\bar p(k,m):=2\frac{m+k}{m+k-2}, \tag{8.2}$$

where

$$e=e(k,n,p)=\frac12\left(\frac12-\frac1p\right)(n+k). \tag{8.3}$$

When $m=n$, Proposition 8.1 gives Theorem 1.5. When $m=n$, we can take $Z=\mathbb{R}^n$ (and $D_Z=1$). If we now choose $A=\bar{A}$ and $p=\bar{p}(k,n)$, then we compute $-e+\frac{1}{2}=0$, and we get the inequality in Theorem 1.5.

We prove Proposition 8.1 by induction. We will do induction on the dimension m, the radius R, and on A. We start by checking the base of the induction. When R is small, we choose the constant $C(K,\varepsilon,m,D_Z)$ sufficiently large and the result follows. So, from now on, we may assume that R is very large compared to K, ε, m and D_Z. To check the case $A=1$, we choose $\bar{A}$ large enough so that $R^{\delta(\log \bar{A}-\log 1)}=R^{10n}$, and the inequality follows because $\|Ef\|_{\mathrm{BL}^p_{k,1}(B_R)}\leqslant |B_R|\,\|f\|_{L^2}$. The base of the induction on m is $m=k-1$. In this case, since $A\geqslant 1$, we have

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}\leqslant \mathrm{RapDec}(R)\|f\|_{L^2}. \tag{8.4}$$

This follows from the definition of $\mathrm{BL}^p_{k,A}$. Recall that

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)}:=\sum_{B_{K^2}\subset B_R}\mu_{Ef}(B_{K^2}),$$

with

$$\mu_{Ef}(B_{K^2}):=\min_{V_1,\dots,V_A}\left(\max_{\tau}\int_{B_{K^2}}|Ef_\tau|^p\right),$$

where the minimum is over $(k-1)$-subspaces of $\mathbb{R}^n$ and the maximum is over all τ such that $\mathrm{Angle}(G(\tau),V_a)>K^{-1}$ for all a.

Fix a ball $B=B_{K^2}\subset N_{R^{1/2+\delta_m}}(Z)\cap B_R$, and let V be the tangent space to Z at some point z in the $R^{1/2+\delta_m}$-neighborhood of the ball B_{K^2}. Notice that the dimension of V is $m=k-1$. If $T_{\theta,v}$ intersects B_{K^2} and if $(\theta,v)\in\mathbb{T}_Z$, then $\mathrm{Angle}(G(\theta),V)\lesssim R^{-1/2+\delta_m}$. So, if τ contains a θ with $(\theta,v)\in\mathbb{T}_Z$ for some v, then $\mathrm{Angle}(G(\tau),V)\leqslant K^{-1}$. On the other hand, if τ does not contain such a θ, then $\|f_\tau\|_{L^2}\leqslant\mathrm{RapDec}(R)\|f\|_{L^2}$, because f is concentrated on wave packets tangent to Z. Therefore,

$$\left(\max_{\tau}\int_{B_{K^2}}|Ef_\tau|^p\right)\leqslant\mathrm{RapDec}(R)\|f\|^p_{L^2}.$$

On the other hand, if $B_{K^2}\subset B_R$ is not contained in $N_{R^{1/2+\delta_m}}(Z)$, then on B_{K^2} we have $|Ef|\leqslant\mathrm{RapDec}(R)\|f\|_{L^2}$. This proves (8.4), establishing the base case $m=k-1$.

For $p=2$, Proposition 8.1 follows quickly from the basic L^2 estimate in Lemma 3.2:

$$\|Ef\|^2_{\mathrm{BL}^2_{k,A}(B_R)}\leqslant\sum_{\tau}\|Ef_\tau\|^2_{L^2(B_R)}\lesssim R\sum_{\tau}\|f_\tau\|^2_{L^2}=R\|f\|^2_{L^2}.$$

Next, we begin the inductive step. We assume that Proposition 8.1 holds if we decrease the dimension m, the radius R, or the value of A. We define p to be

$$p = \begin{cases} \bar{p}(k,m), & \text{if } m > k, \\ \bar{p}(m,m)+\delta, & \text{if } m = k. \end{cases}$$

We will check the main estimate (8.1) for this value of p. Once we check (8.1) for this value of p, we can get the whole range in (8.2) by interpolation between 2 and our value of p, using the Hölder inequality for BL in Lemma 4.2.

There are two cases, depending on whether or not the mass of μ is concentrated into a small neighborhood of a lower-dimensional variety. We let $D(\varepsilon, D_Z)$ be a function that we will define later. We say that we are in the *algebraic case* if there is a transverse complete intersection $Y^l \subset Z^m$ of dimension $l<m$, defined using polynomials of degree $\leqslant D(\varepsilon, D_Z)$, so that

$$\mu_{Ef}(N_{R^{1/2+\delta_m}}(Y) \cap B_R) \gtrsim \mu_{Ef}(B_R).$$

Otherwise, we say that we are in the *non-algebraic case*, or the cellular case.

8.1. The non-algebraic case

We begin with the non-algebraic case. In this case, we will use polynomial partitioning. To set up the polynomial partitioning, we first locate a significant piece of $N_{R^{1/2+\delta_m}}(Z) \cap B_R$ where the tangent space of Z is not changing too fast. We say that a ball $B(x_0, R^{1/2+\delta_m}) \subset N_{R^{1/2+\delta_m}}(Z) \cap B_R$ is regular if, on each connected component of $Z \cap B(x_0, 10R^{1/2+\delta_m})$, the tangent space TZ is constant up to angle $\frac{1}{100}$. Let $w \in \Lambda^m \mathbb{R}^n$. Recall that $Z_w \subset Z$ is defined in (5.2). For generic w, $Z_w \subset Y$ is a transverse complete intersection of dimension $m-1$, defined using polynomials of degree $\lesssim D_Z$. We can choose a set of $\lesssim 1$ values of w so that on each connected component of $Z \setminus \bigcup_w Z_w$, the tangent plane TZ is constant up to angle $\frac{1}{100}$. Since we are in the non-algebraic case,

$$\mu_{Ef}\left(\bigcup_w N_{10R^{1/2+\delta_m}}(Z_w) \cap B_R\right) \leqslant \frac{1}{100}\mu_{Ef}(B_R).$$

Each ball $B(x_0, R^{1/2+\delta_m}) \subset N_{R^{1/2+\delta_m}}(Z) \cap B_R$ which does not intersect

$$\bigcup_w N_{10R^{1/2+\delta_m}}(Z_w)$$

is a regular ball. So, the regular balls contain most of the mass of μ.

For each regular ball

$$B = B_{R^{1/2+\delta_m}} \subset \widetilde{N}_{R^{1/2+\delta_m}}(Z) \cap B_R,$$

we pick a point $z \in Z \cap B_{R^{1/2+\delta_m}}$ and we define V_B to be the m-plane $T_z Z$. For an m-plane V, we define $\mathfrak{B}_V$ to be the set of regular balls so that $\mathrm{Angle}(V_B, V) \leqslant \frac{1}{100}$. By the pigeonhole principle, we may choose a plane V such that

$$\mu_{Ef}\Big(\bigcup_{B \in \mathfrak{B}_V} B\Big) \gtrsim \mu_{Ef}(B_R).$$

We define $N_1 \subset N_{R^{1/2+\delta_m}}(Z) \cap B_R$ to be the union of the balls $B \in \mathfrak{B}_V$. We let μ_1 be the restriction of μ_{Ef} to N_1, and we note that $\mu_1(N_1) \sim \mu_{Ef}(B_R)$.

We are now ready to do polynomial partitioning. We let $P_V : V \to \mathbb{R}$ denote a polynomial defined on V. We let $\pi : \mathbb{R}^n \to V$ be the orthogonal projection. We now apply polynomial partitioning (Theorem 5.5) to the push-forward measure $\pi_* \mu_1$ on V, using the degree $D = D(\varepsilon, D_Z)$. Theorem 5.5 gives us a non-zero polynomial P_V of degree at most D such that $V \setminus Z(P_V) = \bigcup_j O_{V,j}$, the number of cells $O_{V,j}$ is $\sim D^m$, and for each cell, $\pi_* \mu_1(O_{V,j})$ is $\sim D^{-m} \pi_* \mu_1(V)$. Moreover, $P_V = \prod_j Q_{V,j}$, where each $Q_{V,j}$ has a little freedom in the constant term, which we can use for transversality purposes.

We extend V to a polynomial P on $\mathbb{R}^n$ by setting $P(x) := P_V(\pi(x))$. We note that $Z(P) = \pi^{-1}(Z(P_V))$. We define $O_j := \pi^{-1} O_{V,j}$, and we note that $\mathbb{R}^n \setminus Z(P) = \bigcup_j O_j$ and that $\mu_1(O_j) = \pi_* \mu_1(O_{V,j}) \sim D^{-m} \mu_1(N_1)$. Similarly, we define $Q_j(x) = Q_{V,j}(\pi(x))$, so that $P = \prod_j Q_j$. Each Q_j is a polynomial of degree at most D on $\mathbb{R}^n$, and we have a little freedom in the constant term of each Q_j. By Lemma 5.1, we can guarantee that, for each j, $Y_j = Z(P_1, ..., P_{n-m}, Q_j)$ is a transverse complete intersection.

We define $W := N_{R^{1/2+\delta}} Z(P)$ and $O_j' := O_j \setminus W$. As each tube $T_{\theta,v}$ has radius $R^{1/2+\delta}$, each tube $T_{\theta,v}$ enters at most D of the cells O_j'. On the other hand, we claim that

$$W \cap N_1 \subset \bigcup_j N_{20R^{1/2+\delta_m}}(Y_j).$$

Here is the proof of the claim. Suppose that $x \in W \cap N_1$. Since $x \in N_1$, x is in a regular ball $B = B(x_0, R^{1/2+\delta_m}) \in \mathfrak{B}_V$. There is a point $z_B \in Z \cap B$ with $\mathrm{Angle}(T_{z_B} Z, V) \leqslant \frac{1}{100}$. Let Z_B be the component of $Z \cap 10B$ containing this point z_B. On Z_B, the tangent plane TZ makes a small angle with V. Let us use coordinates (v, w), where $v \in V$ and $w \in V^\perp$. Since the tangent plane of TZ_B makes a small angle with V, we have that Z_B is the graph of a function, $w = h(v)$, where h has small Lipschitz constant $\leqslant \frac{1}{100}$. Since $x \in W$, there must be a point $(v_0, w_0) \in Z(P) \cap B(x, R^{1/2+\delta})$. Since $P(v, w) = P_V(v)$, we see that $P_V(v_0) = 0$. Now the point $(v_0, h(v_0))$ lies both in $Z(P)$ and in Z_B, and so it lies in Y_j for some j. Since $(v_0, h(v_0)) \in Z_B \subset B(x_0, 10R^{1/2+\delta_m})$ and $x \in B(x_0, R^{1/2+\delta_m})$, it follows that $x \in N_{20R^{1/2+\delta_m}}(Y_j)$, as desired.

Since we are in the non-algebraic case, $\mu_1(N_{20R^{1/2+\delta_m}}(Y_j))$ is negligible for each j, and so $\mu_1(W)=\mu_1(W\cap N_1)$ is negligible. Therefore, we have

$$\mu_1(O_j')\sim D^{-m}\mu_1(N_1)\sim D^{-m}\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \quad \text{for the vast majority of } j. \tag{8.5}$$

Now, for each index j, we define a function f_j which only includes the wave packets that enter O_j'. More precisely, $f_j=\sum_{(\theta,v)\in\mathbb{T}_j} f_{\theta,v}$, where

$$\mathbb{T}_j := \{(\theta,v) : T_{\theta,v}\cap O_j' \neq \varnothing\}. \tag{8.6}$$

Each function f_j is also concentrated on wave packets in $\mathbb{T}_Z$. Moreover, there are $\sim D^m$ cells O_j' for which

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim D^m\mu_1(O_j') \lesssim D^m\mu_{Ef_j}(O_j') \lesssim D^m\|Ef_j\|^p_{\mathrm{BL}^p_{k,A}(B_R)}.$$

On the other hand, we have good control on the L^2 norms of f_j. Because each tube $T_{\theta,v}$ enters $\lesssim D$ cells O_j', each pair (θ,v) belongs to $\mathbb{T}_j$ for $\lesssim D$ values of j, and so

$$\sum_j \|f_j\|^2_{L^2} \sim \sum_j \sum_{(\theta,v)\in\mathbb{T}_j} \|f_{\theta,v}\|^2_{L^2} \lesssim D\sum_{(\theta,v)} \|f_{\theta,v}\|^2_{L^2} \sim D\|f\|^2_{L^2}.$$

In summary, there are $\sim D^m$ choices of j such that

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim D^m\|Ef_j\|^p_{\mathrm{BL}^p_{k,A}(B_R)}. \tag{8.7}$$

$$\|f_j\|^2_{L^2} \lesssim D^{1-m}\|f\|^2_{L^2}. \tag{8.8}$$

(For later reference, we also record here: for each j and each θ,

$$\|f_{j,\theta}\|^2_{L^2} \lesssim \sum_{\theta'\cap\theta\neq\varnothing} \|f_{\theta'}\|^2_{L^2}. \tag{8.9}$$

This inequality does not appear in the proof of Theorem 1.5, but it could be useful in some later refinements.)

Using these estimates, we can now prove (8.1) by induction on the radius. To make the computation clearer, we abbreviate

$$E = \overline{C}(K,\varepsilon,m,D_Z)R^{m\varepsilon}R^{\delta(\log\bar{A}-\log A)}R^{-e+\frac{1}{2}},$$

so (8.1) reduces to

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant E\|f\|_{L^2}.$$

Since we assume that (8.1) holds on balls of radius $R/2$, it follows that it also holds on balls of radius R up to a constant factor loss. So we can assume that

$$\|Ef_j\|_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant CE\|f_j\|_{L^2}. \tag{8.10}$$

Now using (8.7) and (8.8), we get

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim D^m \|Ef_j\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim D^m E^p \|f_j\|^p_{L^2} \lesssim D^{m-(m-1)p/2} E^p \|f\|^p_{L^2}.$$

By our choice of p, $p>2m/(m-1)$, and so the exponent of D is negative. If $m>k$, the exponent is negative with pretty large absolute value; if $m=k$, then the exponent is $-\delta$. If we choose $D=D(\varepsilon, D_Z)$ sufficiently large, the power of D dominates the implicit constant, and we get $\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant E\|f\|_{L^2}$, which closes the induction in the non-algebraic case.

8.2. The algebraic case

Next, we turn to the algebraic case. By definition, we know that there is a transverse complete intersection Y^l of dimension $l<m$, defined using polynomials of degree $\leqslant D=D(\varepsilon, D_Z)$, such that

$$\mu_{Ef}(N_{R^{1/2+\delta_m}}(Y)) \gtrsim \mu_{Ef}(B_R). \tag{8.11}$$

In the algebraic case, we subdivide B_R into smaller balls B_j of radius ϱ, chosen so that

$$\varrho^{1/2+\delta_l} = R^{1/2+\delta_m}. \tag{8.12}$$

For each j, we define $f_j = \sum_{(\theta,v)\in\mathbb{T}_j} f_{\theta,v}$, where

$$\mathbb{T}_j := \{(\theta, v) : T_{\theta,v} \cap N_{R^{1/2+\delta_m}}(Y) \cap B_j \neq \varnothing\}.$$

On $N_{R^{1/2+\delta_m}}(Y)\cap B_j$, we have $Ef_j = Ef + \mathrm{RapDec}(R)\|f\|_{L^2}$. Therefore,

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim \sum_j \|Ef_j\|^p_{\mathrm{BL}^p_{k,A}(B_j)} + \mathrm{RapDec}(R)\|f\|^p_{L^2}.$$

We further subdivide $\mathbb{T}_j$ into tubes that are tangential to Y and tubes that are transverse to Y. As in definition 6.1, we say that $T_{\theta,v}$ is tangent to Y in B_j if

$$T_{\theta,v} \cap 2B_j \subset N_{R^{1/2+\delta_m}}(Y) \cap 2B_j = N_{\varrho^{1/2+\delta_l}}(Y) \cap 2B_j \tag{8.13}$$

and, for any $x\in T_{\theta,v}$ and $y\in Y\cap 2B_j$ with $|x-y| \lesssim R^{1/2+\delta_m} = \varrho^{1/2+\delta_l}$,

$$\mathrm{Angle}(G(\theta), TY_y) \lesssim \varrho^{-1/2+\delta_l}. \tag{8.14}$$

We define the tangential wave packets by

$$\mathbb{T}_{j,\mathrm{tang}} = \{(\theta, v) \in \mathbb{T}_j : T_{\theta,v} \text{ is tangent to } Y \text{ in } B_j\}.$$

And we define the transverse wave packets by

$$\mathbb{T}_{j,\mathrm{trans}} := \mathbb{T}_j \setminus \mathbb{T}_{j,\mathrm{trans}}.$$

We define $f_{j,\mathrm{tang}} = \sum_{(\theta,v)\in\mathbb{T}_{j,\mathrm{tang}}} f_{\theta,v}$ and $f_{j,\mathrm{trans}} = \sum_{(\theta,v)\in\mathbb{T}_{j,\mathrm{trans}} f_{\theta,v}}$, so

$$f_j = f_{j,\mathrm{tang}} + f_{j,\mathrm{trans}}.$$

Therefore, we have

$$\sum_j \|Ef_j\|^p_{\mathrm{BL}^p_{k,A}(B_j)} \lesssim \sum_j \|Ef_{j,\mathrm{tang}}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)} + \sum_j \|Ef_{j,\mathrm{trans}}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)}. \tag{8.15}$$

We will control the contribution of the tangential wave packets by induction on the dimension m, and we will control the contribution of the transverse wave packets by induction on the radius R.

8.3. The tangential subcase

Suppose first that the tangential wave packets dominate the right-hand side of (8.15). In order to apply induction to $Ef_{j,\mathrm{tang}}$ on B_j, we redo the wave packet decomposition at a scale appropriate to B_j, as in §7. For brevity, during this discussion, we let $g = f_{j,\mathrm{tang}}$.

$$\tilde{g} = \sum_{\tilde{\theta},\tilde{v}} \tilde{g}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(R)\|f\|_{L^2}.$$

Before applying induction, we need to check that this wave packet decomposition is concentrated on pairs $(\tilde{\theta}, \tilde{v})$ that are tangent to Y on B_j, in the sense of Definition 6.1; in other words, on pairs $(\tilde{\theta}, \tilde{v})$ such that

$$T_{\tilde{\theta},\tilde{v}} \subset N_{\varrho^{1/2+\delta_l}}(Y) \cap B_j, \tag{8.16}$$

and, for any $x \in T_{\tilde{\theta},\tilde{v}}$ and $y \in Y \cap B_j$ with $|x-y| \lesssim \varrho^{1/2+\delta_l}$,

$$\mathrm{Angle}(G(\tilde{\theta}), T_y Y) \lesssim \varrho^{-1/2+\delta_l}. \tag{8.17}$$

We know that $g = f_{j,\mathrm{tang}}$ is concentrated on wave packets $(\theta, v) \in \mathbb{T}_{j,\mathrm{tang}}$, which obey (8.13) and (8.14). These tell us that $T_{\theta,v} \cap B_j$ lies in the desired neighborhood of $Y \cap B_j$

and makes a good angle with T_yY. For any (θ, v), $(f_{\theta,v})^\sim$ is concentrated on wave packets $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$, by Lemma 7.1. For $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$, (7.2) and (7.3) tell us that $T_{\tilde{\theta},\tilde{v}}$ lies in a small neighborhood of $T_{\theta,v} \cap B_j$ and makes a small angle with $T_{\theta,v}$. So, if $(\theta, v) \in \mathbb{T}_{j,\mathrm{tang}}$ and $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$, then $T_{\tilde{\theta},\tilde{v}}$ obeys (8.16) and (8.17). We have now checked the hypotheses of Proposition 8.1 for $\tilde{g}$ with the variety Y on the ball B_j, and so we may apply induction on the dimension.

By induction on the dimension, we get the following inequality:

$$\begin{aligned}\|Ef_{j,\mathrm{tang}}\|_{\mathrm{BL}^p_{k,A/2}(B_j)} &= \|E\tilde{g}\|_{\mathrm{BL}^p_{k,A/2}(B_\varrho)} \\ &\leqslant C(K,\varepsilon,l,D(\varepsilon,D_Z))\varrho^{l\varepsilon}\varrho^{\delta(\log \bar{A}-\log(A/2))}\varrho^{-e+1/2}\|f_{j,\mathrm{tang}}\|_{L^2}\end{aligned}$$

for

$$2 \leqslant p \leqslant \bar{p}(k,l) := 2\frac{l+k}{l+k-2},$$

where

$$e = e(k,n,p) = \frac{1}{2}\left(\frac{1}{2}-\frac{1}{p}\right)(n+k).$$

Since $l<m$, we have $\bar{p}(k,m)<\bar{p}(k,l)$, and so our value of p is in the range $2\leqslant p\leqslant \bar{p}(k,l)$ and the bound above applies. The number of balls B_j is $\lesssim R^{O(\delta_l)}$. Summing brutally over the balls, we see that

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim R^{O(\delta_l)}C(K,\varepsilon,l,D(\varepsilon,D_Z))\varrho^{l\varepsilon}\varrho^{\delta(\log \bar{A}-\log(A/2))}\varrho^{-e+1/2}\|f\|_{L^2}.$$

We note that the exponent $-e+\frac{1}{2}$ may well be negative. Nevertheless,

$$\varrho^{-e+1/2} \leqslant R^{O(\delta_l)}R^{-e+1/2}.$$

Also,

$$\varrho^{\delta(\log \bar{A}-\log A/2)} \leqslant R^{\delta(\log \bar{A}-\log A/2)} = R^\delta R^{\delta(\log \bar{A}-\log A)}.$$

Therefore, the last expression is

$$\leqslant R^{O(\delta_l)}C(K,\varepsilon,l,D(\varepsilon,D_Z))R^{l\varepsilon}R^{\delta(\log \bar{A}-\log(A/2))}R^{-e+1/2}\|f\|_{L^2}.$$

Since δ_l is much smaller than ε, $R^{O(\delta_l)}R^{l\varepsilon} \leqslant R^{m\varepsilon}$, and the induction closes. (We have to choose $C(K,\varepsilon,m,D_Z)$ larger than $C(K,\varepsilon,l,D(\varepsilon,D_Z))$.) This finishes the discussion of the tangential algebraic case.

8.4. The transverse subcase

Suppose now that the transverse wave packets dominate (8.15). First, we note that

$$\sum_j \|f_{j,\mathrm{trans}}\|_{L^2}^2 = \sum_{\theta,v} \#\{j : (\theta,v) \in \mathbb{T}_{j,\mathrm{trans}}\} \|f_{\theta,v}\|_{L^2}^2.$$

Next, we claim that, for each (θ,v), $\#\{j:(\theta,v)\in\mathbb{T}_{j,\mathrm{trans}}\}\lesssim_{\varepsilon,D_Z} 1$. In the discussion, we just abbreviate this as $\lesssim 1$. This follows from Lemma 5.7, which controls the transverse intersections between a tube and an algebraic variety. Let T be the cylinder with the same center as $T_{\tilde\theta,\tilde v}$ and with radius $r=R^{1/2+\delta_m}=\varrho^{1/2+\delta_l}$, and let $\alpha=\varrho^{-1/2+\delta_l}$. Let ℓ denote the central axis of T and recall that $Y_{>\alpha}$ is the set $\{y\in Y:\mathrm{Angle}(T_yY,\ell)>\alpha\}$. If $(\theta,v)\in\mathbb{T}_{j,\mathrm{trans}}$, then $T\cap Y_{>\alpha}\cap 2B_j$ must be non-empty. However, Lemma 5.7 tells us that $T\cap Y_{>\alpha}$ is contained in $\leqslant CD_Y^n$ balls of radius $\lesssim r\alpha^{-1}\sim\varrho$. Here, Y is defined by polynomials of degree $D_Y\leqslant D(\varepsilon,D_Z)\lesssim 1$. Therefore, $(\theta,v)\in\mathbb{T}_{j,\mathrm{trans}}$ for $\lesssim 1$ values of j. Plugging this into the last equation, we get

$$\sum_j \|f_{j,\mathrm{trans}}\|_{L^2}^2 \lesssim \|f\|_{L^2}^2. \tag{8.18}$$

Next, we would like to study $Ef_{j,\mathrm{trans}}$ on each ball B_j, by doing induction on the radius. In order to do so, we redo the wave packet decomposition at a scale appropriate to B_j, as in §7. For brevity, during this discussion, we let $g=f_{j,\mathrm{trans}}$. We see that

$$\tilde g = \sum_{\tilde\theta,\tilde v} \tilde g_{\tilde\theta,\tilde v} + \mathrm{RapDec}(R)\|f\|_{L^2}.$$

We recall a couple definitions from §7. For any $b\in B_{R^{1/2+\delta_m}}$, we define

$$\widetilde{\mathbb{T}}_{Z+b} := \{(\tilde\theta,\tilde v) : T_{\tilde\theta,\tilde v} \text{ is tangent to } Z+b \text{ in } B_j\},$$
$$\tilde g_b := \sum_{(\tilde\theta,\tilde v)\in\widetilde{\mathbb{T}}_{Z+b}} \tilde g_{\tilde\theta,\tilde v}.$$

For each b, $\tilde g_b$ is concentrated in wave packets tangent to $Z+b$ in the ball B_j, and so we will be able to apply induction on the radius to study $E\tilde g_b$. By (7.12), if y_j is the center of B_j and $x=y_j+\tilde x\in B_j$, then

$$|E\tilde g_b(\tilde x)| \sim \chi_{N_{\varrho^{1/2+\delta_m}}(Z+b)}(x)|Eg(x)|. \tag{8.19}$$

We define $f_{j,\mathrm{trans},b}$ so that $(f_{j,\mathrm{trans},b})^\sim=\tilde g_b$ (in other words, $f_{j,\mathrm{trans},b}=e^{-i\psi_y(\omega)}\tilde g_b$). In this language, the last equation becomes

$$|Ef_{j,\mathrm{trans},b}(x)| \sim \chi_{N_{\varrho^{1/2+\delta_m}}(Z+b)}(x)|Ef_{j,\mathrm{trans}}(x)|. \tag{8.20}$$

Next, we choose a set of vectors $b \in B_{R^{1/2+\delta_m}}$. The number of vectors b that we choose is related to the geometry of Z. We cover $N_{R^{1/2+\delta_m}}(Z) \cap B_j$ with disjoint balls of radius $R^{1/2+\delta_m}$. We dyadically pigeonhole the volume of $B \cap N_{\varrho^{1/2+\delta_m}}(Z)$: for each s, we consider

$$\mathfrak{B}_s := \{B(x_0, R^{1/2+\delta_m}) \subset N_{R^{1/2+\delta_m}}(Z) \cap B_j : |B(x_0, R^{1/2+\delta_m}) \cap N_{\varrho^{1/2+\delta_m}}(Z)| \sim 2^s\}.$$

We select a value of s such that

$$\|Ef_{j,\mathrm{trans}}\|_{\mathrm{BL}^p_{k,A/2}(\bigcup_{B \in \mathfrak{B}_s} B)} \gtrsim \frac{1}{\log R} \|Ef_{j,\mathrm{trans}}\|_{\mathrm{BL}^p_{k,A/2}(B_j)}.$$

Next, we prune $\mathbb{T}_{j,\mathrm{trans}}$ a little: we include (θ, v) only if $T_{\theta,v}$ intersects one of the balls of $\mathfrak{B}_s$. To avoid making the notation even heavier, we do not make a separate notation for the pruned set. This pruning can only decrease $\|f_{j,\mathrm{trans}}\|_{L^2}$, and it changes $\|Ef_{j,\mathrm{trans}}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)}$ by at most a factor of $\log R$.

Now, we are ready to choose our set of translations $\{b\}$. We choose a random set of $|B_{R^{1/2+\delta_m}}|/2^s$ vectors $b \in B_{R^{1/2+\delta_m}}$. For a typical ball $B(x_0, R^{1/2+\delta_m}) \in \mathfrak{B}_s$, the union $\bigcup_b N_{\varrho^{1/2+\delta_m}}(Z+b)$ covers a definite fraction of the ball (in a random way). Therefore, with high probability, we get

$$\|Ef_{j,\mathrm{trans}}\|^p_{\mathrm{BL}^p_{k,A/2}(B_j)} \lesssim \sum_b \|Ef_{j,\mathrm{trans},b}\|^p_{\mathrm{BL}^p_{k,A/2}(N_{\varrho^{1/2+\delta_m}}(Z+b) \cap B_j)}. \tag{8.21}$$

On the other hand, a typical point of $B(x_0, R^{1/2+\delta_m})$ lies in $\lesssim 1$ of the sets

$$N_{\varrho^{1/2+\delta_m}}(Z+b).$$

Using this geometric fact, we will show that

$$\sum_b \|\tilde{g}_b\|^2_{L^2} \lesssim \|g\|^2_{L^2}. \tag{8.22}$$

To see (8.22), we decompose $g = \sum_{\tilde{\theta},w} g_{\tilde{\theta},w}$ as in §7. If $g_{\tilde{\theta},w}$ is not negligible, then $T_{\tilde{\theta},w}$ must intersect one of the balls $B(x_0, R^{1/2+\delta_m}) \in \mathfrak{B}_s$. Since the sets

$$N_{\varrho^{1/2+\delta_m}}(Z+b) \cap B(x_0, R^{1/2+\delta_m})$$

are essentially disjoint, Lemma 7.4 tells us that

$$\sum_b \|(g_{\tilde{\theta},w})\tilde{}_b\|^2_{L^2} \lesssim \|g_{\tilde{\theta},w}\|^2_{L^2}.$$

But, as we saw in the proof of Lemma 7.6,

$$\tilde{g}_b = \sum_{\tilde{\theta},w} (g_{\tilde{\theta},w})_b^{\sim}$$

is an orthogonal decomposition, and $g = \sum_{\tilde{\theta},w} g_{\tilde{\theta},w}$ is an orthogonal decomposition, and so

$$\sum_b \|\tilde{g}_b\|_{L^2}^2 \sim \sum_{b,\tilde{\theta},w} \|(g_{\tilde{\theta},w})_b^{\sim}\|_{L^2}^2 \lesssim \sum_{\tilde{\theta},w} \|g_{\tilde{\theta},w}\|_{L^2}^2 \sim \|g\|_{L^2}^2.$$

We now have all the estimates that we need in the transverse case, and we collect them here.

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}^p \lesssim \log R \sum_{j,b} \|Ef_{j,\mathrm{trans},b}\|_{\mathrm{BL}^p_{k,A/2}(B_j)}^p, \tag{8.23}$$

$$\sum_{j,b} \|f_{j,\mathrm{trans},b}\|_{L^2}^2 \lesssim \|f\|_{L^2}^2. \tag{8.24}$$

By Lemma 7.6,

$$\max_b \|f_{j,\mathrm{trans},b}\|_{L^2}^2 \leqslant R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|f_{j,\mathrm{trans}}\|_{L^2}^2. \tag{8.25}$$

(For later reference, we also record here: for each j, b and $\tilde{\theta}$,

$$\|f_{j,\mathrm{trans},b}\|_{L^2(\tilde{\theta})}^2 \lesssim R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)} \|f\|_{L^2(2\tilde{\theta})}^2. \tag{8.26}$$

This follows from Lemma 7.6. This inequality does not appear in the proof of Theorem 1.5, but it could be useful in some later refinements.)

By induction on the radius, we know that

$$\begin{aligned}\|Ef_{j,\mathrm{trans},b}\|_{\mathrm{BL}^p_{k,A/2}(B_j)} &\leqslant C(K,\varepsilon,m,D_Z)\varrho^{m\varepsilon}\varrho^{\delta(\log \bar{A} - \log(A/2))}\varrho^{-e+1/2}\|f_{j,\mathrm{trans},b}\|_{L^2} \\ &\leqslant C(K,\varepsilon,m,D_Z)R^{\delta}\varrho^{m\varepsilon}R^{\delta(\log \bar{A} - \log A)}\varrho^{-e+1/2}\|f_{j,\mathrm{trans},b}\|_{L^2}.\end{aligned}$$

Using these estimates, we can now prove (8.1) by induction on the radius.

$$\begin{aligned}\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}^p &\lesssim (\log R) \sum_{j,b} \|Ef_{j,\mathrm{trans},b}\|_{\mathrm{BL}^p_{k,A/2}(B_j)}^p \\ &\leqslant R^{O(\delta)}(C(K,\varepsilon,m,D_Z)\varrho^{m\varepsilon}R^{\delta(\log \bar{A} - \log(A/2))}\varrho^{-e+1/2})^p \sum_{j,b} \|f_{j,\mathrm{trans},b}\|_{L^2}^p.\end{aligned}$$

Using (8.25) and (8.24),

$$\sum_{j,b} \|f_{j,\mathrm{trans},b}\|_{L^2}^p \lesssim R^{O(\delta_m)} \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)(p/2-1)} \|f\|_{L^2}^p.$$

Now, plugging in, we get

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)}$$
$$\lesssim R^{O(\delta_m)} (C(K,\varepsilon,m,D_Z)\varrho^{m\varepsilon} R^{\delta(\log \bar{A} - \log A)} \varrho^{-e+1/2})^p \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)(p/2-1)} \|f\|_{L^2}^p.$$

At the exponent $p = \bar{p}(k,m)$,

$$(\varrho^{-e+1/2})^p \left(\frac{R^{1/2}}{\varrho^{1/2}}\right)^{-(n-m)(p/2-1)} = (R^{-e+1/2})^p.$$

(If $m=k$, so that $p=\bar{p}(k,m)+\delta$, then this is true up to a factor $R^{O(\delta)}$.) So, plugging in our values of p and e, and multiplying out, we get

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \leqslant C R^{O(\delta_m)} \left(\frac{R}{\varrho}\right)^{-m\varepsilon} (C(K,\varepsilon,m,D_Z) R^{m\varepsilon} R^{\delta(\log \bar{A} - \log A)} R^{-e+1/2})^p \|f\|_{L^2}^p.$$

The constant C on the right-hand side depends on ε, D_Z and the dimension n. We have to check that $CR^{O(\delta_m)}(R/\varrho)^{-m\varepsilon} \leqslant 1$. Note that $R/\varrho = R^{\theta(\delta_l)}$. We choose the δ's such that $\delta_m \ll \varepsilon \delta_{m-1}$, and so $(R/\varrho)^{-m\varepsilon}$ dominates the other terms. Therefore, the induction closes in the transverse algebraic case also.

This finishes the proof of Theorem 1.5.

9. Going from k-broad estimates to regular estimates

The paper [BG] introduces a technique to go from multilinear estimates to regular L^p estimates. In this section, we follow this technique to go from k-broad estimates to regular L^p estimates.

Proposition 9.1. *Suppose that, for all K and ε, the operator E obeys the k-broad inequality*:

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim_{K,\varepsilon} R^{\varepsilon} \|f\|_{L^q}. \tag{9.1}$$

(*Here the quantities k, A, p and q are fixed, and the inequality holds for all R.*)

If $p \leqslant q \leqslant \infty$ and p is in the range

$$2\frac{2n-k+2}{2n-k} \leqslant p \leqslant 2\frac{k-1}{k-2}, \tag{9.2}$$

then E obeys

$$\|Ef\|_{L^p(B_R)} \lesssim_{\varepsilon} R^{\varepsilon} \|f\|_{L^q}. \tag{9.3}$$

Remarks. The lower bound on p is important. The upper bound is less important, and it could probably be improved.

Theorem 1.5 together with Proposition 9.1 implies Theorem 1.1. If n is even, we use $k=\frac{1}{2}n+1$. By Theorem 1.5, we have $\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}\lesssim R^\varepsilon\|f\|_{L^2}\lesssim R^\varepsilon\|f\|_{L^p}$ for p slightly bigger than

$$\bar{p}(k,n)=2\frac{n+k}{n+k-2}.$$

With our choice of k, we also have

$$\bar{p}(k,n)=2\frac{2n-k+2}{2n-k}.$$

Applying Proposition 9.1, we get $\|Ef\|_{L^p(B_R)}\lesssim R^\varepsilon\|f\|_{L^p}$ for p slightly bigger than $\bar{p}(k,n)$. Interpolating with the trivial L^∞ bound gives this estimate for all $p>\bar{p}(k,n)$. Finally, applying ε-removal ([T1]) gives Theorem 1.1. If n is odd, we use $k=\frac{1}{2}(n+1)$. The argument is the same (but in this case, $\bar{p}(k,n)>2(2n-k+2)/(2n-k)$).

Proof. By hypothesis, we have an inequality of the form

$$\sum_{B_{K^2}\subset B_R}\min_{V_1,\dots,V_A}\max_{\tau\notin V_a}\int_{B_{K^2}}|Ef_\tau|^p\leqslant C(K,\varepsilon)R^{p\varepsilon}\|f\|_{L^q}^p. \tag{9.4}$$

We recall that here $V_1,\dots,V_A$ are $(k-1)$-planes, and we write $\tau\notin V_a$ as an abbreviation for $\mathrm{Angle}(G(\tau),V_a)>K^{-1}$.

For each B_{K^2}, we fix a choice of $V_1,\dots,V_A$ achieving the minimum above. Then, we can write

$$\int_{B_{K^2}}|Ef|^p\lesssim K^{O(1)}\max_{\tau\notin V_a}\int_{B_{K^2}}|Ef_\tau|^p+\sum_{a=1}^A\int_{B_{K^2}}\Big|\sum_{\tau\in V_a}Ef_\tau\Big|^p. \tag{9.5}$$

The first term is the "broad" part, and it can be controlled by the k-broad estimate. We handle the second term, the "narrow" part, by a decoupling-type argument. We work with B_{K^2} so that we can cleanly apply the decoupling theorem from [Bo4]. (The paper [BG] contains a different but closely related argument.)

THEOREM 9.2. ([Bo4]) *Let $g:\mathbb{R}^m\to\mathbb{C}$ be such that $\hat{g}$ is supported in the K^{-2}-neighborhood of the truncated paraboloid. Divide this neighborhood into slabs τ with $m-1$ long directions of length K^{-1} and one short direction of length K^{-2}. Write $g=\sum_\tau g_\tau$, where $\hat{g}_\tau=\chi_\tau\hat{g}$. Then, on any ball of radius K^2, for $2\leqslant p\leqslant 2m/(m-1)$,*

$$\|g\|_{L^p(B_{K^2})}\lesssim_\delta K^\delta\Big(\sum_\tau\|g_\tau\|^2_{L^p(W_{B_{K^2}})}\Big)^{1/2}, \tag{9.6}$$

where $W_{B_{K^2}}$ is a weight measure, approximately the volume measure on B_{K^2} and rapidly decaying.

Applying this decoupling estimate with $m=k-1$, we get the following lemma.

LEMMA 9.3. *We have*

$$\Big(\int_{B_{K^2}}\Big|\sum_{\tau\in V_a}Ef_\tau\Big|^p\Big)^{1/p}\lesssim_\delta K^\delta\Big(\sum_{\tau\in V_a}\Big(\int W_{B_{K^2}}|Ef_\tau|^p\Big)^{2/p}\Big)^{1/2}.$$

Proof. On B_{K^2}, we use coordinates (u,v), where v is parallel to V_a and u is perpendicular to V_a. We write B_{u,K^2} for a ball of radius K^2 in the u coordinates and B_{v,K^2} for a ball of radius K^2 in the v coordinates. If we restrict Ef_τ to the $(k-1)$-plane $\{u\}\times\mathbb{R}^{k-1}$ (parallel to V_a), then its Fourier transform is supported in the K^{-2}-neighborhood of a cap τ' in the K^{-2}-neighborhood of a paraboloid. By Theorem 9.2, we get

$$\Big\|\sum_{\tau\in V_a}Ef_\tau\Big\|_{L^p(\{u\}\times B_{v,K^2})}\lesssim_\delta K^\delta\Big(\sum_{\tau\in V_a}\|Ef_\tau\|^2_{L^p(\{u\}\times W_{B_{v,K^2}})}\Big)^{1/2}.$$

Using Fubini's and Minkowski's theorems, we then get

$$\Big\|\sum_{\tau\in V_a}Ef_\tau\Big\|_{L^p(B_{u,K^2}\times B_{v,K^2})}\lesssim_\delta K^\delta\Big(\sum_{\tau\in V_a}\|Ef_\tau\|^2_{L^p(W_{B_{K^2}})}\Big)^{1/2}.\qquad\square$$

The number of $\tau\in V_a$ is $\lesssim K^{k-2}$. Applying Hölder's inequality, we see that

$$\int_{B_{K^2}}\Big|\sum_{\tau\in V_a}Ef_\tau\Big|^p\lesssim_\delta K^\delta K^{(k-2)(p/2-1)}\sum_{\tau\in V_a}\int W_{B_{K^2}}|Ef_\tau|^p.\tag{9.7}$$

At this point, we have gotten as much as we can from the knowledge that $\tau\in V_a$, and we relax this estimate to

$$\int_{B_{K^2}}\Big|\sum_{\tau\in V_a}Ef_\tau\Big|^p\lesssim_\delta K^\delta K^{(k-2)(p/2-1)}\sum_{\tau}\int W_{B_{K^2}}|Ef_\tau|^p.\tag{9.8}$$

Next, we sum this inequality over all $a=1,...,A$ and over all $B_{K^2}\subset B_R$. We let $W=\sum_{B_{K^2}\subset B_R}W_{B_{K^2}}$. We see that

$$\sum_{B_{K^2}\subset B_R}\sum_{a=1}^A\int_{B_{K^2}}\Big|\sum_{\tau\in V_a}Ef_\tau\Big|^p\lesssim K^\delta K^{(k-2)(p/2-1)}\sum_{\tau}\int W|Ef_\tau|^p.\tag{9.9}$$

We note that $W\lesssim 1$ on B_{2R} and $W\leqslant\mathrm{RapDec}(R)$ outside B_{2R}. Therefore, we get

$$\sum_{B_{K^2}\subset B_R}\sum_{a=1}^A\int_{B_{K^2}}\Big|\sum_{\tau\in V_a}Ef_\tau\Big|^p\lesssim K^\delta K^{(k-2)(p/2-1)}\sum_{\tau}\int_{B_{2R}}|Ef_\tau|^p+\mathrm{RapDec}(R)\|f\|_{L^q}^p.\tag{9.10}$$

Combining this estimate for the narrow part with our estimate for the broad part, we have

$$\int_{B_R} |Ef|^p \leqslant C(K,\varepsilon) R^{p\varepsilon} \|f\|_{L^q}^p + CK^{\delta} K^{(k-2)(p/2-1)} \sum_{\tau} \int_{B_{2R}} |Ef_\tau|^p. \tag{9.11}$$

With this inequality in hand, we will prove (9.3) by induction on the radius, where we use the induction assumption in order to handle the contribution of the f_τ terms. The inequality we wish to prove is

$$\int_{B_R} |Ef|^p \leqslant \overline{C}(\varepsilon) R^{p\varepsilon} \|f\|_{L^q}^p. \tag{9.12}$$

By induction on the radius, we may assume that (9.12) holds for radii less than $\frac{1}{2}R$. We use this induction and parabolic rescaling to handle the contribution of each f_τ.

On the ball τ we introduce new coordinates. Let ω_τ be the center of τ, and recall that the radius of τ is K^{-1}. Then, we introduce a new coordinate $\widetilde{\omega} \in B^{n-1}$ by

$$\widetilde{\omega} = K(\omega - \omega_\tau). \tag{9.13}$$

We rewrite the phase in these coordinates:

$$x_1\omega_1 + \ldots + x_{n-1}\omega_{n-1} + x_n|\omega_n|^2 = \mathrm{Fcn}(x) + \tilde{x}_1\widetilde{\omega}_1 + \ldots + \tilde{x}_{n-1}\widetilde{\omega}_{n-1} + \tilde{x}_n|\widetilde{\omega}|^2,$$

where $\mathrm{Fcn}(x)$ denotes a function of x only,

$$\tilde{x}_j = K^{-1}(x_j + 2\omega_{\tau,j}x_n) \text{ for } 1 \leqslant j \leqslant n-1 \quad \text{and} \quad \tilde{x}_n = K^{-2}x_n.$$

Here, $\omega_{\tau,j}$ denotes the jth coordinate of ω_τ. Note that the linear transformation $x \mapsto \tilde{x}$ sends B_R into $B_{CRK^{-1}}$ and has Jacobian $\sim K^{-(n+1)}$.

We define

$$\tilde{f}_\tau(\widetilde{\omega}) = f_\tau(\omega) = f_\tau(K^{-1}\widetilde{\omega} + \omega_\tau),$$

so that

$$|Ef_\tau(x)| = K^{-(n-1)} |E\tilde{f}_\tau(\tilde{x})|.$$

By induction on the radius, we may assume that (9.12) holds for $\tilde{f}_\tau$ on a ball of radius CRK^{-1}:

$$\int_{B_{CRK^{-1}}} |E\tilde{f}_\tau|^p \lesssim \overline{C}(\varepsilon) R^{p\varepsilon} K^{-p\varepsilon} \|\tilde{f}_\tau\|_{L^q}^p.$$

By change of variables, we have

$$\begin{aligned}\int_{B_{2R}} |Ef_\tau|^p &\lesssim K^{(n+1)} K^{-(n-1)p} \int_{B_{CRK^{-1}}} |E\tilde{f}_\tau|^p \\ &\lesssim \overline{C}(\varepsilon)\left(\frac{R}{K}\right)^{p\varepsilon} K^{(n+1)-(n-1)p} \|\tilde{f}_\tau\|_{L^q}^p \\ &= \overline{C}(\varepsilon) R^{p\varepsilon} K^{(n+1)-(n-1)p-p\varepsilon} K^{(n-1)p/q} \|f_\tau\|_{L^q}^p.\end{aligned}$$

Plugging this bound into (9.11), we get

$$\begin{aligned}&\int_{B_R} |Ef|^p \\ &\leqslant C(K,\varepsilon) R^{p\varepsilon} \|f\|_{L^q}^p + C\overline{C}(\varepsilon) R^{p\varepsilon} K^{(k-2)(p/2-1)+(n+1)-(n-1)p-p\varepsilon+\delta} K^{(n-1)p/q} \sum_\tau \|f_\tau\|_{L^q}^p.\end{aligned} \tag{9.14}$$

There are K^{n-1} different $\tau \subset B^{n-1}$. Since $p \leqslant q$, we may apply Hölder's inequality to see that

$$\sum_\tau \|f_\tau\|_{L^q}^p \leqslant \left(\sum_\tau \|f_\tau\|_{L^q}^q\right)^{p/q} K^{(n-1)(1-p/q)} = \|f\|_{L^q}^p K^{(n-1)(1-p/q)}.$$

Plugging this into the last inequality, we see that the dependence on q drops out, and we are left with

$$\begin{aligned}&\int_{B_R} |Ef|^p \\ &\leqslant C(K,\varepsilon) R^{p\varepsilon} \|f\|_{L^q}^p + C\overline{C}(\varepsilon) R^{p\varepsilon} K^{(k-2)(p/2-1)+(n+1)-(n-1)p+(n-1)-p\varepsilon+\delta} \|f\|_{L^q}^p.\end{aligned} \tag{9.15}$$

We can close the induction as long as the exponent of K is negative. (First we choose K large enough so that the second term is bounded by $\frac{1}{2}\overline{C}(\varepsilon) R^{p\varepsilon} \|f\|_{L^p}^p$. Then, we choose $\overline{C}(\varepsilon)$ sufficiently large so that the first term is also bounded by $\frac{1}{2}\overline{C}(\varepsilon) R^{p\varepsilon} \|f\|_{L^p}^p$.) Given $\varepsilon > 0$, we can choose $\delta < \varepsilon$. So the induction closes as long as

$$(k-2)\left(\tfrac{1}{2}p-1\right) + (n+1) - (n-1)p + (n-1) \leqslant 0.$$

This is equivalent with the lower bound for p in our hypothesis (9.2):

$$2\frac{2n-k+2}{2n-k} \leqslant p.$$

□

Appendix A. Keeping track of parameters

We have several small parameters. In this appendix, we try to provide a reference to help the reader keep track of the parameters. We list all the parameters, how they compare to each other, and where they appear in the argument.

We begin with the small parameters. For each $\varepsilon>0$, there is a sequence of small parameters

$$\delta \ll \delta_{n-1} \ll \delta_{n-2} \ll ... \ll \delta_1 \ll \delta_0 \ll \varepsilon.$$

In this sequence, each parameter is far smaller than the next. For instance, we will use that $\delta_m < \varepsilon \delta_{m-1}$.

The parameter δ appears in the wave packet decomposition. The tubes $T_{\theta,v}$ in the wave packet decomposition have thickness $R^{1/2+\delta}$. The parameter δ_m appears in the m-dimensional case of the main inductive estimate (Proposition 8.1): in this estimate, we suppose that $f = \sum_{\theta,v} f_{\theta,v}$ is concentrated on wave packets that are $R^{-1/2+\delta_m}$-tangent to Z on the ball B_R.

Another geometric parameter that appears is the radius ϱ. In the transverse algebraic case, we decompose B_R into smaller balls B_j. If we are working on tubes that cluster in the neighborhood of an m-dimensional variety Z, and are transverse to an l-dimensional variety Y, then the radius of each B_j is ϱ given by

$$\varrho^{1/2+\delta_l} = R^{1/2+\delta_m}.$$

The quotient R/ϱ has size $R^{\theta(\delta_l)}$, which dominates $R^{O(\delta_m)}$.

Then, there are positive parameters K and A. We have

$$1 \ll A \ll K.$$

We need $A = A(\varepsilon)$ sufficiently large to run the proof of Theorem 1.5, and the broad inequality is most useful when K is much larger than A.

Given ε, we then fix the small parameters δ and the larger parameters A and K. Then, we consider $R \to \infty$. In the statement of Theorem 1.5, the constant depends on ε and on K. By choosing this constant large enough, the theorem holds trivially unless R is very large compared to all these fixed parameters.

Appendix B. Further directions

B.1. Honest k-linear estimates

Our main result (Theorem 1.5) is a weak version of the k-linear restriction estimate from Conjecture 1.4. For some purposes, we have seen that Theorem 1.5 is a good substitute

for a k-linear restriction estimate, but there are surely other situations where an honest k-linear estimate is better. When I tried to prove Conjecture 1.4 using this method, I ran into the following problem. There are k different functions Ef_j to consider. It may happen that for some of these k functions, the wave packets of Ef_j are tangent to a variety Z, and for others of these k functions, the wave packets of Ef_j are transverse to the variety Z. I did not find a good way to deal with this scenario. The k-broad norm, $\mathrm{BL}^p_{k,A}$, is designed to get around this situation.

B.2. Kakeya-type estimates for low degree varieties

We now return to the extension operator for the paraboloid. Theorem 1.5 gives essentially sharp broad L^p estimates of the form

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim R^{\varepsilon}\|f\|_{L^2}. \tag{B.1}$$

We have seen that this estimate holds if and only if $p \geqslant \bar{p}(k,n)$. What if we consider other norms on the right-hand side? For some q larger than 2, can we prove an estimate

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim R^{\varepsilon}\|f\|_{L^q}, \tag{B.2}$$

for some $p < \bar{p}(k,n)$?

In the introduction, we mentioned some sharp examples for Theorem 1.5. The first question to ask ourselves is whether an inequality of the form (B.2) may hold for these examples. In the examples we considered, the wave packets $Ef_{\theta,v}$ concentrate in the $R^{1/2}$-neighborhood of a low degree variety Z. Let us consider the set of caps θ that can appear in such an example. Define $\Theta(Z)$ as

$$\Theta(Z) := \{\theta : T_{\theta,v} \subset N_{R^{1/2+\delta}}(Z) \text{ for some } v\}.$$

In such an example, the function f must be supported in $\bigcup_{\theta\in\Theta(Z)}\theta$. If the volume of this union is much less than 1, then, for $q>2$, $\|f\|_{L^q}$ will be much bigger than $\|f\|_{L^2}$, and so our special class of examples will obey an inequality of the form (B.2).

In fact, if we had good estimates for $|\Theta(Z)|$, then I believe we could input them into the proof of Theorem 1.5 to get some further estimates of the form (B.2), roughly following the argument in [G].

If Z is an m-dimensional plane, then it is easy to check that $|\Theta(Z)| \sim (R^{1/2})^{m-1}$, and so $|\Omega(Z)| \sim (R^{1/2})^{m-1}(R^{1/2})^{-(n-1)}$. It seems reasonable to conjecture that a similar bound holds for any m-dimensional variety Z of small degree.

Conjecture B.1. If Z is an m-dimensional variety in $\mathbb{R}^n$ of degree at most D, then

$$|\Theta(Z)| \leqslant C(n, D, \varepsilon)(R^{1/2})^{m-1+\varepsilon}. \tag{B.3}$$

Conjecture B.1 is a very special case of the Kakeya conjecture. One variant of the Kakeya conjecture goes as follows.

Conjecture B.2. (Kakeya conjecture) Suppose that $X \subset B^n(1)$. Suppose that $T_j \subset X$ are tubes of length 1 and radius δ, pointing in δ-separated directions. Then

$$\text{number of tubes} \leqslant C(\varepsilon)\delta^{-\varepsilon}\frac{\text{Vol}(X)}{\text{Vol(tube)}}.$$

(I have not seen this exact version of the Kakeya conjecture in print before, but it is straightforward to check that the maximal function version of the Kakeya conjecture implies Conjecture B.2, which implies the Minkowski dimension version of the Kakeya conjecture.) Now Conjecture B.1 is just the special case of Conjecture B.2 where the set X is the δ-neighborhood of a low-degree algebraic variety.

Conjecture B.1 also came up in ongoing joint work with Josh Zahl on the Kakeya problem in $\mathbb{R}^4$. I think it is a basic issue that comes up in trying to apply polynomial methods to the restriction problem or to the Kakeya problem.

References

[Be1] Bejenaru, I., The optimal trilinear restriction estimate for a class of hypersurfaces with curvature. *Adv. Math.*, 307 (2017), 1151–1183.

[Be2] — Optimal multilinear restriction estimates for a class of surfaces with curvature. Preprint, 2016. arXiv:1606.02634 [math.CA].

[BCT] Bennett, J., Carbery, A. & Tao, T., On the multilinear restriction and Kakeya conjectures. *Acta Math.*, 196 (2006), 261–302.

[Bo1] Bourgain, J., Besicovitch type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.*, 1 (1991), 147–187.

[Bo4] — Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces. *Israel J. Math.*, 193 (2013), 441–458.

[BD] Bourgain, J. & Demeter, C., The proof of the l^2 decoupling conjecture. *Ann. of Math.*, 182 (2015), 351–389.

[BG] Bourgain, J. & Guth, L., Bounds on oscillatory integral operators based on multilinear estimates. *Geom. Funct. Anal.*, 21 (2011), 1239–1295.

[CKW] Chen, X., Kayal, N. & Wigderson, A., Partial derivatives in arithmetic complexity and beyond. *Found. Trends Theor. Comput. Sci.*, 6 (2010), 1–138.

[GP] Guillemin, V. & Pollack, A., *Differential Topology*. AMS Chelsea Publishing, Providence, RI, 2010.

[G] Guth, L., A restriction estimate using polynomial partitioning. *J. Amer. Math. Soc.*, 29 (2016), 371–413.

[GHI] Guth, L., Hickman, J. & Iliopoulou, M., Sharp estimates for oscillatory integral operators via polynomial partitioning. Preprint, 2017. arXiv:1710.10349 [math.CA].

[GK] Guth, L. & Katz, N. H., On the Erdős distinct distances problem in the plane. *Ann. of Math.*, 181 (2015), 155–190.

[H] Hörmander, L., Oscillatory integrals and multipliers on FL^p. *Ark. Mat.*, 11 (1973), 1–11.

[L] Lee, S., Linear and bilinear estimates for oscillatory integral operators related to restriction to hypersurfaces. *J. Funct. Anal.*, 241 (2006), 56–98.

[OW] Ou, Y. & Wang, H., A cone restriction estimate using polynomial partitioning. Preprint, 2017. arXiv:1704.05485 [math.CA].

[S1] Stein, E. M., Some problems in harmonic analysis, in *Harmonic Analysis in Euclidean Spaces* (Williamstown, MA, 1978), Part 1, Proc. Sympos. Pure Math., pp. XXXV, 3–20. Amer. Math. Soc., Providence, RI, 1979.

[ST] Stone, A. H. & Tukey, J. W., Generalized "sandwich" theorems. *Duke Math. J.*, 9 (1942), 356–359.

[T1] Tao, T., The Bochner–Riesz conjecture implies the restriction conjecture. *Duke Math. J.*, 96 (1999), 363–375.

[T2] — A sharp bilinear restrictions estimate for paraboloids. *Geom. Funct. Anal.*, 13 (2003), 1359–1384.

[TVV] Tao, T., Vargas, A. & Vega, L., A bilinear approach to the restriction and Kakeya conjectures. *J. Amer. Math. Soc.*, 11 (1998), 967–1000.

[Wi] Wisewell, L., Kakeya sets of curves. *Geom. Funct. Anal.*, 15 (2005), 1319–1362.

[Wol3] Wolff, T., A sharp bilinear cone restriction estimate. *Ann. of Math.*, 153 (2001), 661–698.

[Won] Wongkew, R., Volumes of tubular neighbourhoods of real algebraic varieties. *Pacific J. Math.*, 159 (1993), 177–184.

Larry Guth
Massachusetts Institute of Technology
Department of Mathematics
77 Massachusetts Avenue
Cambridge, MA 02139-4307
U.S.A.
lguth@math.mit-edu

Received November 6, 2017
Received in revised form April 11, 2018