Sharp estimates for oscillatory integral operators via polynomial partitioning

by

LARRY GUTH

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

JONATHAN HICKMAN

University of Chicago
Chicago, IL, U.S.A.

MARINA ILIOPOULOU

University of California, Berkeley
Berkeley, CA, U.S.A.

1. Introduction

1.1. Statement of results

Let B^d denote the unit ball in $\mathbb{R}^d$ and $\Sigma \colon B^{n-1} \to \mathbb{R}^n$ be a smooth[1] parametrisation of a hypersurface. Further, let $a \in C_c^\infty(\mathbb{R}^{n-1})$ be non-negative and supported in B^{n-1} and suppose that Σ has non-vanishing Gaussian curvature on the support $\operatorname{supp} a$ of a. Analytically, this means that Σ satisfies the following conditions:

(E1) $\operatorname{rank} \partial_\omega \Sigma(\omega) = n-1$ for all $\omega \in B^{n-1}$;

(E2) defining the Gauss map $G \colon B^{n-1} \to S^{n-1}$ by

$$G(\omega) := \frac{G_0(\omega)}{|G_0(\omega)|},$$

where

$$G_0(\omega) := \bigwedge_{j=1}^{n-1} \partial_{\omega_j} \Sigma(\omega),$$

the curvature condition

$$\det \partial_{\omega\omega}^2 \langle \Sigma(\omega), G(\omega_0) \rangle|_{\omega=\omega_0} \neq 0$$

holds for all $\omega_0 \in \operatorname{supp} a$.

---

[1] In view of the methods of the present article it is convenient to work in the C^∞ category, but the forthcoming definitions and questions certainly make sense at lower levels of regularity (in particular, in the C^2 class).

Here, the wedge product of $n-1$ vectors in $\mathbb{R}^n$ is identified with a vector in the usual manner.

A central problem in harmonic analysis is to understand the Lebesgue space mapping properties of the extension operator E associated with such a parameterised hypersurface. This operator is defined by the formula

$$Ef(x) := \int_{B^{n-1}} e^{2\pi i \langle x, \Sigma(\omega) \rangle} a(\omega) f(\omega) \, d\omega \tag{1.1}$$

for all integrable $f \colon B^{n-1} \to \mathbb{C}$. Thus, E is an oscillatory integral operator with associated phase function

$$\phi(x; \omega) := \langle x, \Sigma(\omega) \rangle. \tag{1.2}$$

Observe that the parametrisation Σ can be recovered from the phase by differentiation; that is,

$$\partial_x \phi(x; \omega) = \Sigma(\omega).$$

Typically, one is interested in proving local estimates for (1.1) of the form[2]

$$\|Ef\|_{L^p(B(0, \lambda))} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^{n-1})}; \tag{1.3}$$

here, the left-hand norm has been localised to a ball of radius $\lambda \geqslant 1$ and the right-hand constant is allowed some weak dependence on λ. In particular, the Fourier restriction conjecture asserts that (1.3) should hold for any $\varepsilon > 0$ in the range $p \geqslant 2n/(n-1)$.

In this article the natural variable coefficient generalisations of such extension operators (1.1) and estimates (1.3) are studied. In particular, here more general oscillatory integral operators are considered whose associated phase function $\phi(x; \omega)$ shares the property of the extension operator that for each x the map $\omega \mapsto \partial_x \phi(x; \omega)$ parameterises a hypersurface of non-vanishing Gaussian curvature. Crucially, however, the choice of hypersurface is now allowed to smoothly vary with x.

To formalise this discussion, let $n \geqslant 2$, $a \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})$ be non-negative and supported in $B^n \times B^{n-1}$ and $\phi \colon B^n \times B^{n-1} \to \mathbb{R}$ be a smooth function which satisfies the following conditions:

(H1) $\operatorname{rank} \partial_{\omega x}^2 \phi(x; \omega) = n-1$ for all $(x; \omega) \in B^n \times B^{n-1}$;

(H2) defining the map $G \colon B^n \times B^{n-1} \to S^{n-1}$ by

$$G(x; \omega) := \frac{G_0(x; \omega)}{|G_0(x; \omega)|},$$

---

[2] Given a (possibly empty) list of objects L, for real numbers $A_p, B_p \geqslant 0$ depending on some Lebesgue exponent p, the notation $A_p \lesssim_L B_p$ or $B_p \gtrsim_L A_p$ signifies that $A_p \leqslant C B_p$ for some constant $C = C_{L,n,p} \geqslant 0$ depending on the objects in the list, n and p. In addition, $A_p \sim_L B_p$ is used to signify that $A_p \lesssim_L B_p$ and $A_p \gtrsim_L B_p$.

where

$$G_0(x;\omega) := \bigwedge_{j=1}^{n-1} \partial_{\omega_j}\partial_x\phi(x;\omega),$$

the curvature condition

$$\det \partial^2_{\omega\omega}\langle\partial_x\phi(x;\omega), G(x;\omega_0)\rangle|_{\omega=\omega_0} \neq 0$$

holds for all $(x;\omega_0) \in \operatorname{supp} a$.

Clearly, (H1) and (H2) agree with (E1) and (E2) when one restricts to phases of the form (1.2), and this definition therefore leads to a generalisation of the operator E introduced above.

Suppose ϕ satisfies (H1) and (H2), for any $\lambda \geqslant 1$ let

$$a^\lambda(x;\omega) := a\Big(\frac{x}{\lambda};\omega\Big) \quad \text{and} \quad \phi^\lambda(x;\omega) := \lambda\phi\Big(\frac{x}{\lambda};\omega\Big),$$

and define the operator T^λ by

$$T^\lambda f(x) := \int_{B^{n-1}} e^{2\pi i\phi^\lambda(x;\omega)} a^\lambda(x;\omega) f(\omega)\, d\omega$$

for all integrable $f : B^{n-1} \to \mathbb{C}$. In this case, T^λ is said to be a *Hörmander-type operator*. Note that the spatial localisation featured in (1.3) is now built into the operator.

THEOREM 1.1. (Stein [27], Bourgain–Guth [9]) *Suppose T^λ is a Hörmander-type operator. For all $\varepsilon > 0$ the estimate*

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_{\varepsilon,\phi,a} \lambda^\varepsilon \|f\|_{L^p(B^{n-1})} \tag{1.4}$$

holds uniformly for $\lambda \geqslant 1$ whenever p satisfies

$$p \geqslant 2\frac{n+1}{n-1} \quad \text{if } n \text{ is odd,}$$
$$p \geqslant 2\frac{n+2}{n} \quad \text{if } n \text{ is even.}$$

The odd-dimensional case is due to Stein [27], who in fact showed that the above estimates are valid for $p \geqslant 2(n+1)/(n-1)$ in all dimensions without the λ^ε-loss. The strengthened results in even dimensions were established much later by Bourgain and the first author [9].([3]) A detailed history of this problem is provided later in the introduction. It is remarked that Theorem 1.1 is sharp, in the sense that there are examples

---

([3]) Strictly speaking, in [9] weaker L^∞-L^p bounds are proven, but the methods can be used to establish the L^p-L^p strengthening: see, for instance, [14, §9] where the L^p-L^p argument appears (although in a slightly disguised form).

of Hörmander-type operators for which (1.4) fails whenever p does not satisfy (1.5). Such examples originate from work of Bourgain [5] and are discussed in detail in §1.3 and §2.

The majority of this work concerns the case where the phase satisfies a strengthened version of (H2), namely

(H2$^+$) for all $(x; \omega_0) \in \operatorname{supp} a$ the matrix

$$\partial^2_{\omega\omega} \langle \partial_x \phi(x; \omega), G(x; \omega_0) \rangle |_{\omega = \omega_0}$$

is positive-definite.

If ϕ satisfies (H1) and (H2$^+$), then T^λ is said to be a *Hörmander-type operator with positive-definite phase*. Geometrically, this condition implies that the principal curvatures of the hypersurface parameterised by $\omega \mapsto \partial_x \phi(x; \omega)$ are everywhere positive. A hypersurface satisfying this condition is said to be *positively-curved*.

Lee [19] observed that for positive-definite phases one may prove estimates beyond the range of Theorem 1.1.([4]) The main result of this article provides sharp estimates in this setting.

THEOREM 1.2. *Suppose T^λ is a Hörmander-type operator with positive-definite phase. For all $\varepsilon > 0$ the estimate*

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_{\varepsilon, \phi, a} \lambda^\varepsilon \|f\|_{L^p(B^{n-1})} \tag{1.5}$$

holds for all $\lambda \geq 1$ whenever p satisfies

$$p \geq 2\frac{3n+1}{3n-3} \quad \text{if } n \text{ is odd,}$$
$$p \geq 2\frac{3n+2}{3n-2} \quad \text{if } n \text{ is even.}$$

This result improves upon the previous best results of Lee [19] and Bourgain and the first author [9]. Moreover, it is sharp in the sense that there are Hörmander-type operators with positive-definite phase for which (1.5) fails whenever p does not satisfy (1.6). Examples of this kind appear, for instance, in [9], [20] and are discussed in detail in §1.3 and §2. It is remarked that range of p obtained in [19], [9] agrees with Theorem 1.2 for $n=3$ and therefore the sharp result in this case is due to Lee [19] and Bourgain and the first author [9]. In all higher dimensions (1.6) is a strictly larger range of p than what was previously known. Finally, away from the endpoint values, one may apply ε-removal techniques to establish (1.5) without the λ^ε-loss in the constant (see §12 below).

([4]) In particular, Lee [19] proved that for positive-definite phases (1.4) holds for $p \geq 2(n+2)/n$ in *all* dimensions, extending the range in Theorem 1.1 when n is odd.

1.2. Applications to the Bochner–Riesz problem

By the standard Carleson–Sjölin reduction [10] (see also [17]), Theorem 1.2 implies new L^p-bounds on Bochner–Riesz multipliers in dimensions $n \geq 4$. Theorem 1.2 also implies bounds for the oscillatory integral operators of Minicozzi–Sogge [20] which arise in relation to the Bochner–Riesz problem on compact manifolds (see [25, Chapter 5]). Moreover, for certain choices of manifold, these estimates are sharp. In this section these well-known applications are briefly reviewed.

Euclidean Bochner–Riesz

For $\alpha \geq 0$ the *Bochner–Riesz multiplier of order* α is the function

$$m^\alpha(\xi) := (1 - |\xi|^2)_+^\alpha$$

where $(t)_+ := t$ if $t > 0$ and zero otherwise.

COROLLARY 1.3. *If* p *satisfies the condition in* (1.6), *then*

$$\|m^\alpha(D)f\|_{L^p(\mathbb{R}^n)} \lesssim_\alpha \|f\|_{L^p(\mathbb{R}^n)} \quad \text{for } \alpha > \alpha(p) := \max\left\{ n\left|\frac{1}{2} - \frac{1}{p}\right| - \frac{1}{2}, 0 \right\}. \tag{1.6}$$

Here, $m^\alpha(D)$ is the multiplier operator associated with m^α, defined a priori by

$$m^\alpha(D)f(x) := \int_{\widehat{\mathbb{R}}^n} e^{2\pi i \langle x, \xi \rangle} m^\alpha(\xi) \hat{f}(\xi) \, d\xi,$$

where $\hat{f}$ denotes the Fourier transform of f.

Recall that the Bochner–Riesz conjecture asserts that condition (1.6) holds for all $p \geq 2n/(n-1)$.[5] The conjecture was resolved for $n = 2$ by Carleson–Sjölin [10] but remains open in all higher dimensions. Corollary 1.3 provides some progress towards this conjecture when $n \geq 4$, improving over earlier results of Fefferman–Stein (see [12]), Lee [18], Bourgain–Guth [9] and others. When $n = 3$, the range in Corollary 1.3 matches that of Lee [19] and Bourgain–Guth [9].

For completeness, here a sketch is provided to show how one may deduce Corollary 1.3 from Theorem 1.2. This follows a standard argument by Carleson and Sjölin [10].

A stationary phase computation shows that the kernel $K^\alpha := (m^\alpha)^\vee$ of $m^\alpha(D)$ is given by

$$K^\alpha(x) = \sum_\pm \frac{a_\pm(x) e^{\pm 2\pi i |x|}}{(1 + |x|)^{(n+1)/2 + \alpha}},$$

---

[5] Once (1.6) is known in the range $p \geq 2n/(n-1)$, it immediately extends to all $1 \leq p \leq \infty$, by duality and interpolation with the trivial $p = 2$ case.

where the $a_\pm$ are symbols of order zero, in the sense that $|\partial_x^\beta a_\pm(x)| \lesssim_\beta |x|^{-|\beta|}$ for all multi-indices $\beta \in \mathbb{N}_0^n$. After applying a dyadic decomposition, the Bochner–Riesz conjecture is therefore reduced to bounding the *Carleson–Sjölin operators*

$$S^\lambda f(x) := \int_{\mathbb{R}^n} e^{2\pi i \lambda |x-y|} a(x,y) f(y)\, dy, \tag{1.7}$$

where $a \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ has compact support bounded away from the diagonal $\{(x,x): x \in \mathbb{R}^n\}$. In particular, Corollary 1.3 is a consequence of the following bound.

COROLLARY 1.4. *If p satisfies the conditions in* (1.6) *and $\varepsilon > 0$, then*

$$\|S^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^{(n-1)/2 - n/p + \varepsilon} \|f\|_{L^p(\mathbb{R}^n)} \quad \textit{for all } \lambda \geqslant 1. \tag{1.8}$$

Theorem 1.2 can be used to prove estimates of the form (1.8). In particular, one may write S^λ as a superposition of operators $T_{y_n}^\lambda$ for which the y_n variable is frozen:

$$S^\lambda f(x) = \int_{\mathbb{R}} T_{y_n}^\lambda f_{y_n}(x)\, dy_n, \quad f_{y_n}(y') := f(y', y_n).$$

It is not difficult to check that each $T_{y_n}^\lambda$ is a Hörmander-type operator with positive-definite phase. Theorem 1.2 can be applied to the $T_{y_n}^\lambda$ with a uniform constant (this uniformity is a consequence of the proof) and Minkowski's inequality can then be used to convert these estimates into bounds for S^λ, yielding Corollary 1.3.

Bochner–Riesz over compact manifolds

The classical Bochner–Riesz multipliers have natural analogues defined over compact Riemannian manifolds (M, g) without boundary. In this setting, one defines the multiplier operator $m^\alpha(D)$ in terms of spectral projectors associated with an eigenbasis for the Laplace–Beltrami operator $-\Delta_g$; see [23] or [25, Chapter 5] for further details. Unlike in the Euclidean case, Theorem 1.2 does not, in general, directly lead to new L^p bounds for the spectral multipliers $m^\alpha(D)$. Nevertheless, as described presently, Theorem 1.2 does provide bounds for certain variants of the operator (1.7) which, in some sense, control the "local behaviour" of the Bochner–Riesz multipliers on (M, g).

When studying Bochner–Riesz multipliers in the manifold setting, one is led to consider certain variants of the Carleson–Sjölin operator (1.7), defined by

$$S_g^\lambda f(x) := \int_M e^{2\pi i \lambda \operatorname{dist}_g(x,y)} a(x,y) f(y)\, dy; \tag{1.9}$$

here, $a \in C^\infty(M \times M)$ is supported away from the diagonal, whilst dist_g and dy are, respectively, the distance function and measure on M induced by the Riemannian metric.

Operators of this form were studied previously by Minicozzi–Sogge [20] (see also [25, Chapter 5]).

Working in local coordinates and then arguing as in the Euclidean case, one may use Theorem 1.2 to prove the following bound.

COROLLARY 1.5. *Suppose* (M, g) *is a compact Riemannian manifold of dimension* $n \geqslant 2$ *without boundary. If* p *satisfies the conditions in* (1.6) *and* $\varepsilon > 0$, *then*

$$\|S_g^\lambda f\|_{L^p(M)} \lesssim_{\varepsilon, g} \lambda^{(n-1)/2 - n/p + \varepsilon} \|f\|_{L^p(M)} \quad \text{for all } \lambda \geqslant 1.$$

An interesting feature of this result is that there are examples of manifolds (M, g) for which the range of exponents (1.2) is sharp; see [20].

There is an analogue of the Carleson–Sjölin reduction in the manifold setting, which relies on the Hadamard parametrix for the wave equation on (M, g). This can be used to prove L^p bounds for Bochner–Riesz multipliers over (M, g) in the restricted range $p \geqslant 2(n+1)/(n-1)$: see [25, Chapter 5]. Unfortunately, the parametrix is only effective for short time intervals and, as a consequence, Corollary 1.5 does not appear to directly imply any new bounds for Bochner–Riesz multipliers over general (M, g). In particular, it seems that it is necessary to combine Corollary 1.5 with global geometric information to fully understand the Bochner–Riesz problem. For $p \geqslant 2(n+1)/(n-1)$ such difficulties can be overcome using L^p eigenfunction estimates of Sogge [24] (see [23] or [25, Chapter 5]), but the method appears to be tied down to this restricted range of exponents.

Finally, it is remarked that new Bochner–Riesz estimates *can* be obtained for certain specific choices of manifold (M, g) which enjoy additional symmetries. The simplest example is the flat torus $\mathbb{T}^n$; indeed, Corollary 1.3 implies similar L^p bounds in the toral setting via the classical multiplier transference principle. A more involved example is the n-dimensional Euclidean sphere S^n.[6] In this case, using the periodicity of the geodesic flow, one may entirely reduce the Bochner–Riesz problem to bounding operators essentially of the form (1.9), as observed in [22] (see also [21]). It may be possible to extend these methods to treat the class of Zoll manifolds, following a line of investigation initiated in [21].

1.3. Historical remarks

The problem of determining L^p estimates for Hörmander-type operators has an interesting history. Hörmander [17] asked whether (1.5) holds for $p > 2n/(n-1)$ (without ε-loss) under the hypotheses (H1) and (H2) only, and proved that this is indeed the case

---

[6] The authors are grateful to Christopher D. Sogge for drawing their attention to this example.

when $n=2$. This numerology agrees with the Fourier restriction conjecture and also the Bochner–Riesz multiplier problem, both of which would follow from a positive answer to Hörmander's question.[7] Stein [27] provided further evidence for this numerology by proving the estimate

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_{\phi,a} \|f\|_{L^2(B^{n-1})} \quad \text{for all } p \geqslant 2\frac{n+1}{n-1},$$

matching what was known about the high-dimensional Fourier restriction and Bochner–Riesz problems at that time. It was therefore somewhat surprising when Bourgain [5] showed that, in general, Stein's theorem is sharp. In particular, he demonstrated that, for every odd dimension $n \geqslant 3$, there exists a Hörmander-type operator for which

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_{\phi,a} \|f\|_{L^\infty(B^{n-1})} \tag{1.10}$$

fails to hold uniformly in $\lambda \geqslant 1$ whenever $p < 2(n+1)/(n-1)$. Aside from answering Hörmander's original question in the negative, Bourgain's work hinted at an interesting divergence between the odd- and even-dimensional theory. Moreover, it was noted in [6, p. 87] that in even dimensions the L^∞-L^p estimates always hold in a wider range than that of Stein's theorem. Thus, in general, the even-dimensional case is better behaved than the odd-dimensional case. This was further highlighted by Bourgain and the first author [9], who showed that in even dimensions (1.10) holds for $p > 2(n+2)/n$. Furthermore, in [9] and also implicitly in the work of Wisewell [35], examples were found in even dimensions which show that (1.10) can fail for $p < 2(n+2)/n$. Thus, the range of exponents in Hörmander's original question is valid only when $n=2$.

At this point, it is useful to describe the nature of the counterexamples of [5], [35] and [9], and provide some explanation for the difference between the odd- and even-dimensional cases. Roughly speaking, in the odd-dimensional case T^λ and f can be chosen so that $|T^\lambda f|$ is concentrated in the 1-neighbourhood of a low-degree algebraic variety Z of dimension $\frac{1}{2}(n+1)$. This is the smallest possible dimension for which such concentration is possible. In the even-dimensional case, $\frac{1}{2}(n+1)$ is no longer an integer, and it transpires that $|T^\lambda f|$ can only be concentrated into the 1-neighbourhood of a variety of relatively large dimension $\frac{1}{2}(n+2)$. These observations are related to Kakeya compression phenomena for sets of space curves (see [35] for a thorough introduction to this topic and [20] for the related problem of Kakeya sets of geodesics in Riemannian manifolds). They also hint at some underlying algebraic structure in the problem.

So far, the discussion has focused on operators satisfying the original (H1) and (H2) hypotheses of Hörmander. Lee [19] observed that, under the positive-definite hypothesis (H2$^+$), one can establish improvements over the range given by Stein's theorem in

---

[7] The connection with Bochner–Riesz multipliers is made via the classical reduction of Carleson–Sjölin [10], [17], as discussed in the previous subsection.

	n odd	n even
(H2)	$2\dfrac{n+1}{n-1}$	$2\dfrac{n+2}{n}$
(H2$^+$)	$2\dfrac{3n+1}{3n-3}$	$2\dfrac{3n+2}{3n-2}$

Table 1. Endpoint values for p for Hörmander-type operators under various hypotheses.

all dimensions. In particular, he showed that, for any Hörmander-type operator with positive-definite phase, (1.5) holds for $p \geqslant 2(n+2)/n$. This coincides with Theorem 1.2 when $n=3$, but is weaker in higher dimensions. Wisewell [35] and Minicozzi–Sogge [20] produced examples (again relying on Kakeya compression phenomena) to show that this result is sharp when $n=3$ (see also [9, §6]).

Comparing the sharp examples under the (H2) and (H2$^+$) hypotheses highlights another important consideration in addition to Kakeya compression phenomena. This feature relates to how the mass of $|T^\lambda f|$ can be distributed in neighbourhoods of low-degree varieties. It accounts for the improved behaviour demonstrated by operators with positive-definite phase and is described in detail in §2.

Given the results of this article, the sharp range of estimates for this problem are now understood under either the (H2) or (H2$^+$) hypothesis. The corresponding endpoint values for p are concisely listed in Table 1.

It is remarked that it is possible to prove estimates beyond the range of Theorem 1.2 under additional assumptions on the phase function. For example, the first author [13] has shown that, for $n=3$ and all $\varepsilon>0$, the extension operator E_{par} associated with the paraboloid satisfies

$$\|E_{\mathrm{par}}f\|_{L^p(B(0,\lambda))} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^2)}$$

for all $p \geqslant 3+\frac{1}{4}$, and this was further improved to $p \geqslant 3+\frac{3}{13}$ by Wang [33]. Furthermore, the aforementioned restriction conjecture asserts that the above inequality should be valid in the wider range $p \geqslant 3$.

1.4. Multilinear estimates

The proof of Theorem 1.2 follows the strategy introduced by the first author in [14]. The argument relies on establishing (weakened versions of) multilinear estimates for Hörmander-type operators. The multilinear approach was introduced in the late 1990s to study oscillatory integral operators (although it was arguably already implicit in many

earlier foundational works in the subject [12], [10]) and has proven an invaluable tool. To describe the k-linear setup, one first requires the notion of transversality.

Definition 1.6. Let $1 \leqslant k \leqslant n$ and $\mathbf{T}^\lambda = (T_1^\lambda, ..., T_k^\lambda)$ be a k-tuple of Hörmander-type operators, where T_j^λ has associated phase ϕ_j^λ, amplitude a_j^λ and generalised Gauss map G_j for $1 \leqslant j \leqslant k$. Then, $\mathbf{T}^\lambda$ is said to be *ν-transverse* for some $0 < \nu \leqslant 1$ if

$$\left| \bigwedge_{j=1}^{k} G_j(x; \omega_j) \right| \geqslant \nu \quad \text{for all } (x; \omega_j) \in \operatorname{supp} a_j \text{ for } 1 \leqslant j \leqslant k.$$

The following conjecture is a natural generalisation of an existent conjecture of Bennett [3] for Fourier extension operators.

CONJECTURE 1.7. (k-linear Hörmander conjecture) *Let $1 \leqslant k \leqslant n$ and let $(T_1^\lambda, ..., T_k^\lambda)$ be a ν-transverse k-tuple of Hörmander-type operators with positive-definite phase functions. For all $p \geqslant \bar{p}(k, n) := 2(n+k)/(n+k-2)$ and $\varepsilon > 0$ the estimate*

$$\left\| \prod_{j=1}^{k} |T_j^\lambda f_j|^{1/k} \right\|_{L^p(\mathbb{R}^n)} \lesssim_{\varepsilon, \nu, \phi} \lambda^\varepsilon \prod_{j=1}^{k} \|f_j\|_{L^2(B^{n-1})}^{1/k} \tag{1.11}$$

holds for all $\lambda \geqslant 1$.

Techniques have been developed by Tao–Vargas–Vega [32] and Bourgain and the first author [9] to convert k-linear into linear inequalities. There are a number of features of the multilinear theory which suggest that it is more approachable than directly tackling the linear estimates. For instance, here the desired inequalities are L^2-based, giving greater scope for orthogonality methods.

Some instances of the conjecture are known.

• The $k=1$ case corresponds to Stein's theorem [27] (which holds without the positive-definite hypothesis).

• The $k=2$ case was established by Lee [19], who then used the method of Tao–Vargas–Vega [32] to derive estimates for the linear problem. This approach yields Theorem 1.2 in the $n=3$ case, but produces strictly weaker results in higher dimensions (see the discussion in §1.3).

• The $k=n$ case was established by Bennett–Carbery–Tao [4] who also gave partial results at all levels of multilinearity (see also [3] for further discussion of this work). Bourgain and the first author [9] later developed a method to deduce improved linear estimates from these multilinear inequalities.

The precise statement of the Bennett–Carbery–Tao theorem [4] is as follows.

THEOREM 1.8. (Bennett–Carbery–Tao [4]) *Let* $2 \leqslant k \leqslant n$ *and suppose that* $(T_1^\lambda, ..., T_k^\lambda)$ *is a ν-transverse k-tuple of Hörmander-type operators. For all $p \geqslant 2k/(k-1)$ and $\varepsilon > 0$ the estimate*

$$\left\| \prod_{j=1}^k |T_j^\lambda f_j|^{1/k} \right\|_{L^p(\mathbb{R}^n)} \lesssim_{\varepsilon, (\phi_j)_{j=1}^k} \nu^{-C_\varepsilon} \lambda^\varepsilon \prod_{j=1}^k \|f_j\|_{L^2(B^{n-1})}^{1/k}$$

holds for all $\lambda \geqslant 1$.

The positive-definite assumption does not appear in the hypotheses of Theorem 1.8.([8]) Combining Theorem 1.8 with the method of [9] leads to the sharp estimates for Hörmander-type operators stated in Theorem 1.1. For completeness, the details of this argument are given in §11.

1.5. k-broad estimates

In [14] it was observed that, in the context of Fourier extension operators, the method of [9] does not require the full power of the k-linear theory, but rather can take as its input inequalities of a weaker form than (1.11) known as *k-broad estimates*. By applying polynomial partitioning techniques, the first author [14] was further able to prove the sharp range of L^2-based([9]) k-broad estimates for the Fourier extension operator associated with the paraboloid. This led to an improvement on the known range of estimates for parabolic restriction in dimensions $n \geqslant 4$. The main goal of this paper is to extend the theory of k-broad estimates to the more general context of Hörmander-type operators with positive-definite phase.

The k-broad setup involves the notion of a k-broad norm, which was introduced in [14]. Decompose B^{n-1} into finitely-overlapping balls τ of radius K^{-1}, where K is a large constant. These balls will be frequently referred to as K^{-1}-*caps*. Given a function $f : B^{n-1} \to \mathbb{C}$, write $f = \sum_\tau f_\tau$ where f_τ is supported in τ. In view of the rescaling ϕ^λ of the phase function, define the rescaled generalised Gauss map

$$G^\lambda(x; \omega) := G\left(\frac{x}{\lambda}; \omega\right) \quad \text{for } (x; \omega) \in \operatorname{supp} a^\lambda.$$

For each $x \in B(0, \lambda)$ there is a range of normal directions associated with the cap τ given by

$$G^\lambda(x; \tau) := \{G^\lambda(x; \omega) : \omega \in \tau \text{ and } (x; \omega) \in \operatorname{supp} a^\lambda\}.$$

([8]) In particular, of the results mentioned above only Lee's bilinear estimate [19] exploits the positive-definite hypothesis.

([9]) These estimates have an L^2-norm appearing on the right-hand side. Relaxing L^2 to L^∞ has led to further improvements on the Fourier restriction problem for the paraboloid [33], [16].

If $V \subseteq \mathbb{R}^n$ is a linear subspace, then let $\angle(G^\lambda(x; \tau), V)$ denote the smallest angle between any non-zero vector $v \in V$ and $v' \in G^\lambda(x; \tau)$.

The spatial ball $B(0, \lambda)$ is also decomposed into relatively small balls B_{K^2} of radius K^2. In particular, fix $\mathcal{B}_{K^2}$ a collection of finitely-overlapping K^2-balls which are centred in and cover $B(0, \lambda)$. For $B_{K^2} \in \mathcal{B}_{K^2}$ centred at some point $\bar{x} \in B(0, \lambda)$, define

$$\mu_{T^\lambda f}(B_{K^2}) := \min_{V_1, \ldots, V_A \in \mathrm{Gr}(k-1, n)} \Big(\max_{\tau : \angle(G^\lambda(\bar{x}; \tau), V_a) > K^{-1} \text{ for } 1 \leqslant a \leqslant A} \|T^\lambda f_\tau\|^p_{L^p(B_{K^2})} \Big); \quad (1.12)$$

here, $\mathrm{Gr}(k-1, n)$ is the Grassmannian manifold of all $(k-1)$-dimensional subspaces in $\mathbb{R}^n$. It will often be notationally convenient to write $\tau \notin V_a$ to mean that $\angle(G^\lambda(\bar{x}; \tau), V_a) > K^{-1}$ (the choice of centre $\bar{x}$ should always be clear from the context); with this notation the above expression becomes

$$\mu_{T^\lambda f}(B_{K^2}) := \min_{V_1, \ldots, V_A \in \mathrm{Gr}(k-1, n)} \Big(\max_{\tau : \tau \notin V_a \text{ for } 1 \leqslant a \leqslant A} \|T^\lambda f_\tau\|^p_{L^p(B_{K^2})} \Big).$$

For $U \subseteq \mathbb{R}^n$ the k-broad norm over U is then defined to be

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U)} := \Big(\sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap U \neq \varnothing}} \mu_{T^\lambda f}(B_{K^2}) \Big)^{1/p}. \quad (1.13)$$

It is remarked that $\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U)}$ is not a norm in the traditional sense, but it does satisfy weak variants of certain key properties of L^p-norms, as discussed below in §6.

Theorem 1.2 will be a consequence of certain estimates for k-broad norms. These estimates are proved under a further technical assumption that the phase is of *reduced form*. The details of this condition are postponed until §4.

THEOREM 1.9. *For $2 \leqslant k \leqslant n$ and all $\varepsilon > 0$, there exists a constant $C_\varepsilon > 1$ and an integer A such that, whenever T^λ is a Hörmander-type operator with positive-definite reduced phase, the estimate*

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(\mathbb{R}^n)} \lesssim_\varepsilon K^{C_\varepsilon} \lambda^\varepsilon \|f\|_{L^2(B^{n-1})} \quad (1.14)$$

holds for all $\lambda \geqslant 1$ and $K \geqslant 1$ whenever $p \geqslant \bar{p}(k, n) := 2(n+k)/(n+k-2)$.

The range of p is sharp for this theorem, as can be seen by considering the extension operator associated with the (elliptic) paraboloid (see [14]). As explained in §6 below, the k-broad estimate (1.14) is weaker than the corresponding k-linear estimate (1.11), and so Theorem 1.9 can be viewed as a weak substitute for Conjecture 1.7.

To derive L^p estimates from Theorem 1.9, roughly speaking, one argues as follows. The global L^p norm is broken up into contributions over balls B_{K^2}; the problem is to

estimate each $\|T^\lambda f\|^p_{L^p(B_{K^2})}$, and then to sum these estimates in B_{K^2}. Fixing one such ball, there exists a collection of $(k-1)$-dimensional subspaces $V_1, ..., V_A$ such that

$$\mu_{T^\lambda f}(B_{K^2}) = \max_{\tau \notin V_a \text{ for } 1 \leqslant a \leqslant A} \|T^\lambda f_\tau\|^p_{L^p(B_{K^2})}.$$

Thus, by the triangle and Hölder's inequalities,

$$\|T^\lambda f\|^p_{L^p(B_{K^2})} \lesssim_A K^{O(1)} \mu_{T^\lambda f}(B_{K^2}) + \sum_{a=1}^{A} \left\| \sum_{\tau \in V_a} T^\lambda f_\tau \right\|^p_{L^p(B_{K^2})}.$$

The k-broad estimate (1.13) effectively controls the first term on the right-hand side of the above display, after summing over all B_{K^2}. The problem of estimating $\|T^\lambda f\|^p_{L^p(B_R)}$ is therefore reduced to studying expressions of the form

$$\left\| \sum_{\tau \in V} T^\lambda f_\tau \right\|^p_{L^p(B_{K^2})}$$

for each B_{K^2}, where the sum is over caps τ which make a small angle with some $(k-1)$-dimensional subspace V. This term can then be controlled using a combination of ℓ^p-decoupling and an induction on scales argument, leading to the proof of Theorem 1.2. The full details of this argument are given in §11.

Structure of the article

The structure of this article is as follows:

- In §2 sharp examples for Theorems 1.1 and 1.2 are discussed in detail.
- In §3 the key features of the problem are identified in order to motivate the forthcoming analysis.
- In §4 some basic reductions are described which allow one to assume that the phase is of a certain *reduced form* in the proof of Theorem 1.2.
- In §5 and §6 some basic analytic tools are introduced. In particular, the wave packet decomposition for Hörmander-type operators is defined and studied, some elementary aspects of the L^2-theory for Hörmander-type operators are reviewed, and there is also a discussion of the basic properties of the k-broad norms and their relation to k-linear estimates.
- In §7 certain algebraic tools from combinatorial geometry are introduced. In particular, polynomial partitioning techniques are reviewed and some important geometric lemmas are proved; these techniques will play a fundamental role in the proof of Theorem 1.9.

• In §8 and §9 *transverse equidistribution estimates* for T^λ are introduced and studied. These estimates rely heavily on the positive-definite hypothesis and partially account for the improved behaviour exhibited by operators satisfying the $(H2^+)$ hypothesis.

• In §10 the proof of k-broad estimates of Theorem 1.9 is given.

• In §11 the linear estimates of Theorem 1.2 are deduced as a consequence of the k-broad estimates of Theorem 1.9. For completeness, the same methods are also applied to deduce Theorem 1.1 as a consequence of Corollary 6.5.

• In §12 standard ε-removal lemmas are generalised to the variable coefficient setting. This allows one to strengthen Theorem 1.2 away from the endpoint by removing the λ^ε-dependence in the constant.

• Appended are some remarks concerning (non)-stationary phase arguments used throughout the paper.

2. Necessary conditions

2.1. An overview of the sharp examples

In this section examples of Hörmander-type operators are studied in view of establishing the necessity of the conditions on the p exponent in the linear estimates of Theorem 1.1 and Theorem 1.2. This analysis will also identify some key features of operators with positive-definite phase which will later be exploited in the proof of Theorem 1.2. As discussed in §1.3, such examples first arose in the work of Bourgain [5], [6] and were later developed by Wisewell [35], Minicozzi–Sogge [20] and Bourgain–Guth [9], amongst others. The presentation in this section follows the lines of [5], [9].

All the examples considered here are of the following general form: for a fixed operator T^λ, a function f is chosen so that $|f|$ is constant whilst $|T^\lambda f|$ is concentrated in $N_{\lambda^\sigma}(Z) \cap B(0,\lambda)$ for some low-degree algebraic variety Z with $\dim Z = m$; here $N_{\lambda^\sigma}(Z)$ is the λ^σ-neighbourhood of Z. In particular, one has

$$\|T^\lambda f\|_{L^2(\mathbb{R}^n)} \sim \|T^\lambda f\|_{L^2(N_{\lambda^\sigma}(Z) \cap B(0,\lambda))}. \tag{2.1}$$

The examples will further be chosen so that

$$\|T^\lambda f\|_{L^2(\mathbb{R}^n)} \sim \lambda^{1/2}\|f\|_{L^2(B^{n-1})}; \tag{2.2}$$

note that, by Hörmander's generalisation of the Hausdorff–Young inequality [17], the inequality $\|T^\lambda f\|_{L^2(\mathbb{R}^n)} \lesssim \lambda^{1/2}\|f\|_{L^2(B^{n-1})}$ always holds (see also §5 below).

Playing (2.1) and (2.2) off against one another yields the necessary conditions on p. Indeed, for f as above,

$$\|f\|_{L^p(B^{n-1})} \sim \|f\|_{L^2(B^{n-1})} \sim \lambda^{-1/2}\|T^\lambda f\|_{L^2(\mathbb{R}^n)} \sim \lambda^{-1/2}\|T^\lambda f\|_{L^2(N_{\lambda^\sigma}(Z)\cap B(0,\lambda))}.$$

Now, assuming that the estimate $\|T^\lambda g\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^\varepsilon \|g\|_{L^p(B^{n-1})}$ holds for all $\varepsilon > 0$ and applying Hölder's inequality, it follows that

$$\|f\|_{L^p(B^{n-1})} \lesssim_\varepsilon |N_{\lambda^\sigma}(Z)\cap B(0,\lambda)|^{1/2-1/p}\lambda^{-1/2}\lambda^\varepsilon\|f\|_{L^p(B^{n-1})}.$$

By a theorem of Wongkew [37] (see Theorem 8.10 below),

$$|N_{\lambda^\sigma}(Z)\cap B(0,\lambda)| \lesssim \lambda^{m+(n-m)\sigma}, \tag{2.3}$$

where the implied constant depends only on n. In fact, for the simple varieties used in the arguments below, (2.3) can be shown by direct inspection. Thus, combining the previous displays and recalling that λ can be taken arbitrarily large, one concludes that

$$p \geqslant 2\frac{\sigma(n-m)+m}{\sigma(n-m)+m-1}. \tag{2.4}$$

This condition depends on the two parameters m and σ, and becomes more restrictive the more m and σ decrease. Therefore, in order to obtain the strongest possible restriction on p for a given phase function, one wishes to find the lowest possible m and σ, over all f, for which the mass of $T^\lambda f$ can concentrate in the λ^σ-neighbourhood of an m-dimensional low-degree algebraic variety.

• The optimal choice of m is $n-\left\lfloor\frac{1}{2}(n-1)\right\rfloor$. This value arises directly from the theory of Kakeya sets of curves, and will be discussed in more detail in the next subsection.

• The optimal choice of σ depends on the signature of the phase. For general Hörmander-type operators, one may find examples for which $\sigma = 0$. If one assumes the positive-definite condition (H2$^+$), then $\sigma = \frac{1}{2}$ is the lowest possible value. This difference in behaviour is governed by *transverse equidistribution estimates* for T^λ, which were introduced in the context of Fourier extension operators in [14]. This will be discussed in detail in §2.4.

The optimal pairs (m,σ) under the various hypotheses are listed in Table 2. Plugging these values into (2.4) gives the corresponding sharp range of estimates for T^λ.

	n odd	n even
(H2)	$\left(\frac{1}{2}(n+1),0\right)$	$\left(\frac{1}{2}n+1,0\right)$
(H2$^+$)	$\left(\frac{1}{2}(n+1),\frac{1}{2}\right)$	$\left(\frac{1}{2}n+1,\frac{1}{2}\right)$

Table 2. Optimal values of (m,σ) for the sharp examples.

2.2. Model operators

The examples described above will arise from operators with phase of the relatively simple form

$$\phi(x;\omega) := \langle x',\omega\rangle + \tfrac{1}{2}\langle \boldsymbol{A}(x_n)\omega,\omega\rangle, \tag{2.5}$$

where $\boldsymbol{A}\colon \mathbb{R}\to \mathrm{Sym}(n-1,\mathbb{R})$ is a polynomial function taking values in the class of real symmetric matrices which satisfies $\boldsymbol{A}(0)=0$. For such a phase, the condition (H1) always holds whilst (H2) (resp. (H2$^+$)) holds if and only if the componentwise derivative $\boldsymbol{A}'(x_n)\in \mathrm{GL}(n-1,\mathbb{R})$ (resp. $\boldsymbol{A}'(x_n)$ is positive-definite) for all relevant $x_n\in[-1,1]$. Observe that, if $\boldsymbol{A}(x_n):=x_n A$ for some fixed $A\in \mathrm{Sym}(n-1,\mathbb{R})\cap \mathrm{GL}(n-1,\mathbb{R})$, then the resulting operator is the extension operator associated with the graph of the non-degenerate quadratic form $\omega\mapsto\tfrac{1}{2}\langle A\omega,\omega\rangle$. For the present purpose, one is interested in examples with higher-order dependence on x_n.

Let T^λ be an operator associated with the phase function (2.5) for some $\boldsymbol{A}$ and a choice of non-negative amplitude function a. The ω-support of a is assumed to lie in $B(0,c)$, where $c>0$ is a small constant. Cover B^{n-1} by finitely-overlapping balls θ of radius $\lambda^{-1/2}$; these balls will frequently be referred to as $\lambda^{-1/2}$-*caps*. Let ψ_θ be a smooth partition of unity adapted to this cover. Consider a *wave packet* of the form

$$f_{\theta,v_\theta}(\omega) := e^{-2\pi i\lambda\langle v_\theta,\omega-\omega_\theta\rangle}\psi_\theta(\omega)$$

for some choice of $v_\theta\in \mathbb{R}^{n-1}$ and ω_θ the centre of the cap θ. To obtain the necessary conditions for L^p-boundedness of T^λ, the operator will be tested against functions given by superpositions of these basic wave packets.

Each localised piece $T^\lambda f_{\theta,v_\theta}$ is concentrated in a tubular region in $\mathbb{R}^n$. In particular, define the curve

$$\gamma_{\theta,v_\theta}(t) := v_\theta - \boldsymbol{A}(t)\omega_\theta \quad \text{for } t\in(-1,1), \tag{2.6}$$

and let T_{θ,v_θ} be the curved tube

$$T_{\theta,v_\theta} := \left\{x\in B(0,\lambda) : \left|x'-\lambda\gamma_{\theta,v_\theta}\left(\frac{x_n}{\lambda}\right)\right| < c\lambda^{1/2}\right\}$$

for $c>0$ as above. It is not difficult to show that

$$|T^\lambda f_{\theta,v_\theta}(x)| \gtrsim \lambda^{-(n-1)/2} \chi_{T_{\theta,v_\theta}}(x) \quad \text{for all } x \in B(0,\lambda), \tag{2.7}$$

provided that c is chosen suitably small. Indeed, let a_θ^λ be the rescaled amplitude

$$a_\theta^\lambda(x;\omega) := a^\lambda(x;\omega_\theta + \lambda^{-1/2}\omega)\psi_\theta(\omega_\theta + \lambda^{-1/2}\omega)$$

and ϕ_θ^λ be the phase function

$$\phi_\theta^\lambda(x;\omega) := \lambda^{-1/2}\Big\langle x' - \lambda\gamma_{\theta,v_\theta}\Big(\frac{x_n}{\lambda}\Big),\omega\Big\rangle + \frac{1}{2}\Big\langle \mathbf{A}\Big(\frac{x_n}{\lambda}\Big)\omega,\omega\Big\rangle,$$

so that, by a linear change of variables,

$$T^\lambda f_{\theta,v_\theta}(x) = \lambda^{-(n-1)/2} e^{2\pi i \phi^\lambda(x;\omega_\theta)} \int_{\mathbb{R}^{n-1}} e^{2\pi i \phi_\theta^\lambda(x;\omega)} a_\theta^\lambda(x;\omega)\,d\omega. \tag{2.8}$$

Taking absolute values in (2.8) and writing $e^{2\pi i \phi_\theta^\lambda(x;\omega)}$ in terms of its real and imaginary parts, one deduces that

$$|T^\lambda f_{\theta,v_\theta}(x)| \gtrsim \lambda^{-(n-1)/2}\left|\int_{\mathbb{R}^{n-1}} \cos(2\pi\phi_\theta^\lambda(x;\omega))a_\theta^\lambda(x;\omega)\,d\omega\right|.$$

Provided c is sufficiently small,

$$|\phi_\theta^\lambda(x;\omega)| \leqslant \tfrac{1}{100} \quad \text{whenever } x \in T_{\theta,v_\theta} \text{ and } (x;\omega) \in \operatorname{supp} a_\theta^\lambda.$$

Thus, if $x \in T_{\theta,v_\theta}$, we have $\cos(2\pi\phi_\theta^\lambda(x;\omega)) \gtrsim 1$, and the desired bound (2.7) follows.

2.3. Kakeya sets of curves

By studying the geometry of the family of tubes T_{θ,v_θ}, one may construct sharp examples for Theorem 1.2. These examples arise owing to Kakeya compression phenomena, whereby the tubes are arranged to lie in a neighbourhood of a low-dimensional set. For $n=3$, the following example appears in Bourgain–Guth [9] (see also [20], [35] for related constructions). Let ϕ be of the form (2.5), where $\mathbf{A}(t)$ is taken to be the $(n-1)\times(n-1)$ block-diagonal matrix

$$\mathbf{A}(t) := \underbrace{\begin{pmatrix} t & t^2 \\ t^2 & t+t^3 \end{pmatrix} \oplus \ldots \oplus \begin{pmatrix} t & t^2 \\ t^2 & t+t^3 \end{pmatrix}}_{\lfloor (n-1)/2 \rfloor\text{-fold}} \oplus (t).$$

Here, it is understood that the final 1×1 block appears only when n is even. Observe that the resulting phase (2.5) satisfies (H1) and (H2$^+$) on $B^n \times B^{n-1}$.

Suppose that T^λ is the associated oscillatory integral operator. The estimate

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^{n-1})} \tag{2.9}$$

is tested against a superposition of wave packets

$$f := \sum_{\theta:\lambda^{-1/2}\text{-cap}} \varepsilon_\theta \cdot f_{\theta,v_\theta},$$

where $\varepsilon_\theta \in \{1, -1\}$ are uniformly distributed independent random signs. By Khintchine's theorem (see, for instance, [26, Appendix D]), the expected value of $|T^\lambda f(x)|$ is given by

$$\mathbf{E}[\,|T^\lambda f(x)|\,] \sim \left(\sum_{\theta:\lambda^{-1/2}\text{-cap}} |T^\lambda f_{\theta,v_\theta}(x)|^2 \right)^{1/2} \gtrsim \lambda^{-(n-1)/2} \left(\sum_{\theta:\lambda^{-1/2}\text{-cap}} \chi_{T_{\theta,v_\theta}}(x) \right)^{1/2}$$

for all $x \in B(0,\lambda)$. Thus, by Hölder's and Minkowski's inequalities,

$$\lambda^{-(n-1)/2} \left(\int \sum_{\theta:\lambda^{-1/2}\text{-cap}} \chi_{T_{\theta,v_\theta}} \right)^{1/2} \lesssim \left| \bigcup_{\theta:\lambda^{-1/2}\text{-cap}} T_{\theta,v_\theta} \right|^{1/2-1/p} \mathbf{E}[\,\|T^\lambda f\|_{L^p(\mathbb{R}^n)}\,].$$

The hypothesis (2.9) together with a direct computation now gives

$$\|f\|_{L^p(B^{n-1})} \lesssim_\varepsilon \left| \bigcup_{\theta:\lambda^{-1/2}\text{-cap}} T_{\theta,v_\theta} \right|^{1/2-1/p} \lambda^{-1/2+\varepsilon} \|f\|_{L^p(B^{n-1})}, \tag{2.10}$$

since $\|f\|_{L^p(B^{n-1})} \sim 1$ is independent of the outcomes of the ε_θ.

Varying v_θ corresponds to translating the tube T_{θ,v_θ} in space in the x' direction. In view of (2.10), one wishes to choose the v_θ in order to arrange the tubes so that their union has small measure. For the above choice of phase, it is in fact possible to select the v_θ so that the tubes all lie in the $\lambda^{1/2}$-neighbourhood of a low-dimensional, low-degree algebraic variety. In particular, let $m := n - \lfloor \frac{1}{2}(n-1) \rfloor$ and $Z := Z(P_1, ..., P_{n-m})$ be the common zero set of the polynomials

$$P_j(x) := \lambda x_{2j} - x_{2j-1} x_n \quad \text{for } 1 \leqslant j \leqslant \lfloor \tfrac{1}{2}(n-1) \rfloor.$$

Thus, Z is an algebraic variety of dimension m and degree $O_n(1)$. If one defines

$$v_{\theta,2j-1} := -\omega_{\theta,2j} \quad \text{and} \quad v_{\theta,2j} = v_{\theta,n-1} = 0 \quad \text{for } 1 \leqslant j \leqslant \lfloor \tfrac{1}{2}(n-1) \rfloor$$

for each cap θ, then a simple computation shows that the curve $t \mapsto (\lambda \gamma_{\theta, v_\theta}(t/\lambda), t)$ lies in Z. Thus,

$$\bigcup_{\theta : \lambda^{-1/2}\text{-cap}} T_{\theta, v_\theta} \subseteq N_{\lambda^{1/2}}(Z) \cap B(0, \lambda),$$

and the desired necessary conditions on p follow from (2.10).

In conclusion, here the necessary conditions arise from the fact that it is possible to compress an $(n-1)$-dimensional family of curves into a set of small dimension m. The value $m = n - \lfloor \frac{1}{2}(n-1) \rfloor$ is optimal for this kind of behaviour, in view of known estimates for associated Kakeya maximal functions: see [35] and [9].

2.4. Mass concentration

It will be useful to contrast the behaviour in the positive-definite and indefinite cases by considering sharp examples for Theorem 1.1 (that is, for the class of operators satisfying (H1) and the weaker hypothesis (H2)). As before, Kakeya compression plays a significant role in the argument, but one can introduce additional interference between the wave packets which leads to stronger necessary conditions. In particular, this interference causes the mass of $|T^\lambda f|$ to concentrate in a tiny $O(1)$-neighbourhood of a variety Z; such tight concentration is not possible under the (H2$^+$) hypothesis (as demonstrated by Theorem 1.2).

The following example was introduced by Bourgain [5] (see also [6]). Once again, the phase is taken to be of the form (2.5). This time $\boldsymbol{A}(t)$ is defined to be the $(n-1) \times (n-1)$ block-diagonal matrix

$$\boldsymbol{A}(t) := \underbrace{\begin{pmatrix} 0 & t \\ t & t^2 \end{pmatrix} \oplus ... \oplus \begin{pmatrix} 0 & t \\ t & t^2 \end{pmatrix}}_{\lfloor (n-1)/2 \rfloor\text{-fold}} \oplus (t). \tag{2.11}$$

Clearly, the corresponding phase satisfies (H1) and (H2), but (H2$^+$) fails. Define the curves $\gamma_{\theta, v_\theta}$ as in (2.6), so that

$$|T^\lambda f_{\theta, v_\theta}(x)| \gtrsim \lambda^{-(n-1)/2} \chi_{T_{\theta, v_\theta}}(x) \quad \text{for all } x \in B(0, \lambda).$$

If one takes

$$v_{\theta, 2j-1} := -\omega_{\theta, 2j-1} \text{ and } v_{\theta, 2j} = v_{\theta, n-1} = 0 \quad \text{for } 1 \leqslant j \leqslant \lfloor \tfrac{1}{2}(n-1) \rfloor,$$

then it follows that the curve $t \mapsto (\lambda \gamma_{\theta, v_\theta}(t/\lambda), t)$ lies in Z for all $\lambda^{-1/2}$-caps θ, where Z is the same variety as that appearing in the previous subsection (see Figure 1).
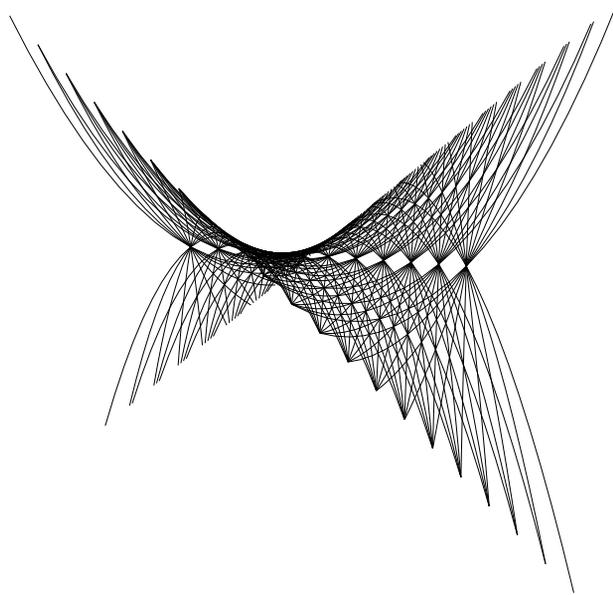
Figure 1. The Kakeya compression phenomenon for the curves arising from the matrix (2.11).

One may repeat the analysis of §2.3 by taking f to be a linear combination of wave packets f_{θ,v_θ} with random signs. This leads to the same necessary conditions as in the positive-definite case. However, certain deterministic choices of f lead to stronger conditions on p. In particular, consider the function

$$\tilde{f} := \sum_{\theta:\lambda^{-1/2}\text{-cap}} e^{2\pi i \lambda Q(\omega_\theta)} f_{\theta,v_\theta}, \tag{2.12}$$

where the v_θ are as defined above and Q is the quadratic polynomial

$$Q(\omega) := \frac{1}{2} \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} \omega_{2j-1}^2.$$

Each modulated wave packet appearing in (2.12) has a phase given by

$$\lambda(Q(\omega_\theta) - \langle v_\theta, \omega - \omega_\theta \rangle) = \lambda Q(\omega) - \frac{\lambda}{2} \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} (\omega_{2j-1} - \omega_{\theta,2j-1})^2.$$

Since the $\lambda(\omega_{2j-1} - \omega_{\theta,2j-1})^2$ terms are bounded functions on the support of ψ_θ, they do not contribute any significant oscillation. One may therefore heuristically identify $\tilde{f}$ with the function

$$f(\omega) := e^{2\pi i \lambda Q(\omega)} \psi(\omega),$$

where ψ is a bump function supported in B^{n-1}. Using a simple stationary phase argument, it was shown in [5] (see also [6], [9]) that

$$|T^\lambda f(x)| \gtrsim \lambda^{-\lfloor n/2 \rfloor/2} \chi_{N_c(Z)}(x) \quad \text{for all } x \in B(0,\lambda). \tag{2.13}$$

Here $0 < c < 1$ is some small, fixed constant (which is independent of λ) and Z is as defined in §2.3. With this estimate, one readily deduces the desired necessary conditions on p. Indeed, testing the inequality

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^{n-1})}$$

against the function f as defined above, it follows from (2.13) that

$$\lambda^{-\lfloor n/2 \rfloor/2} |N_c(Z) \cap B(0,\lambda)|^{1/p} \lesssim_\varepsilon \lambda^\varepsilon.$$

Since $|N_c(Z) \cap B(0,\lambda)| \sim \lambda^{n-\lfloor (n-1)/2 \rfloor}$ and $n = \lfloor \frac{1}{2} n \rfloor + \lfloor \frac{1}{2}(n-1) \rfloor + 1$, one deduces that

$$\lambda^{-\lfloor n/2 \rfloor/2 + (\lfloor n/2 \rfloor + 1)/p} \lesssim_\varepsilon \lambda^\varepsilon.$$

Taking $\lambda \geqslant 1$ large and $0 < \varepsilon < 1$ small, this forces

$$\begin{cases} -\dfrac{n-1}{2} + \dfrac{n+1}{p} \leqslant 0, & \text{if } n \text{ is odd,} \\ -\dfrac{n}{2} + \dfrac{n+2}{p} \leqslant 0, & \text{if } n \text{ is even,} \end{cases}$$

corresponding to the constraints on p from Theorem 1.1.

For completeness, the details of the argument used in [5], [6] to prove (2.13) are reviewed. Observe that $T^\lambda f(x)$ is an oscillatory integral with smooth amplitude ψ and phase

$$\sum_{j=1}^{\lfloor (n-1)/2 \rfloor} x_{2j-1}\omega_{2j-1} + x_{2j}\omega_{2j} + \frac{\lambda}{2}(\omega_{2j-1} + \lambda^{-1}x_n\omega_{2j})^2 + \delta_e\left(x_{n-1}\omega_{n-1} + \frac{x_n}{2}\omega_{n-1}^2\right), \tag{2.14}$$

where

$$\delta_e = \begin{cases} 0, & \text{if } n \text{ is odd,} \\ 1, & \text{if } n \text{ is even.} \end{cases}$$

Introduce new variables

$$z_j := \omega_{2j-1} + \lambda^{-1}x_n\omega_{2j} \quad \text{for } 1 \leqslant j \leqslant \lfloor \tfrac{1}{2}(n-1) \rfloor.$$

If $x \in Z$, then the phase function (2.14) can be expressed as

$$\sum_{j=1}^{\lfloor (n-1)/2 \rfloor} x_{2j-1} z_j + \frac{\lambda}{2} z_j^2 + \delta_e \left(x_{n-1} \omega_{n-1} + \frac{x_n}{2} \omega_{n-1}^2 \right).$$

The integral $T^\lambda f$ can now be reduced to a product of $\lfloor \frac{1}{2}(n-1) \rfloor + \delta_e = \lfloor \frac{1}{2}n \rfloor$ integrals, each in a single variable. For $x \in Z$, the inequality (2.13) follows as a consequence of standard stationary phase estimates applied to each of these integrals (see, for instance, [28, Chapter VIII, Proposition 3]). This lower bound can then be extended to some c-neighbourhood of Z via a simple estimate on the gradient of $T^\lambda f(x)$.

3. Key features of the analysis

The examples of the previous section highlight several key features of Hörmander-type operators. All these features are exploited in the proofs of the linear and k-broad estimates.

(1) **Algebraic structure.** The sharp examples were given by arranging collections of wave packets to lie in a relatively small neighbourhood of a low-degree, low-dimensional algebraic variety Z. It turns out that this is an essential feature of both the linear and k-broad problems. To exploit this underlying algebraic structure, the proof of Theorem 1.9 will rely on a variant of the polynomial partitioning method introduced by Katz and the first author [15]. Roughly speaking, this method allows one to reduce to the case where $|T^\lambda f|$ concentrates around some low-degree, low-dimensional variety, as in the sharp examples. This can be thought of as a dimensional reduction and, indeed, the proof of Theorem 1.9 will proceed by an induction on dimension. Polynomial partitioning has played an increasingly important role in the theory of oscillatory integral operators, beginning with the work on the restriction problem in [13], [14] and, more recently, in [33], [16].

(2) **Non-concentration/transverse equidistribution.** Suppose one does not assume the phase is positive-definite. The example of §2.4 then shows that interference between the wave packets can cause $|T^\lambda f|$ to be concentrated in a tiny neighbourhood of Z. In order to prove the sharp range of estimates in the positive-definite case, one must rule out the possibility of such concentration. This is achieved by extending the theory of so-called *transverse equidistribution estimates* introduced in [14] to the variable coefficient setting. These estimates can be interpreted as showing that $|T^\lambda f(x)|$ is morally constant along transverse directions to Z in a $\lambda^{1/2}$-neighbourhood of the variety. Consequently, $|T^\lambda f(x)|$ cannot concentrate in a smaller neighbourhood.

(3) **Parity of the dimension.** Another key feature of the examples discussed in the previous section is their dependence on the parity of the ambient dimension n. Recall that this is directly related to the minimal dimension

$$d(n) := \begin{cases} \frac{1}{2}(n+1), & \text{if } n \text{ is odd,} \\ \frac{1}{2}(n+2), & \text{if } n \text{ is even,} \end{cases}$$

of Kakeya sets of curves in $\mathbb{R}^n$. The parity of the dimension does not play an overt role in the proof of the k-broad estimates, but it becomes a noticeable feature when one wishes to pass from k-broad to linear estimates in the proof of Theorem 1.2. In particular, for each fixed value of $2 \leqslant k \leqslant n$, the method of §11 shows that the k-broad estimates imply a (possibly empty/trivial) range of linear estimates. It transpires that, to optimise the range of linear estimates obtained in this manner, one should choose k to correspond to the dimension $d(n)$ from the Kakeya problem.

The proof of the k-broad estimates follows the same general scheme as that used to study Fourier extension operators in [14], and heavily exploits the the features (1) and (2) of the problem highlighted above. A detailed sketch of the argument in the extension context is provided in [14]; this sketch is likely to be beneficial to readers new to these ideas.

4. Reductions

4.1. Basic reductions

The prototypical example of a positive-definite phase function is given by

$$\phi_{\mathrm{par}}(x;\omega) := \langle x', \omega \rangle + \tfrac{1}{2} x_n |\omega|^2. \tag{4.1}$$

This is the phase associated with the extension operator for the (elliptic) paraboloid. In general, to prove Theorem 1.2, it suffices to only consider phases which are given by small perturbations of ϕ_{par}. Observations of this kind have been used previously in the theory of oscillatory integral operators and the arguments of this section are inspired by [17], [19].

To understand why such a reduction is possible, first recall that the class of operators under consideration are those of the form

$$T^\lambda f(x) := \int_{\mathbb{R}^{n-1}} e^{2\pi i \phi^\lambda(x;\omega)} a^\lambda(x;\omega) f(\omega) \, d\omega,$$

where ϕ satisfies (H1) and (H2$^+$). In addition, one may assume a number of fairly stringent conditions on the form of ϕ on the support of a.

LEMMA 4.1. *To prove Theorem* 1.2 *for some fixed* $\varepsilon > 0$, *it suffices to consider the case where* a *is supported on* $X \times \Omega$, *where* $X := X' \times X_n$ *and* $X' \subset B^{n-1}$, $X_n \subset B^1$ *and* $\Omega \subset B^{n-1}$ *are small balls centred at zero upon which the phase* ϕ *has the form*

$$\phi(x; \omega) = \langle x', \omega \rangle + x_n h(\omega) + \mathcal{E}(x; \omega). \tag{4.2}$$

Here h *and* $\mathcal{E}$ *are smooth functions,* h *is quadratic in* ω *and* $\mathcal{E}$ *is quadratic in* x *and* ω.[10] *Furthermore, letting* $c_{\mathrm{par}} > 0$ *be a small constant, which may depend on the admissible parameters* n, p *and* ε, *one may assume that*

$$\|\partial^2_{\omega x'} \phi(x; \omega) - \mathrm{I}_{n-1}\|_{\mathrm{op}} < c_{\mathrm{par}}, \quad |\partial_\omega \partial_{x_n} \phi(x; \omega)| < c_{\mathrm{par}}, \tag{4.3}$$

$$\|\partial^2_\omega \partial_{x_k} \phi(x; \omega) - \delta_{kn} \mathrm{I}_{n-1}\|_{\mathrm{op}} < c_{\mathrm{par}} \tag{4.4}$$

for all $(x; \omega) \in X \times \Omega$ *and* $1 \leqslant k \leqslant n$.

Here, I_{n-1} denotes the $(n-1) \times (n-1)$ identity matrix, δ_{ij} the Kronecker δ-function and $\| \cdot \|_{\mathrm{op}}$ the operator norm.

It is perhaps useful to provide a brief interpretation of the formula (4.2). Since h is quadratic,

$$h(\omega) = \tfrac{1}{2} \langle \partial^2_{\omega\omega} h(0) \omega, \omega \rangle + O(|\omega|^3).$$

Although unnecessary for the forthcoming analysis, by rotating the ω coordinates, one may further assume that

$$h(\omega) = \tfrac{1}{2} |\omega|^2 + O(|\omega|^3).$$

Thus, the phase in (4.2) is given by

$$\phi(x; \omega) = \phi_{\mathrm{par}}(x; \omega) + \text{higher-order terms,}$$

and is therefore a perturbation of the prototypical example ϕ_{par}.

The proof of the lemma is based upon three elementary principles:

Localisation. If a property P of a phase holds locally on $\mathrm{supp}\, a$, then typically one may assume P holds on the whole of $\mathrm{supp}\, a$ by applying a partition of unity, the triangle inequality and shifting coordinates.

Parametrisation invariance. By the change of variables formula, one may compose ϕ with a smooth change of either the x or ω variables.

---

[10] Explicitly, if $(\alpha, \beta) \in \mathbb{N}_0 \times \mathbb{N}_0^{n-1}$ is a pair of multi-indices, then
(i) $\partial^\beta_\omega h(0) = \partial^\beta_\omega \partial^\alpha_x \mathcal{E}(x; 0) = 0$ whenever $x \in X$ and $|\beta| \leqslant 1$;
(ii) $\partial^\beta_\omega \partial^\alpha_x \mathcal{E}(0; \omega) = 0$ whenever $\omega \in \Omega$ and $|\alpha| \leqslant 1$.

Modulation invariance. One is free to add smooth functions to the phase which depend only on either the x or on the ω variables. In particular, ϕ can be replaced by

$$\phi(x;\omega)+\phi(0;0)-\phi(0;\omega)-\phi(x;0),$$

and therefore one may assume that

$$\partial_x^\alpha \phi(x;0)=0 \quad \text{and} \quad \partial_\omega^\beta \phi(0;\omega)=0 \tag{4.5}$$

for all multi-indices $(\alpha,\beta)\in\mathbb{N}_0^n\times\mathbb{N}_0^{n-1}$.

The following argument provides an example of these three principles working together. Rotating the x-coordinates, one may assume that $\partial_\omega\partial_{x_n}\phi(0;0)=0$. By (H2), it follows that

$$\det \partial^2_{x'\omega}\phi(0;0)\neq 0.$$

The inverse function theorem now implies the existence of local inverses to the functions $\omega\mapsto\partial_{x'}\phi(x;\omega)$ and $x'\mapsto\partial_\omega\phi(x;\omega)$ in a neighbourhood of 0zero. Thus, by localisation, one may assume that $\operatorname{supp} a$ is contained in $X\times\Omega$, where $X=X'\times X_n$ for $X'\subset B^{n-1}$, $X_n\subset B^1$ and $\Omega\subset B^{n-1}$ small balls centred at zero, and that there exist smooth functions Φ and Ψ taking values in X and Ω, respectively, such that

$$\partial_{x'}\phi(x;\Psi(x;u))=u \quad \text{and} \quad \partial_\omega\phi(\Phi(z',x_n;\omega),x_n;\omega)=z'. \tag{4.6}$$

The former identity can be thought of as a generalisation of the fact that any hypersurface can be locally parameterised as a graph. The latter identity features in the proof of Lemma 4.1 and it is useful to highlight some further properties of Φ. For each $(x_n,\omega)\in X_n\times\Omega$ the map $z'\mapsto\Phi(z',x_n;\omega)$ is a diffeomorphism from its domain onto X'; this provides a useful change of variables on X'. Furthermore, it is easy to see that zero lies in the domain of this map when $x_n=0$, $\omega=0$, and that

$$\Phi(0;0)=0, \quad \partial_{x_n}\Phi(0;0)=0 \quad \text{and} \quad \partial_{x'}\Phi(0;0)=\partial^2_{x'\omega}\phi(0;0)^{-1}. \tag{4.7}$$

Indeed, the first identity follows directly from (4.5), whilst the remaining identities are obtained by differentiating the defining expression for Φ from (4.6).

Proof of Lemma 4.1. By (4.5), one may assume that

$$\phi(x;\omega)=\langle\partial_\omega\phi(x;0),\omega\rangle+\varrho(x;\omega),$$

where ϱ is quadratic in ω and satisfies $\varrho(0;\omega)=0$ for all $\omega\in\Omega$. Let Φ be the function defined in (4.6) and, using localisation and parametrisation invariance, perform the change of variables $x'\mapsto\Phi(x',x_n;0)$ on X', so that the phase becomes

$$\phi(x;\omega)=\langle x',\omega\rangle+\varrho(\Phi(x;0),x_n;\omega).$$

By (4.7), one has $\partial_{x_n}\Phi(x',0;0)=O(|x|)$ and, taking a Taylor expansion of ϱ in x_n,

$$\varrho(\Phi(x;0),x_n;\omega) = \varrho(\Phi(x',0;0),0;\omega)+(\partial_{x_n}\varrho)(\Phi(x',0;0),0;\omega)x_n+O(|x|^2).$$

Note that the first expression on the right-hand side satisfies

$$\varrho(\Phi(x',0;0),0;\omega) = \langle\partial_{x'\omega}\phi(0;0)^{-1}x',(\partial_{x'}\varrho)(0;\omega)\rangle+O(|x|^2)$$

whilst, Taylor-expanding now in x', it follows that

$$(\partial_{x_n}\varrho)(\Phi(x',0;0),0;\omega) = (\partial_{x_n}\varrho)(0;\omega)+O(|x|).$$

Combining these observations, and noting, for instance, that (4.5) implies that

$$\partial_x^\alpha\varrho(x;0) = 0 \quad \text{for all } \alpha\in\mathbb{N}_0^n \text{ and } x\in B^n,$$

one deduces that

$$\phi(x;\omega) = \langle x',\omega+\partial_{x'\omega}\phi(0;0)^{-\top}(\partial_{x'}\varrho)(0;\omega)\rangle+x_n(\partial_{x_n}\varrho)(0;\omega)+O(|x|^2|\omega|^2);$$

Here the symbol $(\cdot)^\top$ is used to denote the matrix transpose and $(\cdot)^{-\top}$ the inverse matrix transpose.

Since ϱ is quadratic in ω, it follows that $\omega\mapsto\omega+\partial_{x'\omega}\phi(0;0)^{-\top}(\partial_{x'}\varrho)(0;\omega)$ is a well-defined change of variables in a neighbourhood of the origin and so, once again by localisation and parametrisation invariance, the problem is reduced to considering phase functions of the from (4.2). By construction, h and $\mathcal{E}$ are quadratic. Finally, the condition (H2$^+$) implies that the matrix $\partial_{\omega\omega}^2\partial_{x_n}\phi(0;0)$ is positive definite. Applying a linear coordinate change, one may therefore suppose that $\partial_{\omega\omega}^2\partial_{x_n}\phi(0;0)=\mathrm{I}_{n-1}$. On the other hand, clearly $\partial_\omega\partial_{x_n}\phi(0;0)=0$, $\partial_{\omega x'}^2\phi(0;0)=\mathrm{I}_{n-1}$ and $\partial_{\omega\omega}^2\partial_{x_k}\phi(0;0)=0_{n-1}$ for $1\leqslant k\leqslant n-1$. By continuity, if the support of a is sufficiently small, then the conditions of (4.3) and (4.4) are valid on the support of a. $\qquad\square$

4.2. Parabolic rescaling

In addition to the reductions of Lemma 4.1, it will be useful to have control over higher-order derivatives of the phase, and also the amplitude function. Such control is made possible using a simple scaling argument.

Consider the constant coefficient case $\phi(x;\omega):=\langle x',\omega\rangle+x_nh(\omega)$, where

$$h(\omega) = \tfrac{1}{2}|\omega|^2+O(|\omega|^3),$$

so that ϕ is a perturbation of the prototypical phase ϕ_{par} defined in (4.1). The corresponding operator

$$Ef(x) := \int_{B^{n-1}} e^{2\pi i \phi(x;\omega)} f(\omega) \, d\omega$$

(that is, the extension operator associated with the graph of h) has a special scaling structure. Let $\bar{\omega} \in B^{n-1}$ and $\varrho \geqslant 1$, and note that

$$\phi(x; \bar{\omega} + \varrho^{-1}\omega) - \phi(x; \bar{\omega}) = \tilde{\phi}(\tilde{x}; \omega),$$

where $\tilde{\phi}$ is defined the same way as ϕ, but with h replaced by

$$\tilde{h}(\omega) := \varrho^2 (h(\bar{\omega} + \varrho^{-1}\omega) - h(\bar{\omega}) - \varrho^{-1} \langle \partial_\omega h(\bar{\omega}), \omega \rangle)$$

and $\tilde{x} := (\varrho^{-1} \partial_\omega \phi(x; \bar{\omega}), \varrho^{-2} x_n)$. The linear map $x \mapsto \tilde{x}$ is referred to as a *parabolic rescaling*, owing to the ϱ^{-1} and ϱ^{-2} scaling factors.

Now consider the special case where $h(\omega) = h_{\mathrm{par}}(\omega) := \frac{1}{2}|\omega|^2$ (so that $\phi = \phi_{\mathrm{par}}$), noting that

$$|\partial_\omega^\beta h_{\mathrm{par}}(\omega)| = 0 \quad \text{for all } \beta \in \mathbb{N}_0^n \text{ with } |\beta| \geqslant 3. \tag{4.8}$$

Write E_{par} for the operator E and observe that $\tilde{h}_{\mathrm{par}} = h_{\mathrm{par}}$ and, consequently, $\tilde{\phi}_{\mathrm{par}} = \phi_{\mathrm{par}}$. Thus, if $\mathrm{supp} f \subseteq B(\bar{\omega}, \varrho^{-1})$, then

$$|E_{\mathrm{par}} f(x)| = |E_{\mathrm{par}} \tilde{f}(\tilde{x})|,$$

where $\tilde{f}(\omega) := \varrho^{-(n-1)} f(\bar{\omega} + \varrho^{-1}\omega)$ is supported in B^{n-1}.

For general h, such a clean scaling identity does not hold. In particular, parabolic rescaling does not preserve E but transforms it into the operator $\widetilde{E}$ associated with the phase $\tilde{\phi}$: that is,

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

A useful feature, however, is that the new phase $\tilde{\phi}$ more closely resembles the prototypical example ϕ_{par} (relative to ϕ). In particular,

$$|\partial_\omega^\beta \tilde{h}(\omega)| \lesssim_\beta \varrho^{-(|\beta|-2)} \quad \text{for all } \beta \in \mathbb{N}_0^n \text{ with } |\beta| \geqslant 3,$$

so that, as ϱ is large, the function $\tilde{h}$ is 'closer' to satisfying (4.8) (relative to h).

These observations can be extended to the variable coefficient setting to prove the following reduction.

LEMMA 4.2. *To prove Theorem* 1.2 *for some fixed* $\varepsilon > 0$, *it suffices to consider the case where, in addition to the properties described in Lemma* 4.1, *the phase satisfies*

$$\|\partial_\omega^\beta \partial_x^\alpha \phi\|_{L^\infty(X \times \Omega)} < c_{\mathrm{par}} \quad \text{for } 1 \leqslant |\alpha| \leqslant N_{\mathrm{par}} \text{ and } 3 \leqslant |\beta| \leqslant N_{\mathrm{par}}$$

for some small constant c_{par} *and large integer* $N_{\mathrm{par}} \in \mathbb{N}$, *which can be chosen to depend on* n, p *and* ε. *Furthermore, one may assume that the amplitude satisfies*

$$\|\partial_\omega^\beta a\|_{L^\infty(X \times \Omega)} \lesssim_\beta 1 \quad \text{for all } 0 \leqslant |\beta| \leqslant N_{\mathrm{par}}.$$

Proof. One may assume that the phase of T^λ is given by

$$\phi^\lambda(x; \omega) := \lambda \phi\left(\frac{x}{\lambda}; \omega\right),$$

where

$$\phi(x; \omega) = \langle x', \omega \rangle + x_n h(\omega) + \mathcal{E}(x; \omega) \quad \text{for } (x; \omega) \in X \times \Omega.$$

Let $\varrho \geqslant 1$, $f \in L^1(B^{n-1})$ and cover B^{n-1} by finitely-overlapping ϱ^{-1}-balls. Provided ϱ is chosen to depend only on ϕ and $\varepsilon > 0$, by the triangle inequality one may assume that f is supported on one such ball, say $B(\bar\omega, \varrho^{-1})$, where $\bar\omega \in B^{n-1}$. Thus, by a linear change of variables,

$$|T^\lambda f(x)| = \left| \int_{B^{n-1}} e^{2\pi i (\phi^\lambda(x; \bar\omega + \varrho^{-1}\omega) - \phi^\lambda(x; \bar\omega))} a^\lambda(x; \bar\omega + \varrho^{-1}\omega) \tilde{f}(\omega) \, d\omega \right|,$$

where $\tilde{f}(\omega) := \varrho^{-(n-1)} f(\bar\omega + \varrho^{-1}\omega)$. The phase function appearing in the above oscillatory integral may be expressed as

$$\varrho^{-1} \langle (\partial_\omega \phi^\lambda)(x; \bar\omega), \omega \rangle + \varrho^{-2}\left(x_n \tilde{h}(\omega) + \lambda \tilde{\mathcal{E}}_1\left(\frac{x}{\lambda}; \omega\right) \right),$$

where

$$\tilde{h}(\omega) := \varrho^2 (h(\bar\omega + \varrho^{-1}\omega) - h(\bar\omega) - \varrho^{-1} \langle \partial_\omega h(\bar\omega), \omega \rangle)$$

and

$$\tilde{\mathcal{E}}_1(x; \omega) := \varrho^2 (\mathcal{E}(x; \bar\omega + \varrho^{-1}\omega) - \mathcal{E}(x; \bar\omega) - \varrho^{-1} \langle \partial_\omega \mathcal{E}(x; \bar\omega), \omega \rangle).$$

These definitions are motivated by the analysis of the constant coefficient case, as in the discussion prior to the statement of the lemma. Defining

$$\tilde{\mathcal{E}}(x; \omega) := \tilde{\mathcal{E}}_1(\Phi(\varrho^{-1}x', x_n; \bar\omega), x_n; \omega),$$

where Φ is the function introduced in (4.6), under the change of variables

$$x' \longmapsto \lambda \Phi\left(\frac{\varrho x'}{\lambda}, \frac{\varrho^2 x_n}{\lambda}; \bar\omega\right) \quad \text{and} \quad x_n \longmapsto \varrho^2 x_n,$$

the phase and amplitude are transformed into $\tilde{\phi}^{\lambda/\varrho^2}(x;\omega)$ and $\tilde{a}^{\lambda/\varrho^2}(x;\omega)$, respectively, where

$$\tilde{\phi}(x;\omega):=\langle x',\omega\rangle+x_n\tilde{h}(\omega)+\tilde{\mathcal{E}}(x;\omega)$$

and

$$\tilde{a}(x;\omega):=a(\Phi(\varrho^{-1}x',x_n;\bar{\omega}),x_n;\bar{\omega}+\varrho^{-1}\omega).$$

In particular, defining

$$\widetilde{T}^{\lambda/\varrho^2}g(x):=\int_{\mathbb{R}^{n-1}}e^{2\pi i\tilde{\phi}^{\lambda/\varrho^2}(x;\omega)}\tilde{a}^{\lambda/\varrho^2}(x;\omega)g(\omega)\,d\omega, \tag{4.9}$$

it follows that

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)}\lesssim\varrho^{(n+1)/p}\|\widetilde{T}^{\lambda/\varrho^2}\tilde{f}\|_{L^p(\mathbb{R}^n)}.$$

It is easy to verify that the phase $\tilde{\phi}$ satisfies the conditions of Lemma 4.1 and, provided ϱ is chosen appropriately (depending on ϕ and a), it also satisfies the additional conditions described in Lemma 4.2. The same is true for the amplitude $\tilde{a}$, except that the ω support has now been enlarged. However, by applying a partition of unity and translation argument, it is possible to assume without loss of generality that $\tilde{a}$ satisfies the desired support condition. This facilitates the reduction. $\qquad\square$

4.3. Controlling higher-order x derivatives

A final, elementary scaling argument allows one to control higher-order derivatives in x.

LEMMA 4.3. *To prove Theorem* 1.2 *for some fixed* $\varepsilon>0$, *it suffices to consider the case where, in addition to the properties described in Lemmas* 4.1 *and* 4.2, *the phase satisfies*

$$\|\partial_\omega^\beta\partial_x^\alpha\phi\|_{L^\infty(X\times\Omega)}<c_{\mathrm{par}}\quad\textit{for }2\leqslant|\alpha|\leqslant N_{\mathrm{par}}\textit{ and }0\leqslant|\beta|\leqslant N_{\mathrm{par}}. \tag{4.10}$$

Furthermore, one may assume that the amplitude satisfies

$$\|\partial_\omega^\beta\partial_x^\alpha a\|_{L^\infty(X\times\Omega)}\lesssim_{\alpha,\beta}1\quad\textit{for all }0\leqslant|\alpha|,|\beta|\leqslant N_{\mathrm{par}}.$$

Proof. Let T^λ be an operator associated with a phase ϕ and amplitude a satisfying the conditions of Lemma 4.2. Let $\varrho\geqslant1$ be a large constant, which will be chosen depending on a, ϕ, n and ε only, and define

$$\tilde{\phi}(x;\omega):=\varrho\phi\Big(\frac{x}{\varrho};\omega\Big)\quad\text{and}\quad\tilde{a}(x;\omega):=a\Big(\frac{x}{\varrho};\omega\Big).$$

One may easily verify that $\tilde{\phi}$ and $\tilde{a}$ satisfy the conditions in Lemma 4.2, except for an enlargement of the x-support which may be dealt with via a partition of unity. Furthermore,

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} = \|\widetilde{T}^{\lambda/\varrho} f\|_{L^p(\mathbb{R}^n)},$$

and so to prove L^p estimates for T^λ it suffices to prove corresponding estimates for $\widetilde{T}^\lambda$. Finally, provided ϱ is suitably chosen, it follows that $\tilde{\phi}$ and $\tilde{a}$ satisfy the additional conditions in the statement of Lemma 4.3. $\qquad\square$

Definition 4.4. Henceforth, $c_{\mathrm{par}} > 0$ and $N_{\mathrm{par}} \in \mathbb{N}$ are assumed to be fixed constants, chosen to satisfy the requirements of the forthcoming arguments. A positive-definite phase satisfying the properties of Lemma 4.3 for this choice of c_{par} and N_{par} is said to be *reduced*.

This notion of reduced positive-definite phase is precisely that which appears, then undefined, in the statement of the k-broad estimates of Theorem 1.9.

4.4. Geometric consequences

Henceforth, unless otherwise stated, all positive-definite phase functions ϕ are assumed to be reduced in the sense described above. The strategy is to obtain uniform estimates over this class of phases.

By the definition of Hörmander-type operators, for each $x \in X$ the map $\omega \mapsto \partial_x \phi(x; \omega)$ parameterises a smooth hypersurface Σ_x. In many respects, these hypersurfaces are geometrically very similar to the paraboloid $\omega \mapsto \left(\omega, \frac{1}{2}|\omega|^2\right)$. To see this, recall that

$$\Psi \colon U \longrightarrow \Omega$$

is a smooth function which satisfies

$$\partial_{x'}\phi(x; \Psi(x; u)) = u \tag{4.11}$$

for all $(x; u) \in U \subset X \times \mathbb{R}^{n-1}$. On each of the fibres $U_x := \{u \in \mathbb{R}^{n-1} : (x; u) \in U\}$ of the domain U, the map $u \mapsto \Psi(x; u)$ is a diffeomorphism. Thus, (4.11) implies that Σ_x is the graph of the function

$$h_x(u) := \partial_{x_n}\phi(x; \Psi(x; u))$$

over the fibre U_x. Each h_x is a perturbation of $\frac{1}{2}|u|^2$ in the following sense.

LEMMA 4.5. *The function h_x satisfies $h_x(0) = 0$, $\partial_u h_x(0) = 0$ and*

$$\|\partial_{uu}^2 h_x(u) - \mathrm{I}_{n-1}\|_{\mathrm{op}} = O(c_{\mathrm{par}})$$

for all $u \in U_x$.

Before proving the lemma, some simple properties of Ψ are recorded. By (4.5), it follows that $\Psi(x;0)=0$. The implicit function theorem implies that

$$\partial_u\Psi(x;u)=\partial^2_{x'\omega}\phi(x;\Psi(x;u))^{-1},$$

so that, by (4.3) and the local Lipschitz continuity of taking matrix inverses,

$$\|\partial_u\Psi(x;u)-\mathrm{I}_{n-1}\|_{\mathrm{op}}=O(c_{\mathrm{par}}). \tag{4.12}$$

As a consequence of this identity (and choosing c_{par} to be sufficiently small),

$$|\Psi(x;u)-\Psi(x;u')|\sim|u-u'| \quad \text{for all } u,u'\in U_x,$$

where the implied constant depends only on n. In addition, if $1\leqslant k\leqslant n-1$, then by twice differentiating $\partial_{x_k}\phi(x;\Psi(x;u))=u_k$ in the u variables, one may deduce that the kth coordinate Ψ_k of Ψ satisfies

$$\|\partial^2_{uu}\Psi_k(x;u)\|_{\mathrm{op}}=O(c_{\mathrm{par}}). \tag{4.13}$$

The stated properties of h_x now easily follow.

Proof of Lemma 4.5. By (4.5), one has

$$h_x(0)=\partial_{x_n}\phi(x;\Psi(x;0))=\partial_{x_n}\phi(x;0)=0.$$

Similarly, $\partial_u h_x(0)=0$ since $\partial_\omega\partial_{x_n}\phi(x;0)=0$. Finally,

$$\partial^2_{uu}h_x(u)=(\partial_u\Psi)^\top(x;u)(\partial^2_{\omega\omega}\partial_{x_n}\phi)(x;\Psi(x;u))\partial_u\Psi(x;u)+E(x;u),$$

where $E(x;\omega)$ is the $(n-1)\times(n-1)$ matrix whose (i,j)-th entry is given by

$$E_{ij}(x;u)=\langle(\partial_\omega\partial_{x_n}\phi)(x;\Psi(x;u)),\partial_{u_iu_j}\Psi(x;u)\rangle.$$

By (4.3) and (4.13), it follows that $\|E(x;u)\|_{\mathrm{op}}=O(c_{\mathrm{par}})$, whilst (4.4) and multiple applications of (4.12) imply that

$$\|(\partial_u\Psi)^\top(x;u)(\partial^2_{\omega\omega}\partial_{x_n}\phi)(x;\Psi(x;u))\partial_u\Psi(x;u)-\mathrm{I}_{n-1}\|_{\mathrm{op}}=O(c_{\mathrm{par}}).$$

This concludes the proof. $\qquad\square$

Similar reasoning can be used to provide useful uniform estimates for the generalised Gauss map associated with T^λ. To state the result, let

$$X^\lambda:=\left\{x\in\mathbb{R}^n:\frac{x}{\lambda}\in X\right\}$$

denote the λ-dilate of X, so that a^λ is supported in $X^\lambda\times\Omega$.

LEMMA 4.6. *For all* $x, \bar{x} \in X^\lambda$ *and* $\omega, \bar{\omega} \in \Omega$, *the estimates*

$$\measuredangle(G^\lambda(x;\omega), G^\lambda(x;\bar{\omega})) \sim |\omega - \bar{\omega}| \quad \text{and} \quad \measuredangle(G^\lambda(x;\omega), G^\lambda(\bar{x};\omega)) \lesssim \lambda^{-1}|x - \bar{x}|$$

hold with implied constants which depend only on the dimension.

The proof, which is an elementary calculus exercise in the style of the proof of Lemma 4.5, is omitted.

If the parameter x is restricted to a relatively small ball, then the second inequality in Lemma 4.6 often allows the Gauss map to be treated as if it were constant in x. This is consistent with the idea that T^λ is a small perturbation of a constant coefficient operator and can therefore be effectively approximated by constant coefficient operators at certain spatial scales.

5. Basic analytic preliminaries

5.1. Wave packet decomposition

Throughout the following sections, $\varepsilon > 0$ is a fixed small parameter and $\delta > 0$ is a tiny number satisfying[11] $\delta \ll \varepsilon$ and $\delta \sim_\varepsilon 1$.

A wave packet decomposition is carried out with respect to some spatial parameter $1 \ll R \ll \lambda$. Cover B^{n-1} by finitely-overlapping balls θ of radius $R^{-1/2}$ and let ψ_θ be a smooth partition of unity adapted to this cover. These θ will frequently be referred to as $R^{-1/2}$-*caps*. Cover $\mathbb{R}^{n-1}$ by finitely-overlapping balls of radius $CR^{(1+\delta)/2}$ centred on points belonging to the lattice $R^{(1+\delta)/2}\mathbb{Z}^{n-1}$. By Poisson summation, one may find a bump function adapted to $B(0, R^{(1+\delta)/2})$ so that the functions $\eta_v(z) := \eta(z - v)$ for $v \in R^{(1+\delta)/2}\mathbb{Z}^{n-1}$ form a partition of unity for this cover. Let $\mathbb{T}$ denote the collection of all pairs (θ, v). Thus, for $f : \mathbb{R}^{n-1} \to \mathbb{C}$ with support in B^{n-1} and belonging to some suitable a-priori class, one has

$$f = \sum_{(\theta,v) \in \mathbb{T}} (\eta_v(\psi_\theta f)\check{})\hat{} = \sum_{(\theta,v) \in \mathbb{T}} \hat{\eta}_v * (\psi_\theta f).$$

For each $R^{-1/2}$-cap θ, let $\omega_\theta \in B^{n-1}$ denote its centre. Choose a real-valued smooth function $\tilde{\psi}$ so that the function $\tilde{\psi}_\theta(\omega) := \tilde{\psi}(R^{1/2}(\omega - \omega_\theta))$ is supported in θ, and $\tilde{\psi}_\theta(\omega) = 1$ whenever ω belongs to a $cR^{-1/2}$ neighbourhood of the support of ψ_θ for some small constant $c > 0$. Finally, define

$$f_{\theta,v} := \tilde{\psi}_\theta \cdot [\hat{\eta}_v * (\psi_\theta f)].$$

---

[11] For $A, B \geqslant 0$ the notation $A \ll B$ or $B \gg A$ is used to denote that A is 'much smaller' than B; a more precise interpretation of this is that $A \leqslant C_\varepsilon^{-1}B$ for some constant $C_\varepsilon \geqslant 1$ which can be chosen to be large depending on n and ε.

Definition 5.1. The notation $\mathrm{RapDec}(R)$ is used to denote any quantity C_R which is rapidly decaying in R. More precisely, $C_R = \mathrm{RapDec}(R)$ if

$$|C_R| \lesssim_\varepsilon R^{-N} \quad \text{for all } N \leqslant \sqrt{N_{\mathrm{par}}}.$$

Here, N_{par} is the large integer appearing in the definition of reduced phase from Definition 4.4. By choosing N_{par} large, one may assume, say,

$$\varepsilon^{-10^n} \ll \sqrt{N_{\mathrm{par}}} \ll \varepsilon^{10^n} N_{\mathrm{par}}.$$

The function $\hat{\eta}_v(\omega)$ is rapidly decaying for $|\omega| \gtrsim R^{-(1+\delta)/2}$ and, consequently,

$$\|f_{\theta,v} - \hat{\eta}_v * (\psi_\theta f)\|_{L^\infty(\mathbb{R}^{n-1})} \leqslant \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}.$$

It follows that

$$\left\| f - \sum_{\theta,v} f_{\theta,v} \right\|_{L^\infty(\mathbb{R}^{n-1})} \leqslant \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}.$$

The functions $f_{\theta,v}$ are almost orthogonal: if $\mathbb{S} \subseteq \mathbb{T}$, then

$$\left\| \sum_{(\theta,v) \in \mathbb{S}} f_{\theta,v} \right\|_{L^2(\mathbb{R}^{n-1})}^2 \sim \sum_{(\theta,v) \in \mathbb{S}} \|f_{\theta,v}\|_{L^2(\mathbb{R}^{n-1})}^2.$$

Let T^λ be an operator with reduced phase ϕ and amplitude a supported in $X \times \Omega$ as in Lemma 4.1. For $(\theta, v) \in \mathbb{T}$ define the curve $\gamma^1_{\theta,v} \colon I^1_{\theta,v} \to \mathbb{R}^{n-1}$ by setting $\gamma^1_{\theta,v}(t) := \Phi(v, t; \omega_\theta)$, where Φ is the function introduced in §4 and

$$I^1_{\theta,v} := \{t \in X_n : \partial_\omega \phi(x', t; \omega_\theta) = v \text{ for some } x' \in X'\}.$$

Thus, $\partial_\omega \phi(\gamma^1_{\theta,v}(t), t; \omega_\theta) = v$ for all $t \in I^1_{\theta,v}$. Moreover, the rescaled curve

$$\gamma^\lambda_{\theta,v}(t) := \lambda \gamma^1_{\theta,v/\lambda}\left(\frac{t}{\lambda}\right)$$

satisfies

$$\partial_\omega \phi^\lambda(\gamma^\lambda_{\theta,v}(t), t; \omega_\theta) = v \quad \text{for all } t \in I^\lambda_{\theta,v} := \left\{ t \in \mathbb{R} : \frac{t}{\lambda} \in I^1_{\theta,v/\lambda} \right\}.$$

Let $\Gamma^\lambda_{\theta,v} \colon I^\lambda_{\theta,v} \to \mathbb{R}^n$ denote the graphing map $\Gamma^\lambda_{\theta,v}(t) := (\gamma^\lambda_{\theta,v}(t), t)$; by an abuse of notation, $\Gamma^\lambda_{\theta,v}$ will also be used to denote the image of this mapping. The geometry of the curves $\Gamma^\lambda_{\theta,v}$ is related to the generalised Gauss map G^λ by the following elementary lemma.

LEMMA 5.2. *The tangent space* $T_{\Gamma^\lambda_{\theta,v}(t)}\Gamma^\lambda_{\theta,v}$ *lies in the direction of the unit vector* $G^\lambda(\Gamma^\lambda_{\theta,v}(t); \omega_\theta)$ *for all* $t \in I^\lambda_{\theta,v}$.

Proof. Differentiating the equation $\partial_\omega\phi(\Gamma^\lambda_{\theta,v}(t);\omega_\theta)=v$, it follows that

$$\partial^2_{\omega x}\phi^\lambda(\Gamma^\lambda_{\theta,v}(t);\omega_\theta)(\Gamma^\lambda_{\theta,v})'(t)=0.$$

Thus, $(\Gamma^\lambda_{\theta,v})'(t)$ must be parallel to $G^\lambda(\Gamma^\lambda_{\theta,v}(t);\omega_\theta)$ since, by definition, the latter vector spans the kernel of $\partial^2_{\omega x}\phi^\lambda(\Gamma^\lambda_{\theta,v}(t);\omega_\theta)$. $\qquad\square$

Define the curved $R^{1/2+\delta}$-tube as

$$T_{\theta,v}:=\{(x',x_n)\in B(0,R):x_n\in I^\lambda_{\theta,v}\text{ and }|x'-\gamma^\lambda_{\theta,v}(x_n)|\leqslant R^{1/2+\delta}\}.$$

The curve $\Gamma^\lambda_{\theta,v}$ is referred to as the *core* of $T_{\theta,v}$. Observe that, since ϕ is of the reduced form defined in §4, one has

$$|x'-\gamma^\lambda_{\theta,v}(x_n)|\sim|\partial_\omega\phi^\lambda(x;\omega_\theta)-v|, \tag{5.1}$$

for all $x=(x',x_n)\in X^\lambda$ with $x_n\in I^\lambda_{\theta,v}$ (uniformly in λ).

Example 5.3. Let $\phi^\lambda(x;\omega):=\langle x',\omega\rangle+x_n h(\omega)$ and observe that $\gamma^\lambda_{\theta,v}(t)=v-t\partial_\omega h(\omega_\theta)$ parameterises a straight line through v in the direction of $\partial_\omega h(\omega_\theta)$. The tube is given by

$$T_{\theta,v}=\{x\in B(0,R):|x'+x_n\partial_\omega h(\omega_\theta)-v|\leqslant R^{1/2+\delta}\},$$

which agrees with those studied in the case of the extension operator.

LEMMA 5.4. *If* $1\ll R\ll\lambda$ *and* $x\in B(0,R)\backslash T_{\theta,v}$, *then*

$$|T^\lambda f_{\theta,v}(x)|\leqslant(1+R^{-1/2}|\partial_\omega\phi^\lambda(x;\omega_\theta)-v|)^{-(n+1)}\,\mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}. \tag{5.2}$$

Proof. Observe that

$$T^\lambda f_{\theta,v}(x)=\int_{\mathbb{R}^{n-1}}\hat{\eta}_v*(\psi_\theta f)\cdot\overline{G_x},$$

where

$$G_x(\omega):=e^{-2\pi i\phi^\lambda(x;\omega)}a^\lambda(x;\omega)\tilde{\psi}_\theta(\omega)$$

and so, by Plancherel,

$$T^\lambda f_{\theta,v}(x)=\int_{\mathbb{R}^{n-1}}\eta_v\cdot(\psi_\theta f)^\vee\cdot\overline{\check{G}_x}.$$

By a simple change of variables, if one defines the phase and amplitude functions

$$\phi^{\lambda,R}_{x,z}(\omega):=R^{1/2}\phi^\lambda(x;\omega_\theta+R^{-1/2}\omega)-\langle z,\omega\rangle,$$
$$a^{\lambda,R}_x(\omega):=a^\lambda(x;\omega_\theta+R^{-1/2}\omega)\tilde{\psi}(\omega),$$

then

$$|\check{G}_x(z)|=R^{-(n-1)/2}\left|\int_{\mathbb{R}^{n-1}}e^{-2\pi iR^{-1/2}\phi^{\lambda,R}_{x,z}(\omega)}a^{\lambda,R}_x(\omega)\,d\omega\right|.$$

This integral is analysed using (non-)stationary phase.

CLAIM. *Fixing $x \in B(0,R) \setminus T_{\theta,v}$, $z \in \operatorname{supp} \eta_v$ and $R \gg 1$, the following estimates hold on* $\operatorname{supp} a_x^{\lambda,R}$:

(i) $R^{-1/2}|\partial_\omega \phi_{x,z}^{\lambda,R}(\omega)| \sim R^{-1/2}|\partial_\omega \phi^\lambda(x;\omega_\theta) - v| \gtrsim R^\delta$;

(ii) $|\partial_\omega^\alpha \phi_{x,z}^{\lambda,R}(\omega)| \lesssim |\partial_\omega \phi_{x,z}^{\lambda,R}(\omega)|$ *for all* $2 \leqslant |\alpha| \leqslant N_{\mathrm{par}}$;

(iii) $|\partial_\omega^\alpha a_x^{\lambda,R}(\omega)| \lesssim_\varepsilon 1$ *for all* $|\alpha| \leqslant N_{\mathrm{par}}$.

Once the claim is established, repeated integration-by-parts yields

$$|\breve{G}_x(z)| = (1 + R^{-1/2}|\partial_\omega \phi^\lambda(x;\omega_\theta) - v|)^{-(n+1)} \operatorname{RapDec}(R). \tag{5.3}$$

Such integration-by-parts or non-stationary phase arguments are standard but, for the reader's convenience (and since arguments of this kind will feature repeatedly in the article), the details are appended. The precise result used here is Lemma A.1. The desired inequality (5.2) is an immediate consequence of (5.3) together with Young's inequality and Plancherel's theorem.

It remains to establish the claim. Here the stated uniformity in the estimates is a consequence of the reductions made in §4. In view of Lemma 4.2, the bound (iii) for the amplitude is immediate, and so it remains to show the bounds for the phase.

Proof of (i)

Computing the derivative of the phase function,

$$\begin{aligned}
\partial_\omega \phi_{x,z}^{\lambda,R}(\omega) &= \partial_\omega \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - z \\
&= [\partial_\omega \phi^\lambda(x;\omega_\theta) - v] + [v - z] + [\partial_\omega \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(x;\omega_\theta)].
\end{aligned}$$

Observe that $|v - z| \lesssim R^{(1+\delta)/2}$ whenever $z \in \operatorname{supp} \eta_v$. Moreover,

$$|\partial_{\omega_j} \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - \partial_{\omega_j}\phi^\lambda(x;\omega_\theta)| \leqslant R^{-1/2} \int_0^1 |\langle \partial_\omega \partial_{\omega_j} \phi^\lambda(x; \omega_\theta + tR^{-1/2}\omega), \omega\rangle|\, dt.$$

Since $\partial_{\omega_i\omega_j}^2 \phi^\lambda(0;\omega) = 0$ for all $\omega \in \Omega$ and $1 \leqslant i,j \leqslant n-1$, it follows that

$$\begin{aligned}
|\partial_{\omega_i\omega_j}^2 \phi^\lambda(x; \omega_\theta + tR^{-1/2}\omega)| &\leqslant \int_0^1 |\langle \partial_x \partial_{\omega_i\omega_j}^2 \phi^\lambda(sx; \omega_\theta + tR^{-1/2}\omega), x\rangle|\, ds \\
&\leqslant \|\partial_x \partial_{\omega_i\omega_j}^2 \phi^\lambda\|_{L^\infty(X^\lambda \times \Omega)} |x| \lesssim R
\end{aligned} \tag{5.4}$$

for $x \in B(0,R) \cap X^\lambda$, where the uniformity in the last inequality is due to Lemma 4.1. Combining these estimates,

$$|\partial_\omega \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(x;\omega_\theta)| \lesssim R^{1/2}.$$

On the other hand, for $x \in B(0,R) \setminus T_{\theta,v}$, it is claimed that

$$|\partial_\omega \phi^\lambda(x; \omega_\theta) - v| \gtrsim R^{1/2+\delta}. \tag{5.5}$$

If $x_n \in I_{\theta,v}^\lambda$, then (5.5) follows directly from the definition of $T_{\theta,v}$ and (5.1). Temporarily assuming that (5.5) holds in general, it follows that the $\partial_\omega \phi^\lambda(x; \omega_\theta) - v$ term dominates in the above expansion of $\partial_\omega \phi_{x,z}^{\lambda,R}(\omega)$. In particular, for all $z \in \operatorname{supp} \eta_v$, one concludes that

$$R^{-1/2} |\partial_\omega \phi_{x,z}^{\lambda,R}(\omega)| \sim R^{-1/2} |\partial_\omega \phi^\lambda(x; \omega_\theta) - v| \gtrsim R^\delta,$$

whenever $R \gg 1$.

It remains to establish (5.5) for $x \in B(0,R) \setminus T_{\theta,v}$ with $x_n \notin I_{\theta,v}^\lambda$. The condition on x_n implies that $v \notin \widetilde{X}'$, where $\widetilde{X}'$ is defined to be the image of X' under the diffeomorphism

$$z' \longmapsto \lambda \partial_\omega \phi\left(z', \frac{x_n}{\lambda}; \omega_\theta\right).$$

The set $\widetilde{X}'$ will contain a ball centred at $\lambda \partial_\omega \phi(0, x_n/\lambda; \omega_\theta)$ of radius $2c\lambda$ for some dimensional constant $c > 0$. Since $|x_n| < R$ and $\partial_\omega \phi(0; \omega_\theta) = 0$, one observes that

$$\lambda \left| \partial_\omega \phi\left(0, \frac{x_n}{\lambda}; \omega_\theta\right) \right| \lesssim R,$$

and so $B(0, c\lambda) \subseteq \widetilde{X}'$, provided $R \ll \lambda$. Consequently, $|v| \gtrsim \lambda$, whilst, on the other hand,

$$|\partial_\omega \phi^\lambda(x; \omega_\theta)| \lesssim R$$

and so, again provided $R \ll \lambda$, the estimate (5.5) immediately follows.([12])

Proof of (ii)

Fix $\alpha \in \mathbb{N}_0^n$ with $2 \leqslant |\alpha| \leqslant N_{\mathrm{par}}$. By arguing as in (5.4), one obtains

$$\begin{aligned} |\partial_\omega^\alpha \phi_{x,z}^{\lambda,R}(\omega)| &= R^{-(|\alpha|-1)/2} |(\partial_\omega^\alpha \phi^\lambda)(x; \omega_\theta + R^{-1/2}\omega)| \\ &\leqslant R^{-(|\alpha|-1)/2} \|\partial_x \partial_\omega^\alpha \phi^\lambda\|_{L^\infty(X^\lambda \times \Omega)} |x| \lesssim R^{1/2}, \end{aligned}$$

where the uniformity in the last inequality is due to Lemmas 4.1 and 4.2. Since

$$|\partial_\omega \phi_{x,z}^{\lambda,R}(\omega)| \geqslant R^{1/2}$$

by (i), this concludes the proof of the claim and of Lemma 5.4. $\square$

([12]) This argument can also be used to show that, for $|v| \lesssim \lambda$, the domain $I_{\theta,v}^\lambda$ contains an interval about zero of length $\sim \lambda$.

5.2. An L^2 estimate

The following standard L^2-bound, which has been mentioned previously in §2, will play a significant role in the forthcoming analysis.

LEMMA 5.5. (Hörmander [17]) *If* $1 \leqslant R \leqslant \lambda$ *and* B_R *is any ball of radius* R, *then*

$$\|T^\lambda f\|_{L^2(B_R)} \lesssim R^{1/2} \|f\|_{L^2(B^{n-1})}.$$

This lemma is a direct corollary of the following lemma which, in turn, is a consequence of Hörmander's generalisation of the Hausdorff–Young inequality [17].

LEMMA 5.6. *For any fixed* $x_n \in \mathbb{R}$, *the estimate*

$$\|T^\lambda f\|_{L^2(\mathbb{R}^{n-1} \times \{x_n\})} \lesssim \|f\|_{L^2(B^{n-1})}$$

holds.

Proof. Defining $Sf(x') := T^\lambda f(\lambda x', x_n)$, the problem is to show that

$$\|Sf\|_{L^2(\mathbb{R}^{n-1})} \lesssim \lambda^{-(n-1)/2} \|f\|_{L^2(\mathbb{R}^{n-1})}. \tag{5.6}$$

Observe that

$$Sf(x') = \int_{\mathbb{R}^{n-1}} e^{2\pi i \lambda \phi(x', x_n/\lambda; \omega)} a\left(x', \frac{x_n}{\lambda}; \omega\right) f(\omega) \, d\omega.$$

The original hypotheses on the phase ϕ imply that

$$\left| \det \partial^2_{x'\omega} \phi\left(x', \frac{x_n}{\lambda}; \omega\right) \right| \gtrsim 1,$$

whilst $(x'; \omega) \mapsto a(x', x_n/\lambda; \omega)$ has support in some bounded subset of $\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$. Both these conditions hold uniformly in x_n and λ. Thus, the operator S satisfies the conditions of Hörmander's generalisation of the Hausdorff–Young inequality [17] (see also, for instance, [28, p. 377]), uniformly in x_n and λ. Applying Hörmander's theorem immediately yields (5.6). $\qquad \square$

5.3. The locally constant property

As a final analytic preliminary, some simple consequences of the uncertainty principle are discussed. It is remarked that the result of this subsection (that is, Lemma 5.8) only plays a role in the proof of Theorem 1.2 much later in the argument (namely, in the parabolic rescaling argument in §11). It does, however, feature in an independent discussion in the following section.

Definition 5.7. A function $\zeta \colon \mathbb{R}^n \to [0, \infty)$ is said to be *locally constant at scale* ϱ for some $\varrho > 0$ if $\zeta(x) \sim \zeta(y)$ for all $x, y \in \mathbb{R}^n$ with $|x - y| \lesssim \varrho$.

Owing to the uncertainty principle, heuristically one expects the following: if f is supported on a ϱ^{-1}-cap, then $|T^\lambda f|$ is essentially constant at scale ϱ. For extension operators this is due to the fact that, under the support hypothesis on the input function, Ef has (distributional) Fourier support inside a ϱ^{-1}-ball. For general Hörmander-type operators T^λ the Fourier transform of $T^\lambda f$ does not necessarily have compact support. It will, however, be concentrated in some ϱ^{-1}-ball and this is sufficient to ensure the locally constant property holds. This discussion is formalised by the following lemma.

LEMMA 5.8. *Let T^λ be a Hörmander-type operator and $1 \leqslant k \leqslant n$. There exists a smooth, rapidly decreasing function $\zeta \colon \mathbb{R}^n \to [0, \infty)$ with the following properties*:
  (1) *ζ is locally constant at scale 1.*
  (2) *If $\delta > 0$ and $1 \leqslant \varrho \leqslant \lambda^{1-\delta}$, then the pointwise inequality*

$$|T^\lambda f|^{1/k} \lesssim |T^\lambda f|^{1/k} * \zeta_\varrho + \operatorname{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}^{1/k}$$

holds whenever f is supported in some ϱ^{-1}-ball. Here, $\zeta_\varrho(x) := \varrho^{-n} \zeta(x/\varrho)$.

It is useful to work with the parameter k here, in order to apply the locally constant property effectively in k-linear settings.

The locally constant property of ζ implies that

$$|T^\lambda f|^{1/k} * \zeta_\varrho(x) \sim |T^\lambda f|^{1/k} * \zeta_\varrho(y) \quad \text{for all } x, y \in \mathbb{R}^n \text{ with } |x - y| \lesssim \varrho;$$

namely, $|T^\lambda f|^{1/k} * \zeta_\varrho$ is locally constant at scale ϱ. This is a rigorous formulation of the locally constant heuristic discussed above.

Proof. Suppose that $\operatorname{supp} f \subset B(\bar{\omega}, \varrho^{-1})$, where $\bar{\omega} \in \Omega$, and observe that

$$[e^{-2\pi i \phi^\lambda(\,\cdot\,;\bar{\omega})} T^\lambda f]\widehat{\;}(\xi) = \int_{\mathbb{R}^{n-1}} K^\lambda(\xi; \omega) f(\omega)\, d\omega,$$

where the function K^λ is given by

$$K^\lambda(\xi; \omega) = \lambda^n \int_{\mathbb{R}^n} e^{-2\pi i \lambda(\langle x, \xi \rangle - \phi(x;\omega) + \phi(x;\bar{\omega}))} a(x; \omega)\, dx.$$

This oscillatory integral is estimated via (non-)stationary phase, using the simple estimate

$$|\partial_x \phi(x; \omega) - \partial_x \phi(x; \bar{\omega})| \lesssim \varrho^{-1} \quad \text{for } (x; \omega) \in X \times \Omega \text{ with } \omega \in \operatorname{supp} f.$$

In particular, if $|\xi| \geqslant C\varrho^{-1}$ for a suitably large constant $C \geqslant 1$, then repeated integration-by-parts, combined with the control on the derivatives of a ensured by Lemma 4.3, shows that

$$|K^\lambda(\xi;\omega)| \leqslant \mathrm{RapDec}(\lambda)(1+|\xi|)^{-(n+1)}.$$

Let η be a Schwartz function on $\mathbb{R}^n$ with $\hat{\eta}(\xi)=1$ for all $|\xi|<C$ for a suitable constant $C \geqslant 1$ and support in $B(0,2C)$. Such a function can further be chosen so that $|\eta|^{1/k}$ admits a smooth, rapidly decreasing majorant ζ which is locally constant at scale 1. From the above observations,

$$[e^{-2\pi i\phi^\lambda(\cdot;\bar\omega)}T^\lambda f]\widehat{\ }(\xi) = [e^{-2\pi i\phi^\lambda(\cdot;\bar\omega)}T^\lambda f]\widehat{\ }(\xi)\hat{\eta}_\varrho(\xi) + E(f,\lambda)(\xi),$$

where $|E(f,\lambda)(\xi)| \leqslant \mathrm{RapDec}(\lambda)(1+|\xi|)^{-(n+1)}\|f\|_{L^2(B^{n-1})}$. Applying Fourier inversion and using the triangle inequality to estimate the error,

$$e^{-2\pi i\phi^\lambda(x;\bar\omega)}T^\lambda f(x) = [e^{-2\pi i\phi^\lambda(\cdot;\bar\omega)}T^\lambda f]*\eta_\varrho(x) + \mathrm{RapDec}(\lambda)\|f\|_{L^2(B^{n-1})}$$

and, in particular,

$$|T^\lambda f(x)| \leqslant \int_{\mathbb{R}^n} |T^\lambda f(x-y)\eta_\varrho(y)|\,dy + \mathrm{RapDec}(\lambda)\|f\|_{L^2(B^{n-1})}.$$

Observe that, for fixed x, the function appearing in absolute values in the above integrand has Fourier support in a ball of radius $O(\varrho^{-1})$. Bernstein's inequality[13] may therefore be applied to dominate the right-hand side by

$$\left(\int_{\mathbb{R}^n} |T^\lambda f(x-y)|^{1/k}\zeta_\varrho(y)\,dy\right)^k + \mathrm{RapDec}(\lambda)\|f\|_{L^2(B^{n-1})},$$

which concludes the proof. $\qquad\square$

6. Properties of the k-broad norms

6.1. k-broad triangle inequality and logarithmic convexity

The functional $f \mapsto \|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U)}$ is not a norm in a literal sense, but it does exhibit some properties similar to those of L^p-norms. For instance, the map $U \mapsto \|T^\lambda f\|^p_{\mathrm{BL}^p_{k,A}(U)}$ behaves similarly to a measure.

---

[13] More precisely, here the proof uses a general form of Bernstein's inequality, valid for exponents less than 1. In particular, if $0<p\leqslant q\leqslant\infty$ and g is an integrable function on $\mathbb{R}^n$ satisfying $\mathrm{supp}\,\hat{g}\subseteq B_r$, then

$$\|g\|_{L^q(\mathbb{R}^n)} \lesssim r^{n(1/p-1/q)}\|g\|_{L^p(\mathbb{R}^n)}.$$

This extension follows from the classical Bernstein inequality (that is, the above estimate in the restricted range $1\leqslant p\leqslant q\leqslant\infty$) in a rather straightforward manner. The classical Bernstein inequality is itself a direct consequence of Young's convolution inequality; see, for instance, [36, §5].

LEMMA 6.1. (Finite (sub-)additivity) *Let $U_1, U_2 \subseteq \mathbb{R}^n$ and $U := U_1 \cup U_2$. If $1 \leqslant p < \infty$ and A is a non-negative integer, then*

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U)}^p \leqslant \|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U_1)}^p + \|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U_2)}^p$$

holds for all integrable $f \colon B^{n-1} \to \mathbb{C}$.

This is an immediate consequence of the definition of the k-broad norms. A slightly less trivial observation is that $\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U)}$ also satisfies weak versions of the triangle and logarithmic convexity inequalities.

LEMMA 6.2. (Triangle inequality [14]) *If $U \subseteq \mathbb{R}^n$, $1 \leqslant p < \infty$ and $A := A_1 + A_2$ for A_1 and A_2 non-negative integers, then*

$$\|T^\lambda (f_1 + f_2)\|_{\mathrm{BL}^p_{k,A}(U)} \lesssim \|T^\lambda f_1\|_{\mathrm{BL}^p_{k,A_1}(U)} + \|T^\lambda f_2\|_{\mathrm{BL}^p_{k,A_2}(U)}$$

holds for all integrable $f_1, f_2 \colon B^{n-1} \to \mathbb{C}$.

LEMMA 6.3. (Logarithmic convexity [14]) *Suppose that $U \subseteq \mathbb{R}^n$, $1 \leqslant p, p_1, p_2 < \infty$ and $0 \leqslant \alpha_1, \alpha_2 \leqslant 1$ satisfy $\alpha_1 + \alpha_2 = 1$ and*

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

If $A := A_1 + A_2$ for A_1 and A_2 non-negative integers, then

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(U)} \lesssim \|T^\lambda f\|_{\mathrm{BL}^{p_1}_{k,A_1}(U)}^{\alpha_1} \|T^\lambda f\|_{\mathrm{BL}^{p_2}_{k,A_2}(U)}^{\alpha_2}$$

holds for all integrable $f \colon B^{n-1} \to \mathbb{C}$.

These estimates are proven in the context of Fourier extension operators in [14]. The arguments are entirely elementary and readily generalise to the variable coefficient case. It is remarked that the parameter A appears in the definition of the k-broad norm to allow for these weak triangle and logarithmic convexity inequalities.

6.2. k-broad versus k-linear

A relationship between k-broad and k-linear estimates is given by the following proposition.

PROPOSITION 6.4. *Let $\mathcal{T}$ be a class of Hörmander-type operators which is closed under translation,*[14] $2 \leqslant p < \infty$, $2 \leqslant k \leqslant n$ *and* $\varepsilon > 0$. *Suppose that, for all* $1 \ll R \leqslant \lambda$ *and R-balls B_R, the k-linear inequality*

$$\left\| \prod_{j=1}^{k} |T_j^\lambda f_j|^{1/k} \right\|_{L^p(B_R)} \lesssim_{\varepsilon, (\phi_j)_{j=1}^k} \nu^{-C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})}$$

holds whenever $(T_1^\lambda, ..., T_k^\lambda) \in \mathcal{T}^k$ is a ν-transverse k-tuple of Hörmander-type operators. Then, for all $1 \ll R \leqslant \lambda$ and R-balls B_R, the k-broad inequality

$$\|T^\lambda f\|_{\mathrm{BL}_{k,1}^p(B_R)} \lesssim_{\varepsilon, \phi} K^{C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})}$$

holds for any $T^\lambda \in \mathcal{T}$.

Recall the notion of ν-transversality was introduced in Definition 1.6. The parameter K in the above theorem is the same as that which appears in the definition of the k-broad norms; the C_ε denote constants, which may vary from line to line, which depend only on n and ε.

The (local version of the) Bennett–Carbery–Tao theorem [4] therefore implies a version of Theorem 1.9 which holds for all Hörmander-type operators (that is, without the positive-definite hypothesis) with a restricted range of p.[15]

COROLLARY 6.5. *Let T^λ be a Hörmander-type operator. For all $2 \leqslant k \leqslant n$, all $p \geqslant 2k/(k-1)$ and all $\varepsilon > 0$ the estimate*

$$\|T^\lambda f\|_{\mathrm{BL}_{k,1}^p(B_R)} \lesssim_{\varepsilon, \phi} K^{C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})}$$

holds for all $\lambda \geqslant 1$.

For completeness, the proof of Proposition 6.4 is given; the result itself will not be used in the proof of Theorem 1.2 and is included mainly for expository purposes. Thus, readers interested only in the proof of Theorem 1.2 may safely skip to the next section.

Proof of Proposition 6.4. Let $\mathcal{Z} \subset B_R$ be a maximal set of points with the property that the balls $B(z, R/2\overline{C}K)$ for $z \in \mathcal{Z}$ are pairwise disjoint. Here, $\overline{C} \geqslant 1$ is a suitable constant, chosen to meet the forthcoming requirements of the proof. Letting

$$B_z := B\left(z, \frac{R}{\overline{C}K}\right) \quad \text{for } z \in \mathcal{Z},$$

---

[14] That is, if $T^\lambda \in \mathcal{T}$ and $a \in \mathbb{R}^n$, then the translated operator T_a^λ defined by $T_a^\lambda f(x) := T^\lambda f(x+a)$ also belongs to $\mathcal{T}$.

[15] The version of the Bennett–Carbery–Tao theorem used here is not explicitly stated in [4] (there, the variable coefficient estimates are only presented at the n-linear level). Nevertheless, k-linear inequalities for Hörmander-type operators are readily obtained by combining the analysis of §5 and §6 of [4]; see [9, §5].

it follows that $\#\mathcal{Z} \lesssim K^n$ and

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,1}(B_R)}^p \leqslant \sum_{z \in \mathcal{Z}} \|T^\lambda f\|_{\mathrm{BL}^p_{k,1}(B_z)}^p.$$

Fixing $z \in \mathcal{Z}$, it therefore suffices to show that

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,1}(B_z)} \lesssim_\varepsilon K^{C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})},$$

since summing the contributions from each choice of $z \in \mathcal{Z}$ only introduces an acceptable K^n factor into the estimate. By introducing a bump function into the definition of the operator, one may further assume that the amplitude a^λ has x-support in the ball concentric to B_z with twice the radius.

Fix a ball $B_{K^2} = B(\bar{x}, K^2) \in \mathcal{B}_{K^2}$, with $B_{K^2} \cap B_z \neq \varnothing$, and suppose that V is a $(k-1)$-dimensional subspace which realises the minimum in $\mu_{T^\lambda f}(B_{K^2})$. Thus, by definition, if τ is a K^{-1}-cap for which

$$\|T^\lambda f_\tau\|_{L^p(B_{K^2})}^p > \mu_{T^\lambda f}(B_{K^2}),$$

then $\tau \in V$, where the inclusion symbol is used in the non-standard sense described in the introduction. Amongst all such subspaces V, choose one which maximises the cardinality of the set

$$\mathcal{T}(V) := \{\tau \in V : \|T^\lambda f_\tau\|_{L^p(B_{K^2})}^p \geqslant \mu_{T^\lambda f}(B_{K^2})\}.$$

By definition, there exists some cap $\tau^* \notin V$ such that $\|T^\lambda f_{\tau^*}\|_{L^p(B_{K^2})}^p = \mu_{T^\lambda f}(B_{K^2})$.

Suppose there exists a $(k-2)$-dimensional subspace $W \subset \mathbb{R}^n$ such that $\tau \in W$ for all $\tau \in \mathcal{T}(V)$. Then, defining $V' := \mathrm{span}(W \cup \{G^\lambda(\bar{x}, \omega_{\tau^*})\})$, where ω_{τ^*} is the centre of τ^*, it follows that $\tau^* \in V'$ and $\tau \in V'$ for all $\tau \in \mathcal{T}(V)$. On the other hand, V' also realises the minimum in the definition in $\mu_{T^\lambda f}(B_{K^2})$, since

$$\|T^\lambda f_\tau\|_{L^p(B_{K^2})}^p \leqslant \mu_{T^\lambda f}(B_{K^2}) \quad \text{for all } \tau \notin V' \text{ with } \tau \in V;$$

this is immediate by the fact that, if $\tau \notin V'$, then $\tau \notin W$, so τ does not belong to $\mathcal{T}(V)$. These observations contradict the maximality of V and, consequently, no such subspace W can exist.

By the preceding discussion, one may find a family of caps $\tau_1^*, ..., \tau_{k-1}^* \in \mathcal{T}(V)$ satisfying

$$\left| \bigwedge_{j=1}^k G^\lambda(\bar{x}, \omega_j) \right| \gtrsim K^{-(k-1)} \quad \text{for all } \omega_j \in \tau_j^*, \, 1 \leqslant j \leqslant k. \tag{6.1}$$

Thus,

$$\mu_{T^\lambda f}(B_{K^2}) \leqslant \prod_{j=1}^{k} \|T^\lambda f_{\tau_j^*}\|_{L^p(B_{K^2})}^{p/k} \tag{6.2}$$

for $\tau_k^* := \tau^*$. To apply the hypothesised multilinear estimate, one wishes to exchange the order of taking the norm and product on the right-hand side of the above expression; that is, one wishes to prove an estimate of the form

$$\prod_{j=1}^{k} \|T^\lambda f_{\tau_j^*}\|_{L^p(B_{K^2})}^{p/k} \lesssim K^{O(1)} \left\| \prod_{j=1}^{k} |T^\lambda f_{\tau_j^*}|^{1/k} \right\|_{L^p(B_{K^2})}^{p}.$$

This is achieved by exploiting the locally constant property of the $T^\lambda f_\tau$, as discussed in §5.3. In particular, by Lemma 5.8 and Hölder's inequality, there exists a non-negative, rapidly decreasing, locally constant function ζ such that

$$|T^\lambda f_\tau|^{p/k} \lesssim |T^\lambda f_\tau|^{p/k} * \zeta_K + \mathrm{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}^{p/k} \tag{6.3}$$

holds for all K^{-1}-caps τ. Since rapidly decaying error terms are entirely harmless, henceforth they will be suppressed in the notation. Observe that, for all $z \in B(\bar{x}, K^2)$ and $y \in \mathbb{R}^n$, one has $\zeta_K(z-y) \lesssim K^{O(1)} w_K(\bar{x}-y)$, where $w_K(y) := (1+|y|/K)^{-N}$ for some suitable choice of large exponent N satisfying $N = O(1)$. Combining these observations with a second application of (6.3) yields

$$\|T^\lambda f_\tau\|_{L^p(B_{K^2})}^{p/k} \lesssim K^{O(1)} \int_{\mathbb{R}^n} |T^\lambda f_\tau|^{p/k} * \zeta_K(y) w_K(\bar{x}-y) \, dy. \tag{6.4}$$

Temporarily fix $x \in B(0,K)$ and note that the locally constant property of ζ implies that

$$|T^\lambda f_\tau|^{p/k} * \zeta_K(y) \lesssim |T^\lambda f_\tau|^{p/k} * \zeta_K(x+y) \quad \text{for all } y \in \mathbb{R}^n. \tag{6.5}$$

Thus, by (6.2), (6.4) and (6.5), one deduces that

$$\mu_{T^\lambda f}(B_{K^2}) \lesssim K^{O(1)} \int_{(\mathbb{R}^n)^k} \prod_{j=1}^{k} |T^\lambda f_{\tau_j^*}|^{p/k} * \zeta_K(x+y_j) w_K(\bar{x}-y_j) \, d\mathbf{y}.$$

Taking the average of both sides of this estimate over all $x \in B(0,K)$ and shifting the y_j variables,

$$\mu_{T^\lambda f}(B_{K^2}) \lesssim K^{O(1)} \int_{(\mathbb{R}^n)^{2k}} \int_{B(\bar{x},K)} \prod_{j=1}^{k} |T^\lambda f_{\tau_j^* \circ \sigma_{y_j,z_j}}(x)|^{p/k} \, dx \, Z_K(\mathbf{y},\mathbf{z}) \, d\mathbf{y} \, d\mathbf{z},$$

where

$$Z_K(\mathbf{y}, \mathbf{z}) := \prod_{j=1}^{k} w_K(y_j) \zeta_K(z_j)$$

and

$$\sigma_{y_j, z_j}(x) := x + y_j - z_j.$$

Since both w_K and ζ_K decay (at least) as rapidly as $|y|^{-N}$ away from $B(0, K)$, one may restrict the integral in $\mathbf{y}$ and $\mathbf{z}$ from the whole space $(\mathbb{R}^n)^{2k}$ to the bounded region $B(0, \lambda/\overline{C}K)^{2k}$ at the expense of an additional harmless error term.

It is possible to localise to a finer scale than $\lambda/\overline{C}K$, but this scale suffices for the purposes of the proof. In particular, if $|y|, |z| < \lambda/\overline{C}K$, then it follows from Lemma 4.6 that

$$|G^\lambda(x, \omega) - G^\lambda(\sigma_{y,z}(x), \omega)| \lesssim \frac{|x - \sigma_{y,z}(x)|}{\lambda} \lesssim \overline{C}^{-1} K^{-1} \quad \text{for all } (x, \omega) \in \operatorname{supp} a. \quad (6.6)$$

If $\overline{C}$ is chosen sufficiently large, then this bound is enough to ensure that pre-composing by $\sigma_{y,z}$ preserves certain transversality properties, as discussed below.

Given a K^{-1}-cap τ, let T_τ^λ be a Hörmander-type operator given by replacing the amplitude a^λ in the definition of T^λ with some amplitude a_τ^λ which has ω-support in a $2K^{-1}$-cap concentric to τ and which satisfies $T_\tau^\lambda f_\tau = T^\lambda f_\tau$. One now wishes to bound

$$\int_{B(0, \lambda/\overline{C}K)^{2k}} \int_{B(\bar{x}, K)} \prod_{j=1}^{k} |T_{\tau_j^*}^\lambda f_{\tau_j^*} \circ \sigma_{y_j, z_j}(x)|^{p/k} \, dx \, Z_K(\mathbf{y}, \mathbf{z}) \, d\mathbf{y} \, d\mathbf{z}. \quad (6.7)$$

For the purposes of this proof, a k-tuple $(\tau_1, ..., \tau_k)$ of K^{-1}-caps is said to be *transverse* if $(T_{\tau_1}^\lambda, ..., T_{\tau_k}^\lambda)$ is a $cK^{-(k-1)}$-transverse k-tuple of Hörmander-type operators, for a suitable choice of small constant $c > 0$. Now, suppose $(x, \omega) \in \operatorname{supp} a$ so that, by the original decomposition, both x and $\bar{x}$ lie in a ball of radius $R/\overline{C}K$. Since $R \leqslant \lambda$, if follows from Lemma 4.6 that

$$|G^\lambda(\bar{x}, \omega) - G^\lambda(x, \omega)| \lesssim \frac{|\bar{x} - x|}{\lambda} \lesssim \overline{C}^{-1} K^{-1} \quad \text{for all } (x, \omega) \in \operatorname{supp} a.$$

Thus, choosing $\overline{C}$ sufficiently large, in addition to (6.1), one may assume that $(\tau_1^*, ..., \tau_k^*)$ is transverse. The expression (6.7) is therefore dominated by

$$\sum_{(\tau_j)_{j=1}^k \text{ trans.}} \int_{B(0, \lambda/\overline{C}K)^{2k}} \int_{B(\bar{x}, K)} \prod_{j=1}^{k} |T_{\tau_j}^\lambda f_{\tau_j} \circ \sigma_{y_j, z_j}(x)|^{p/k} \, dx \, Z_K(\mathbf{y}, \mathbf{z}) \, d\mathbf{y} \, d\mathbf{z},$$

where the sum is over all choices of transverse k-tuples of caps. Summing both sides of this inequality over all $B(\bar{x}, K^2) \in \mathcal{B}_{K^2}$ with $B(\bar{x}, K^2) \cap B_z \neq \varnothing$, it suffices to show that

$$\int_{\mathbb{R}^n} \prod_{j=1}^{k} |T_{\tau_j}^{\lambda} f_{\tau_j} \circ \sigma_{y_j, z_j}(x)|^{p/k} \, dx \lesssim R^{\varepsilon} \|f\|_{L^2(B^{n-1})}^p$$

for any choice of $K^{-(k-1)}$-transverse tuple $(\tau_1, ..., \tau_k)$ and any $\mathbf{y}, \mathbf{z} \in B(0, \lambda/\overline{C}K)^k$. However, defining $T_j^{\lambda} f(x) := T_{\tau_j}^{\lambda} f \circ \sigma_{y_j, z_j}(x)$ and again choosing $\overline{C}$ to be sufficiently large, it follows from (6.6) that these operators are $\sim K^{-(k-1)}$-transverse in the sense of Definition 1.6. Thus, the desired estimate is an immediate consequence of the hypothesised multilinear inequality. $\qquad \square$

7. Algebraic preliminaries

7.1. Basic definitions and results

Let $0 \leqslant m \leqslant n$ and consider a collection of real polynomials $P_j \in \mathbb{R}[X_1, ..., X_n]$, $1 \leqslant j \leqslant n-m$. Let $Z(P_1, ..., P_{n-m})$ denote their zero locus; that is,

$$Z(P_1, ..., P_{n-m}) := \{x \in \mathbb{R}^n : P_j(x) = 0 \text{ for } 1 \leqslant j \leqslant n-m\}.$$

Any such set is referred to as a *variety* and the *maximum degree* of $Z(P_1, ..., P_{n-m})$ is defined to be the number

$$\overline{\deg} \, Z(P_1, ..., P_{n-m}) := \max_{1 \leqslant j \leqslant n-m} \deg P_j.$$

Remark 7.1. The notion of maximum degree is unnatural from a geometric perspective: it is not an intrinsic quantity associated with the variety, but depends on the choice of defining polynomials $P_1, ..., P_{n-m}$. Nevertheless, it is a convenient quantity to work with for the purposes of this article.

Throughout this article, it will be convenient to work with varieties which satisfy the additional property that the $n \times (n-m)$ matrix $(\nabla P_1(z) ... \nabla P_{n-m}(z))$ has full rank whenever $z \in Z(P_1, ..., P_{n-m})$. In this case, $Z(P_1, ..., P_{n-m})$ is said to be a *transverse complete intersection*. Clearly, any transverse complete intersection is a smooth m-dimensional submanifold of $\mathbb{R}^n$.

For zero-dimensional transverse complete intersections, the following well-known variant of the classical Bézout theorem holds (see, for instance, [11]).

THEOREM 7.2. (Bézout's theorem) *Suppose that* $Z = Z(P_1, ..., P_n)$ *is a transverse complete intersection. Then, Z is finite and* $\#Z \leqslant \prod_{j=1}^{n} \deg P_j$.

A key tool in the present analysis of Hörmander-type operators is the following polynomial partitioning result, which is a variant of the polynomial partitioning theorem introduced in [15], and is based on the classical polynomial ham-sandwich theorem of Stone and Tukey [29].

THEOREM 7.3. (Polynomial partitioning [14]) *Suppose that* $W \in L^1(\mathbb{R}^n)$ *is non-negative. For any degree* $D \in \mathbb{N}$, *there is a polynomial* P *of degree* $\deg P \lesssim D$ *such that the following hold*:

(i) *The set* $Z(P)$ *is a finite union of* $\sim \log D$ *transverse complete intersections*;

(ii) *If* $\{O_i\}_{i \in \mathcal{I}}$ *denotes the set of connected components of* $\mathbb{R}^n \setminus Z(P)$, *then* $\#\mathcal{I} \lesssim D^n$ *and*

$$\int_{O_i} W \sim D^{-n} \int_{\mathbb{R}^n} W \quad \text{for all } i \in \mathcal{I}.$$

The connected components O_i of the set $\mathbb{R}^n \setminus Z(P)$ are referred to as *cells*. It is remarked that in [14] a stronger version of the above theorem is stated and proved, which provides further structural information about the polynomial P (in particular, the full result is stable under certain small perturbations of P). Whilst the methods of this article will require this strengthened version of Theorem 7.3, the full statement of the result is not reproduced here (it is only needed to address certain technical aspects of the analysis).

It was observed in recent work of the first author [13], [14] that polynomial partitioning is a useful tool for studying oscillatory integral operators. Roughly speaking, Theorem 7.3 can be used to effectively reduce the problem to situations where the mass of $T^\lambda f$ is concentrated in the neighbourhood of some low-degree algebraic variety; note that this is precisely the setup in the sharp examples discussed in §2.

7.2. Polynomial approximation

Recall that the operators T^λ are defined with respect to data belonging to the C^∞ category. In order to apply algebraic methods to the problem, one must approximate certain C^∞ functions by polynomials. This applies, in particular, to the core curves $\Gamma^\lambda_{\theta,v}$ which appear in the definition of the wave packets in §5. Similar issues were addressed in [9], [38] via a Jackson-type approximation theorem (see, for instance, [1]); for the present purpose, an entirely elementary Taylor approximation argument is all that is required.

Let $\varepsilon > 0$ be a small parameter and define $N = N_\varepsilon := \lceil 1/2\varepsilon \rceil \in \mathbb{N}$. Suppose that

$$\Gamma: (-1, 1) \longrightarrow \mathbb{R}^n$$

is a smooth curve satisfying

$$\|\Gamma\|_{C^{N+1}(-1,1)} := \max_{0 \leqslant k \leqslant N+1} \sup_{|t|<1} |\Gamma^{(k)}(t)| \lesssim 1.$$

The following lemma implies that

$$\|\Gamma^1_{\theta,v}\|_{C^{N+1}(-1,1)} \lesssim 1,$$

revealing further properties of the core curves of the tubes defined in §5.

LEMMA 7.4. *The curves* $\Gamma^1_{\theta,v}$ *satisfy*

$$|(\Gamma^1_{\theta,v})'(t)| \sim 1 \quad \text{for all } t \in I^1_{\theta,v},$$

and

$$\sup_{t \in I^1_{\theta,v}} |(\Gamma^1_{\theta,v})^{(k)}(t)| \lesssim c_{\mathrm{par}} \quad \text{for } 2 \leqslant k \leqslant N.$$

Proof. This follows from the reductions made in §4. Indeed, recall that $\Gamma^1_{\theta,v}(t) = (\gamma^1_{\theta,v}(t), t)$ satisfies $\partial_\omega \phi(\gamma^1_{\theta,v}(t), t; \omega_\theta) = v$. Differentiating this identity yields

$$(\gamma^1_{\theta,v})'(t) = -\partial^2_{\omega x'}\phi(\gamma^1_{\theta,v}(t), t; \omega_\theta)^{-1} \partial_\omega \partial_{x_n} \phi(\gamma^1_{\theta,v}(t), t; \omega_\theta).$$

The bounds now follow from (4.3) and (4.10), provided N_{par} is chosen to be sufficiently large. $\qquad\square$

Let $[\Gamma]_\varepsilon \colon \mathbb{R} \to \mathbb{R}^n$ denote the polynomial curve given by the degree-N Taylor approximation of Γ around zero. Observe that

$$\|[\Gamma]_\varepsilon\|_{C^\infty(-2,2)} \leqslant e^2 \|\Gamma\|_{C^N(-1,1)} \lesssim 1.$$

Given $\lambda \gg 1$, noting that $\lambda^{-\varepsilon N} \leqslant \lambda^{-1/2}$, Taylor's theorem yields

$$|\Gamma^{(i)}(t) - [\Gamma]^{(i)}_\varepsilon(t)| \lesssim_\varepsilon \lambda^{-1/2}|t|^{1-i} \quad \text{for all } |t| \lesssim_\varepsilon \lambda^{-\varepsilon} \text{ and } i = 0, 1.$$

Letting $\Gamma^\lambda \colon (-\lambda, \lambda) \to \mathbb{R}^n$ denote the rescaled curve $\Gamma^\lambda(t) := \lambda \Gamma(t/\lambda)$, the above inequalities trivially imply that

$$\|[\Gamma^\lambda]'_\varepsilon\|_{C^\infty(-2\lambda,2\lambda)} \lesssim 1 \quad \text{and} \quad \|[\Gamma^\lambda]''_\varepsilon\|_{C^\infty(-2\lambda,2\lambda)} \lesssim \lambda^{-1}, \tag{7.1}$$

and

$$|(\Gamma^\lambda)^{(i)}(t) - ([\Gamma^\lambda]_\varepsilon)^{(i)}(t)| \lesssim_\varepsilon \lambda^{-1/2}|t|^{1-i} \quad \text{for all } |t| \lesssim_\varepsilon \lambda^{1-\varepsilon} \text{ and } i = 0, 1. \tag{7.2}$$

Combining the $i = 1$ case of the above estimate with the elementary inequality[16]

$$|x \wedge y| \leqslant \min\{|x|, |y|\}|x-y| \quad \text{for all } x, y \in \mathbb{R}^n \setminus \{0\},$$

together with the fact that $|(\Gamma^\lambda)'(t)| \sim |[\Gamma^\lambda]'_\varepsilon(t)| \sim 1$, one observes that the tangent spaces to the curves Γ^λ and $[\Gamma^\lambda]_\varepsilon$ have a small angular separation; more precisely,

$$\measuredangle(T_{\Gamma^\lambda(t)}\Gamma^\lambda, T_{[\Gamma^\lambda]_\varepsilon(t)}[\Gamma^\lambda]_\varepsilon) \lesssim_\varepsilon \lambda^{-1/2} \quad \text{for all } |t| \lesssim_\varepsilon \lambda^{1-\varepsilon}. \tag{7.3}$$

---

[16] This follows by estimating the area of a triangle with sides of length $|x|$, $|y|$ and $|x-y|$, taking the length of the 'base' to be $\min\{|x|, |y|\}$ and bounding the 'perpendicular height' by $|x-y|$.

7.3. Transverse interactions between curved tubes and varieties

Let $Z = Z(P_1, ..., P_{n-m})$ be a transverse complete intersection and fix a polynomial curve $\Gamma \colon \mathbb{R} \to \mathbb{R}^n$. The purpose of this subsection is to study transverse interactions between Γ and an r-neighbourhood of Z; that is, roughly speaking, one wishes to understand the set of points at which the curve Γ enters $\mathcal{N}_r Z$ at a large angle. More precisely, given $\alpha, r > 0$, the problem is to estimate the size of the set

$$Z_{>\alpha,r,\Gamma} := \{ z \in Z : \text{there exists } x \in \Gamma \text{ with } |x - z| < r \text{ and } \measuredangle(T_z Z, T_x \Gamma) > \alpha \}.$$

It will be convenient to assume that Γ is a *polynomial graph*, by which it is meant that the curve can be rotated so that it is given by $\Gamma(t) = (\gamma(t), t)$ for some polynomial mapping $\gamma \colon \mathbb{R} \to \mathbb{R}^{n-1}$.

LEMMA 7.5. *Let $n \geqslant 2$, $1 \leqslant m \leqslant n$ and $Z = Z(P_1, ..., P_{n-m}) \subseteq \mathbb{R}^n$ be a transverse complete intersection. Suppose that $\Gamma \colon \mathbb{R} \to \mathbb{R}^n$ is a polynomial graph satisfying*

$$\|\Gamma'\|_{L^\infty(-2\lambda, 2\lambda)} \lesssim 1 \quad and \quad \|\Gamma''\|_{L^\infty(-2\lambda, 2\lambda)} \leqslant \delta \tag{7.4}$$

for some $\lambda, \delta > 0$. There exists a dimensional constant $\overline{C} > 0$ such that, for all $\alpha > 0$ and $0 < r < \lambda$ satisfying $\alpha \geqslant \overline{C} \delta r$, the set $Z_{>\alpha,r,\Gamma} \cap B(0,\lambda)$ is contained in a union of

$$O((\overline{\deg} Z \cdot \deg \Gamma)^n)$$

balls of radius r/α.

The case of interest is given by taking $\Gamma := [\Gamma^\lambda_{\theta,v}]_\varepsilon$ to be the polynomial approximant of the curve $\Gamma^\lambda_{\theta,v}$ introduced in the previous subsection. Here, $\deg \Gamma \lesssim_\varepsilon 1$ and, by (7.1), the condition (7.4) holds with $\delta \sim_\varepsilon 1/\lambda$; thus, Lemma 7.5 implies that, for $\alpha > 0$ and $0 < r < \lambda$ satisfying $\alpha \gtrsim r/\lambda$, the set $Z_{>\alpha,r,\Gamma} \cap B(0,\lambda)$ is contained in a union of $O_\varepsilon((\overline{\deg} Z)^n)$ balls of radius r/α.

Using Bézout's theorem (that is, Theorem 7.2), Lemma 7.5 was established in the case where Γ is a line by the first author in [14, Lemma 5.7]. If $\Gamma = \ell$ is a line, then the condition (7.4) holds for any $\lambda > 0$ and any $\delta > 0$, and therefore the lemma implies that, for any $\alpha, r > 0$, the set $Z_{>\alpha,r,\ell}$ is contained in a union of $O((\overline{\deg} Z)^n)$ balls of radius r/α. The result for general curves Γ is, in fact, a rather straightforward consequence of the special case of lines.

Proof. Since the problem is rotationally invariant, one may assume that

$$\Gamma(t) = (\gamma(t), t),$$

where $\gamma\colon\mathbb{R}\to\mathbb{R}^{n-1}$ is a polynomial mapping.

The function $\Upsilon\colon\mathbb{R}^n\to\mathbb{R}^n$ given by $\Upsilon(x',x_n):=(x'-\gamma(x_n),x_n)$ is clearly a diffeomorphism which maps bijectively between Γ and the vertical line $\ell=\mathrm{span}\{e_n\}$. Furthermore, it easily follows that the image set

$$\widetilde{Z}:=\Upsilon(Z(P_1,...,P_{n-m}))=Z(P_1\circ\Upsilon^{-1},...,P_{n-m}\circ\Upsilon^{-1})$$

is a transverse complete intersection of maximum degree $\overline{\deg}\,\widetilde{Z}\leqslant\overline{\deg}\,Z\cdot\deg\Gamma$.

Let λ, α and r satisfy the hypotheses of the lemma for some suitably large-dimensional constant $\overline{C}\geqslant 1$. The key observation is as follows.

CLAIM. *There exist dimensional constants $0<c\leqslant 1$ and $C\geqslant 1$ such that*

$$Z_{>\alpha,r,\Gamma}\cap(\mathbb{R}^{n-1}\times(-\lambda,\lambda))\subseteq\Upsilon^{-1}(\widetilde{Z}_{>c\alpha,Cr,\ell}).$$

Once this claim is verified, Lemma 7.5 easily follows. Indeed, one may apply the special case of Lemma 7.5 for lines (which, as previously remarked, is proved in [14, Lemma 5.7]) to conclude that $\widetilde{Z}_{>c\alpha,Cr,\ell}$ is contained in a union of $O((\overline{\deg}\,Z\cdot\deg\Gamma)^n)$ balls of radius r/α. On the other hand, as a consequence of the first hypothesis in (7.4),

$$|\Upsilon(x)-\Upsilon(x')|\sim|x-x'| \quad \text{for all } x,x'\in\mathbb{R}^{n-1}\times(-\lambda,\lambda). \tag{7.5}$$

Combining these observations, it follows that the set $Z_{>\alpha,r,\Gamma}\cap B(0,\lambda)$ can be covered by $O((\overline{\deg}\,Z\cdot\deg\Gamma)^n)$ balls of radius r/α, as required.

Turning to the proof of the claim, let $z\in Z_{>\alpha,r,\Gamma}\cap B(0,\lambda)$ and note that there exists some $x=\Gamma(x_n)\in\Gamma$ with $|x-z|<r$ and $\measuredangle(T_zZ,T_x\Gamma)>\alpha$. Defining

$$\tilde{z}:=\Upsilon(z)\in\widetilde{Z} \quad \text{and} \quad \tilde{x}:=\Upsilon(x)\in\ell,$$

it follows from (7.5) that $|\tilde{x}-\tilde{z}|\lesssim r$. Thus, the problem is reduced to showing that $\measuredangle(T_{\tilde{z}}\widetilde{Z},e_n)=\measuredangle(T_{\tilde{z}}\widetilde{Z},T_{\tilde{x}}\ell)\gtrsim\alpha$.

Observe that, provided $\overline{C}$ is sufficiently large depending only on n,

$$\measuredangle(T_zZ,T_{\Gamma(z_n)}\Gamma)>\tfrac{1}{2}\alpha. \tag{7.6}$$

Indeed, $|z_n|<\lambda$ and $|x_n|<\lambda+r<2\lambda$, and so, by the second condition in (7.4),

$$|\Gamma'(x_n)-\Gamma'(z_n)|\leqslant\delta|x_n-z_n|<\delta r<\overline{C}^{-1}\alpha.$$

Thus, if $\overline{C}$ is appropriately chosen, then

$$\measuredangle(T_{\Gamma(x_n)}\Gamma,T_{\Gamma(z_n)}\Gamma)<\tfrac{1}{2}\alpha,$$

which immediately yields (7.6).

Combining the observations of the previous paragraphs, the claim follows, provided one can show that $\measuredangle(T_{\tilde{z}}\widetilde{Z}, e_n) \sim \measuredangle(T_z Z, T_{\Gamma(z_n)}\Gamma)$. Let η be a smooth curve in Z containing z and define $\tilde{\eta} := \Upsilon(\eta)$; thus, $\tilde{\eta}$ is a smooth curve in $\widetilde{Z}$ containing $\tilde{z}$. The problem is now reduced to proving that

$$\measuredangle(T_{\tilde{z}}\tilde{\eta}, e_n) \sim \measuredangle(T_z\eta, T_{\Gamma(z_n)}\Gamma). \tag{7.7}$$

If $T_z\eta$ lies in the hyperplane $e_n^{\perp}$ orthogonal to e_n, then the above estimate easily follows. Indeed, the tangent space $T_{\Gamma(z_n)}\Gamma$ is spanned by $\Gamma'(z_n)=(\gamma'(z_n), 1)$ and therefore, by (7.4), one has $\measuredangle(T_z\eta, T_{\Gamma(z_n)}\Gamma) \sim 1$. On the other hand, it is clear from the definition of Υ that $T_{\tilde{z}}\tilde{\eta}$ also lies in $e_n^{\perp}$, and so $\measuredangle(T_{\tilde{z}}\tilde{\eta}, e_n) = \frac{1}{2}\pi$. Thus, (7.7) holds in this case.

If $T_z\eta$ does not lie in the hyperplane $e_n^{\perp}$, then η can be locally parameterised as a graph over the x_n-variable. By an abuse of notation, let η denote this graph parametrisation and $\tilde{\eta} := \Upsilon \circ \eta$ so that $\eta(z_n)=z$ and $\tilde{\eta}(z_n)=\tilde{z}$. One may easily verify that

$$|\tilde{\eta}'(z_n) \wedge e_n| = |\eta'(z_n) \wedge \Gamma'(z_n)|,$$

and so

$$\sin\measuredangle(T_{\tilde{z}}\tilde{\eta}, e_n)|\tilde{\eta}'(z_n)| = \sin\measuredangle(T_z\eta, T_{\Gamma(z_n)}\Gamma)|\eta'(z_n)|\,|\Gamma'(z_n)|.$$

By the first hypothesis in (7.4), one has

$$|\tilde{\eta}'(z_n)| \sim |\eta'(z_n)| \quad \text{and} \quad |\Gamma'(z_n)| \sim 1,$$

and (7.7) follows. $\square$

8. Transverse equidistribution estimates

8.1. Tangential wave packets and transverse equidistribution

In this section the theory of transverse equidistribution estimates, as introduced in [14], is extended to the variable coefficient setting. This is a key step in the proof of Theorem 1.9, and here the positive-definite hypothesis (H2$^+$) plays a crucial role in the argument.

The first step is to give a precise definition of what it means for a wave packet to be 'tangential' to a transverse complete intersection Z. Throughout this section, let T^{λ} be a Hörmander-type operator with reduced positive-definite phase ϕ and for some $R \ll \lambda$ define the (curved) tubes $T_{\theta,v}$ as in §5. Furthermore, let δ_m denote a small parameter satisfying $0 < \delta \ll \delta_m \ll 1$ (here δ is the same parameter as that which appears in the definition of the wave packets).

Definition 8.1. Suppose $Z = Z(P_1, ..., P_{n-m})$ is a transverse complete intersection. A tube $T_{\theta,v}$ is $R^{-1/2+\delta_m}$-tangent to Z in $B(0, R)$ if

$$T_{\theta,v} \subseteq N_{R^{1/2+\delta_m}}(Z)$$

and

$$\angle(G^\lambda(x; \omega_\theta), T_z Z) \leqslant \bar{c}_{\mathrm{tang}} R^{-1/2+\delta_m}$$

for any $x \in T_{\theta,v}$ and $z \in Z \cap B(0, 2R)$ with $|x - z| \leqslant \overline{C}_{\mathrm{tang}} R^{1/2+\delta_m}$.

Here, $\bar{c}_{\mathrm{tang}} > 0$ (resp. $\overline{C}_{\mathrm{tang}} \geqslant 1$) is a dimensional constant, chosen to be sufficiently small (resp. large) for the purposes of the following arguments.

Definition 8.2. If $\mathbb{S} \subseteq \mathbb{T}$, then f is said to be concentrated on wave packets from $\mathbb{S}$ if

$$f = \sum_{(\theta,v) \in \mathbb{S}} f_{\theta,v} + \mathrm{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

One wishes to study functions concentrated on wave packets from the collection

$$\mathbb{T}_Z := \{(\theta, v) \in \mathbb{T} : T_{\theta,v} \text{ is } R^{-1/2+\delta_m}\text{-tangent to } Z \text{ in } B(0, R)\}.$$

Let $B \subseteq \mathbb{R}^n$ be a fixed ball of radius $R^{1/2+\delta_m}$ with centre $\bar{x} \in B(0, R)$. Throughout this section, the analysis will be essentially confined to a spatially localised operator $\eta_B \cdot T^\lambda g$, where η_B is a suitable choice of Schwartz function concentrated on B. For any (θ, v), a stationary phase argument shows that the Fourier transform of $\eta_B \cdot T^\lambda g_{\theta,v}$ is concentrated near the surface

$$\Sigma := \{\Sigma(\omega) : \omega \in \Omega\}, \quad \text{where } \Sigma(\omega) := \partial_x \phi^\lambda(\bar{x}; \omega). \tag{8.1}$$

Now, consider the refined set of wave packets

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

If $(\theta, v) \in \mathbb{T}_{Z,B}$, then the direction $G^\lambda(\bar{x}; \omega_\theta)$ of the curved tube $T_{\theta,v}$ on the ball B must make a small angle with each of the tangent spaces $T_z Z$ for all $z \in Z \cap B$. It transpires that this essentially constrains the frequency $\Sigma(\omega_\theta)$ to lie in a small neighbourhood of some fixed (depending on the choice of ball B) m-dimensional manifold S_ξ (here, $m = \dim Z$).[17] In the case of the parabolic extension operator E_{par}, which is studied in [14], the relationship between the normal direction $G^\lambda(\bar{x}; \omega_\theta)$ and the frequency $\Sigma(\omega_\theta)$ is particularly simple. Here, $\Sigma(\omega_\theta) = (\omega_\theta, \frac{1}{2}|\omega_\theta|^2)$ is constrained to lie in roughly the $R^{-1/2}$-neighbourhood of some *affine* subspace A_ξ.

---

[17] The subscript ξ is used here to indicate that S_ξ lies in the ξ parameter space (that is, $\widehat{\mathbb{R}}^n$). In particular, it does *not* denote a dependence on some choice of ξ. Variants of this notation (such as A_ξ, S_ω, A_ω, etc.) feature throughout this section with the obvious corresponding intended meaning.

Example 8.3. Suppose, for simplicity, that Z is an m-dimensional affine plane so that $T_z Z = V$ for all $z \in Z$, where V is the m-dimensional linear subspace parallel to Z. To avoid degenerate situations, also assume that V makes a small angle with the e_n direction. For the prototypical case of the parabolic extension operator E_{par}, the (unnormalised) Gauss map is an affine map: $G_0(\omega) = (-\omega, 1)$. Consequently,

$$A_\omega := \{ \omega \in \mathbb{R}^{n-1} : G_0(\omega) \in V \}$$

is an affine subspace of $\mathbb{R}^{n-1}$ of dimension $m-1$. Thus, if $G_0(\omega) \in V$, then

$$\Sigma(\omega) \in A_\xi := A_\omega \times \mathbb{R}.$$

Note that, for general Hörmander-type operators, the condition $G^\lambda(\bar{x}; \omega) \in V$ defines a (possibly curved) submanifold rather than an affine subspace.

In view of this frequency concentration in the case of E_{par}, the uncertainty principle then suggests that, if g is concentrated on wave packets from $\mathbb{T}_{Z,B}$, then the function $|E_{\mathrm{par}} g(x)|$ is morally constant as one varies x by $R^{1/2}$ in directions perpendicular to A_ξ. Furthermore, it can be shown that the affine subspace A_ξ makes a small angle with the tangent planes $T_z Z$ for $z \in Z \cap B$, and so $|E_{\mathrm{par}} g(x)|$ is morally constant as one varies x by $R^{1/2}$ in directions transverse to $Z \cap B$.

One wishes to extend the above observations for E_{par} to the variable coefficient setting; that is, for g concentrated on wave packets from $\mathbb{T}_{Z,B}$,[18] the problem is to show that $|T^\lambda g|$ is morally constant in directions transverse to $Z \cap B$. More precisely, one wishes to establish an inequality roughly of the form

$$\fint_{N_{\varrho^{1/2+\delta_m}}(Z) \cap B} |T^\lambda g|^2 \lesssim \fint_B |T^\lambda g|^2 \tag{8.2}$$

for $0 < \varrho < R$; this would show that the L^2 mass of $T^\lambda g$ is unable to concentrate in a small neighbourhood of $Z \cap B$. For the parabolic extension operator the observations of the previous paragraph can be used to prove (8.2) (up to a rapidly decaying error term). The general case is more complicated, however. First of all, the surface S_ξ described above is no longer necessarily an affine subspace and may possess curvature. One way to circumvent this issue is to introduce a further constraint on the family of wave packets. Let $R^{1/2} < \varrho \ll R$ and, throughout this section, let $\tau \subset \mathbb{R}^{n-1}$ be a fixed cap of radius $O(\varrho^{-1/2+\delta_m})$ centred at a point in B^{n-1}. Now define

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

---

[18] In fact, in the general case, a more stringent hypothesis on g is required, as discussed below.

The frequencies $\Sigma(\omega_\theta)$ for $(\theta, v) \in \mathbb{T}_{Z,B,\tau}$ are further constrained to lie in a small region of Σ upon which the curved space S_ξ can be effectively approximated by an affine space A_ξ. Consequently, one can carry out a similar analysis as in the parabolic extension case.

The second issue is to ensure that the resulting affine space A_ξ makes a small angle with the tangent spaces $T_z Z$ for $z \in Z \cap B$. This is crucial to ensure that the morally constant property holds in directions transverse to Z. For general Hörmander-type operators this property can fail (a simple example is given by the extension operator associated with the hyperbolic paraboloid, as discussed below). In order to ensure the angle condition, one needs to exploit the additional positive-definite hypothesis (H2$^+$).

In practice, the rigorous formulation of these heuristics is somewhat messier than (8.2), and it is convenient to state the key estimate in the following manner.

LEMMA 8.4. *With the above setup, if* $\overline{\dim}\, Z \lesssim_\varepsilon 1$ *and* g *is concentrated on wave packets from* $\mathbb{T}_{Z,B,\tau}$, *then*

$$\int_{N_{\varrho^{1/2+\delta_m}}(Z)\cap B} |T^\lambda g|^2 \lesssim R^{1/2+O(\delta_m)}\Big(\frac{\varrho}{R}\Big)^{(n-m)/2}\|g\|_{L^2(B^{n-1})}^2. \tag{8.3}$$

The inequality (8.3) is related to the heuristic inequality (8.2) via Hörmander's L^2 bound

$$\|T^\lambda g\|_{L^2(B)}^2 \lesssim R^{1/2+\delta_m}\|g\|_{L^2(B^{n-1})}^2.$$

The estimate is presented in this way (rather than in a form more closely resembling (8.2)) as it provides a relatively clean statement and, moreover, (8.3) happens to be the precise bound required later in the proof.

8.2. Uncertainty principle preliminaries

If $G: \mathbb{R}^n \to \mathbb{C}$ is frequency supported on a ball of radius $r > 0$, then the uncertainty principle dictates that G should be essentially constant at spatial scale r^{-1}. In particular, the L^2-mass of G cannot be highly concentrated in any ball of radius $\varrho < r^{-1}$, and so one has

$$\fint_{B(x_0,\varrho)} |G|^2 \lesssim \fint_{B(x_0,r^{-1})} |G|^2.$$

Strictly speaking, for this inequality to hold, the right-hand integral should be taken with respect to a rapidly decaying weight function, rather than over the compact region $B(x_0, r^{-1})$ (see, for instance, [14, §6]). There is a variant of this estimate which is effective in cases where G has the property that $\widehat{G}$ is merely concentrated in (rather than supported in) an r-ball.

LEMMA 8.5. *If $r^{-1/2} \leqslant \varrho \leqslant r^{-1}$, then for any ball $B(x_0, \varrho)$, $\xi_0 \in \widehat{\mathbb{R}}^n$ and $\delta > 0$ one has*

$$\fint_{B(x_0,\varrho)} |G|^2 \lesssim_\delta \|\widehat{G} w_{B(\xi_0,r)}^{-1}\|_\infty^{2\delta/(1+\delta)} \frac{1}{|B(0,r^{-1})|} \left(\int_{\mathbb{R}^n} |G|^2 \right)^{1/(1+\delta)}.$$

Here, $w_{B(\xi_0,r)}$ is a weight concentrated on $B(\xi_0, r)$ given by

$$w_{B(\xi_0,r)}(\xi) := (1 + r^{-1}|\xi - \xi_0|)^{-N} \tag{8.4}$$

for some large $N = N_\delta \in \mathbb{N}$.

Hence, if $\widehat{G}$ is concentrated in $B(\xi_0, r)$ in the sense that $|\widehat{G}(\xi)| \lesssim M w_{B(\xi_0,r)}(\xi)$ for some controllable constant $M \geqslant 0$, then the lemma produces a favourable estimate.

Remark 8.6. Lemma 8.5 is *not* sharp in terms of the dependence on the r and ϱ parameters. It differs, however, from the sharp inequality only by $O(\delta)$ powers and such losses are negligible for the purposes of this article.

Proof of Lemma 8.5. Define ψ_ϱ by $(\psi_\varrho)^\vee(x) := \check{\psi}(\varrho^{-1}(x - x_0))$, where ψ is a Schwartz function which satisfies $|\check{\psi}(x)| \gtrsim 1$ on $B(0, 1)$. Thus, by Plancherel,

$$\fint_{B(x_0,\varrho)} |G|^2 \lesssim \int |\check{\psi}_\varrho G|^2 = \int |\psi_\varrho * \widehat{G}|^2.$$

Using the rapid decay of ψ, one deduces that

$$|\psi_\varrho * \widehat{G}(\xi)| \lesssim_\delta \varrho^n \int_{\widehat{\mathbb{R}}^n} w_{B(\xi,\varrho^{-1})}(\eta)^{\delta/(1+\delta)} |\widehat{G}(\eta)| \, d\eta$$

for all $\xi \in \widehat{\mathbb{R}}^n$. By expressing the right-hand integral as

$$\int_{\widehat{\mathbb{R}}^n} (w_{B(\xi,\varrho^{-1})}(\eta) w_{B(\xi_0,r)}(\eta)^{1/2})^{\delta/(1+\delta)} (|\widehat{G}(\eta)|^{1+\delta} w_{B(\xi_0,r)}(\eta)^{-\delta/2})^{1/(1+\delta)} \, d\eta$$

and applying Hölder's inequality, it follows that

$$|\hat{\psi} * \widehat{G}(\xi)| \lesssim |B(x_0, \varrho)| \cdot \mathrm{I}(\xi)^{\delta/(1+\delta)} \cdot \mathrm{II}(\xi)^{1/(1+\delta)},$$

where

$$\mathrm{I}(\xi) := \int_{\widehat{\mathbb{R}}^n} w_{B(\xi,\varrho^{-1})}(\eta) w_{B(\xi_0,r)}(\eta)^{1/2} \, d\eta,$$

$$\mathrm{II}(\xi) := \int_{\widehat{\mathbb{R}}^n} |\widehat{G}(\eta)|^{1+\delta} w_{B(\xi_0,r)}(\eta)^{-\delta/2} \, d\eta.$$

To estimate I(ξ), first perform the variable shift $\eta \mapsto \eta + \xi_0$, and then decompose the range of integration into the regions $|\eta| < \frac{1}{2}|\xi - \xi_0|$ and $|\eta| \geqslant \frac{1}{2}|\xi - \xi_0|$. Since

$$w_{B(\xi - \xi_0, \varrho^{-1})}(\eta) \lesssim_\delta w_{B(\xi_0, \varrho^{-1})}(\xi) \quad \text{for } |\eta| < \tfrac{1}{2}|\xi - \xi_0|$$

and

$$w_{B(0, \varrho^{-1})}(\eta) \lesssim_\delta w_{B(\xi_0, \varrho^{-1})}(\xi) \quad \text{for } |\eta| \geqslant \tfrac{1}{2}|\xi - \xi_0|,$$

it follows that

$$\mathrm{I}(\xi) \lesssim_\delta |B(x_0, \varrho)|^{-1} w_{B(\xi_0, \varrho^{-1})}(\xi)^{1/2}.$$

To estimate II(ξ) note that, provided N_δ is chosen sufficiently large, by Cauchy–Schwarz and Plancherel's theorem one has

$$\mathrm{II}(\xi) \leqslant \|\widehat{G} w_{B(\xi_0, r)}^{-1}\|_\infty^\delta \int |\widehat{G}(\eta)| w_{B(\xi_0, r)}(\eta)^{\delta/2} \, d\eta$$
$$\lesssim_\delta |B(0, r^{-1})|^{-1/2} \|\widehat{G} w_{B(\xi_0, r)}^{-1}\|_\infty^\delta \|G\|_{L^2(\mathbb{R}^n)}.$$

Combining these observations, one obtains the desired estimate but with an additional factor of $(\varrho r^{1/2})^{-2n\delta/(1+\delta)}$ on the right-hand side. Since $1 \leqslant \varrho r^{1/2}$, the result immediately follows. $\qquad\square$

8.3. Wave packets tangential to linear subspaces

Here, as a step towards Lemma 8.4, transverse equidistribution estimates are proven for functions concentrated on wave packets tangential to some fixed linear subspace $V \subseteq \mathbb{R}^n$. As before, let B be a ball of radius $R^{1/2+\delta_m}$ with centre $\bar{x} \in \mathbb{R}^n$, and define

$$\mathbb{T}_{V,B} := \{(\theta, v) : \measuredangle(G^\lambda(\bar{x}, \omega_\theta), V) \lesssim R^{-1/2+\delta_m} \text{ and } T_{\theta,v} \cap B \neq \varnothing\}.$$

Let $R^{1/2} < \varrho < R$ and, for $\tau \subset \mathbb{R}^{n-1}$ a ball of radius $O(\varrho^{-1/2+\delta_m})$ centred at a point in B^{n-1}, define

$$\mathbb{T}_{V,B,\tau} := \{(\theta, v) \in \mathbb{T}_{V,B} : \theta \cap \left(\tfrac{1}{10} \cdot \tau\right) \neq \varnothing\},$$

where $\left(\frac{1}{10}\tau\right)$ is the cap concentric to τ but with $\frac{1}{10}$th of the radius.

The key estimate is the following.

LEMMA 8.7. *If* $V \subseteq \mathbb{R}^n$ *is a linear subspace, then there exists a linear subspace* V' *with the following properties*:

(1) $\dim V + \dim V' = n$;

(2) V *and* V' *are quantitatively transverse, in the sense that there exists a uniform constant* $c_{trans} > 0$ *such that*

$$\angle(v, v') \geqslant 2c_{trans} \quad \text{for all non-zero vectors } v \in V \text{ and } v' \in V';$$

(3) *if* g *is concentrated on wave packets from* $\mathbb{T}_{V,B,\tau}$, Π *is any plane parallel to* V' *and* $x_0 \in \Pi \cap B$, *then the inequality*

$$\int_{\Pi \cap B(x_0, \varrho^{1/2+\delta_m})} |T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)} \left(\frac{\varrho}{R}\right)^{\dim V'/2} \|g\|^{2\delta/(1+\delta)}_{L^2(B^{n-1})} \left(\int_{\Pi \cap 2B} |T^\lambda g|^2\right)^{1/(1+\delta)}$$

holds, up to the inclusion of a $\mathrm{RapDec}(R) \|g\|_{L^2(B^{n-1})}$ *term on the right-hand side.*

Proof. The argument is presented in a number of stages.

Constructing the subspace V'

Recall that

$$\omega \longmapsto G^\lambda(\bar{x}; \omega) := \frac{G_0^\lambda(\bar{x}; \omega)}{|G_0^\lambda(\bar{x}; \omega)|}$$

is the Gauss map associated with the hypersurface Σ, defined in (8.1). Since $G(x; 0) = e_n$ for all $x \in X$, Lemma 4.6 implies that

$$\angle(G^\lambda(\bar{x}; \omega), e_n) \sim |\omega| \quad \text{for all } \omega \in \Omega.$$

Consequently, by choosing $\mathrm{diam}\,\Omega$ to be sufficiently small in the initial reductions, one may assume that the Gauss map $\omega \mapsto G^\lambda(\bar{x}; \omega)$ always makes a wide angle with the hyperplane $e_n^\perp = \mathbb{R}^{n-1} \times \{0\}$. In particular,

$$\angle(G^\lambda(\bar{x}; \omega), e_n^\perp) \gtrsim 1 \quad \text{for all } \omega \in \Omega.$$

Since the situation is trivial if $\mathbb{T}_{V,B} = \varnothing$, one may assume that

$$\overline{\angle}(V, e_n^\perp) := \max_{v \in V \cap S^{n-1}} \angle(v, e_n^\perp) \gtrsim 1. \tag{8.5}$$

Define $S_\omega \subset \mathbb{R}^{n-1}$ by

$$S_\omega := \{\omega \in \Omega : G^\lambda(\bar{x}; \omega) \in V\}.$$

Fixing an orthonormal basis $\{N_1, ..., N_{n-\dim V}\}$ for $V^\perp$, one has

$$S_\omega = \{\omega \in \Omega : \langle G_0^\lambda(\bar{x}; \omega), N_k \rangle = 0 \text{ for } 1 \leqslant k \leqslant n - \dim V\}.$$

CLAIM 1. *If $S_\omega \neq \varnothing$, then S_ω is a smooth surface in $\mathbb{R}^{n-1}$ of dimension $\dim V - 1$.*

Proof. Let $\omega \in S_\omega$ and note that each N_k is tangential to Σ at $\Sigma(\omega)$. Hence, one may write

$$N_k = \sum_{j=1}^{n-1} N_k^{(j)}(\omega) \partial_{\omega_j} \partial_x \phi^\lambda(\bar{x}; \omega)$$

for some choice of coefficients $N_k^{(j)}(\omega) \in \mathbb{R}$. A computation now shows that

$$\partial_{\omega_i} \langle G_0^\lambda(\bar{x}; \omega), N_k \rangle = -\sum_{j=1}^{n-1} \langle \partial_{\omega_i \omega_j}^2 \partial_x \phi^\lambda(\bar{x}; \omega), G_0^\lambda(\bar{x}; \omega) \rangle N_k^{(j)}(\omega).$$

The condition (H2) implies the invertibility of the $(n-1) \times (n-1)$ matrix whose (i, j) entry is given by

$$\langle \partial_{\omega_i \omega_j}^2 \partial_x \phi^\lambda(\bar{x}; \omega), G_0^\lambda(\bar{x}; \omega) \rangle.$$

Thus, the Jacobian of $\omega \mapsto (\langle G_0^\lambda(\bar{x}; \omega), N_k \rangle)_{k=1}^{n-\dim V}$ has maximal rank, and the claim follows by the implicit function theorem. $\square$

At this point, it is convenient to switch to a graph parametrisation of Σ via the change of variables $u \mapsto \Psi^\lambda(\bar{x}; u)$, where Ψ^λ is the (appropriate λ-rescaling of the) function introduced in §4. For convenience, let $\Psi: U \to \Omega$ denote this mapping; that is,

$$\Psi(u) := \Psi^\lambda(\bar{x}; u).$$

Recall that the hypersurface Σ coincides with the graph of the function

$$\bar{h}: U \to \mathbb{R}, \quad \bar{h}(u) := \partial_{x_n} \phi^\lambda(\bar{x}; \Psi(u)). \tag{8.6}$$

If $S_\omega \cap \tau = \varnothing$, then it follows by Lemma 4.6 that

$$\angle(G^\lambda(\bar{x}; \theta), V) \gtrsim \varrho^{-1/2 + \delta_m} > R^{-1/2 + \delta_m}$$

whenever $\theta \cap (\frac{1}{10} \cdot \tau) \neq \varnothing$. Consequently, $\mathbb{T}_{V,B,\tau} = \varnothing$ and the situation is trivial. Thus, one may assume without loss of generality that $S_\omega \cap \tau \neq \varnothing$ and so, letting

$$S_u := \Psi^{-1}(S_\omega) = \{u \in U : G_0^\lambda(\bar{x}; \Psi(u)) \in V\},$$

it follows that $S_u \cap \Psi^{-1}(\tau) \neq \varnothing$. The properties of the mapping Ψ discussed in §4 imply that $\Psi^{-1}(\tau)$ is roughly a ball of radius $O(\varrho^{-1/2 + \delta_m})$.

Fix some $u_0 \in S_u \cap \Psi^{-1}(\tau)$ and let A_u denote the tangent plane to S_u at u_0. Here, the tangent plane is interpreted as a $(\dim V - 1)$-dimensional affine subspace of $\mathbb{R}^{n-1}$ through u_0. Now define $A_\xi := A_u \times \mathbb{R} \subseteq \mathbb{R}^n$, so that $\dim A_\xi = \dim V$, and let V_u and V_ξ be the linear subspaces parallel to A_u and A_ξ, respectively. Finally, let $V' := V_\xi^\perp$, so that $\dim V + \dim V' = n$.

Verifying the transverse equidistribution estimate in (3)

Suppose $\Pi \subseteq \mathbb{R}^n$ is an affine subspace parallel to V' which intersects B, and $x_0 \in \Pi \cap B$. Let $\eta_B(x) := \eta((x - \bar{x})/R^{1/2 + \delta_m})$ where η is a Schwartz function which satisfies $\eta(x) = 1$ for $x \in B(0, 2)$ and, for any $(\theta, v) \in \mathbb{T}$, consider

$$(\eta_B \cdot T^\lambda g_{\theta,v}|_\Pi)^\frown(\xi) = e^{-2\pi i \langle x_0, \xi \rangle} R^{\dim V'(1/2 + \delta_m)} \int_{B^{n-1}} K^{\lambda,R}(\xi; \omega) g_{\theta,v}(\omega) \, d\omega,$$

where the kernel $K^{\lambda,R}$ is given by

$$K^{\lambda,R}(\xi; \omega) := \int_{V'} e^{2\pi i \phi_\omega^{\lambda,R}(z)} a_\omega^{\lambda,R}(z) \, dz$$

for the phase an amplitude functions

$$\phi_\omega^{\lambda,R}(z) := \phi^\lambda(x_0 + R^{1/2+\delta_m} z; \omega) - R^{1/2+\delta_m} \langle z, \xi \rangle,$$
$$a_\omega^{\lambda,R}(z) := a^\lambda(x_0 + R^{1/2+\delta_m} z; \omega) \tilde{\eta}(z),$$

and

$$\tilde{\eta}(z) := \eta\left(z + \frac{x_0 - \bar{x}}{R^{1/2+\delta_m}}\right).$$

CLAIM 2. *Fixing* $\omega \in \Omega$, $\xi \in \widehat{\mathbb{R}}^n$ *such that* $|\xi - \mathrm{proj}_{V'} \Sigma(\omega)| \gtrsim R^{-1/2+\delta_m}$ *and* $R \gg 1$, *the following estimates hold on* $\mathrm{supp}\, a_z^{\lambda,R}$:

(i) $|\partial_z \phi_\omega^{\lambda,R}(z)| \sim R^{1/2+\delta_m} |\xi - \mathrm{proj}_{V'} \Sigma(\omega)| \gtrsim R^{2\delta_m}$;
(ii) $|\partial_z^\alpha \phi_\omega^{\lambda,R}(z)| \lesssim |\partial_z \phi_\omega^{\lambda,R}(z)|$ *for all* $2 \leqslant |\alpha| \leqslant N_{\mathrm{par}}$;
(iii) $|\partial_z^\alpha a_\omega^{\lambda,R}(z)| \lesssim_\varepsilon 1$ *for all* $|\alpha| \leqslant N_{\mathrm{par}}$.
Here, $\Sigma(\omega) := \partial_x \phi^\lambda(\bar{x}; \omega)$ *is as defined in* (8.1).

Once the claim is established, repeated integration-by-parts (see Lemma A.1) shows that $K^{\lambda,R}$ is rapidly decaying whenever $|\xi - \mathrm{proj}_{V'} \Sigma(\omega)| \gtrsim R^{-1/2+\delta_m}$ and, in particular,

$$|K^{\lambda,R}(\xi; \omega)| \lesssim_\varepsilon (1 + R^{1/2} |\xi - \mathrm{proj}_{V'} \Sigma(\omega)|)^{-N} \quad \text{for all } N \leqslant N_{\mathrm{par}}.$$

Proof. The uniformity in the estimates is due to the reductions from §4. The bound (iii) for the amplitude immediately follows from Lemma 4.3, and it remains to prove the bounds for the phase.

Proof of (i)

The z-gradient of the phase $\phi_\omega^{\lambda,R}$ is equal to

$$R^{1/2+\delta_m} (\mathrm{proj}_{V'}[(\partial_x \phi^\lambda)(x_0 + R^{1/2+\delta_m} z; \omega) - (\partial_x \phi^\lambda)(\bar{x}; \omega)] - [\xi - \mathrm{proj}_{V'} \Sigma(\omega)]),$$

where, by Lemma 4.3, the first term satisfies

$$|\mathrm{proj}_{V'}[(\partial_x \phi^\lambda)(x_0 + R^{1/2+\delta_m} z; \omega) - (\partial_x \phi^\lambda)(\bar{x}; \omega)]| \lesssim \frac{R^{1/2+\delta_m}}{\lambda} \ll R^{-1/2+\delta_m}.$$

Thus, if $|\xi - \mathrm{proj}_{V'} \Sigma(\omega)| \gtrsim R^{-1/2+\delta_m}$, then the desired bound immediately follows.

Proof of (ii)

Fix $\alpha \in \mathbb{N}_0^n$ with $2 \leqslant |\alpha| \leqslant N_{\text{par}}$. It follows that

$$\partial_z^\alpha |\phi_\omega^{\lambda,R}(z)| \leqslant \lambda \Big(\frac{R^{1/2+\delta_m}}{\lambda} \Big)^{|\alpha|} \|\partial_x^\alpha \phi\|_{L^\infty(X \times \Omega)} \lesssim R^{2\delta_m},$$

and the desired bound now follows from (i). $\qquad\square$

If $\omega \in \operatorname{supp} g_{\theta,v}$, then $|\omega - \omega_\theta| < R^{-1/2}$ and so $|\Sigma(\omega) - \xi_\theta| \lesssim R^{-1/2}$, where $\xi_\theta := \Sigma(\omega_\theta)$. Consequently,

$$|(\eta_B \cdot T^\lambda g_{\theta,v}|_\Pi)^\frown(\xi)| \lesssim_N R^{O(1)} w_{B(\operatorname{proj}_{V'} \xi_\theta, R^{-1/2})}(\xi) \|g_{\theta,v}\|_{L^2(B^{n-1})}, \qquad (8.7)$$

where the definition of the weight function

$$w_{B(\operatorname{proj}_{V'} \xi_\theta, R^{-1/2})}(\xi) := (1 + R^{1/2}|\xi - \xi_\theta|)^{-N}$$

agrees with that of Lemma 8.5 (although here the weights are thought of as functions on V'), and so $N = N_\delta$ is a large integer, depending on δ.

The following geometric observation is key to the proof of property (3).

CLAIM 3. *If* $(\theta, v) \in \mathbb{T}_{B,\tau,V}$, *then* $\operatorname{dist}(\xi_\theta, A_\xi) \lesssim R^{-1/2+\delta_m}$.

Temporarily assume this claim, and recall that $V' := V_\xi^\perp$, where V_ξ is the linear subspace parallel to the affine subspace A_ξ. Thus, if $(\theta, v) \in \mathbb{T}_{B,\tau,V}$, then $\operatorname{proj}_{V'} \xi_\theta$ lies in some fixed ball of radius $O(R^{-1/2+\delta_m})$. Letting $\xi_* \in V'$ denote the centre of this ball, it follows that $w_{B(\operatorname{proj}_{V'} \xi_\theta, R^{-1/2})} \lesssim_\delta w_{B(\xi_*, R^{-1/2+\delta_m})}$, and so

$$\sum_{(\theta,v) \in \mathbb{T}_{V,B,\tau}} w_{B(\operatorname{proj}_{V'} \xi_\theta, R^{-1/2})} \lesssim_\delta R^{O(1)} w_{B(\xi_*, R^{-1/2+\delta_m})}.$$

Recalling (8.7),

$$\|(\eta_B \cdot T^\lambda g|_\Pi)^\frown w_{B(\xi_*, R^{-1/2+\delta_m})}^{-1}\|_\infty \lesssim R^{O(1)} \|g\|_{L^2(B^{n-1})}$$

and, applying Lemma 8.5, one concludes that

$$\int_{B(x_0, \varrho^{1/2+\delta_m}) \cap \Pi} |T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)} \Big(\frac{\varrho^{1/2}}{R^{1/2}} \Big)^{\dim V'} \|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)} \Big(\int_\Pi |T^\lambda g|^2 |\eta_B|^2 \Big)^{1/(1+\delta)}.$$

If $\xi \notin 2B$, then $\xi \notin \bigcup_{(\theta,v) \in \mathbb{T}_B} T_{\theta,v}$, and so

$$|T^\lambda g_{\theta,v}(\xi)| = \operatorname{RapDec}(R) \|g\|_{L^2(B^{n-1})} \quad \text{for all } (\theta, v) \in \mathbb{T}_{V,B,\tau}.$$

Hence,

$$\int_\Pi |T^\lambda g|^2 |\eta_B|^2 \leqslant \int_{2B \cap \Pi} |T^\lambda g|^2 + \operatorname{RapDec}(R) \|g\|_{L^2(B^{n-1})},$$

completing the proof of property (3) under the assumption that the above claim holds.

Proof of Claim 3. Fix $(\theta, v) \in \mathbb{T}_{B,\tau,V}$ and let $u_\theta := \mathrm{proj}_{x_n^\perp} \Sigma(\omega_\theta)$. Recalling that $A_\xi = A_u \times \mathbb{R}$ and applying triangle inequality, one deduces that

$$\mathrm{dist}(\xi_\theta, A_\xi) = \mathrm{dist}(u_\theta, A_u) \leqslant \mathrm{dist}(u_\theta, S_u \cap \Psi^{-1}(\tau)) + \sup_{u_* \in S_u \cap \Psi^{-1}(\tau)} \mathrm{dist}(u_*, A_u).$$

Furthermore, by Lemma 4.6,

$$\mathrm{dist}(u_\theta, S_u \cap \Psi^{-1}(\tau)) \sim \mathrm{dist}(\omega_\theta, S_\omega \cap \tau) \lesssim \measuredangle(G^\lambda(\bar{x}; \omega_\theta), V) \lesssim R^{-1/2 + \delta_m},$$

where the last inequality is by the definition of $\mathbb{T}_{B,\tau,V}$. On the other hand, fixing $u_* \in S_u \cap \Psi^{-1}(\tau)$, one wishes to estimate $\mathrm{dist}(u_*, A_u)$. Provided ϱ is sufficiently large (so that $\mathrm{diam}\,\tau$ is sufficiently small), the surface $S_u \cap \Psi^{-1}(\tau)$ can be parameterised as the graph of some function $\psi \colon \mathcal{W} \to V_u^\perp \subseteq \mathbb{R}^{n-1}$, where $\mathcal{W} \subset V_u$ is an open set about the origin of diameter $O(\varrho^{-1/2 + \delta_m})$. In particular, one may write

$$S_u \cap \Psi^{-1}(\tau) = \{w + \psi(w) : w \in \mathcal{W}\} + u_0,$$

where $\psi(0) = 0$ and $\partial_w \psi_j(0) = 0$ for $1 \leqslant j \leqslant n - \dim V_u$. Thus, $u_* = w_* + \psi(w_*) + u_0$ for some $w_* \in \mathcal{W}$ and, since $w_* \in V_u + u_0 = A_u$, it follows that $\mathrm{dist}(u_*, A_u) \leqslant |\psi(w_*)|$. By Taylor's theorem (here using the hypothesis that $R^{1/2} \leqslant \varrho$),

$$|\psi_j(w_*)| \leqslant \int_0^1 (1-t) |\langle \partial_{ww}^2 \psi_j(tw_*) w_*, w_* \rangle|\, dt \lesssim |w_*|^2 \lesssim \varrho^{-1 + 2\delta_m} \leqslant R^{-1/2 + \delta_m}$$

for $1 \leqslant j \leqslant n - \dim V_u$, and combining these observations yields the desired estimate. $\square$

Verifying the transversality condition in (2)

In the prototypical case of the parabolic extension operator E_{par}, the transversality condition holds by a straightforward computation and the minimum angle can be explicitly computed (see [14, Sublemma 6.6] and Figure 2). To establish the result for variable coefficient operators, one uses the localisation to the cap τ and ball B to show that the situation is only a slight perturbation of the prototypical case. This argument is carried out in detail below.

The transversality of the planes V and V' heavily relies upon the positive-definite hypothesis (H2$^+$); the property does not hold in general if one only assumes the weaker condition (H2).

(a) $V_\omega = V \cap e_n^\perp$. (b) $\angle(v, v') \gtrsim 1$ for $v \in V$, $v' \in V'$ non-zero.
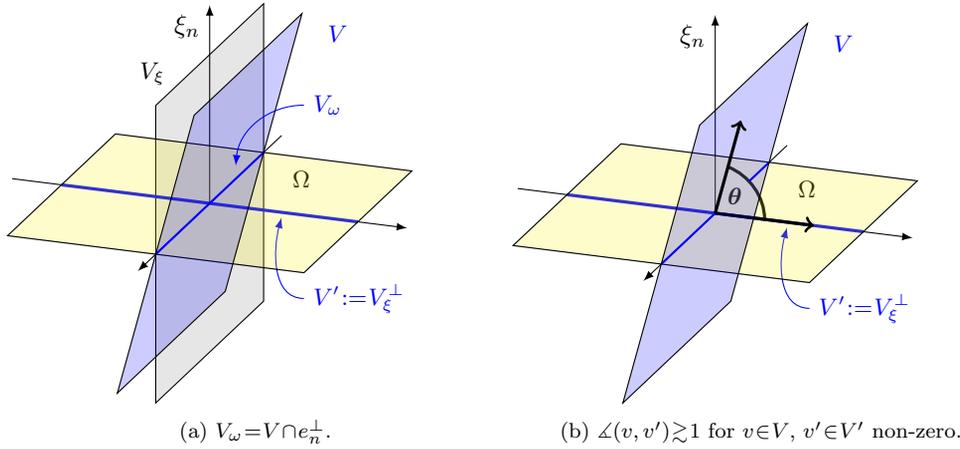
Figure 2. The transversality condition in the prototypical case of the extension operator E_{par}. Here S_ω is an affine subspace and V_ω is defined to be the linear subspace parallel to S_ω. Moreover, V_ω coincides with V_u and is given by the intersection of V with the horizontal slice $e_n^\perp = \mathbb{R}^{n-1} \times \{0\}$. From the right-hand diagram it is clear that

$$\theta := \min_{v \in V \setminus \{0\}, v' \in V' \setminus \{0\}} \angle(v, v') = \overline{\angle}(V, e_n^\perp) \gtrsim 1;$$

see [14, Sublemma 6.6] for a formal proof of this fact. For general operators T^λ with positive-definite phase, Claim 4 guarantees that V_u is a slight perturbation of the horizontal slice $V \cap e_n^\perp$ (see Figure 3) so that the angle condition still holds.

Example 8.8. For $\phi_{\mathrm{hyp}}(x; \omega) := \langle x', \omega \rangle + x_3 \omega_1 \omega_2$ for $(x; \omega) \in \mathbb{R}^3 \times \mathbb{R}^2$ the oscillatory integral operator

$$E_{\mathrm{hyp}} f(x) := \int_{B^2} e^{2\pi i \phi_{\mathrm{hyp}}(x;\omega)} f(\omega) \, d\omega$$

is the extension operator associated with the hyperbolic paraboloid. This is the prototypical example of a Hörmander-type operator for which (H2$^+$) fails. Here,

$$G_0(\omega) = (-\omega_2, -\omega_1, 1)^\top,$$

and therefore, if $V := \{x \in \mathbb{R}^3 : x_1 = 0\}$, then

$$S_\omega := \{\omega \in B^2 : G_0(\omega) \in V\} = \{\omega \in B^2 : \omega_2 = 0\}.$$

It follows that $V_\xi = \{\xi \in \widehat{\mathbb{R}}^3 : \xi_2 = 0\}$, and so $V' := V_\xi^\perp$ is a vector subspace of V. Clearly, in this situation, the desired transversality condition completely fails.

The present analysis concerns Hörmander-type operators with reduced positive-definite phase ϕ^λ, so that ϕ^λ is a small perturbation of ϕ_{par}. Such phases do not exhibit the phenomenon observed in the above example: the following claim is key to understanding this.

CLAIM 4. *Let c_{par} be the constant defined in §4. Then,*

$$\max_{v^* \in V \cap (S^{n-2} \times \{0\})} \measuredangle(v^*, V_u) = O(c_{\mathrm{par}}).$$

Here, V_u is identified with a subspace of $e_n^\perp = \mathbb{R}^{n-1} \times \{0\}$ in the natural manner.

Example 8.9. Returning to the example of the hyperbolic paraboloid with

$$V := \{x \in \mathbb{R}^3 : x_1 = 0\},$$

the spaces $V \cap (\mathbb{R}^2 \times \{0\})$ and $V_u := \{x \in \mathbb{R}^3 : x_2 = x_3 = 0\}$ are mutually orthogonal, and so the claim does not hold in this case.

Provided $c_{\mathrm{par}} > 0$ is chosen sufficiently small, the claim implies the transversality condition. Indeed, let $\{v_1^*, ..., v_{\dim V - 1}^*\}$ be an orthonormal basis for $V \cap e_n^\perp$. Fix a unit vector $v_{\dim V}^* \in V$ which is perpendicular to $V \cap e_n^\perp$, so that $\{v_1^*, ..., v_{\dim V}^*\}$ forms an orthonormal basis for V. By the above claim, there exist $v_k \in V_u \cap S^{n-2} \subset e_n^\perp$ such that

$$\measuredangle(v_k^*, v_k) = O(c_{\mathrm{par}}) \quad \text{for } 1 \leqslant k \leqslant \dim V - 1.$$

Applying the Gram–Schmidt process, one may further assume that $\{v_1, ..., v_{\dim V - 1}\}$ forms an orthonormal basis of V_u; adjoining e_n to this set then gives an orthonormal basis of V_ξ. Given $v \in V \cap S^{n-1}$ and writing

$$v = \sum_{k=1}^{\dim V - 1} \langle v, v_k^* \rangle v_k + \langle v, v_{\dim V}^* \rangle v_{\dim V}^* + \sum_{k=1}^{\dim V - 1} \langle v, v_k^* \rangle (v_k^* - v_k),$$

since $\sin \measuredangle(v, V') = |\mathrm{proj}_{V_\xi} v|$, it follows that

$$\sin \measuredangle(v, V') = \left(\sum_{k=1}^{\dim V - 1} |\langle v, v_k^* \rangle|^2 + |\langle v, v_{\dim V}^* \rangle|^2 |\langle v_{\dim V}^*, e_n \rangle|^2 \right)^{1/2} - O(c_{\mathrm{par}})$$

$$\geqslant \left(\sum_{k=1}^{\dim V} |\langle v, v_k^* \rangle|^2 \right)^{1/2} \cdot |\langle v_{\dim V}^*, e_n \rangle| - O(c_{\mathrm{par}}).$$

Consequently, provided that c_{par} is chosen to be sufficiently small,

$$\sin \measuredangle(v, V') \geqslant |\langle v_{\dim V}^*, e_n \rangle| - O(c_{\mathrm{par}}) \gtrsim 1;$$

indeed, the last inequality holds, since (8.5) implies that

$$|\langle v_{\dim V}^*, e_n \rangle| = \measuredangle(v_{\dim V}^*, e_n^\perp) = \overline{\measuredangle}(V, e_n^\perp) \gtrsim 1;$$

see Figure 3. This concludes the proof of the transversality condition, conditional on the above claim.
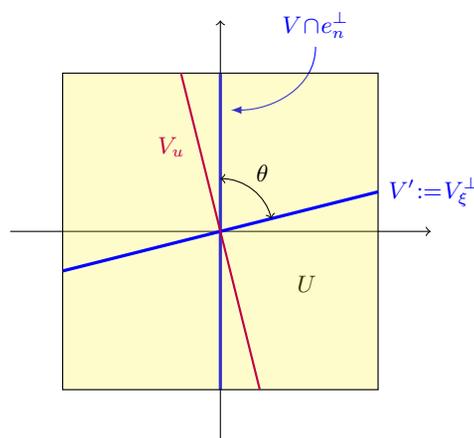
Figure 3. In the variable coefficient case, Claim 4 guarantees that the subspace V_u makes a small angle with the horizontal slice $V \cap e_n^\perp$. This implies that the angle θ is always large (that is, bounded below by an absolute constant) and thereby ensures that V and V' are transverse.

Proof of Claim 4. Fix $v^* \in V \cap (S^{n-2} \times \{0\})$ and let $v \in V_u \cap S^{n-2}$ denote the unit normalisation of the vector $\mathrm{proj}_{V_u} v^*$. It suffices to show that $\measuredangle(v^*, v) = O(c_{\mathrm{par}})$. Since

$$v^* - v = (|\mathrm{proj}_{V_u} v^*| - 1)v + \mathrm{proj}_{V_u^\perp} v^*$$

and $\big||\mathrm{proj}_{V_u} v^*| - 1\big| \leqslant |\mathrm{proj}_{V_u^\perp} v^*|$, the problem is further reduced to proving that

$$|\mathrm{proj}_{V_u^\perp} v^*| = O(c_{\mathrm{par}}). \tag{8.8}$$

Recall that

$$
\begin{aligned}
S_u &:= \{u \in U : G_0^\lambda(\bar{x}; \Psi(u)) \in V\} \\
&= \{u \in U : \langle G_0^\lambda(\bar{x}; \Psi(u)), N_k \rangle = 0 \text{ for } 1 \leqslant k \leqslant n - \dim V\},
\end{aligned}
$$

where, as above, $\{N_1, ..., N_{n-\dim V}\}$ is a choice of orthonormal basis for $V^\perp$. If $\bar{h}$ is the function introduced in (8.6), then $u \mapsto (u, \bar{h}(u))$ is a graph parametrisation of the surface Σ and $u \mapsto G_0^\lambda(\bar{x}; \Psi(u))$ is the (unnormalised) Gauss map associated with this parametrisation. Thus, the surface S_u is defined by the equations

$$-\langle \partial_u \bar{h}(u), N_k' \rangle + N_{k,n} = 0 \quad \text{for } 1 \leqslant k \leqslant n - \dim V,$$

where $N_k = (N_k', N_{k,n}) \in \mathbb{R}^{n-1} \times \mathbb{R}$. By differentiating these expressions, one deduces that a basis for $V_u^\perp$ is given by $\{M_1, ..., M_{n-\dim V}\}$, where

$$M_k := \partial_{uu}^2 \bar{h}(u_0) N_k' \quad \text{for } 1 \leqslant k \leqslant n - \dim V.$$

Let $1 \leqslant k \leqslant n - \dim V$, and recall from Lemma 4.5 that

$$\|\partial_{uu}^2 \bar{h}(u_0) - I_{n-1}\|_{op} = O(c_{par}).\tag{8.9}$$

Consequently,

$$|M_k - N_k'| = O(c_{par})\tag{8.10}$$

and, combining this with the fact that $\langle v^*, N_k' \rangle = 0$ for $1 \leqslant k \leqslant n - \dim V$ (where v^* is identified with a vector in $\mathbb{R}^{n-1}$ in the natural manner), it follows that

$$\langle v^*, M_k \rangle = \langle v^*, M_k - N_k' \rangle = O(c_{par}).\tag{8.11}$$

Let $\mathbf{M}$ be the $(n-1) \times (n - \dim V)$ matrix whose kth column is given by the vector M_k. The orthogonal projection of v^* onto the subspace $V_u^\perp$ can be expressed in terms of $\mathbf{M}$ via the formula

$$\mathrm{proj}_{V_u^\perp} v^* := \mathbf{M}(\mathbf{M}^\top \mathbf{M})^{-1} \mathbf{M}^\top v^*.$$

By (8.11), the components of the vector $\mathbf{M}^\top v^*$ are all $O(c_{par})$. Thus, to prove (8.8) (and thereby establish the claim) it suffices to show that $\|\mathbf{M}(\mathbf{M}^\top \mathbf{M})^{-1}\|_{op} \lesssim 1$, which would in turn follow from

$$\|\mathbf{M}\|_{op} \lesssim 1 \quad \text{and} \quad \|(\mathbf{M}^\top \mathbf{M})^{-1}\|_{op} \lesssim 1.$$

The bound for $\mathbf{M}$ is an immediate consequence of the definition of the M_k and (8.9). The remaining estimate would follow if one could show that, provided c_{par} is sufficiently small, $|\lambda| \gtrsim 1$ for every eigenvalue λ of the symmetric matrix $\mathbf{M}^\top \mathbf{M}$. By (8.10) and continuity of eigenvalues, it suffices to show that the matrix $\mathbf{N}^\top \mathbf{N}$ satisfies the same property, where $\mathbf{N}$ is the $(n-1) \times (n - \dim V)$ matrix whose kth column is given by the vector N_k'. By (8.5), the vectors $N_1', ..., N_{n-\dim V}' \in \mathbb{R}^{n-1}$ are linearly independent and, moreover, satisfy

$$|\det \mathbf{N}^\top \mathbf{N}| = |N_1' \wedge ... \wedge N_{n-\dim V}'|^2 \gtrsim 1.$$

Therefore, the desired condition on the eigenvalues holds if the spectral radius of $\mathbf{N}^\top \mathbf{N}$ is $O(1)$. But the latter property is an obvious consequence of the Newton–Girard identity

$$\sum_{i=1}^m a_i^2 = \left(\sum_{i=1}^m a_i\right)^2 - 2 \sum_{1 \leqslant i_1 < i_2 \leqslant m} a_{i_1} a_{i_2}, \quad a_i \in \mathbb{R} \text{ for } 1 \leqslant i \leqslant m,$$

and the fact that the entries of $\mathbf{N}^\top \mathbf{N}$ are all $O(1)$. $\qquad\square$

This concludes the proof of Lemma 8.7. $\qquad\square$

8.4. The proof of the transverse equidistribution estimate

It remains to demonstrate how to pass from Lemma 8.7 to Lemma 8.4. At this stage, the proof is very similar to the argument found in [14], but the details are nevertheless included for completeness.

There are two additional ingredients needed for the proof of Lemma 8.4. The first is the following theorem of Wongkew [37] (see also [13], [39]), which controls the size of a neighbourhood of a variety.

THEOREM 8.10. (Wongkew [37]) *Suppose* $Z = Z(P_1, ..., P_{n-m})$ *is an m-dimensional transverse complete intersection in $\mathbb{R}^n$ with $\overline{\deg}\, Z \leqslant D$. For any $0 < \varrho \leqslant R$ and R-ball B_R, the neighbourhood $N_\varrho(Z \cap B_R)$ can be covered by $O_D((R/\varrho)^m)$ balls of radius ϱ.*

The second ingredient is a geometric lemma concerning planar slices of neighbourhoods of varieties. The statement of this result requires a general quantitative notion of transversality for pairs of linear subspaces in $\mathbb{R}^n$. Any m-dimensional linear subspace V can be expressed as a transverse complete intersection $V = Z(P_{N_1}, ..., P_{N_{n-m}})$, where $\{N_1, ..., N_{n-m}\}$ forms an orthonormal basis of $V^\perp$ and $P_{N_j}(x) := \langle x, N_j \rangle$. Suppose V_1 and V_2 are linear subspaces in $\mathbb{R}^n$ satisfying

$$\dim V_1 + \dim V_2 \geqslant n. \tag{8.12}$$

It is easy to verify that the subspace $V_1 \cap V_2$ is a transverse complete intersection if and only if

$$\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - n$$

(of course, the inequality $\dim(V_1 \cap V_2) \geqslant \dim V_1 + \dim V_2 - n$ always holds, so the latter condition says that $V_1 \cap V_2$ is as small as possible).

Definition 8.11. A pair (V_1, V_2) of linear subspaces in $\mathbb{R}^n$ satisfying (8.12) is said to be *quantitatively transverse* if the following hold:
  (i) $\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - n$;
  (ii) $\angle(v_1, v_2) \geqslant c_{\text{trans}}$ for all non-zero vectors $v_j \in (V_1 \cap V_2)^\perp \cap V_j$, $j = 1, 2$.

Remark 8.12. In the special case where $\dim V_1 + \dim V_2 = n$, it follows that the pair (V_1, V_2) is quantitatively transverse if and only if $\angle(v_1, v_2) \geqslant c_{\text{trans}}$ for all non-zero vectors $v_1 \in V_1$ and $v_2 \in V_2$. Thus, up to the minor disparity between the choice of constant c_{trans}, this agrees with the transversality condition appearing in the statement of Lemma 8.7.

LEMMA 8.13. *There exists some dimensional constant $C > 0$ such that the following holds. Let $B_r \subseteq \mathbb{R}^n$ be an r-ball, $V \subseteq \mathbb{R}^n$ be a linear subspace, Z be a transverse complete intersection and suppose that $\dim Z + \dim V \geqslant n$ and $(T_z Z, V)$ is a quantitatively*

transverse pair for all $z \in Z \cap 2B_r$. Then,

$$V \cap B_r \cap N_\varrho(Z) \subseteq N_{C\varrho}(V \cap Z)$$

for all $0 < \varrho \ll r$.

The proof of this simple lemma is postponed until the end of this section.

Proof of Lemma 8.4. If $T_{\theta,v} \cap N_{R^{1/2+\delta_m}}(Z) \cap B = \varnothing$, then it follows that

$$|T^\lambda g_{\theta,v}(x)| = \mathrm{RapDec}(R)\|g\|_{L^2(B^{n-1})} \quad \text{for all } x \in N_{\varrho^{1/2+\delta_m}}(Z) \cap B.$$

Consequently, one may assume that g is concentrated on only those wave packets from $\mathbb{T}_{B,\tau,Z}$ for which $T_{\theta,v}$ intersects $N_{R^{1/2+\delta_m}}(Z) \cap B$ non-trivially. Suppose $(\theta,v) \in \mathbb{T}_{B,\tau,Z}$ has this property and let $x \in T_{\theta,v} \cap N_{R^{1/2+\delta_m}}(Z) \cap B$. If $z \in Z \cap 2B$, then $|x-z| \lesssim R^{1/2+\delta_m}$ and, by the $R^{1/2+\delta_m}$-tangent condition,

$$\measuredangle(G^\lambda(x;\theta), T_z Z) \lesssim R^{-1/2+\delta_m}.$$

Since $|G^\lambda(\bar{x};\theta) - G^\lambda(x;\theta)| \lesssim |\bar{x}-x|/\lambda \lesssim R^{-1/2+\delta_m}$, one concludes that

$$\measuredangle(G^\lambda(\bar{x};\theta), T_z Z) \lesssim R^{-1/2+\delta_m} \quad \text{for all } z \in Z \cap 2B.$$

Thus, there exists a subspace $V \subseteq \mathbb{R}^n$ of minimal dimension $\dim V \leqslant \dim Z$ such that

$$\measuredangle(G^\lambda(\bar{x};\theta), V) \lesssim R^{-1/2+\delta_m}$$

for all $(\theta,v) \in \mathbb{T}_{B,\tau,Z}$ for which $T_{\theta,v} \cap N_{R^{1/2+\delta_m}}(Z) \cap B \neq \varnothing$. In particular, g is concentrated on wave packets from $\mathbb{T}_{B,\tau,V}$. One may apply Lemma 8.7 to find a subspace V' of dimension $n - \dim V$ such that

$$\measuredangle(v,v') \geqslant 2c_{\mathrm{trans}} \quad \text{for all non-zero vectors } v \in V \text{ and } v' \in V', \tag{8.13}$$

and

$$\int_{\Pi \cap B(x_0, \varrho^{1/2+\delta_m})} |T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)} \left(\frac{\varrho}{R}\right)^{\dim V'/2} \|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)} \left(\int_{\Pi \cap 2B} |T^\lambda g|^2\right)^{1/(1+\delta)} \tag{8.14}$$

for every affine subspace Π parallel to V'. More precisely, the above estimate holds, up to the inclusion of some additional rapidly decreasing term. This small error will propagate through the remainder of the argument, but in the end it will be harmless and is therefore suppressed in the notation.

It is claimed that, for each $z \in Z \cap 2B$, the tangent space $T_z Z$ forms a quantitatively transverse pair $(T_z Z, V')$ with V'. Indeed, if this fails, then it is easy to see that, for some $z \in Z \cap 2B$, there exists a subspace $W \subseteq T_z Z$ of dimension

$$\dim W > \dim Z - \dim V$$

with the property that $\measuredangle(w, V') < c_{\text{trans}}$ for all $w \in W \setminus \{0\}$. Consequently, the crucial angle condition (8.13) guarantees that

$$\measuredangle(w, V) \geqslant c_{\text{trans}} \quad \text{for all } w \in W \setminus \{0\}.$$

This implies that there exists 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 $\|L\|_{\text{op}} \lesssim 1$. Recall that, for each $(\theta, v) \in \mathbb{T}_{B,Z}$, one has $\measuredangle(G^\lambda(\bar{x}, \theta), V) \lesssim R^{-1/2 + \delta_m}$, and so

$$\sup_{\omega, \omega' \in \theta} |L(G^\lambda(\bar{x}, \omega)) - G^\lambda(\bar{x}, \omega')| \lesssim R^{-1/2 + \delta_m}.$$

On the other hand, $G^\lambda(\bar{x}, \theta) \subset N_{C_1 R^{-1/2 + \delta_m}}(T_z Z) \cap S^{n-1}$, and so $L(G^\lambda(\bar{x}, \theta))$ lies in

$$L(N_{C_1 R^{-1/2 + \delta_m}}(T_z Z) \cap S^{n-1}) \subseteq N_{C_2 R^{-1/2 + \delta_m}}(L(T_z Z)).$$

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

$$\measuredangle(G^\lambda(\bar{x}, \theta), L(T_z Z)) \lesssim R^{-1/2 + \delta_m}.$$

Since L vanishes on W, by rank-nullity $L(T_z Z)$ is a subspace of dimension at most

$$\dim Z - \dim W < \dim V.$$

This contradicts the minimality of V, and so $(T_z Z, V')$ is a quantitatively transverse pair for all $z \in Z \cap 2B$.

By Lemma 8.13, one deduces that

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

Since $\Pi \cap Z$ is a transverse complete intersection of dimension $\dim V' + \dim Z - n$, Wongkew's theorem now implies that $\Pi \cap N_{\varrho^{1/2 + \delta_m}}(Z) \cap B$ can be covered by

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

balls of radius $\varrho^{1/2+\delta_m}$. Applying the estimate (8.14) to each of these balls and summing, one deduces that

$$\int_{\Pi\cap N_{\varrho^{1/2+\delta_m}}(Z)\cap B}|T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)}\Big(\frac{\varrho}{R}\Big)^{(n-m)/2}\|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)}\Big(\int_{\Pi\cap 2B}|T^\lambda g|^2\Big)^{1/(1+\delta)}.$$

Integrating over planes Π parallel to V' and applying Hölder's inequality (recalling that $\delta\ll\delta_m$), it follows that

$$\int_{N_{\varrho^{1/2+\delta_m}}(Z)\cap B}|T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)}\Big(\frac{\varrho}{R}\Big)^{(n-m)/2}\|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)}\Big(\int_{2B}|T^\lambda g|^2\Big)^{1/(1+\delta)}.$$

Finally, recalling Hörmander's bound,

$$\Big(\int_{2B}|T^\lambda g|^2\Big)^{1/(1+\delta)} \lesssim R^{1/2+O(\delta_m)}\Big(\int_{B^{n-1}}|g|^2\Big)^{1/(1+\delta)}$$

and absorbing the implied rapidly decaying error into the main term, one concludes that

$$\int_{N_{\varrho^{1/2+\delta_m}}(Z)\cap B}|T^\lambda g|^2 \lesssim_\delta R^{1/2+O(\delta_m)}\Big(\frac{\varrho}{R}\Big)^{(n-m)/2}\|g\|_{L^2(B^{n-1})}^2,$$

which is the desired estimate. $\qquad\square$

It remains to prove Lemma 8.13.

Proof of Lemma 8.13. Applying a rotation, one may assume that V is the span of the coordinate vectors $e_1,...,e_{\dim V}$. For the purposes of this proof, $\gamma_V:=(\gamma_1,...,\gamma_{\dim V})$ and $\gamma_{V^\perp}=(\gamma_{\dim V+1},...,\gamma_n)$ will denote the orthogonal projections of a space curve γ onto V and $V^\perp$, respectively.

Suppose that $x\in V\cap B_r\cap N_\varrho(Z)$ and fix some $z_0\in Z\cap N_\varrho(B_r)$ with $0<|x-z_0|<\varrho$. Let $\gamma:\mathbb{R}\to\mathbb{R}^n$ be the constant speed parametrisation of the line through $\gamma(0):=z_0$ and $\gamma(1):=x$. To prove the lemma, it suffices to show that there exists a curve $\widetilde\gamma:[0,1]\to\mathbb{R}^n$ such that, for all $t\in[0,1]$, the following hold:

(1) $\widetilde\gamma(0)=\gamma(0)=z_0$;
(2) $\widetilde\gamma(t)\in Z$;
(3) $\widetilde\gamma_{V^\perp}(t)=\gamma_{V^\perp}(t)$;
(4) $|\widetilde\gamma'(t)|\leqslant\overline{C}|\widetilde\gamma'_{V^\perp}(t)|$ where $\overline{C}:=(\sin c_{\mathrm{trans}})^{-1}$.

Indeed, once this is established, observe that $z_1:=\widetilde\gamma(1)\in Z\cap V$ by properties (2) and (3). Furthermore, (3) and (4) ensure that

$$|\widetilde\gamma'(t)|\leqslant\overline{C}|\widetilde\gamma'_{V^\perp}(t)|\leqslant\overline{C}|\gamma'(t)|<\overline{C}\varrho,$$

and so, combining this observation with (1),

$$|x - \widetilde{\gamma}(t)| \leqslant |x - z_0| + |z_0 - \widetilde{\gamma}(t)| < (1 + \overline{C}t)\varrho \tag{8.15}$$

for all $t \in [0, 1]$. In particular, $|x - z_1| \lesssim \varrho$, giving the desired conclusion.

The transversality condition implies that the distribution (in the sense of Frobenius; see, for instance, [34, Chapter 1])

$$W_z := (T_z Z \cap V)^\perp \cap T_z Z$$

has rank $n - \dim V$ on $Z \cap 2B_r$ and, moreover, $\mathrm{proj}_{W_z^\perp}|_V : V \to W_z^\perp$ is an isomorphism for all $z \in Z \cap 2B_r$. Smoothly extend W_z to a small neighbourhood U of $Z \cap 2B_r$, so that

$$\mathrm{proj}_{W_x^\perp}|_V : V \to W_x^\perp \text{ is an isomorphism} \quad \text{for all } x \in U. \tag{8.16}$$

The curve $\widetilde{\gamma}$ will be chosen so that its tangent always lies in this distribution. Given that $\widetilde{\gamma}_{V^\perp}$ is already defined by property (3), to satisfy this condition $\widetilde{\gamma}_V$ must be a solution to the ordinary differential equation (ODE)

$$\begin{cases} \mathrm{proj}_{W_{x(t)}^\perp}(y'(t), \gamma'_{V^\perp}(t)) = 0, \\ y(0) = \mathrm{proj}_V z_0, \end{cases}$$

where $x(t) := (y(t), \gamma_{V^\perp}(t))$. By (8.16), solving the above ODE is equivalent to solving a system of the form

$$\begin{cases} y' = g(t, y), \\ y(0) = \mathrm{proj}_V z_0, \end{cases} \tag{8.17}$$

for a smooth function g defined on $\{(t, y) \in \mathbb{R} \times V : (y, \gamma_{V^\perp}(t)) \in U\}$. Note that g can be described explicitly in terms of the inverse of $\mathrm{proj}_{W_x^\perp}|_V$ and, provided U is appropriately chosen, the derivatives of g are bounded.

The Picard–Lindelöf existence theorem implies that the system (8.17) has a solution $\widetilde{\gamma}_V$ defined on an interval $[0, T]$ for some $T > 0$ such that $\widetilde{\gamma} := (\widetilde{\gamma}_V, \gamma_{V^\perp})$ satisfies $\widetilde{\gamma}(t) \in 2B_r$ for all $t \in [0, T]$. It can be checked that on this interval the curve $\widetilde{\gamma}$ further satisfies (1) and (3) and, by the tangency condition which motivated the definition of the ODE, (2) also holds. If $t \in [0, T]$, then it follows that $\widetilde{\gamma}(t) \in Z \cap 2B_r$ and $\widetilde{\gamma}'(t) \in W_{\widetilde{\gamma}(t)}$, and so the transversality hypothesis implies that

$$\overline{C}^{-1} = \sin c_{\mathrm{trans}} \leqslant \sin \measuredangle(\widetilde{\gamma}'(t), V) = \frac{|\widetilde{\gamma}'_{V^\perp}(t)|}{|\widetilde{\gamma}'(t)|}.$$

Rearranging, one concludes that properties (1)–(4) all hold on $[0, T]$.

It remains to show that T can be chosen to satisfy $T \geqslant 1$. If $\mathrm{dist}(\widetilde{\gamma}(T), U^c) \gtrsim 1$, then the regularity of g implies that the interval of existence can be extended by a fixed increment. Thus, one may assume that at least one of the following holds: $T \geqslant 1$ or $|\widetilde{\gamma}(T) - x| \geqslant \frac{1}{2} r$.[19] Supposing the latter holds, by the choice of T and (8.15), one deduces that

$$\tfrac{1}{2} r \leqslant |\widetilde{\gamma}(T) - x| \leqslant (1 + \overline{C} T) \varrho.$$

Provided r is chosen to be sufficiently large compared to ϱ, the desired bound immediately follows. $\qquad\square$

9. Comparing wave packets at different spatial scales

9.1. Wave packet decomposition at scale ϱ

The proof of Theorem 1.9 relies on a multi-scale analysis, and for this it is necessary to compare wave packets at different scales.

Let $1 \ll R \ll \lambda$, and recall the decomposition

$$T^\lambda f(x) = \sum_{(\theta, v) \in \mathbb{T}} T^\lambda f_{\theta, v}(x) + \mathrm{RapDec}(R) \|f\|_{L^2(B^{n-1})}$$

described in §5. Consider a smaller spatial scale[20] $R^{1/2} \leqslant \varrho \leqslant R$, and fix $B(y, \varrho) \subset B(0, R)$ with centre $y \in X^\lambda$. Each of the $T^\lambda f_{\theta, v}$ can be further decomposed into wave packets at scale ϱ over $B(y, \varrho)$. To do this, first apply a transformation to recentre $B(y, \varrho)$ at the origin. For $g \colon B^{n-1} \to \mathbb{C}$ integrable, define $\tilde{g} := e^{2\pi i \phi^\lambda(y; \cdot)} g$, so that

$$T^\lambda g(x) = \widetilde{T}^\lambda \tilde{g}(\tilde{x}) \quad \text{for } \tilde{x} = x - y, \tag{9.1}$$

where $\widetilde{T}^\lambda$ is the Hörmander-type operator with phase $\tilde{\phi}^\lambda$ and amplitude $\tilde{a}^\lambda$ given by

$$\tilde{\phi}(x; \omega) := \phi\Big(x + \frac{y}{\lambda}; \omega\Big) - \phi\Big(\frac{y}{\lambda}; \omega\Big) \quad \text{and} \quad \tilde{a}(x; \omega) := a\Big(x + \frac{y}{\lambda}; \omega\Big). \tag{9.2}$$

Applying this identity to the wave packet decomposition above,

$$T^\lambda f(x) = \sum_{(\theta, v) \in \mathbb{T}} \widetilde{T}^\lambda ((f_{\theta, v})^\curvearrowright)(\tilde{x}) + \mathrm{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

---

[19] Indeed, suppose both conditions fail for T. The failure of the latter condition implies that $\mathrm{dist}(\widetilde{\gamma}(T), (2B_r)^c) \geqslant \frac{1}{2} r$. Since $\widetilde{\gamma}(T) \in Z$ by property (2), one concludes that $\widetilde{\gamma}(T)$ is far from U^c, and thus the interval of existence for $\widetilde{\gamma}$ can be extended by a fixed increment. One may redefine T to be some value in the interval of existence incrementally larger than the original value of T, and repeat this procedure until at least one of the stated conditions hold.

[20] Later it will be useful to assume the more stringent condition $R^{1/2} \leqslant \varrho \leqslant R^{1-2\delta}$.

Each $T^\lambda f_{\theta,v}$ is (spatially) concentrated on the curved $R^{1/2+\delta}$-tube $T_{\theta,v}$ and, consequently, each $\widetilde{T}^\lambda (f_{\theta,v})^\sim$ is concentrated on the translate $T_{\theta,v}-y$. Since[21]

$$\partial_\omega \tilde{\phi}^\lambda((\gamma_{\omega,v}^\lambda(t),t)-y;\omega) = \partial_\omega \phi^\lambda((\gamma_{\omega,v}^\lambda(t),t);\omega) - \partial_\omega \phi^\lambda(y;\omega) = v - \partial_\omega \phi^\lambda(y;\omega), \qquad (9.3)$$

the core curve $\Gamma_{\theta,v}^\lambda - y$ of $T_{\theta,v}-y$ is equal to $\Gamma_{\theta,v-\bar{v}(y;\omega_\theta)}$, where

$$\bar{v}(y;\omega) := \partial_\omega \phi^\lambda(y;\omega).$$

Now, repeat the construction of the wave packets for each $\widetilde{T}^\lambda(f_{\theta,v})^\sim$, but at scale ϱ. In particular, cover Ω by finitely overlapping caps $\tilde{\theta}$ of radius $\varrho^{-1/2}$, and $\mathbb{R}^{n-1}$ by finitely-overlapping balls of radius $\varrho^{(1+\delta)/2}$ centered at vectors $\tilde{v} \in \varrho^{(1+\delta)/2}\mathbb{Z}^{n-1}$. Let $\widetilde{\mathbb{T}}$ denote the set of all pairs $(\tilde{\theta},\tilde{v})$. For each $(\theta,v) \in \mathbb{T}$ one may decompose

$$(f_{\theta,v})^\sim = \sum_{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}} (f_{\theta,v})^\sim_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})},$$

as in §5. The significant contributions to this sum arise from pairs $(\tilde{\theta},\tilde{v})$ belonging to

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

as demonstrated by the following lemma.

LEMMA 9.1. *If $R^{1/2} \leqslant \varrho \leqslant R$, then, with the above definitions, the function $(f_{\theta,v})^\sim$ is concentrated on wave packets from $\widetilde{\mathbb{T}}_{\theta,v}$; that is,*

$$(f_{\theta,v})^\sim = \sum_{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}} (f_{\theta,v})^\sim_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}.$$

Proof. Since $(f_{\theta,v})^\sim$ is supported in θ, clearly the wave packets of $(f_{\theta,v})^\sim$ at scale ϱ are all supported in

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

Thus, it suffices to show that, for each $(\theta,v) \in \mathbb{T}$ and $\varrho^{-1/2}$-cap $\tilde{\theta}$, one has

$$\sum_{\tilde{v}:\,|v-\bar{v}(y;\omega_\theta)-\tilde{v}| \gtrsim R^{(1+\delta)/2}} (f_{\theta,v})^\sim_{\tilde{\theta},\tilde{v}} = \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}.$$

---

[21] For every fixed ω and v, here $\gamma_{\omega,v}^\lambda$ is used to denote the curve satisfying $\partial_\omega \phi^\lambda((\gamma_{\omega,v}^\lambda(t),t);\omega)=v$ for all (admissible) $t \in (-R,R)$. In the notation of §5, $\gamma_{\omega,v}^\lambda = \gamma_{\theta,v}^\lambda$ for a cap θ with $\omega=\omega_\theta$.

Fixing $(\theta, v) \in \mathbb{T}$ and $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}$ with $|v - \bar{v}(y; \omega_\theta) - \tilde{v}| \gtrsim R^{(1+\delta)/2}$, the above estimate would follow if one could show that

$$(f_{\theta,v})_{\tilde{\theta},\tilde{v}}^{\smile} = (1 + R^{-1/2}|v - \bar{v}(y; \omega_\theta) - \tilde{v}|)^{-(n+1)} \operatorname{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

By definition, $(f_{\theta,v})_{\tilde{\theta},\tilde{v}}^{\smile} = \tilde{\psi}_{\tilde{\theta}} \cdot [(\eta_{\tilde{v}})^{\frown} * (\psi_{\tilde{\theta}} f_{\theta,v})]$ for the bump functions $\tilde{\psi}_{\tilde{\theta}}$, $\eta_{\tilde{v}}$ and $\psi_{\tilde{\theta}}$ as defined in §5. Thus, Fourier inversion yields the pointwise bound

$$|(f_{\theta,v})_{\tilde{\theta},\tilde{v}}^{\smile}(\omega)| \leqslant \|(\tilde{\psi}_{\tilde{\theta}})^{\smile}\|_{L^1(\mathbb{R}^n)} \|\eta_{\tilde{v}}(\psi_{\tilde{\theta}}(f_{\theta,v})^{\frown})^{\frown}\|_{L^\infty(\mathbb{R}^n)}$$

$$\lesssim \|(\psi_{\tilde{\theta}})^{\smile} * ((f_{\theta,v})^{\frown})^{\frown}\|_{L^\infty(B(\tilde{v}, C\varrho^{(1+\delta)/2}))}$$

for all $\omega \in \mathbb{R}^{n-1}$. Since $(\psi_{\tilde{\theta}})^{\smile}$ is concentrated in $B(0, \varrho^{1/2})$, the problem is further reduced to showing that

$$((f_{\theta,v})^{\frown})^{\frown}(z) = (1 + R^{-1/2}|z - v + \bar{v}(y; \omega_\theta)|)^{-(n+1)} \operatorname{RapDec}(R) \|f\|_{L^2(B^{n-1})}$$

whenever $|z - v + \bar{v}(y; \omega_\theta)| \gtrsim R^{(1+\delta)/2}$.

Let $\tilde{\tilde{\psi}}$ be a Schwartz function on $\mathbb{R}^{n-1}$ satisfying $\tilde{\tilde{\psi}}(\omega) = 1$ for $\omega \in B^{n-1}$, and define $\tilde{\tilde{\psi}}_\theta(\omega) := \tilde{\tilde{\psi}}(R^{1/2}(\omega - \omega_\theta))$ so that

$$((f_{\theta,v})^{\frown})^{\smile} = (\tilde{\tilde{\psi}}_\theta e^{2\pi i \phi^\lambda(y; \cdot)})^{\smile} * (f_{\theta,v})^{\smile}.$$

On the one hand, since $\operatorname{supp} \eta_v \subset B(v, CR^{(1+\delta)/2})$, it is not difficult to show that

$$|(f_{\theta,v})^{\smile}(z)| = (1 + R^{-1/2}|z - v|)^{-(n+1)} \operatorname{RapDec}(R) \|f\|_{L^2(B^{n-1})} \tag{9.4}$$

whenever $|z - v| \gtrsim R^{(1+\delta)/2}$. On the other hand, it is claimed that

$$|(\tilde{\tilde{\psi}}_\theta e^{2\pi i \phi^\lambda(y; \cdot)})^{\smile}(z)| = (1 + R^{-1/2}|z + \bar{v}(y; \omega_\theta)|)^{-(n+1)} \operatorname{RapDec}(R) \tag{9.5}$$

whenever $|z + \bar{v}(y; \omega_\theta)| \gtrsim R^{(1+\delta)/2}$. A routine argument then shows that (9.4) and (9.5) imply the desired estimate for $((f_{\theta,v})^{\frown})^{\smile}$, and so it only remains to prove the claim.

The inverse Fourier transform in (9.5) can be expressed as

$$R^{-(n-1)/2} \int_{\mathbb{R}^{n-1}} e^{2\pi i (\langle z, \omega_\theta + R^{-1/2}\omega \rangle + \phi^\lambda(y; \omega_\theta + R^{-1/2}\omega))} \tilde{\tilde{\psi}}(\omega) \, d\omega,$$

where the ω-gradient of the phase is given by

$$R^{-1/2}([\partial_\omega \phi^\lambda(y; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(y; \omega_\theta)] + [\bar{v}(y; \omega_\theta) + z]). \tag{9.6}$$

Using the fact that $\partial^2_{\omega\omega} \phi^\lambda(0; \omega) = 0$, the first term in (9.6) can thus be estimated by

$$|\partial_\omega \phi^\lambda(y; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(y; \omega_\theta)| \lesssim R^{-1/2}|y| \leqslant R^{1/2}.$$

Consequently, if $z \notin B(-\bar{v}(y; \omega_\theta), R^{(1+\delta)/2})$, then the second term in (9.6) dominates the ω-gradient of the phase and (9.6) is $\gtrsim R^{\delta/2}$ in norm. Repeated integration by parts now implies (9.4), concluding the proof. □

9.2. Tangency properties

Let Z be a transverse complete intersection of dimension m, and suppose that h is a function which is concentrated on wave packets from

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

What can be said about the scale ϱ wave packets of $\tilde{h}$? In particular, do the lower scale wave packets inherit the tangency property; namely, is $\tilde{h}$ concentrated on scale ϱ wave packets which are $\varrho^{-1/2+\delta_m}$-tangent to some variety? It transpires that this is *not* true in general. It is true, however, that $\tilde{h}$ can be broken up into pieces which are each made up of scale ϱ wave packets tangential to some translate of Z.

In particular, while all the scale ϱ wave packets in question form very small angle with $Z-y$, they can be traced all the way up to distance $\sim R^{1/2+\delta_m}$ (rather than $\lesssim \varrho^{1/2+\delta_m}$) from $Z-y$, which means that they generally live too far from $Z-y$ to be tangential to it at scale ϱ. Translations of $Z-y$ however, up to distance $\sim R^{1/2+\delta_m}$, are tangential to such remote wave packets.

To make the above discussion precise, let $\widetilde{\gamma}_{\omega,v}^{\lambda}$ be the curve defined by

$$\partial_\omega \tilde{\phi}^\lambda(\widetilde{\gamma}_{\omega,v}^\lambda(t),t;\omega) = v \quad \text{for } t \in (-\varrho,\varrho). \tag{9.7}$$

It is remarked that (9.3) implies the relation

$$\gamma_{\omega,v}^\lambda(t) = \widetilde{\gamma}_{\omega,v-\bar{v}(y;\omega)}^\lambda(t-y_n)+y'. \tag{9.8}$$

Let $\widetilde{T}_{\omega,v}$ be the $\varrho^{1/2+\delta}$-tube with core curve $\widetilde{\Gamma}_{\omega,v}^\lambda = (\widetilde{\gamma}_{\omega,v}^\lambda(t),t)$ (defined analogously to the $R^{1/2+\delta}$-tube $T_{\theta,v}$), and for $b \in \mathbb{R}^n$ define

$$\widetilde{\mathbb{T}}_{Z+b} := \{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}} : \widetilde{T}_{\tilde{\theta},\tilde{v}} \text{ is } \varrho^{-1/2+\delta_m}\text{-tangent to } Z+b \text{ in } B(0,\varrho)\}.$$

The key observation is as follows.

PROPOSITION 9.2. *Let* $R^{1/2} \leqslant \varrho \leqslant R^{1-\delta}$ *and* $Z \subseteq \mathbb{R}^n$ *be a transverse complete intersection.*

(1) *Let* $(\theta,v) \in \mathbb{T}_Z$ *and* $b \in B(0,2R^{1/2+\delta_m})$. *If* $(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$ *satisfies*

$$\widetilde{T}_{\tilde{\theta},\tilde{v}} \cap N_{\varrho^{1/2+\delta_m}/2}(Z-y+b) \neq \varnothing,$$

then $(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_{Z-y+b}$.

(2) *If* h *is concentrated on wave packets in* $\mathbb{T}_{Z,B(y,\varrho)}$, *then* $\tilde{h}$ *is concentrated on wave packets in*

$$\bigcup_{|b| \lesssim R^{1/2+\delta_m}} \widetilde{\mathbb{T}}_{Z-y+b}.$$

In view of the forthcoming analysis, before proving these statements a simple application is discussed. Under the hypotheses of the proposition, if one defines

$$\widetilde{\mathbb{T}}_b := \left\{ (\tilde{\theta}, \tilde{v}) \in \bigcup_{(\theta,v) \in \mathbb{T}_{Z,B(y,\varrho)}} \widetilde{\mathbb{T}}_{\theta,v} : \widetilde{T}_{\tilde{\theta},\tilde{v}} \cap N_{\varrho^{1/2+\delta_m}/2}(Z-y+b) \neq \varnothing \right\},$$

then it follows that $\widetilde{\mathbb{T}}_b \subseteq \widetilde{\mathbb{T}}_{Z-y+b}$. Given a function h concentrated on wave packets in $\mathbb{T}_{Z,B(y,\varrho)}$, consider a function of the form

$$\tilde{h}_b := \sum_{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}_b} \tilde{h}_{\tilde{\theta},\tilde{v}}. \tag{9.9}$$

Expressions of the form (9.9) will play an important role in later arguments. Proposition 9.2 implies that

$$\widetilde{T}^\lambda \tilde{h}_b(\tilde{x}) = T^\lambda h_b(x) \chi_{N_{\varrho^{1/2+\delta_m}}(Z+b)}(x) + \mathrm{RapDec}(R) \|h\|_{L^2(B^{n-1})} \tag{9.10}$$

for all $x = \tilde{x} + y \in B(y, \varrho)$.

The proof of Proposition 9.2 relies on the following lemma.

LEMMA 9.3. *If* $(\theta, v) \in \mathbb{T}$ *and* $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$, *then*

$$|\widetilde{\Gamma}^\lambda_{\tilde{\theta},\tilde{v}}(t) - (\Gamma^\lambda_{\theta,v}(t+y_n) - y)| \lesssim R^{(1+\delta)/2} \quad \text{for all } t \in (-\varrho, \varrho).$$

Proof. By the identity (9.8) and the definition of $\widetilde{\mathbb{T}}_{\theta,v}$, it suffices to show that, if $(\omega_1, v_1), (\omega_2, v_2) \in \Omega \times \mathbb{R}^{n-1}$ satisfy $|\omega_1 - \omega_2| \lesssim \varrho^{-1/2}$ and $|v_1 - v_2| \lesssim R^{(1+\delta)/2}$, then

$$|\widetilde{\gamma}^\lambda_{\omega_1,v_1}(t) - \widetilde{\gamma}^\lambda_{\omega_2,v_2}(t)| \lesssim R^{(1+\delta)/2} \quad \text{for all } t \in (-\varrho, \varrho).$$

Fixing $t \in (-\varrho, \varrho)$, let $x_t := (\widetilde{\gamma}^\lambda_{\omega_1,v_1}(t), t)$ and $v_t := \partial_\omega \tilde{\phi}^\lambda(x_t; \omega_2)$, and note that, since the value of $\widetilde{\gamma}_{\omega,v}(t)$ is uniquely determined by (9.7), $x_t = (\widetilde{\gamma}^\lambda_{\omega_2,v_t}(t), t)$. Observe that

$$|v_1 - v_t| = |\partial_\omega \tilde{\phi}^\lambda(x_t; \omega_1) - \partial_\omega \tilde{\phi}^\lambda(x_t; \omega_2)| \lesssim |\omega_1 - \omega_2| \, |x_t| \lesssim \varrho^{1/2}.$$

Since $|v_1 - v_2| \lesssim R^{(1+\delta)/2}$, it follows that $|v_t - v_2| \lesssim R^{(1+\delta)/2}$. Therefore,

$$|\widetilde{\gamma}^\lambda_{\omega_1,v_1}(t) - \widetilde{\gamma}^\lambda_{\omega_2,v_2}(t)| = |\widetilde{\gamma}^\lambda_{\omega_2,v_t}(t) - \widetilde{\gamma}^\lambda_{\omega_2,v_2}(t)| \sim |v_t - v_2| \lesssim R^{(1+\delta)/2},$$

which establishes the lemma. $\square$

One may now turn to the proof of Proposition 9.2.

Proof of Proposition 9.2. The proof is broken into stages.

The angle condition.

Fix $(\theta, v) \in \mathbb{T}_Z$ and $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta, v}$. Motivated by the definition of tangency, let $x \in \widetilde{T}_{\tilde{\theta}, \tilde{v}}$ and suppose $z \in Z$ and $b \in B(0, 2R^{1/2+\delta_m})$ are such that

$$z - y + b \in B(0, 4\varrho) \quad \text{and} \quad |x - (z - y + b)| \leqslant \overline{C}_{\text{tang}} \varrho^{1/2+\delta_m}.$$

It is claimed that

$$\measuredangle(\widetilde{G}^\lambda(x; \omega_{\tilde{\theta}}), T_{z-y+b}(Z-y+b)) \leqslant \overline{c}_{\text{tang}} \varrho^{-1/2+\delta_m}, \tag{9.11}$$

where $\widetilde{G}^\lambda$ is the generalised Gauss map associated with the phase $\tilde{\phi}^\lambda$. It is easy to see that $\widetilde{G}^\lambda(x; \omega) = G^\lambda(x + y; \omega)$ and $T_{z-y+b}(Z - y + b) = T_z Z$, so the above estimate can be written as

$$\measuredangle(G^\lambda(x + y; \omega_{\tilde{\theta}}), T_z Z) \leqslant \overline{c}_{\text{tang}} \varrho^{-1/2+\delta_m}. \tag{9.12}$$

By Lemma 9.3, the definition of $\widetilde{T}_{\tilde{\theta}, \tilde{v}}$ and the hypothesis $\varrho \leqslant R^{1-\delta}$, it follows that

$$|x + y - \Gamma_{\theta, v}^\lambda(x_n + y_n)| \lesssim R^{(1+\delta)/2}, \tag{9.13}$$

which, by Lemma 4.6, implies that

$$\measuredangle(G^\lambda(x + y; \omega_{\tilde{\theta}}), T_z Z) \lesssim \measuredangle(G^\lambda(\Gamma_{\theta, v}^\lambda(x_n + y_n); \omega_\theta), T_z Z) + \varrho^{-1/2}.$$

Finally,

$$\Gamma_{\theta, v}^\lambda(x_n + y_n) \in T_{\theta, v},$$

and this tube is $R^{-1/2+\delta_m}$-tangent to Z. Note that $z \in Z \cap B(0, 2R)$ whilst, recalling (9.13), one has

$$|\Gamma_{\theta, v}^\lambda(x_n + y_n) - z| \leqslant |x + y - \Gamma_{\theta, v}^\lambda(x_n + y_n)| + |x - (z - y + b)| + |b| \lesssim R^{1/2+\delta_m}.$$

Thus, if the constant $\overline{C}_{\text{tang}}$ in Definition 8.1 is appropriately chosen, then the tangency of $T_{\theta, v}$ implies that

$$\measuredangle(G^\lambda(\Gamma_{\theta, v}^\lambda(x_n + y_n); \omega_\theta), T_z Z) \leqslant \overline{c}_{\text{tang}} R^{-1/2+\delta_m}$$

and, provided R is sufficiently large, (9.12) (and therefore (9.11)) immediately follows.

Containment properties.

The angle condition (9.11) implies that any tube $\widetilde{T}_{\tilde{\theta}, \tilde{v}}$ which intersects

$$N_{\varrho^{1/2+\delta_m}/2}(Z - y + b) \cap B(0, \varrho)$$

is actually contained in $N_{\varrho^{1/2+\delta_m}}(Z-y+b)$. To demonstrate this containment property, continue with the setup of the previous stage, but now assume the slightly stronger conditions that $z-y+b \in B(0,\varrho)$ and $|x-(z-y+b)| \leqslant \frac{1}{2}\varrho^{1/2+\delta_m}$. Define a time-dependent vector field $X_{\tilde{\theta},\tilde{v}} \colon (-\varrho,\varrho) \times Z \cap B(0,2R) \to \mathbb{R}^m$ on $Z \cap B(0,2R)$ by

$$X_{\tilde{\theta},\tilde{v}}(t,z) := \operatorname{proj}_{T_z Z}(\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}})'(t) \quad \text{for all } (t,z) \in [-\varrho,\varrho] \times Z \cap B(0,2R).$$

This can be smoothly extended to a map on $[-\varrho,\varrho] \times U$, where $U \subseteq \mathbb{R}^n$ is a small open neighbourhood of $Z \cap B(0,2R)$. By the Picard–Lindelöf existence theorem for ODEs, there exists some smooth mapping $Z_{\tilde{\theta},\tilde{v}} \colon (-\varrho,\varrho) \to Z$ such that $Z_{\tilde{\theta},\tilde{v}}(x_n) = z$ and

$$Z'_{\tilde{\theta},\tilde{v}}(t) = X_{\tilde{\theta},\tilde{v}}(t, Z_{\tilde{\theta},\tilde{v}}(t)) \quad \text{for all } t \in (-\varrho,\varrho).$$

Here, $x = (x', x_n) \in \widetilde{T}_{\tilde{\theta},\tilde{v}}$ are the points fixed above.

Observe that

$$|\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}}(x_n) - (z-y+b)| \leqslant |\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}}(x_n) - x| + |x - (z-y+b)| < \tfrac{2}{3}\varrho^{1/2+\delta_m}.$$

Let I denote the set of all $t \in (-\varrho,\varrho)$ such that $t \geqslant x_n$ and

$$|\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}}(s) - (Z_{\tilde{\theta},\tilde{v}}(s)-y+b)| \leqslant \tfrac{9}{10}\varrho^{1/2+\delta_m} \quad \text{for all } x_n \leqslant s \leqslant t.$$

It is claimed that $t_* := \sup I = \varrho$. To see this, first note that, if $t_* < \varrho$, then

$$\tfrac{9}{10}\varrho^{1/2+\delta_m} = |\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}}(t_*) - (Z_{\tilde{\theta},\tilde{v}}(t_*)-y+b)|.$$

The angle condition (9.11) implies that

$$\angle((\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}})'(t), T_{Z_{\tilde{\theta},\tilde{v}}(t)}Z) \leqslant \bar{c}_{\mathrm{tang}}\varrho^{-1/2+\delta_m} \quad \text{for all } x_n \leqslant t \leqslant t_*.$$

Combining the previous two displays with the identity

$$\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}}(t_*) - Z_{\tilde{\theta},\tilde{v}}(t_*) = \int_{x_n}^{t_*} \operatorname{proj}_{(T_{Z_{\tilde{\theta},\tilde{v}}(t)}Z)^{\perp}}(\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}})'(t)\,dt + (\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}}(x_n) - z),$$

one concludes that

$$\tfrac{9}{10}\varrho^{1/2+\delta_m} < \int_{x_n}^{t_*} \sin\angle((\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}})'(t), T_{Z_{\tilde{\theta},\tilde{v}}(t)}Z)|(\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}})'(t)|\,dt + \tfrac{2}{3}\varrho^{1/2+\delta_m}$$

$$\leqslant 2\bar{c}_{\mathrm{tang}}\varrho^{-1/2+\delta_m}|t_* - x_n| + \tfrac{2}{3}\varrho^{1/2+\delta_m}.$$

Since $|t_* - x_n| \leqslant 2\varrho$, if $\bar{c}_{\mathrm{tang}}$ is appropriately chosen, then this yields a contradiction and, consequently, $[x_n,\varrho] \subseteq I$. A similar argument shows that $(-\varrho, x_n] \subseteq I$, and so

$$\widetilde{\Gamma}^{\lambda}_{\tilde{\theta},\tilde{v}}((-\varrho,\varrho)) \subseteq N_{(9/10)\varrho^{1/2+\delta_m}}(Z-y+b).$$

One therefore concludes that $\widetilde{T}_{\tilde{\theta},\tilde{v}} \subseteq N_{\varrho^{1/2+\delta_m}}(Z-y+b)$.

Proof of Proposition 9.2 (1)

Let $b \in B(0, 2R^{1/2+\delta_m})$ and suppose that $\widetilde{T}_{\tilde{\theta},\tilde{v}} \cap N_{\varrho^{1/2+\delta}/2}(Z - y + b) \cap B(0, \varrho) \neq \varnothing$; the problem is to show that $\widetilde{T}_{\tilde{\theta},\tilde{v}}$ is $\varrho^{-1/2+\delta_m}$-tangential to $Z - y + b$. By hypothesis, there exists some $x \in \widetilde{T}_{\tilde{\theta},\tilde{v}}$ and $z \in Z$ such that $z - y + b \in B(0, \varrho)$ and $|x - (z - y + b)| \leqslant \frac{1}{2}\varrho^{1/2+\delta_m}$. The preceding analysis therefore implies that $\widetilde{T}_{\tilde{\theta},\tilde{v}} \subseteq N_{\varrho^{1/2+\delta_m}}(Z - y + b)$, which is the desired containment condition for tangency. On the other hand, the angle condition for tangency always holds by (9.11), and so the proof of part (1) is complete.

Proof of Proposition 9.2 (2)

By Lemma 9.1, it suffices to prove that

$$\bigcup_{(\theta,v)\in\mathbb{T}_{Z,B(y,\varrho)}} \widetilde{\mathbb{T}}_{\theta,v} \subseteq \bigcup_{|b|\lesssim R^{1/2+\delta_m}} \widetilde{\mathbb{T}}_{Z-y+b}.$$

Fixing $(\theta, v) \in \mathbb{T}_{Z,B(y,\varrho)}$ and $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$, by (9.11) the problem is further reduced to showing that there exists some $|b| \lesssim R^{1/2+\delta_m}$ such that $\widetilde{T}_{\tilde{\theta},\tilde{v}} \subseteq N_{\varrho^{1/2+\delta_m}}(Z - y + b)$. Lemma 9.3 implies that $\widetilde{\Gamma}^\lambda_{\tilde{\theta},\tilde{v}}(t) \in N_{CR^{1/2+\delta_m}}(Z - y)$ for $t \in [-\varrho, \varrho]$. Consequently, fixing $t_0 \in [-\varrho, \varrho]$, there exists some $|b| \lesssim R^{1/2+\delta_m}$ such that $\widetilde{\Gamma}^\lambda_{\tilde{\theta},\tilde{v}}(t_0) \in Z - y + b$. The desired inclusion now follows from the containment property discussed earlier in the proof. $\square$

9.3. Sorting the wave packets

If $(\theta, v) \in \mathbb{T}$ and $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\theta,v}$, then Lemma 9.3 implies that[22]

$$\mathrm{dist}_H(T_{\theta,v} \cap B(y, \varrho), \widetilde{T}_{\tilde{\theta},\tilde{v}} + y) \lesssim R^{1/2+\delta}. \tag{9.14}$$

In particular, if a pair of wave packets $(\theta_1, v_1), (\theta_2, v_2) \in \mathbb{T}$ are such that

$$\widetilde{\mathbb{T}}_{\theta_1,v_1} \cap \widetilde{\mathbb{T}}_{\theta_2,v_2} \neq \varnothing,$$

then the tubes T_{θ_1,v_1} and T_{θ_2,v_2} are approximately equal on $B(y, \varrho)$.[23] This suggests sorting the scale R wave packets $(\theta, v) \in \mathbb{T}$ into disjoint sets for which the associated tubes essentially agree on $B(y, \varrho)$.

Let $\mathcal{T}$ denote the collection of all pairs $(\tilde{\theta}, w)$ formed by a $\varrho^{-1/2}$-cap $\tilde{\theta}$ and $w \in R^{(1+\delta)/2}\mathbb{Z}^{n-1}$. For each $(\tilde{\theta}, w) \in \mathcal{T}$, choose some

$$\mathcal{T}_{\tilde{\theta},w} \subseteq \{(\theta, v) \in \mathbb{T} : \mathrm{dist}(\theta, \tilde{\theta}) \lesssim \varrho^{-1/2} \text{ and } |v - \bar{v}(y; \omega_\theta) - w| \lesssim R^{(1+\delta)/2}\}$$

---

[22] Here, dist_H denotes the Hausdorff distance.
[23] More precisely, enlarging the radius of either one of the T_{θ_j,v_j} by a constant factor produces a tube which contains $(T_{\theta_1,v_1} \cup T_{\theta_2,v_2}) \cap B(y, \varrho)$.
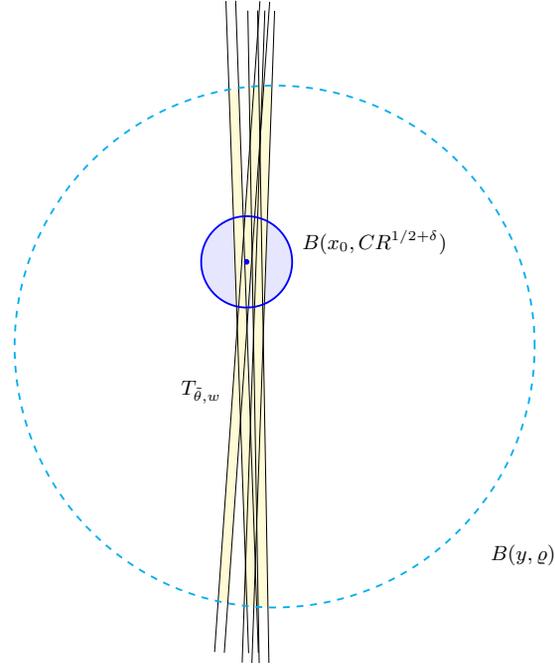
Figure 4. The set $T_{\tilde{\theta},w} := \bigcup_{(\theta,v)\in\mathcal{T}_{\tilde{\theta},w}} T_{\theta,v} \cap B(y,\varrho)$ is highlighted in yellow. Fixing $x_0 \in T_{\tilde{\theta},w}$, for every $(\theta,v)\in\mathcal{T}_{\tilde{\theta},w}$ the tube $T_{\theta,v}$ intersects the ball $B(x_0, CR^{1/2+\delta})$.

so that the family $\{\mathcal{T}_{\tilde{\theta},w} : (\tilde{\theta},w)\in\mathcal{T}\}$ forms a covering of $\mathbb{T}$ by disjoint sets. Defining

$$T_{\tilde{\theta},w} := \bigcup_{(\theta,v)\in\mathcal{T}_{\tilde{\theta},w}} T_{\theta,v} \cap B(y,\varrho),$$

one obtains the following corollary to Lemma 9.3.

COROLLARY 9.4. *If* $(\tilde{\theta},w)\in\mathcal{T}$ *and* $(\theta,v)\in\mathcal{T}_{\tilde{\theta},w}$, *then*

$$\mathrm{dist}_H(T_{\theta,v}\cap B(y,\varrho), T_{\tilde{\theta},w}) \lesssim R^{1/2+\delta}.$$

Let $g\colon B^{n-1}\to\mathbb{C}$ be integrable and define

$$g_{\tilde{\theta},w} := \sum_{(\theta,v)\in\mathcal{T}_{\tilde{\theta},w}} g_{\theta,v} \quad \text{for all } (\tilde{\theta},w)\in\mathcal{T}.$$

Since the $\mathcal{T}_{\tilde{\theta},w}$ cover $\mathbb{T}$ and are disjoint, it follows that

$$g = \sum_{(\tilde{\theta},w)\in\mathcal{T}} g_{\tilde{\theta},w} + \mathrm{RapDec}(R)\|g\|_{L^2(B^{n-1})}; \tag{9.15}$$

furthermore, the functions $g_{\tilde{\theta},w}$ are almost orthogonal and, consequently,

$$\|g\|_{L^2(B^{n-1})}^2 \sim \sum_{(\tilde{\theta},w)\in\mathcal{T}} \|g_{\tilde{\theta},w}\|_{L^2(B^{n-1})}^2. \tag{9.16}$$

By Lemma 9.1, the function $(g_{\tilde{\theta},w})^{\sim}$ is concentrated on scale ϱ wave packets belonging to $\bigcup_{(\theta,v)\in\mathcal{T}_{\tilde{\theta},w}} \widetilde{\mathbb{T}}_{\theta,v}$. This union is contained in

$$\widetilde{\mathcal{T}}_{\tilde{\theta},w} := \{(\tilde{\theta}',\tilde{v})\in\widetilde{\mathbb{T}} : \mathrm{dist}(\tilde{\theta}',\tilde{\theta})\lesssim\varrho^{-1/2} \text{ and } |\tilde{v}-w|\lesssim R^{(1+\delta)/2}\},$$

and therefore each $(g_{\tilde{\theta},w})^{\sim}$ is concentrated on wave packets from $\widetilde{\mathcal{T}}_{\tilde{\theta},w}$. The family

$$\{\widetilde{\mathcal{T}}_{\tilde{\theta},w} : (\tilde{\theta},w)\in\mathcal{T}\}$$

forms a covering of $\widetilde{\mathbb{T}}$ by almost disjoint sets. This implies almost orthogonality between the scale ϱ wave packets of the different functions $(g_{\tilde{\theta},w})^{\sim}$. A particular consequence of this observation is that

$$\left\|\sum_{(\tilde{\theta},w)\in\mathcal{T}} (g_{\tilde{\theta},w})^{\sim}_{\tilde{b}}\right\|_{L^2(B^{n-1})}^2 \sim \sum_{(\tilde{\theta},w)\in\mathcal{T}} \|(g_{\tilde{\theta},w})^{\sim}_{\tilde{b}}\|_{L^2(B^{n-1})}^2, \tag{9.17}$$

where $\tilde{h}_b$ is defined for a given function h as in (9.9).

9.4. Transverse equidistribution revisited

Let Z be an m-dimensional transverse complete intersection, $(\tilde{\theta},w)\in\mathcal{T}$ and h be a function concentrated on wave packets in $\mathbb{T}_{Z\cap B(y,\varrho)}\cap\mathcal{T}_{\tilde{\theta},w}$. Here, the key example to have in mind is $h=g_{\tilde{\theta},w}$, for some function g concentrated on wave packets in $\mathbb{T}_{Z,B(y,\varrho)}$.

Every scale R wave packet of h intersects $B(y,\varrho)$ on the set $T_{\tilde{\theta},w}$ which, by Corollary 9.4, is comparable to $T_{\theta,v}\cap B(y,\varrho)$ for any $(\theta,v)\in\mathcal{T}_{\tilde{\theta},w}$.[24] Consequently, if $x_0\in T_{\tilde{\theta},w}$, then all the scale R wave packets of h intersect $B(x_0,CR^{1/2+\delta_m})$ (see Figure 4). Moreover, (9.14) implies that the scale ϱ wave packets of $\tilde{h}$ will intersect $B(x_0-y,CR^{1/2+\delta_m})$. Under these conditions, a useful reverse-type version of Hörmander's L^2 bound holds.

LEMMA 9.5. *Let T^λ be a Hörmander-type operator with phase ϕ^λ given by a translate of a reduced phase in the sense of* (9.2) *and $1\leqslant R^{1/2+\delta}\leqslant r\lesssim\lambda^{1/2}$. There exists a family of Hörmander-type operators $\mathbf{T}^\lambda$ all with phase ϕ^λ such that the following hold:*

(i) each $T^\lambda\in\mathbf{T}^\lambda$ is again an operator with phase given by a translate of a reduced phase in the sense of (9.2) *(in particular, all the relevant bounds from §4 hold on the support of the amplitude)*;

---

[24] Here, 'the scale R wave packets of h' refers to the scale R wave packets upon which h is concentrated.

(ii) $\#\mathbf{T}^\lambda = O(1)$;

(iii) *if f is concentrated on wave packets (with respect to T^λ) which intersect some* $B(\bar{x}, r) \subseteq B(0, R)$, *then*

$$\|f\|_{L^2(B^{n-1})}^2 \lesssim r^{-1} \|T_*^\lambda f\|_{L^2(B(\bar{x}, Cr))}^2$$

holds for some $T_*^\lambda \in \mathbf{T}^\lambda$.

Lemma 9.5 can be proven for extension operators fairly directly via Plancherel's theorem (see [14, §3]). Establishing the general (variable coefficient) version of Lemma 9.5 involves a number of additional technicalities, and the proof is therefore postponed until the end of the section.

For h as above, $x_0 \in T_{\tilde{\theta}, w}$ and $|b| \lesssim R^{1/2 + \delta_m}$, the preceding discussion implies that the function $\tilde{h}_b$, as defined in (9.9), is a sum of wave packets which intersect

$$B(x_0 - y, C R^{1/2 + \delta_m}).$$

Lemma 9.5 can therefore be applied at scale ϱ with $r \sim R^{1/2 + \delta_m}$ to yield

$$\|\tilde{h}_b\|_{L^2(B^{n-1})}^2 \lesssim R^{-1/2 - \delta_m} \|\widetilde{T}_*^\lambda \tilde{h}_b\|_{L^2(B(x_0 - y, C R^{1/2 + \delta_m}))}^2.$$

The wave packets defined by T^λ and T_*^λ will have identical geometric properties (since these properties are essentially independent of the precise choice of amplitude). By (9.10), one concludes that

$$\|\tilde{h}_b\|_{L^2(B^{n-1})}^2 \lesssim R^{-1/2 - \delta_m} \|T_*^\lambda h_b\|_{L^2(N_{\varrho^{1/2 + \delta_m}}(Z + b) \cap B(x_0, C R^{1/2 + \delta_m}))}^2. \tag{9.18}$$

This observation has several useful consequences. First of all, by applying Hörmander's L^2 bound, one obtains the following.

LEMMA 9.6. *Let h be concentrated on wave packets from $\mathbb{T}_{Z \cap B(y, \varrho)} \cap \mathcal{T}_{\tilde{\theta}, w}$, for some* $(\tilde{\theta}, w) \in \mathcal{T}$. *Let $\mathfrak{B} \subset B(0, C R^{1/2 + \delta_m})$ be such that the sets*

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

are pairwise disjoint over all $b \in \mathfrak{B}$. Then,

$$\sum_{b \in \mathfrak{B}} \|\tilde{h}_b\|_{L^2(B^{n-1})}^2 \lesssim \|h\|_{L^2(B^{n-1})}^2.$$

A further consequence of (9.18) is the following transverse equidistribution result.

LEMMA 9.7. *Let* $(\tilde{\theta}, w) \in \mathcal{T}$, $|b| \lesssim R^{1/2+\delta_m}$ *and* Z *be an* m-*dimensional transverse complete intersection with* $\overline{\deg}\, Z \lesssim_{\varepsilon} 1$. *If* h *is concentrated on wave packets from*

$$\mathbb{T}_{Z \cap B(y,\varrho)} \cap \mathcal{T}_{\tilde{\theta},w},$$

then

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

Proof. The transverse equidistribution estimate in Lemma 8.4 implies that

$$\|T_*^\lambda \tilde{h}_b\|_{L^2(N_{\varrho^{1/2+\delta_m}}(Z+b) \cap B(x_0, CR^{1/2+\delta_m}))}^2 \lesssim R^{1/2+O(\delta_m)} \Big(\frac{\varrho}{R}\Big)^{(n-m)/2} \|h\|_{L^2(B^{n-1})}^2.$$

Combining this with (9.18) completes the proof. $\qquad\square$

Let g be concentrated on wave packets of $\mathbb{T}_{Z,B(y,\varrho)}$. For each $(\tilde{\theta}, w) \in \mathcal{T}$ the function $g_{\tilde{\theta},w}$ is concentrated on wave packets in $\mathbb{T}_{Z \cap B(y,\varrho)} \cap \mathcal{T}_{\tilde{\theta},w}$. It follows that Lemma 9.6 and Lemma 9.7 hold for $h = g_{\tilde{\theta},w}$. Combining the contributions from distinct $\mathcal{T}_{\tilde{\theta},w}$, one obtains the following.

LEMMA 9.8. *Let* $|b| \lesssim R^{1/2+\delta_m}$ *and* Z *be an* m-*dimensional transverse complete intersection with* $\overline{\deg}\, Z \lesssim_{\varepsilon} 1$. *If* g *is concentrated on wave packets from* $\mathbb{T}_{Z,B(y,\varrho)}$, *then*

$$\|\tilde{g}_b\|_{L^2(B^{n-1})} \lesssim R^{O(\delta_m)} \Big(\frac{\varrho}{R}\Big)^{(n-m)/4} \|g\|_{L^2(B^{n-1})}.$$

Proof. By (9.15) and the linearity of the map $h \mapsto \tilde{h}_b$, it follows that

$$\tilde{g}_b = \sum_{(\tilde{\theta},w) \in \mathcal{T}} (g_{\tilde{\theta},w})_{\tilde{b}} + \mathrm{RapDec}(R)\|g\|_{L^2(B^{n-1})}.$$

The almost orthogonality relation (9.17) between the $(g_{\tilde{\theta},w})_{\tilde{b}}$ implies that

$$\|\tilde{g}_b\|_{L^2(B^{n-1})}^2 \lesssim \sum_{(\tilde{\theta},w) \in \mathcal{T}} \|(g_{\tilde{\theta},w})_{\tilde{b}}\|_{L^2(B^{n-1})}^2 + \mathrm{RapDec}(R)\|g\|_{L^2(B^{n-1})}^2.$$

By Lemma 9.7, the right-hand side of the above display is in turn dominated by

$$R^{O(\delta_m)} \Big(\frac{\varrho}{R}\Big)^{(n-m)/2} \sum_{(\tilde{\theta},w) \in \mathcal{T}} \|g_{\tilde{\theta},w}\|_{L^2(B^{n-1})}^2 + \mathrm{RapDec}(R)\|g\|_{L^2(B^{n-1})}^2.$$

An application of (9.16) yields the desired estimate. $\qquad\square$

9.5. The proof of the reverse L^2 bound

Proof of Lemma 9.5. One may assume without loss of generality that $\bar{x}=0$ and $r=R^{1/2+\delta}$. Indeed, the first reduction follows from the formula (9.1), which can be used to replace T^λ and f, respectively, by $\widetilde{T}^\lambda$ and $\tilde{f}$, here defined with $y:=\bar{x}$. Lemma 9.1 and the identity (9.8) imply that $\tilde{f}$ is concentrated on scale R wave packets associated with $\widetilde{T}^\lambda$ which intersect $B(0,r)$. For the second reduction, suppose the result is known for $r=R^{1/2+\delta}$, and let $R^{1/2+\delta}\leqslant r\leqslant R$ and f be as in the statement of the theorem. For $a\in\mathbb{R}$ consider the slab

$$S_a := \mathbb{R}^{n-1}\times[a-R^{1/2+\delta},a+R^{1/2+\delta}]\cap B(0,Cr),$$

where $C\geqslant 2$ is a constant, chosen large enough for the purposes of the argument. Cover S_a with a collection $\{B_j\}_{j\in\mathcal{J}}$ of finitely-overlapping $R^{1/2+\delta}$-balls satisfying $B_j\cap S_a\neq\varnothing$ for all $j\in\mathcal{J}$. By the initial reductions, any tube $T_{\theta,v}$ makes a small angle with the e_n direction, and thus intersects at most $O(1)$ of these balls. Orthogonality of the wave packets together with the hypothesised estimate therefore imply that

$$\|g\|_{L^2(B^{n-1})}^2 \lesssim R^{-1/2-\delta}\|T^\lambda g\|_{L^2(N_{CR^{1/2+\delta}}(S_a))}^2 \tag{9.19}$$

for any g concentrated on wave packets at scale R which intersect S_a. If $(\theta,v)\in\mathbb{T}_{B(0,r)}$, then the aforementioned angle condition implies that $T_{\theta,v}$ intersects every slab S_a for which $S_a\cap B(0,r)\neq\varnothing$. Hence, if f is concentrated on wave packets from $\mathbb{T}_{B(0,r)}$, then one may sum (9.19) over a collection of $\sim r/R^{1/2+\delta}$ slabs which cover $B(0,r)$ to obtain the desired result.

Fix a function f satisfying the hypotheses of the lemma with $\bar{x}=0$ and $r=R^{1/2+\delta}$, and note that

$$\|f\|_{L^2(B^{n-1})} \sim \left\|\sum_{(\theta,v):T_{\theta,v}\cap B(0,R^{1/2+\delta})\neq\varnothing} f_{\theta,v}\right\|_{L^2(B^{n-1})}.$$

Recall that the change of variables $u\mapsto\Psi(\bar{x}/\lambda;u)$, where Ψ is the function introduced in §4, reparametrises the surface $\{\partial_x\phi^\lambda(\bar{x};\omega):\omega\in\Omega\}$ as the graph of the function $\partial_{x_n}\phi^\lambda(\bar{x};\Psi(\bar{x}/\lambda;u))$, for any $\bar{x}\in X^\lambda$. With an abuse of notation for the sake of simplicity, let Ψ denote the above change of variables for $\bar{x}=0$; that is, $\Psi(u):=\Psi(0;u)$. Thus, $\Psi:U\to\Omega$ is a diffeomorphism that reparametrises the surface $\{\partial_x\phi^\lambda(0;\omega):\omega\in\Omega\}$ as the graph of the function $\bar{h}(u):=\partial_{x_n}\phi^\lambda(0;\Psi(u))$; in particular, $\partial_{x'}\phi^\lambda(0;\Psi(u))=u$ for all $u\in U$.

Applying the change of variables $u \mapsto \Psi(u)$, denoting by J_Ψ the absolute value of the determinant of the corresponding Jacobian matrix and recalling that $J_\Psi \sim 1$ by (4.12), one obtains

$$\|g\|_{L^2(B^{n-1})} \sim \|g \circ \Psi \cdot J_\Psi^{1/2}\|_{L^2(B^{n-1})} \sim \|g \circ \Psi \cdot J_\Psi\|_{L^2(B^{n-1})}$$

for all $g \in L^2(B^{n-1})$. It follows that

$$\|f\|_{L^2(B^{n-1})} \sim \|f_\Psi\|_{L^2(B^{n-1})},$$

where

$$f_\Psi := \left(\sum_{(\theta,v): T_{\theta,v} \cap B(0,R^{1/2+\delta}) \neq \varnothing} f_{\theta,v} \circ \Psi \right) \cdot J_\Psi.$$

Let E denote the extension operator

$$Eg(x) := \int_U e^{2\pi i(\langle x', u \rangle + x_n \bar{h}(u))} g(u) \, du$$

associated with the graph $u \mapsto (u, \bar{h}(u))$. For any $x_n \in \mathbb{R}$ and a square integrable function g supported on U, Plancherel's theorem implies that

$$\|g\|_{L^2(B^{n-1})}^2 = \|e^{2\pi i x_n \bar{h}} g\|_{L^2(B^{n-1})}^2 = \|Eg\|_{L^2(\mathbb{R}^{n-1} \times \{x_n\})}^2.$$

Hence, averaging this estimate over $|x_n| < R^{1/2+\delta}$, one obtains

$$\|g\|_{L^2(B^{n-1})}^2 \sim R^{-1/2-\delta} \|Eg\|_{L^2(\mathbb{R}^{n-1} \times (-R^{1/2+\delta}, R^{1/2+\delta}))}^2. \tag{9.20}$$

The key observation is that the hypothesis on f implies that the right-hand L^2-norm can be localised.

CLAIM. *If $|x_n| < R^{1/2+\delta}$ and $x' \notin B(0, CR^{1/2+\delta})$, then*

$$Ef_\Psi(x) = (1 + R^{-1/2}|x'|)^{-(n+1)} \operatorname{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

This concentration property is very similar to that detailed in Lemma 5.4, the main difference being that the condition $(\theta, v) \in \mathbb{T}_{B(0,R^{1/2+\delta})}$ is defined with respect to the operator T^λ, whilst the above identity concerns the linearised version E. The proof is a minor adaptation of the stationary phase analysis used to establish Lemma 5.4, and is therefore omitted.

For the specific choice of function $g = f_\Psi$, the claim implies that (9.20) may be strengthened to

$$\|f\|_{L^2(B^{n-1})}^2 \sim \|f_\Psi\|_{L^2(B^{n-1})}^2 \sim R^{-1/2-\delta} \|Ef_\Psi\|_{L^2(B(0,CR^{1/2+\delta}))}^2. \tag{9.21}$$

This is easily seen to imply the lemma. Indeed, reversing the earlier change of variables,

$$Ef_\Psi(z) = \int_\Omega e^{2\pi i \phi^\lambda(z;\omega)} e^{-2\pi i \lambda \mathcal{E}(z/\lambda;\omega)} f(\omega)\, d\omega + \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}$$

for all $z \in B(0, CR^{1/2+\delta})$, where $\mathcal{E}$ is the error term in the Taylor expansion

$$\phi(z;\omega) = \phi(0;\omega) + \langle \partial_x \phi(0;\omega), z\rangle + \mathcal{E}(z;\omega).$$

Were it not for the factor $e^{-2\pi i \lambda \mathcal{E}(z/\lambda;\omega)}$, the functions $Ef_\Psi(z)$ and $T^\lambda f(z)$ would be equal, up to a negligible error term, and (9.21) would directly imply the desired estimate. This unwanted additional factor can be removed via a Fourier series decomposition.

More precisely, it holds by the integral form of the remainder that

$$\partial_\omega^\beta \mathcal{E}(z;\omega) = \sum_{|\gamma|=2} \frac{2}{\alpha!} \int_0^1 (1-t) \partial_x^\gamma \partial_\omega^\beta \phi(tz;\omega)\, dt\, z^\gamma \quad \text{for all } \beta \in \mathbb{N}_0^{n-1}. \tag{9.22}$$

Applying the uniform bounds on the derivatives of the reduced phase function ϕ as described in §4, and recalling the hypothesis $R^{1/2+\delta} \lesssim \lambda^{1/2}$, (9.22) implies that

$$\left| \partial_\omega^\beta \partial_z^\alpha \lambda \mathcal{E}\left(\frac{R^{1/2+\delta} z}{\lambda}; \omega \right) \right| \lesssim c_{\mathrm{par}} \quad \text{for } 0 \leqslant |\alpha|, |\beta| \leqslant N_{\mathrm{par}} \text{ and } |z| \lesssim 1. \tag{9.23}$$

Let $\psi \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})$ be supported on $X \times \Omega$ and equal to 1 on $\mathrm{supp}\, a$. By forming the Fourier series expansion in both the x and ω variables, one obtains

$$e^{-2\pi i \lambda \mathcal{E}(z/\lambda;\omega)} \psi\left(\frac{z}{R^{1/2+\delta}}; \omega \right) = \sum_{\substack{k \in \mathbb{Z}^n \times \mathbb{Z}^{n-1} \\ k=(k_1,k_2)}} (1+|k|)^{-2n} c_k e^{2\pi i(\langle z/R^{1/2+\delta}, k_1\rangle + \langle \omega, k_2\rangle)},$$

where the c_k are weighted Fourier coefficients. Observe that (9.23) implies that $|c_k| \lesssim 1$ for all $k \in \mathbb{Z}^n \times \mathbb{Z}^{n-1}$. Thus, (9.21) now yields

$$\|f\|_{L^2(B^{n-1})} \lesssim R^{-(1/4+\delta/2)} \sum_{\substack{k \in \mathbb{Z}^n \times \mathbb{Z}^{n-1} \\ k=(k_1,k_2)}} (1+|k|)^{-2n} \|T^\lambda(e^{2\pi i \langle \cdot, k_2\rangle} f)\|_{L^2(B(0,CR^{1/2+\delta}))}.$$

The above sum is split into a sum over $k=(k_1, k_2)$ satisfying $|k| > C_n$ and a sum over the remaining k, where C_n is a dimensional constant, chosen large enough for the present purpose. The sum over large k is bounded above by

$$A \sum_{|k|>C_n} (1+|k|)^{-2n} \|f\|_{L^2(B^{n-1})} \tag{9.24}$$

for some dimensional constant in A; indeed, this follows by applying Hörmander's L^2-estimate from Lemma 5.5 to each of the summands (the constant in Hörmander's theorem can be made uniform over the class of reduced phases). By choosing C_n to be sufficiently large so that $A \sum_{|k|>C_n} (1+|k|)^{-2n} \leqslant \frac{1}{2}$, the term (9.24) can be absorbed into the left-hand side of the inequality. Thus, one obtains

$$\|f\|_{L^2(B^{n-1})} \lesssim R^{-(1/4+\delta/2)} \sum_{\substack{k \in \mathbb{Z}^{n-1} \\ |k| \leqslant C_n}} \|T^\lambda(e^{2\pi i \langle \cdot, k \rangle} f)\|_{L^2(B(0,CR^{1/2+\delta}))}.$$

Finally, define the class of operators

$$\overline{\mathbf{T}}^\lambda := \{T_k^\lambda : k \in \mathbb{Z}^{n-1} \text{ and } |k| \leqslant C_n\},$$

where T_k^λ has phase ϕ^λ and amplitude a_k^λ, for

$$a_k(z;\omega) := \psi(z;\omega) e^{2\pi i \langle \omega, k \rangle}.$$

It is easy to see that each such amplitude a_k can be written as a linear combination of $O(1)$ amplitudes satisfying the conditions of Lemma 4.3, with complex coefficients of order of magnitude $O(1)$. Defining $\mathbf{T}^\lambda$ to be the union, over all $|k| \leqslant C_n$, of the operators with phase ϕ^λ and the corresponding rescaled amplitudes, it follows by the pigeonhole principle that there exists at least one operator $T_*^\lambda \in \mathbf{T}^\lambda$ for which the desired inequality holds. $\qquad \square$

10. Proof of the k-broad estimate

10.1. A more general result

In this section the proof of the k-broad estimate in Theorem 1.9 is given. In order to facilitate an inductive argument, a more general result will be established, which is described presently.

Throughout this section, T^λ denotes an arbitrary choice of a translate of a Hörmander-type operator with reduced positive-definite phase. That is, T^λ is of the form of the operator $\widetilde{T}^\lambda$ discussed in the previous section, with phase and amplitude of the type described in (9.2). Many of the estimates involving T^λ, such as (10.1) below, are understood to hold uniformly for the entire class of such operators; it is important to work with the whole class rather than a single choice of T^λ in order to run certain induction arguments.

In order to state the result, first define the exponent

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

and note that $-e_{k,n}(p) + \frac{1}{2} \leqslant 0$ if and only if $p \geqslant \bar{p}(k,n)$, where

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

is the exponent appearing in Theorem 1.9.

PROPOSITION 10.1. *Given $\varepsilon > 0$ sufficiently small and $1 \leqslant m \leqslant n$, there exist*

$$0 < \delta \ll \delta_{n-1} \ll \delta_{n-2} \ll ... \ll \delta_1 \ll \varepsilon$$

and constants $\overline{C}_\varepsilon$ and $\bar{A}_\varepsilon$ dyadic, $D_{m,\varepsilon} \lesssim_\varepsilon 1$ and $\vartheta_m < \varepsilon$ such that the following holds.

Suppose $Z = Z(P_1, ..., P_{n-m})$ is a transverse complete intersection with $\overline{\deg}\, Z \leqslant D_{m,\varepsilon}$. For all $2 \leqslant k \leqslant n$, $1 \leqslant A \leqslant \bar{A}_\varepsilon$ dyadic and $1 \leqslant K \leqslant R \leqslant \lambda$, the inequality

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim_\varepsilon K^{\overline{C}_\varepsilon} R^{\vartheta_m + \delta(\log \bar{A}_\varepsilon - \log A) - e_{k,n}(p) + 1/2} \|f\|_{L^2(B^{n-1})} \tag{10.1}$$

holds whenever f is concentrated on wave packets from $\mathbb{T}_Z$ and

$$2 \leqslant p \leqslant \bar{p}_0(k,m) := \begin{cases} \bar{p}(k,m), & \text{if } k < m, \\ \bar{p}(m,m) + \delta, & \text{if } k = m. \end{cases}$$

Here, $\mathbb{T}_Z$ is defined as in §8; that is,

$$\mathbb{T}_Z := \{(\theta, v) \in \mathbb{T} : T_{\theta,v} \text{ is } R^{-1/2 + \delta_m}\text{-tangent to } Z \text{ in } B(0,R)\}.$$

Proposition 10.1 immediately yields the desired k-broad estimate.

Proof of Theorem 1.9. Theorem 1.9 is a special case of Proposition 10.1. Indeed, $Z = \mathbb{R}^n$ is a transverse complete intersection of dimension n, and $\mathbb{T}_Z$ contains all wave packets in $B(0,R)$. Thus, taking $A = \bar{A}_\varepsilon$ and $p = \bar{p}(k,n)$ yields the endpoint case of Theorem 1.9. The general result follows by interpolating with the trivial $p = \infty$ estimate via the logarithmic convexity of the k-broad norms (that is, Lemma 6.3). $\qquad\square$

10.2. Reducing to $R \lesssim \lambda^{1-\varepsilon}$.

Turning to the proof of Proposition 10.1, the first step is a technical reduction on the radii R which is needed to facilitate certain polynomial partitioning arguments. In particular,

it will be necessary to approximate the curves $\Gamma^\lambda_{\theta,v}$ by polynomial curves of degree independent of R; by the observations of §7.2, this is possible when $1 \leqslant R \lesssim_\varepsilon \lambda^{1-\varepsilon}$, and thus (10.1) will first be proved for this restricted range of R. The result can then be extended to the full $R \leqslant \lambda$ range by a triangle inequality argument (incurring a permissible $R^{O(\varepsilon)}$ loss). The concentration hypothesis on f in Proposition 10.1 creates some difficulties here, which are addressed by the following lemma.

LEMMA 10.2. *Let $1 \leqslant \varrho \leqslant R \leqslant \lambda$ and assume that, for any transverse complete intersection $Z = Z(P_1, ..., P_{n-m})$ with $\overline{\deg} Z \leqslant D$, the inequality*

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,\varrho))} \leqslant E \|f\|_{L^2(B^{n-1})} \tag{10.2}$$

holds with some constant $E > 0$ whenever f is concentrated on wave packets from $\widetilde{\mathbb{T}}_Z$. Then for all Z as above, the inequality

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim \left(\frac{R}{\varrho}\right)^{O(1)} E \|f\|_{L^2(B^{n-1})}$$

holds for all functions f concentrated on wave packets in $\mathbb{T}_Z$.

Here, $\widetilde{\mathbb{T}}_Z$ denotes the collection of wave packets at scale ϱ that are $\varrho^{1/2+\delta_m}$-transverse to Z in $B(0, \varrho)$; this notation is consistent with that used in §9.

Proof. Let f be a function concentrated on wave packets in $\mathbb{T}_Z$ for some transverse complete intersection Z, as in the statement of the lemma. Fix a cover $\mathcal{B}_\varrho$ of $B(0, R)$ by finitely overlapping ϱ-balls. By the sub-additivity of the k-broad norms and Lemma 5.4, there exists some $B = B(y, \varrho) \in \mathcal{B}_\varrho$ such that

$$\|T^\lambda f\|^p_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim \left(\frac{R}{\varrho}\right)^{O(1)} \|T^\lambda h\|^p_{\mathrm{BL}^p_{k,A}(B)} + \mathrm{RapDec}(R)\|f\|^p_{L^2(B^{n-1})},$$

where $h := \sum_{(\theta,v) \in \mathbb{T}_{Z,B}} f_{\theta,v}$; here and below, the notation $\mathbb{T}_{Z,B}$ is consistent with that used in §8. As in §9, write $T^\lambda h(x+y) = \widetilde{T}^\lambda \tilde{h}(x)$ so that, suppressing the harmless rapidly decaying term in the notation, one has[25]

$$\|T^\lambda f\|^p_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim \left(\frac{R}{\varrho}\right)^{O(1)} \|\widetilde{T}^\lambda \tilde{h}\|_{\mathrm{BL}^p_{k,A}(B(0,\varrho))}. \tag{10.3}$$

In general, $\tilde{h}$ is not concentrated on wave packets which are $\varrho^{-1/2+\delta_m}$-tangential to a suitable variety inside $B(0, \varrho)$; thus, hypothesis (10.2) cannot be applied directly to estimate the right-hand side of (10.3). Rather, one approaches the problem via the methods

---

[25] Strictly speaking, in order for (10.3) to hold, the k-broad norm on the right-hand side should be defined with respect to a translate of the family of balls $\mathcal{B}_{K^2}$. Since the estimates will be uniform over all choices of families $\mathcal{B}_{K^2}$ of bounded multiplicity, this slight technicality does not affect the argument.

of §9. By Proposition 9.2, $\tilde{h}$ is concentrated on wave packets in $\bigcup_{b\in\mathfrak{B}}\widetilde{\mathbb{T}}_{Z-y+b}$, where $\mathfrak{B}$ is a discrete set of cardinality $\lesssim(R/\varrho)^{O(1)}$ such that $|b|\lesssim R^{1/2+\delta_m}$ for all $b\in\mathfrak{B}$. Consequently, by the sub-additivity of the k-broad norms and Lemma 5.4 and Proposition 9.2, there exists some $b\in\mathfrak{B}$ such that

$$\|\widetilde{T}^\lambda\tilde{h}\|^p_{\mathrm{BL}^p_{k,A}(B(0,\varrho))}\lesssim\left(\frac{R}{\varrho}\right)^{O(1)}\|\widetilde{T}^\lambda\tilde{h}_b\|^p_{\mathrm{BL}^p_{k,A}(B(0,\varrho))}, \qquad (10.4)$$

for $\tilde{h}_b$ as defined in §9. Recall from Proposition 9.2 that $\tilde{h}_b$ is concentrated on wave packets in $\widetilde{\mathbb{T}}_{Z-y+b}$ and satisfies

$$\|\tilde{h}_b\|_{L^2(B^{n-1})}\lesssim\|f\|_{L^2(B^{n-1})}. \qquad (10.5)$$

Combining (10.3) and (10.4), the desired estimate now follows by applying hypothesis (10.2) to the function $\tilde{h}_b$, and then using (10.5) to bound the resulting right-hand expression. $\qquad\square$

10.3. Setting up the induction argument

Henceforth it is assumed that $1\leqslant R\lesssim_\varepsilon\lambda^{1-\varepsilon}$. Under this hypothesis, given $\varepsilon>0$ sufficiently small, Proposition 10.1 will be established for the following choice of parameters:

$$D_{m,\varepsilon}:=\varepsilon^{-\delta^{-(2n-m)}}, \quad \vartheta(\varepsilon):=\varepsilon-c_n\delta_m, \quad \bar{A}_\varepsilon:=\lceil e^{10n/\delta}\rceil, \qquad (10.6)$$
$$\delta_i=\delta_i(\varepsilon):=\varepsilon^{2i+1} \quad\text{for all } i=1,...,n-1.$$

Here, $0<\delta=\delta(\varepsilon)\ll\delta_{n-1}(\varepsilon)$ and $c_n>0$ is a fixed dimensional constant.

The proof is by induction on the radius R and the dimension m; presently the base cases for this induction are established.

Base case for the radius: $R\lesssim_\varepsilon K^n$

Provided that the implied constant in (10.1) and C_ε are chosen to be sufficiently large, in this case Proposition 10.1 follows immediately from the trivial inequality

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))}\lesssim R^{n/p}\|f\|_{L^2(B^{n-1})}, \qquad (10.7)$$

valid for all $A\in\mathbb{N}$ and $1\leqslant p\leqslant\infty$.

Base case for the dimension: $m \leqslant k-1$

Assuming (without loss of generality) that $K \lesssim_\varepsilon R^{1/n}$, one can show in this case that

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))} = \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}. \tag{10.8}$$

Indeed, fix a ball $B_{K^2} \in \mathcal{B}_{K^2}$ with $B_{K^2} \cap B(0,R) \neq \varnothing$; here, $\mathcal{B}_{K^2}$ denotes the collection of K^2-balls featured in the definition of the k-broad norm (1.13). Let $\mathfrak{T}_{B_{K^2}}$ denote the collection of all K^{-1}-caps τ for which there exists some $(\theta, v) \in \mathbb{T}_{Z,B_{K^2}}$ with $\theta \cap \tau \neq \varnothing$. Observe that, if $\tau \notin \mathfrak{T}_{B_{K^2}}$, then

$$\int_{B_{K^2}} |T^\lambda f_\tau|^p = \mathrm{RapDec}(R)\|f\|^p_{L^2(B^{n-1})}, \tag{10.9}$$

since f is concentrated on wave packets in $\mathbb{T}_Z$.

It is claimed that there exists some $V \in \mathrm{Gr}(k-1, n)$ such that

$$\angle(G^\lambda(\bar{x}; \tau), V) \leqslant K^{-1} \quad \text{for all } \tau \in \mathfrak{T}_{B_{K^2}}, \tag{10.10}$$

where $\bar{x} \in \mathbb{R}^n$ denotes the centre of B_{K^2}. Indeed, by (10.9), one may assume without loss of generality that $\mathfrak{T}_{B_{K^2}} \neq \varnothing$, and hence $\mathbb{T}_{Z,B_{K^2}} \neq \varnothing$. Thus, there exists $z \in Z \cap B(0, R)$ with $|z - \bar{x}| \lesssim R^{1/2 + \delta_m}$ and, taking $V \in \mathrm{Gr}(k-1, n)$ to be any subspace that contains $T_z Z$, the claim is easily deduced from the definition of $R^{1/2 + \delta_m}$-tangency (see Definition 8.1), together with the hypothesis $K \lesssim R^{1/n}$.

Recalling the definition of $\mu_{T^\lambda f}(B_{K^2})$ from (1.12), it follows from (10.10) that

$$\mu_{T^\lambda f}(B_{K^2}) \leqslant \max_{\tau \notin V} \int_{B_{K^2}} |T^\lambda f_\tau|^p \leqslant \max_{\tau \notin \mathfrak{T}_{B_{K^2}}} \int_{B_{K^2}} |T^\lambda f_\tau|^p,$$

and the desired estimate (10.8) is now a consequence of (10.9).

Reduction to $A \geqslant 2$

Recall that $\bar{A}_\varepsilon \geqslant e^{10n/\delta}$, so that $\delta \log \bar{A}_\varepsilon \geqslant 10n$. Although the argument does not require one to induct on A, it is useful to note that (10.7) implies that Proposition 10.1 holds for $A = 1$. This allows one to assume $A \geqslant 2$ throughout the following argument, and therefore permits the use of the k-broad triangle and logarithmic-convexity inequalities.

10.4. An overview of the inductive step

Let $2 \leqslant k \leqslant n-1$, $k \leqslant m \leqslant n$ and $R \gtrsim_\varepsilon K^n$. Assume, by way of induction hypothesis, that (10.1) holds whenever the dimension of the transverse complete intersection Z is at most $m-1$ or the radial parameter is at most $\frac{1}{2}R$.

Fix $\varepsilon > 0$, $1 < A \leqslant \bar{A}_\varepsilon$ and a transverse complete intersection $Z = Z(P_1, ..., P_{n-m})$ with $\overline{\deg}\, Z \leqslant D_{m,\varepsilon}$, where the parameters $\bar{A}_\varepsilon$ and $D_{m,\varepsilon}$ are as defined in (10.6). Let f be concentrated on wave packets from $\mathbb{T}_Z$.

It suffices to show that (10.1) holds at the endpoint $p = \bar{p}_0(k, m)$. Indeed, observe that Lemma 5.5 implies the L^2-bound

$$\|T^\lambda f\|_{\mathrm{BL}_{k,A}^2(B(0,R))}^2 \lesssim \sum_{\tau : K^{-1}\text{-cap}} \|T^\lambda f_\tau\|_{L^2(B(0,R))}^2 \lesssim R\|f\|_{L^2(B^{n-1})}^2.$$

Once (10.1) is established for $p = \bar{p}_0(k, m)$, one may use the logarithmic convexity of the k-broad norms to interpolate the $p = \bar{p}_0(k, m)$ estimate against the above inequality, and thereby obtain (10.1) in the desired range.

The analysis proceeds by considering two different cases.

The algebraic case

There exists a transverse complete intersection $Y^l \subseteq Z$ of dimension $1 \leqslant l \leqslant m-1$ of maximum degree at most $(D_{m,\varepsilon})^n$ such that

$$\|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(N_{R^{1/2+\delta_m/4}}(Y^l) \cap B(0,R))}^p \geqslant c_{\mathrm{alg}} \|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(B(0,R))}^p. \tag{10.11}$$

Here $c_{\mathrm{alg}} > 0$ is a constant depending only on n and ε which is chosen to be sufficiently small to suit the needs of the forthcoming argument.

The cellular case

The negation of the algebraic case: for every transverse complete intersection $Y^l \subseteq Z$ of dimension $1 \leqslant l \leqslant m-1$ and maximum degree at most $(D_{m,\varepsilon})^n$, the inequality

$$\|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(N_{R^{1/2+\delta_m/4}}(Y^l) \cap B(0,R))}^p < c_{\mathrm{alg}} \|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(B(0,R))}^p \tag{10.12}$$

holds.

The cellular case is the simplest situation and will be treated first. Here a polynomial partitioning argument is employed which splits the mass of the k-broad norm into small pieces; these pieces can then be treated individually using the radial induction hypothesis. The algebraic case is the most involved situation; it encapsulates the kind of behaviour exhibited by the sharp examples in §2. In this case, $T^\lambda f$ can be thought of as concentrated near a low-dimensional and low-degree variety Y^l (in a k-broad sense). If the wave packets of f are also tangential to this variety, then one may use induction on the dimension to

conclude the argument. This might not be the case, however, and if many of the wave packets of f are transverse to Y^l, then an alternative argument is required. Thus, the algebraic case naturally splits into two sub-cases, a tangential and a transverse sub-case, which are discussed in detail below. Lemma 7.5 can be applied to show that a given tube $T_{\theta,v}$ can only interact transversely with the variety Y^l inside a small number of finitely overlapping balls of some fixed radius $\varrho \ll R$ (more precisely, the radius is chosen to satisfy $\varrho^{1/2+\delta_l} = R^{1/2+\delta_m}$); this eventually allows one to also close the induction in the transverse situation.

10.5. The cellular case

The cellular case can be treated using polynomial partitioning. Roughly speaking, since by hypothesis $T^\lambda f$ is concentrated in a neighbourhood of an m-dimensional transverse complete intersection, for any $D \geqslant 1$ Theorem 7.3 can be applied in m dimensions to show that there exists a non-zero polynomial P of degree at most D such that, letting $\{O_i\}_{i \in \mathcal{I}}$ denote the connected components of $\mathbb{R}^n \backslash Z(P)$ (which, recall, are referred to as *cells*), one has $\#\{O_i : i \in \mathcal{I}\} \sim D^m$ and

$$\|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(O_i)}^p \sim D^{-m} \|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(B(0,R))}^p \quad \text{for all } i \in \mathcal{I}. \tag{10.13}$$

Thus, the mass of the k-broad norm is essentially equally distributed amongst the cells. Moreover, using the hypothesis of the cellular case, one can construct P so that very little of the mass lies near the cell wall

$$W := N_{2R^{1/2+\delta}}(Z(P)) \cap B(0,R).$$

In particular, the estimate (10.13) essentially still holds if the O_i are replaced with the shrunken cells $O_i' := O_i \backslash W$. The O_i' can be thought of as well separated,[26] and this facilitates a divide-and-conquer-style argument. More precisely, the fact that a non-zero univariate polynomial of degree at most D has at most D roots quickly leads to the following observation.

LEMMA 10.3. *If P is a polynomial of degree $\deg P \leqslant D$ and $\{O_i'\}_{i \in \mathcal{I}}$ are defined as above, then each tube $T_{\theta,v}$ enters at most D/ε of the cells O_i'.*

It is important to note that, in general, Lemma 10.3 does not hold if the O_i' are replaced with the cells O_i.

---

[26] In particular, the distance between a pair of distinct O_i' is wider than the width $R^{1/2+\delta}$ of any tube $T_{\theta,v}$.

Proof. Let $[\Gamma_{\theta,v}^\lambda]_\varepsilon : \mathbb{R} \to \mathbb{R}^{n-1}$ denote the polynomial approximant of $\Gamma_{\theta,v}^\lambda$, as defined in §7.2. Thus, $\deg[\Gamma_{\theta,v}^\lambda]_\varepsilon \leqslant \lceil 1/2\varepsilon \rceil$ and (7.2) implies that

$$|[\Gamma_{\theta,v}^\lambda]_\varepsilon(t) - \Gamma_{\theta,v}^\lambda(t)| \leqslant R^{1/2} \quad \text{for all } t \in (-R, R).$$

Suppose that $x \in O_i' \cap T_{\theta,v}$. By the definition of O_i', the ball $B\left(x, 2R^{1/2+\delta}\right)$ contains no points of $Z(P)$, and is therefore contained in O_i. On the other hand, we have $\operatorname{dist}(x, \Gamma_{\theta,v}^\lambda) < R^{1/2+\delta}$, and therefore $\operatorname{dist}(x, [\Gamma_{\theta,v}^\lambda]_\varepsilon) < 2R^{1/2+\delta}$. Consequently, $[\Gamma_{\theta,v}^\lambda]_\varepsilon$ enters $B(x, 2R^{1/2+\delta}) \subseteq O_i$. Thus, if $T_{\theta,v}$ enters a cell O_i', then the polynomial curve $[\Gamma_{\theta,v}^\lambda]_\varepsilon$ enters O_i whilst, by the simple property of univariate polynomials quoted above, $[\Gamma_{\theta,v}^\lambda]_\varepsilon$ can enter at most $\deg P \cdot \deg[\Gamma_{\theta,v}^\lambda]_\varepsilon + 1 \leqslant D/\varepsilon$ cells O_i. $\qquad\square$

Some aspects of the discussion prior to Lemma 10.3 are not entirely precise; for instance, to apply the polynomial partitioning theorem, one must work with an L^1 function, rather than a k-broad norm. In view of this, let μ denote the measure on $\mathbb{R}^n$ with Radon–Nikodym derivative

$$\sum_{B_{K^2} \in \mathcal{B}_{K^2}} \mu_{T^\lambda f}(B_{K^2}) \frac{1}{|B_{K^2}|} \chi_{B_{K^2}}$$

with respect to the Lebesgue measure. One may easily verify that

$$\mu(U) \leqslant \|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(U)}^p \quad \text{and} \quad \|T^\lambda f\|_{\mathrm{BL}_{k,A}^p(B(0,R))}^p \leqslant \mu(B(0, 2R)) \qquad (10.14)$$

for all Lebesgue measurable sets $U \subseteq \mathbb{R}^n$. These inequalities ensure that the measure μ acts as an effective surrogate for the k-broad norm in the polynomial partitioning argument.

By combining the cellular hypothesis with Theorem 7.3, one obtains the following partitioning result.

LEMMA 10.4. (Polynomial partitioning [14]) *There exists a polynomial P of degree $\deg P \leqslant D_{m,\varepsilon}$ such that, if $\{O_i\}_{i \in \mathcal{I}}$ and W are defined as above and $O_i' := O_i \setminus W$ for all $i \in \mathcal{I}$, then $\#\mathcal{I} \lesssim (D_{m,\varepsilon})^m$ and*

$$\mu(O_i') \sim (D_{m,\varepsilon})^{-m} \mu(B(0, 2R)) \qquad (10.15)$$

for at least 99% of the cells O_i'.

This lemma is contained in the work of the first author [14, §8.1], and the details of the proof are not reproduced here. The basic idea is as follows: by hypothesis, the mass of μ is concentrated in $N_{R^{1/2+\delta_m}}(Z)$, where Z is an m-dimensional algebraic variety; this

allows one to apply Theorem 7.3 in m-dimensions to construct a polynomial P which satisfies the desired properties, with O_i in place of O_i'. The hypothesis of the cellular case implies that the mass of μ cannot concentrate in a neighbourhood of a certain type of algebraic variety, and this can be used to show, in particular, that the mass cannot concentrate around the cell wall W. Provided the constant c_{alg} is chosen to be sufficiently small, this allows one to pass to the shrunken cells O_i' in (10.15) (at least for 99% of the cells).

There are a number of technicalities involved in rigorously carrying out this argument. In particular, one must justify the application of Theorem 7.3 in dimension m; this requires locally approximating Z by some tangent plane $T_z Z$ and applying the theorem to the push-forward of μ onto $T_z Z$ under orthogonal projection. The partitioning variety in $T_z Z$ is lifted to a variety $\widetilde{Z}$ in $\mathbb{R}^n$ by taking the pre-image under the orthogonal projection; it is possible to define $\widetilde{Z}$ in this way, so that it is transverse to Z. The cells O_i are then defined with respect to $\widetilde{Z}$.([27])

Presently, it is shown how together Lemmas 10.3 and 10.4 easily yield the proof of Proposition 10.1 in the cellular case. Applying Lemma 10.4 one obtains a partition of $\mathbb{R}^n \backslash W$ into disjoint cells $\{O_i\}_{i \in \mathcal{I}}$. For each $i \in \mathcal{I}$ let

$$\mathbb{T}_i := \{(\theta, v) \in \mathbb{T}_Z : T_{\theta,v} \cap O_i' \neq \varnothing\} \quad \text{and} \quad f_i := \sum_{(\theta,v) \in \mathbb{T}_i} f_{\theta,v}.$$

By Lemma 5.4 one has

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(O_i')}^p \leqslant \|T^\lambda f_i\|_{\mathrm{BL}^p_{k,A}(O_i')}^p + \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}^p.$$

Combining this inequality with (10.14) and Lemma 10.4, one deduces that at least 99% of the cells O_i' have the property that

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))}^p \lesssim (D_{m,\varepsilon})^m \|T^\lambda f_i\|_{\mathrm{BL}^p_{k,A}(O_i')}^p + \mathrm{RapDec}(R)\|f\|_{L^2(B^{n-1})}^p. \qquad (10.16)$$

On the other hand, by Lemma 10.3 and orthogonality between the $f_{\theta,v}$, one has

$$\sum_{i \in \mathcal{I}} \|f_i\|_{L^2(B^{n-1})}^2 \sim \sum_{(\theta,v) \in \mathbb{T}_Z} \#\{i \in \mathcal{I} : (\theta,v) \in \mathbb{T}_i\} \|f_{\theta,v}\|_{L^2(B^{n-1})}^2 \lesssim_\varepsilon D_{m,\varepsilon} \|f\|_{L^2(B^{n-1})}^2.$$

Since there are roughly $(D_{m,\varepsilon})^m$ cells in total, Markov's inequality shows that at least 99% of the cells O_i' have the property that

$$\|f_i\|_{L^2(B^{n-1})}^2 \lesssim_\varepsilon D_{m,\varepsilon}^{-(m-1)} \|f\|_{L^2(B^{n-1})}^2. \qquad (10.17)$$

([27]) To carry out this argument rigorously, one must further ensure that all the relevant varieties are transverse complete intersections of dimension at most $m-1$ and controlled degree in order to invoke (10.12). Such technicalities account for the choice of maximum degree $(D_{m,\varepsilon})^n$ in the definition of the algebraic and cellular cases.

Therefore, there exists some cell O_i' for which (10.16) and (10.17) simultaneously hold; henceforth, attention is focused on a single such cell O_i'.

Let $E_{m,A}(R)$ denote the constant appearing on the right-hand side of (10.1); namely,

$$E_{m,A}(R) := C_{m,\varepsilon} K^{\bar{C}_\varepsilon} R^{\varepsilon - c_n \delta_m + \delta \left(\log \bar{A}_\varepsilon - \log A \right) - e_{k,n}(p) + 1/2}.$$

By the radial induction hypothesis, Proposition 10.1 holds for the radius $\frac{1}{2}R$. For the fixed choice of i as above, cover O_i' with $O(1)$ balls of radius $\varrho = \frac{1}{2}R$. Applying Lemma 10.2 to f_i on each of these balls, one obtains

$$\|T^\lambda f_i\|_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim E_{m,A}\left(\tfrac{1}{2}R\right)\|f_i\|_{L^2(B^{n-1})} \lesssim E_{m,A}(R)\|f_i\|_{L^2(B^{n-1})}.$$

Combining the above estimate with (10.16) and (10.17), one deduces that

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))} \leqslant C_\varepsilon (D_{m,\varepsilon})^{-(m-1)/2+m/p} E_{m,A}(R)\|f\|_{L^2(B^{n-1})}$$

for some constant $C_\varepsilon > 0$. The $D_{m,\varepsilon}$ exponent is negative if and only if $p > 2m/(m-1)$; note this is the case for the choice of exponent $p = \bar{p}_0(k,m)$.[28] Thus, recalling the definition of $D = D_{m,\varepsilon}$ and assuming ε is sufficiently small depending on n, it follows that $C_\varepsilon (D_{m,\varepsilon})^{-(m-1)/2+m/p} \leqslant 1$. This establishes the desired estimate (10.1) and closes the induction in the cellular case.

10.6. The algebraic case

Fix a transverse complete intersection Y^l of dimension $1 \leqslant l \leqslant m-1$ and maximum degree $\overline{\deg}\, Y^l \leqslant (D_{m,\varepsilon})^n$ which satisfies (10.11). Let $R^{1/2} \ll \varrho \ll R$ be such that $\varrho^{1/2+\delta_l} = R^{1/2+\delta_m}$, and note that

$$R \leqslant R^{2\delta_l} \varrho \quad \text{and} \quad \varrho \leqslant R^{-\delta_l/2} R. \tag{10.18}$$

Let $\mathcal{B}_\varrho$ be a finitely-overlapping cover of $B(0,R)$ by ϱ-balls, and for each $B \in \mathcal{B}_\varrho$ define

$$\mathbb{T}_B := \{(\theta,v) \in \mathbb{T} : T_{\theta,v} \cap N_{R^{1/2+\delta_m}/4}(Y^l) \cap B \neq \varnothing\}$$

and

$$f_B := \sum_{(\theta,v) \in \mathbb{T}_B} f_{\theta,v}.$$

Thus, by (10.11) and Lemma 5.4,

$$\|T^\lambda f\|^p_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim \sum_{B \in \mathcal{B}_\varrho} \|T^\lambda f_B\|^p_{\mathrm{BL}^p_{k,A}(N_{R^{1/2+\delta_m}/4}(Y^l) \cap B)}$$

---

[28] It is for this reason that one works with the modified exponent $\bar{p}_0(k,m)$ rather than $\bar{p}(k,m)$.

holds, up to the inclusion of a rapidly decreasing error term. In what follows, such error terms, which are harmless, are suppressed in the notation.

For $B = B(y, \varrho) \in \mathcal{B}_\varrho$, let $\mathbb{T}_{B, \text{tang}}$ denote the set of all $(\theta, v) \in \mathbb{T}_B$ with the property that, whenever $x \in T_{\theta, v}$ and $z \in Y^l \cap B(y, 2\varrho)$ satisfy $|x - z| \leqslant 2\overline{C}_{\text{tang}} \varrho^{1/2 + \delta_l}$, it follows that

$$\angle(G^\lambda(x; \omega_\theta), T_z Y^l) \leqslant \tfrac{1}{2} \bar{c}_{\text{tang}} \varrho^{-1/2 + \delta_l},$$

where $\overline{C}_{\text{tang}}$ and $\bar{c}_{\text{tang}}$ are the constants appearing in the definition of tangency. Furthermore, let $\mathbb{T}_{B, \text{trans}} := \mathbb{T}_B \setminus \mathbb{T}_{B, \text{tang}}$ and define

$$f_{B, \text{tang}} := \sum_{(\theta, v) \in \mathbb{T}_{B, \text{tang}}} f_{\theta, v} \quad \text{and} \quad f_{B, \text{trans}} := \sum_{(\theta, v) \in \mathbb{T}_{B, \text{trans}}} f_{\theta, v}.$$

It follows that $f_B = f_{B, \text{tang}} + f_{B, \text{trans}}$ and, by the triangle inequality for broad norms (that is, Lemma 6.2), one concludes that

$$\|T^\lambda f\|_{\text{BL}^p_{k,A}(B(0,R))}^p \lesssim \sum_{B \in \mathcal{B}_\varrho} \|T^\lambda f_{B, \text{tang}}\|_{\text{BL}^p_{k,A/2}(B)}^p + \sum_{B \in \mathcal{B}_\varrho} \|T^\lambda f_{B, \text{trans}}\|_{\text{BL}^p_{k,A/2}(B)}^p.$$

Either the tangential or transverse contribution to the above sum dominates, and each case is treated separately.

The tangential sub-case

Suppose that the tangential term dominates; that is,

$$\|T^\lambda f\|_{\text{BL}^p_{k,A}(B(0,R))}^p \lesssim \sum_{B \in \mathcal{B}_\varrho} \|T^\lambda f_{B, \text{tang}}\|_{\text{BL}^p_{k,A/2}(B)}^p. \tag{10.19}$$

Each term in the right-hand sum is bounded using the dimensional induction hypothesis. Fix $B = B(y, \varrho) \in \mathcal{B}_\varrho$ and, as in §9, let

$$\widetilde{T}^\lambda (f_{B, \text{tang}})^\sim (x) = T^\lambda f_{B, \text{tang}}(x + y),$$

so that[29]

$$\|T^\lambda f_{B, \text{tang}}\|_{\text{BL}^p_{k,A/2}(B(y, \varrho))} = \|\widetilde{T}^\lambda (f_{B, \text{tang}})^\sim\|_{\text{BL}^p_{k,A/2}(B(0, \varrho))}. \tag{10.20}$$

Since $\overline{\deg} \, Y^l \leqslant D_{l, \varepsilon}$, in order to apply the induction hypothesis, one must verify that $(f_{B, \text{tang}})^\sim$ is concentrated on scale ϱ wave packets that are $\varrho^{-1/2 + \delta_l}$-tangent to $Y^l - y$ in $B(0, \varrho)$. By Lemma 9.1, $(f_{B, \text{tang}})^\sim$ is concentrated on scale ϱ wave packets from

$$\widetilde{\mathbb{T}}_{B, \text{tang}} := \bigcup_{(\theta, v) \in \mathbb{T}_{B, \text{tang}}} \widetilde{\mathbb{T}}_{\theta, v},$$

---

where the $\widetilde{\mathbb{T}}_{\theta,v}$ are as defined in §9. Fix $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{B,\mathrm{tang}}$ and recall from (9.14) that

$$\mathrm{dist}_H(\widetilde{T}_{\tilde{\theta},\tilde{v}}, (T_{\theta,v}-y) \cap B(0,\varrho)) \lesssim R^{1/2+\delta} \ll \varrho^{1/2+\delta_l}. \tag{10.21}$$

Combining this with the definition of $\mathbb{T}_{B,\mathrm{tang}}$, it is easy to see that $\widetilde{T}_{\tilde{\theta},\tilde{v}}$ satisfies the angle condition for tangency, and it remains to verify the containment property

$$\widetilde{T}_{\tilde{\theta},\tilde{v}} \subseteq N_{\varrho^{1/2+\delta_l}}(Y^l-y).$$

The definition of $\mathbb{T}_B$ and (10.21) imply that

$$\widetilde{T}_{\tilde{\theta},\tilde{v}} \cap N_{\varrho^{1/2+\delta_l}/2}(Y^l-y) \cap B(0,\varrho) \neq \varnothing,$$

and so the containment property follows from the angle condition, as in the proof of Proposition 9.2.

Thus, the dimensional induction hypothesis may be applied to $(f_{B,\mathrm{tang}})^\sim$, and one therefore deduces that

$$\|\widetilde{T}^\lambda(f_{B,\mathrm{tang}})^\sim\|_{\mathrm{BL}^p_{k,A/2}(B(0,\varrho))} \leqslant E_{l,A/2}(\varrho)\|f_{B,\mathrm{tang}}\|_{L^2(B^{n-1})}.$$

Combining this estimate with (10.19) and (10.20), one concludes that

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))} \leqslant R^{O(\delta_l)} E_{l,A/2}(\varrho)\|f\|^2_{L^2(B^{n-1})}.$$

To close the induction in this case, it remains to show that

$$R^{O(\delta_l)} E_{l,A/2}(\varrho) \leqslant E_{m,A}(R).$$

By (10.18),

$$\varrho^{\delta(\log \bar{A}_\varepsilon - \log(A/2))} \leqslant R^{O(\delta_l)} R^{\delta(\log \bar{A}_\varepsilon - \log A)},$$

$$\varrho^{-e_{k,n}(p)+1/2} \leqslant R^{O(\delta_l)} R^{-e_{k,n}(p)+1/2}.$$

Combining these observations, the problem is further reduced to showing that

$$\varrho^{\varepsilon-c_n\delta_l} \leqslant R^{-c\delta_l} R^{\varepsilon-c_n\delta_m},$$

where $c>1$ is a suitably large-dimensional constant. By (10.18), one may ensure that this inequality holds by choosing the constant c_n in (10.6) at the outset to be large relative to c.

The transverse sub-case

Now suppose the transverse term dominates; that is,

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))}^p \lesssim \sum_{B\in\mathcal{B}_\varrho} \|T^\lambda f_{B,\mathrm{trans}}\|_{\mathrm{BL}^p_{k,A/2}(B)}^p. \tag{10.22}$$

The idea here is somewhat similar to that used in the cellular case. Recall, in the cellular case the number of cells a given tube can enter is controlled by Bézout's theorem. In the transverse case, the number of balls $B\in\mathcal{B}_\varrho$ inside which a given tube can be transverse to Y^l is again controlled due to Bézout's theorem, this time by Lemma 7.5. This yields the following key inequality.

CLAIM.

$$\sum_{B\in\mathcal{B}_\varrho} \|f_{B,trans}\|_{L^2(B^{n-1})}^2 \lesssim_\varepsilon \|f\|_{L^2(B^{n-1})}^2. \tag{10.23}$$

Proof. This is a fairly direct consequence of the hypothesis of the transverse case together with Lemma 7.5. Indeed, note that

$$\sum_{B\in\mathcal{B}_\varrho} \|f_{B,\mathrm{trans}}\|_{L^2(B^{n-1})}^2 \sim \sum_{(\theta,v)\in\mathbb{T}} \#\{B\in\mathcal{B}_\varrho : (\theta,v)\in\mathbb{T}_{B,\mathrm{trans}}\}\|f_{\theta,v}\|_{L^2(B^{n-1})}^2,$$

and so, to prove (10.23), it suffices to fix an arbitrary $(\theta,v)\in\mathbb{T}_{B,\mathrm{trans}}$ and show that

$$\#\{B\in\mathcal{B}_\varrho : (\theta,v)\in\mathbb{T}_{B,\mathrm{trans}}\} \lesssim_\varepsilon 1. \tag{10.24}$$

Let $\Gamma:=[\Gamma^\lambda_{\theta,v}]_\varepsilon:\mathbb{R}\to\mathbb{R}^n$ be the polynomial approximant of the core curve $\Gamma^\lambda_{\theta,v}$ defined in §7.2. Thus, $\deg\Gamma\lesssim_\varepsilon 1$ and, recalling that $R\lesssim_\varepsilon\lambda^{1-\varepsilon}$, property (7.2) of the approximant implies that

$$|\Gamma(t)-\Gamma^\lambda_{\theta,v}(t)| \leqslant R^{1/2} \quad \text{for all } t\in(-R,R). \tag{10.25}$$

Let $u\in T_{\theta,v}$ and $x\in\Gamma\cap B(0,R)$ with $|u-x|\lesssim R^{1/2+\delta}$. It follows from the definition of $T_{\theta,v}$ and (10.25) that there exists some $t\in(-R,R)$ such that

$$|u-\Gamma^\lambda_{\theta,v}(t)|\lesssim R^{1/2+\delta} \quad \text{and} \quad |x-\Gamma(t)|\lesssim R^{1/2+\delta}.$$

Consequently, recalling Lemma 4.6,

$$\measuredangle(G^\lambda(u;\omega_\theta),\mathrm{T}_x\Gamma) \lesssim \measuredangle(T_{\Gamma^\lambda_{\theta,v}(t)}\Gamma^\lambda_{\theta,v},T_{\Gamma(t)}\Gamma)+\frac{R^{1/2+\delta}}{\lambda},$$

and therefore, by property (7.3) of the approximant,

$$\measuredangle(G^\lambda(u;\omega_\theta),\mathrm{T}_x\Gamma) \lesssim_\varepsilon \lambda^{-1/2}+\frac{R^{1/2+\delta}}{\lambda} < \frac{\bar{c}_{\mathrm{tang}}}{4}\varrho^{-1/2+\delta_l}.$$

Using the above inequality, one may easily verify that, if $B=B(y,\varrho)\in\mathcal{B}_\varrho$ and $(\theta,v)\in$ $\mathbb{T}_{B,\mathrm{trans}}$, then $Y_{>\alpha,r,\Gamma}^l\cap B(y,2\varrho)\neq\varnothing$ for $\alpha\sim\varrho^{-1/2+\delta_l}$ and $r\sim\varrho^{1/2+\delta_l}$. Here, $Y_{>\alpha,r,\Gamma}^l$ is as defined in §7; that is

$$Y_{>\alpha,r,\Gamma}^l:=\{z\in Y^l:\text{there exists }x\in\Gamma\text{ with }|x-z|<r\text{ and }\measuredangle(T_zY^l,T_x\Gamma)>\alpha\}.$$

By Lemma 7.5, the number of balls $B=B(y,\varrho)\in\mathcal{B}_\varrho$ for which $B(y,2\varrho)$ intersects $Y_{>\alpha,r,\Gamma}^l$ non-trivially is at most $O((\deg\Gamma)^n\cdot(\overline{\deg}\,Y^l)^n)=O_\varepsilon(1)$. Combining these observations, one immediately deduces (10.24), as required. $\qquad\square$

In view of (10.23), the strategy in the transverse case is to use the radial induction hypothesis to show that, for some constant $\bar{c}_\varepsilon>0$, one has

$$\|T^\lambda f_{B,\mathrm{trans}}\|_{\mathrm{BL}_{k,A/2}^p(B)}\leqslant\bar{c}_\varepsilon E_{m,A}(R)\|f_{B,\mathrm{trans}}\|_{L^2(B^{n-1})}\quad\text{for all }B\in\mathcal{B}_\varrho.\qquad(10.26)$$

Indeed, provided $\bar{c}_\varepsilon>0$ is sufficiently small, depending only on n and ε, the preceding inequality may be combined with (10.22), (10.23) and the simple estimate

$$\|f_{B,\mathrm{trans}}\|_{L^2(B^{n-1})}\lesssim\|f\|_{L^2(B^{n-1})},$$

to yield

$$\|T^\lambda f\|_{\mathrm{BL}_{k,A/2}^p(B(0,R))}\lesssim_\varepsilon\bar{c}_\varepsilon E_{m,A}(R)\|f\|_{L^2(B^{n-1})}^{1-2/p}\bigg(\sum_{B\in\mathcal{B}_\varrho}\|f_{B,\mathrm{trans}}\|_{L^2(B^{n-1})}^2\bigg)^{1/p}$$
$$\leqslant E_{m,A}(R)\|f\|_{L^2(B^{n-1})},$$

closing the induction in this case.

The main obstacle in carrying out this programme is that the $f_{B,\mathrm{trans}}$ do not, in general, satisfy the hypothesis of Proposition 10.1 at scale ϱ, and therefore one cannot directly apply the radial induction hypothesis to these functions. However, one can appeal to the theory developed in §9, which essentially allows $f_{B,\mathrm{trans}}$ to be broken into pieces $f_{B,\mathrm{trans},b}$ which do satisfy the hypothesis of the proposition at scale ϱ. Here is a sketch of how the argument works. By choosing a suitable set of translates $\mathfrak{B}$, one may essentially write

$$\|T^\lambda f_{B,\mathrm{trans}}\|_{\mathrm{BL}_{k,A/2}^p(B)}\lesssim\bigg(\sum_{b\in\mathfrak{B}}\|T^\lambda f_{B,\mathrm{trans},b}\|_{\mathrm{BL}_{k,A/2}^p(B)}^p\bigg)^{1/p},\qquad(10.27)$$

where each piece $f_{B,\mathrm{trans},b}$ is defined so that it is concentrated on scale ϱ wave packets which are tangential to some translate $Z-y+b$ of Z. By the theory of transverse equidistributions developed in §8 and §9, the $f_{B,\mathrm{trans},b}$ satisfy favourable L^2 estimates and, in

particular, the inequality (10.30) below holds. The radial induction hypothesis is applied to each of the $T^\lambda f_{B,\text{trans},b}$. To close the induction, one must estimate the resulting sum

$$
\left(\sum_{b \in \mathfrak{B}} \| f_{B,\text{trans},b} \|^2_{L^2(B^{n-1})} \right)^{1/2}
$$

in terms of $\| f_{B,\text{trans}} \|_{L^2(B^{n-1})}$. Here, the gain in ϱ/R in (10.30), afforded by transverse equidistribution, is crucial to the argument: it allows one to sum up the contributions from the individual pieces $f_{B,\text{trans},b}$ without any (significant) loss in R. It is this gain which accounts for the improved range of estimates for the k-broad inequalities under the positive-definite hypothesis (recall, the proof of the transverse equidistribution lemma relied heavily on the positive-definite condition).

As part of this argument, to ensure that the $f_{B,\text{trans},b}$ form a reasonable decomposition of $f_{B,\text{trans}}$ so that (10.27) essentially holds, the set of translates $\mathfrak{B}$ must be chosen so that $\bigcup_{b \in \mathfrak{B}} N_{\varrho^{1/2+\delta_m}}(Z - y + b)$ covers $N_{R^{1/2+\delta_m}}(Z)$ (recall, by hypothesis $f_{B,\text{trans}}$ is concentrated on wave packets in $\mathbb{T}_Z$, and so the mass of $T^\lambda f_{B,\text{trans}}$ is concentrated in $N_{R^{1/2+\delta_m}}(Z)$), and so that the $N_{\varrho^{1/2+\delta_m}}(Z - y + b)$ are essentially disjoint. This can be achieved using a probabilistic construction. More precisely, fixing $B = B(y, \varrho) \in \mathcal{B}_\varrho$, one may show the following.

LEMMA 10.5. *There exist a finite set $\mathfrak{B} \subset B(0, 2R^{1/2+\delta_m})$ and a collection*

$$
\mathcal{B}' \subseteq \{ B_{K^2} \in \mathcal{B}_{K^2} : B_{K^2} \cap B(y, \varrho) \neq \varnothing \}
$$

such that, up to inclusion of a rapidly decreasing error term,

$$
\| T^\lambda f_{B,trans} \|_{\mathrm{BL}^p_{k,A/2}(B)} \lesssim (\log R)^2 \left(\sum_{B_{K^2} \in \mathcal{B}'} \mu_{T^\lambda f_{B,trans}}(B_{K^2}) \right)^{1/p}, \tag{10.28}
$$

and for each $B_{K^2} \in \mathcal{B}'$ the following hold:
(i) *there exists some $b \in \mathfrak{B}$ such that*

$$
B_{K^2} \subset N_{\varrho^{1/2+\delta_m}/2}(Z + b); \tag{10.29}
$$

(ii) *there exist at most $O(1)$ vectors $b \in \mathfrak{B}$ for which*

$$
B_{K^2} \cap N_{\varrho^{1/2+\delta_m}}(Z + b) \neq \varnothing.
$$

The proof of the lemma, which is slightly technical, is postponed until the end of the section. Temporarily assuming this result, one may argue as follows to complete the proof of Proposition 10.1.

For each $b \in \mathfrak{B}$, let $\mathcal{B}'_b$ denote the collection of all $B_{K^2} \in \mathcal{B}'$ for which (10.29) holds. Thus, by (10.28) and property (i) in the lemma,

$$\|T^\lambda f_{B,\mathrm{trans}}\|_{\mathrm{BL}^p_{k,A/2}(B)} \lesssim (\log R)^2 \bigg(\sum_{b \in \mathfrak{B}} \sum_{B_{K^2} \in \mathcal{B}'_b} \mu_{\widetilde{T}^\lambda(f_{B,\mathrm{trans}})^\sim}(B_{K^2} - y) \bigg)^{1/p},$$

up to a rapidly decreasing error term.

Define the collection of wave packets

$$\widetilde{\mathbb{T}}'_b := \bigg\{ (\tilde{\theta}, \tilde{v}) \in \bigcup_{(\theta,v) \in \mathbb{T}_{B(y,\varrho),\mathrm{trans}}} \widetilde{\mathbb{T}}_{\theta,v} : \widetilde{T}_{\tilde{\theta},\tilde{v}} \cap \bigcup_{B_{K^2} \in \mathcal{B}'_b} (B_{K^2} - y) \neq \varnothing \bigg\};$$

note this set is a subset of the collection $\widetilde{\mathbb{T}}_b$ defined in §9 and so, by Proposition 9.2, one has $\widetilde{\mathbb{T}}'_b \subseteq \widetilde{\mathbb{T}}_{Z-y+b}$. Therefore, if $f_{B,\mathrm{trans},b}$ is defined by

$$(f_{B,\mathrm{trans},b})^\sim = \sum_{(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{T}}'_b} (f_{B,\mathrm{trans}})^\sim_{\tilde{\theta},\tilde{v}},$$

then $(f_{B,\mathrm{trans},b})^\sim$ is concentrated on wave packets that are $\varrho^{-1/2+\delta_m}$-tangent to $Z-y+b$. Furthermore, again up to a rapidly decreasing error term, one has

$$\|T^\lambda f_{B,\mathrm{trans}}\|_{\mathrm{BL}^p_{k,A/2}(B)} \lesssim (\log R)^2 \bigg(\sum_{b \in \mathfrak{B}} \|\widetilde{T}^\lambda(f_{B,\mathrm{trans},b})^\sim\|^p_{\mathrm{BL}^p_{k,A/2}(B(0,\varrho))} \bigg)^{1/p}.$$

The function $(f_{B,\mathrm{trans},b})^\sim$ satisfies the hypotheses of Proposition 10.1 at scale ϱ, and therefore the radial induction hypothesis yields

$$\bigg(\sum_{b \in \mathfrak{B}} \|\widetilde{T}^\lambda(f_{B,\mathrm{trans},b})^\sim\|^p_{\mathrm{BL}^p_{k,A/2}(B(0,\varrho))} \bigg)^{1/p} \leqslant E_{m,A/2}(\varrho) \bigg(\sum_{b \in \mathfrak{B}} \|f_{B,\mathrm{trans},b}\|^p_{L^2(B^{n-1})} \bigg)^{1/p}.$$

On the other hand, it is claimed that

$$\bigg(\sum_{b \in \mathfrak{B}} \|f_{B,\mathrm{trans},b}\|^p_{L^2(B^{n-1})} \bigg)^{1/p} \lesssim R^{O(\delta_m)} \Big(\frac{\varrho}{R} \Big)^{(n-m)(1/4-1/2p)} \|f_{B,\mathrm{trans}}\|_{L^2(B^{n-1})}.$$

Clearly, it suffices to prove the above inequality for $p=2$ and $p=\infty$; the desired estimate for $p=\bar{p}_0(k,m)$ then follows by interpolation (via Hölder's inequality).

$p=2$

Observe that, by the orthogonality between the wave packets,

$$\sum_{b\in\mathfrak{B}} \|f_{B,\text{trans},b}\|^2_{L^2(B^{n-1})} \sim \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}} \#\mathfrak{B}_{\tilde{\theta},\tilde{v}}\cdot\|(f_{B,\text{trans}})_{\tilde{\theta},\tilde{v}}\|^2_{L^2(B^{n-1})},$$

where $\mathfrak{B}_{\tilde{\theta},\tilde{v}}:=\{b\in\mathfrak{B}:(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}'_b\}$. Fixing $(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}$, it suffices to show that $\#\mathfrak{B}_{\tilde{\theta},\tilde{v}}\lesssim1$. Supposing $\mathfrak{B}_{\tilde{\theta},\tilde{v}}\neq\varnothing$, there exists some $B_{K^2}\in\mathcal{B}'$ with $\widetilde{T}_{\tilde{\theta},\tilde{v}}\cap(B_{K^2}-y)\neq\varnothing$. For any $b\in\mathfrak{B}_{\tilde{\theta},\tilde{v}}$, it follows that $(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}_b$, and so $\widetilde{T}_{\tilde{\theta},\tilde{v}}\subseteq N_{\varrho^{1/2+\delta_m}}(Z-y+b)$ by Proposition 9.2. Consequently, $B_{K^2}\cap N_{\varrho^{1/2+\delta_m}}(Z+b)\neq\varnothing$ for all $b\in\mathfrak{B}_{\tilde{\theta},\tilde{v}}$, and so the desired bound follows from property (ii) of Lemma 10.5.

$p=\infty$

In this case, the estimate is a direct consequence of the transverse equidistribution estimates established in §8 and §9. In particular, the function $f_{B,\text{trans}}$ is concentrated on wave packets belonging to $\mathbb{T}_{Z,B}$ and so, by Lemma 9.8, one deduces that

$$\|f_{B,\text{trans},b}\|_{L^2(B^{n-1})} \lesssim R^{O(\delta_m)}\Big(\frac{\varrho}{R}\Big)^{(n-m)/4}\|f_{B,\text{trans}}\|_{L^2(B^{n-1})}, \tag{10.30}$$

as required.

The preceding analysis shows that $\|T^\lambda f_{B,\text{trans}}\|_{\text{BL}^p_{k,A/2}(B(0,R))}$ is bounded above by

$$R^{O(\delta_m)}E_{m,A}(\varrho)\Big(\frac{\varrho}{R}\Big)^{(n-m)(1/4-1/2p)}\|f_{B,\text{trans}}\|_{L^2(B^{n-1})}$$

and therefore, to prove (10.26) and thereby close the induction argument in this case, it suffices to show that

$$R^{O(\delta_m)}E_{m,A}(\varrho)\Big(\frac{\varrho}{R}\Big)^{(n-m)(1/4-1/2p)} \leqslant \bar{c}_\varepsilon E_{m,A}(R). \tag{10.31}$$

For the exponent $p=\bar{p}(k,m)$, one has

$$\varrho^{-e_{k,n}(p)+1/2}\Big(\frac{\varrho}{R}\Big)^{(n-m)(1/4-1/2p)} \leqslant R^{-e_{k,n}(p)+1/2},$$

whilst for the perturbed exponent $p=\bar{p}_0(k,m)$ the same inequality holds up to a $R^{O(\delta)}$ factor. Thus, the left-hand side of (10.31) is dominated by

$$R^{O(\delta_m)}\Big(\frac{\varrho}{R}\Big)^\varepsilon E_{m,A}(R).$$

Recalling (10.18) and the choice of parameters δ_l and δ_m, one obtains the desired inequality.

The probabilistic argument

The above argument establishes Proposition 10.1, except for the details of the probabilistic argument used to prove Lemma 10.5.

Proof of Lemma 10.5. Before commencing the argument proper, a few technical reductions are necessary. By a standard dyadic pigeonholing argument, one may assume that

$$\|T^\lambda f_{B,\text{trans}}\|_{\text{BL}^p_{k,A/2}(B)} \lesssim \log R \bigg(\sum_{B_{K^2} \in \mathcal{B}''} \mu_{T^\lambda f_{B,\text{trans}}}(B_{K^2}) \bigg)^{1/p} \qquad (10.32)$$

for some sub-collection $\mathcal{B}'' \subseteq \mathcal{B}_{K^2}$ with the property that

$$\mu_{T^\lambda f_{B,\text{trans}}}(B_{K^2}) \sim \mu_{T^\lambda f_{B,\text{trans}}}(\overline{B}_{K^2}) \quad \text{for all } B_{K^2}, \overline{B}_{K^2} \in \mathcal{B}''. \qquad (10.33)$$

Since $f_{B,\text{trans}}$ is concentrated on wave packets from $\mathbb{T}_{Z,B}$, one may further assume that $B_{K^2} \cap B(y, \varrho) \cap N_{R^{1/2+\delta_m}}(Z) \neq \varnothing$ for all $B_{K^2} \in \mathcal{B}''$, at the cost of a rapidly decaying term on the right-hand side of (10.32).

A set of translates $\mathfrak{B}$ will be selected at random from $\mathbb{R}^n$ according to a probability measure $\mathbf{P}$. The distribution $\mathbf{P}$ is taken to be a mollified version of the uniform probability distribution $\mathbf{P}_{\text{unif}}$ on $B(0, R^{1/2+\delta_m})$. In particular, let $\omega \colon \mathbb{R}^n \to [0, \infty)$ be given by[30]

$$\omega(x) := \exp \left(\frac{-(|x| - R^{1/2+\delta_m})_+}{\varrho^{1/2+\delta_m}} \right) \quad \text{for all } x \in \mathbb{R}^n,$$

and $\mathbf{P}$ be the continuous probability measure on $\mathbb{R}^n$ with Radon–Nikodym derivative $(\int_{\mathbb{R}^n} \omega)^{-1}\omega$ (with respect to Lebesgue measure). This measure approximates $\mathbf{P}_{\text{unif}}$, in the sense that

$$\mathbf{P}(\mathbb{R}^n \setminus B(0, 2R^{1/2+\delta_m})) = \text{RapDec}(R). \qquad (10.34)$$

The motivation behind the definition of $\mathbf{P}$ is that, in contrast with $\mathbf{P}_{\text{unif}}$, it enjoys the doubling property

$$\mathbf{P}(B(x, 2r)) \lesssim \mathbf{P}(B(x, r)) \quad \text{for all } x \in \mathbb{R}^n \text{ and } 0 < r \lesssim \varrho^{1/2+\delta_m}.$$

Consequently, by the Vitali covering lemma, for any $E \subseteq \mathbb{R}^n$ one has

$$\mathbf{P}(N_{2r}(E)) \lesssim \mathbf{P}(N_r(E)) \quad \text{for all } 0 < r \lesssim \varrho^{1/2+\delta_m}. \qquad (10.35)$$

Recall, if $B(x, K^2) \in \mathcal{B}''$, then $B(x, K^2) \cap N_{R^{1/2+\delta_m}}(Z) \neq \varnothing$, and so

$$|B(0, R^{1/2+\delta_m}) \cap N_{\varrho^{1/2+\delta_m}}(Z - x)| \gtrsim |B(0, \varrho^{1/2+\delta_m})|,$$

---

[30] Here $(u)_+ := \begin{cases} u, & \text{if } u \geqslant 0, \\ 0, & \text{if } u < 0, \end{cases}$ for all $u \in \mathbb{R}$.

which implies that

$$\mathbf{P}(N_{\varrho^{1/2+\delta_m}}(Z-x)) \gtrsim \frac{|B(0,\varrho^{1/2+\delta_m})|}{|B(0,R^{1/2+\delta_m})|}.$$

For any $s \in \mathbb{N}$ with $2^s \gtrsim |B(0,\varrho^{1/2+\delta_m})|$, define

$$\mathcal{B}^s := \left\{ B(x,K^2) \in \mathcal{B}'' : \mathbf{P}(N_{\varrho^{1/2+\delta_m}}(Z-x)) \sim \frac{2^s}{|B(0,R^{1/2+\delta_m})|} \right\}.$$

By a further pigeonholing argument, there exists some value of s as above such that (10.28) holds with $\mathcal{B}^s$ in place of $\mathcal{B}'$.

Let $\overline{C} \geqslant 1$ be a dimensional constant, chosen to be sufficiently large for the purposes of the following argument, and define $N := \lceil \overline{C} 2^{-s} |B(0,R^{1/2+\delta_m})| \rceil \in \mathbb{N}$. Recalling (10.18), it follows that

$$N \lesssim \frac{|B(0,R^{1/2+\delta_m})|}{|B(0,\varrho^{1/2+\delta_m})|} \lesssim R^{2n\delta_l}. \tag{10.36}$$

Suppose that $\mathfrak{B} = \{b_1, ..., b_N\}$ is a sequence of vectors in $\mathbb{R}^n$ formed by choosing each term independently at random, according to the probability distribution $\mathbf{P}$. The problem is to show that $\mathfrak{B}$ satisfies each of the desired properties with high probability.

The containment property $\mathfrak{B} \subset B(0, 2R^{1/2+\delta_m})$

Recalling (10.34) and (10.36), it follows that

$$\mathbf{P}(\mathfrak{B} \subseteq B(0,2R^{1/2+\delta_m})) = 1 + \sum_{k=1}^{N} \binom{N}{k}(-1)^k \mathbf{P}(\mathbb{R}^n \setminus B(0,2R^{1/2+\delta_m}))^k$$
$$= 1 + \mathrm{RapDec}(R).$$

Indeed, for the second equality we use the elementary bound

$$\left| \sum_{k=1}^{N} \binom{N}{k}(-1)^k u^k \right| = |(1-u)^N - 1| \leqslant N|u| \quad \text{for all } 0 \leqslant u \leqslant 1,$$

which follows from the mean value theorem. Thus, if $R \geqslant 1$ is sufficiently large depending only on n and ε, then

$$\mathbf{P}(\mathfrak{B} \subseteq B(0,2R^{1/2+\delta_m})) \geqslant \tfrac{99}{100}, \tag{10.37}$$

which verifies that the desired containment property holds with high probability.

Property (i)

Let $B(x, K^2) \in \mathcal{B}^s$ and observe that

$$\mathbf{P}\bigg(B(x, K^2) \subseteq \bigcup_{j=1}^{N} N_{\varrho^{1/2+\delta_m}/2}(Z+b_j) \bigg) \geqslant \mathbf{P}\bigg(x \in \bigcup_{j=1}^{N} N_{\varrho^{1/2+\delta_m}/4}(Z+b_j) \bigg)$$

$$= 1 - (1 - \mathbf{P}(N_{\varrho^{1/2+\delta_m}/4}(Z-x)))^N.$$

By the definition of $\mathcal{B}^s$ and the doubling property (10.35) of $\mathbf{P}$, it follows that

$$\mathbf{P}(N_{\varrho^{1/2+\delta_m}/4}(Z-x)) = \frac{c\overline{C}}{N}$$

for some dimensional constant $c>0$ and, consequently,

$$\mathbf{P}\bigg(B(x, K^2) \subseteq \bigcup_{j=1}^{N} N_{\varrho^{1/2+\delta_m}/2}(Z+b_j) \bigg) \geqslant 1 - \bigg(1 - \frac{c\overline{C}}{N} \bigg)^N > 1 - e^{-c\overline{C}}.$$

Let X denote the random variable that counts the number of $B_{K^2} \in \mathcal{B}^s$ for which $B_{K^2} \subseteq N_{\varrho^{1/2+\delta_m}/2}(Z+b)$ for some $b \in \mathfrak{B}$. If $\overline{C}$ is suitably chosen, then the above inequality implies that the expected value of X satisfies $\mathbf{E}[X] \geqslant (1-10^{-4})\#\mathcal{B}^s$. By Markov's inequality,

$$\mathbf{P}\bigg(X > \frac{99}{100}\#\mathcal{B}^s \bigg) \geqslant 1 - \frac{100}{\#\mathcal{B}^s}\mathbf{E}[\#\mathcal{B}^s - X] \geqslant \frac{99}{100}, \tag{10.38}$$

which verifies that property (i) of the lemma holds with high probability.

Property (ii)

For each $x \in \mathbb{R}^n$, let M_x denote the random variable that counts the number of sets $N_{\varrho^{1/2+\delta_m}}(Z+b_j)$ containing x; that is,

$$M_x(b_1, ..., b_N) := \sum_{j=1}^{N} \chi_{N_{2\varrho^{1/2+\delta_m}}(Z+b_j)}(x).$$

If $B(x, K^2) \in \mathcal{B}^s$, then

$$\mathbf{E}[M_x] = \sum_{j=1}^{N} \mathbf{P}(N_{2\varrho^{1/2+\delta_m}}(Z-x)) \sim N\frac{2^s}{|B(0, R^{1/2+\delta_m})|} \sim 1.$$

Now, let $C \geqslant 1$ be a dimensional constant and Y denote the random variable that counts the number of $B(x, K^2) \in \mathcal{B}^s$ for which $M_x \leqslant C$. By a 2-fold application of Markov's inequality, if C is chosen to be sufficiently large, then

$$\mathbf{P}\bigg(Y > \frac{99}{100}\#\mathcal{B}^s \bigg) \geqslant \mathbf{P}\bigg(\#\mathcal{B}^s - \frac{1}{C}\sum_{B(x, K^2) \in \mathcal{B}^s} M_x > \frac{99}{100}\#\mathcal{B}^s \bigg) \geqslant \frac{99}{100}, \tag{10.39}$$

which verifies that property (ii) of the lemma holds with high probability.

In view of (10.37), (10.38) and (10.39), there exists a choice of $\mathfrak{B} \subseteq B(0, 2R^{1/2+\delta_m})$ and a subset $\mathcal{B}' \subseteq \mathcal{B}^s$ of cardinality comparable to that of $\mathcal{B}^s$ for which the desired properties (i) and (ii) hold. Finally, by (10.33), the inequality (10.28) also holds for the sub-collection $\mathcal{B}'$. $\qquad\square$

11. Going from k-broad to linear estimates

11.1. Applying the Bourgain–Guth method

Theorem 1.2 can be deduced as a consequence of the k-broad estimates via the method of [9]. The key proposition is as follows.

PROPOSITION 11.1. *Suppose that for all $K \geqslant 1$ and all $\varepsilon > 0$ any Hörmander-type operator T^λ with reduced positive-definite phase obeys the k-broad inequality*

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim_\varepsilon K^{C_\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})} \tag{11.1}$$

for some fixed k, A, p, q, C_ε and all $R \geqslant 1$. If

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

then any Hörmander-type operator T^λ with positive-definite phase satisfies

$$\|T^\lambda f\|_{L^p(B(0,R))} \lesssim_{\phi,\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})}.$$

Theorem 1.2 is now a direct consequence of Proposition 11.1 and Theorem 1.9.

Proof of Theorem 1.2. Theorem 1.9 implies that, for each $2 \leqslant k \leqslant n$, the estimate (11.1) is valid for all $p \geqslant \bar{p}(n, k)$. Thus, for each k satisfying the constraint

$$2\frac{2n-k+2}{2n-k} \leqslant \bar{p}(n, k) = 2\frac{n+k}{n+k-2}, \tag{11.2}$$

one may apply Proposition 11.1 with $\bar{p}(n, k) \leqslant p \leqslant 2(k-1)/(k-2)$ to obtain a (potentially empty) range of estimates for the linear problem. Since $\bar{p}(n, k)$ is a decreasing function of k, the optimal estimate is given by applying Proposition 11.1 as above with k chosen to be as large as possible, subject to (11.2). Rearranging (11.2) yields $k \leqslant \frac{1}{2}n+1$. Defining

$$k_* := \begin{cases} \frac{1}{2}n+1, & \text{if } n \text{ is even,} \\ \frac{1}{2}(n+1), & \text{if } n \text{ is odd,} \end{cases}$$

one may readily verify that

$$\bar{p}(n, k_*) \leqslant 2\frac{k_*-1}{k_*-2},$$

and so the linear estimate holds for all $p \geqslant \bar{p}(n, k_*)$. A simple computation shows that this corresponds to the range of estimates stated in Theorem 1.2. $\qquad\square$

For contrast, it is noted that there is also a version of Proposition 11.1 which holds without the positive-definite assumption. This can be combined with the multilinear estimates of Bennett–Carbery–Tao [4] to prove Theorem 1.1 (this is essentially the argument used in [9]).

PROPOSITION 11.2. *Suppose that, for all $K \geqslant 1$ and all $\varepsilon > 0$, any Hörmander-type operator T^λ with reduced phase*[31] *obeys the k-broad inequality*

$$\|T^\lambda f\|_{\mathrm{BL}^p_{k,A}(B(0,R))} \lesssim_\varepsilon K^{C_\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})}$$

for some fixed k, A, p, q, C_ε and all $R \geqslant 1$. If

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

then any Hörmander-type operator T^λ satisfies

$$\|T^\lambda f\|_{L^p(B(0,R))} \lesssim_{\phi,\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})}.$$

Theorem 1.1 is now a direct consequence of Proposition 11.2 and the Bennett–Carbery–Tao theorem.

Proof of Theorem 1.1. The proof is precisely the same as that of Theorem 1.2 above, but with the exponent $2k/(k-1)$ from the Bennett–Carbery–Tao theorem (that is, Theorem 1.8 or, more precisely, the k-broad version given by Corollary 6.5) playing the role of $\bar{p}(n,k)$. □

Remark 11.3. From the above, the narrow range of exponents in Theorem 1.1 compared with Theorem 1.2 can be broadly attributed to:

(1) The weaker k-broad estimates coming from the Bennett–Carbery–Tao theorem compared with Theorem 1.9. One cannot work with stronger k-broad estimates than those given by Corollary 6.5, due to the failure of transverse equidistribution in the mixed-signature case.

(2) The more stringent constraints on p in Proposition 11.2 compared with Proposition 11.1. These additional constraints arise due to the fact that hyperbolic paraboloids contain linear subspaces, as discussed below.

---

[31] The notation of a *reduced* phase under a general signature hypothesis has not been introduced but is almost identical to that used in the positive-definite case. Indeed, the only difference is that the first condition in (4.3) is suitably modified, with I_{n-1} replaced by a diagonal matrix with diagonal entries 1 and -1.

To establish the main result, Theorem 1.2, it remains to prove Proposition 11.1. Both Propositions 11.1 and 11.2 can be established using very similar arguments: in fact, the proofs differ only at one (crucial) point. To highlight the essential differences between the positive-definite and mixed-signature cases, at the end of this subsection it is indicated how one may adapt the proof of Proposition 11.1 to yield Proposition 11.2.

The proof of Proposition 11.1 is an induction on scales argument. The induction quantity is defined as follows.

Definition 11.4. For $1 \leqslant p \leqslant \infty$ and $R \geqslant 1$ let $Q_p(R)$ denote the infimum over all constants C for which the estimate

$$\|T^\lambda f\|_{L^p(B(0,r))} \leqslant C \|f\|_{L^p(B^{n-1})}$$

holds for $1 \leqslant r \leqslant R$ and all Hörmander-type operators T^λ with reduced positive-definite phase and all $\lambda \geqslant R$.

With this definition, the problem is now to show that, under the hypotheses of Proposition 11.1, one has

$$Q_p(R) \lesssim_\varepsilon R^\varepsilon \tag{11.3}$$

for all $\varepsilon > 0$ and $1 \leqslant R \leqslant \lambda$. Indeed, this establishes the linear estimates in the case of reduced phases, and then the arguments of §4 extend the result to general Hörmander-type operators with positive-definite phase.

It is useful to introduce some of the ingredients of the proof of (11.3). Decompose $B(0, R)$ into balls B_{K^2} of radius K^2 and consider $\|T^\lambda f\|_{L^p(B_{K^2})}$ for some fixed B_{K^2} with centre $\bar{x}$. To bound this quantity, one expresses f as a sum of two terms: a "narrow" and a "broad" term. The narrow term is of the form

$$\sum_{\tau \in V_a \text{ for some } a} f_\tau, \tag{11.4}$$

consisting of contributions to f from caps for which $G^\lambda(\bar{x}; \tau)$ makes a small angle with some member of a family of $(k-1)$-planes. The broad term consists of the contributions to f from all the remaining caps. One may choose the planes $V_1, ..., V_A$ so that the broad term can be bounded by the k-broad inequality from the hypothesis. Thus, the problem is roughly reduced to studying the case where f is of the form (11.4). To treat this case, the first step is to apply an ℓ^p-decoupling inequality to isolate the contributions of the different f_τ.

THEOREM 11.5. *Suppose that T^λ is a Hörmander-type operator with reduced positive-definite phase. If $V \subseteq \mathbb{R}^n$ is an m-dimensional linear subspace, then for $2 \leqslant p \leqslant 2m/(m-1)$*

and $\delta > 0$ *one has*

$$\left\| \sum_{\tau \in V} T^\lambda g_\tau \right\|_{L^p(B_{K^2})} \lesssim_\delta K^{(m-1)(1/2-1/p)+\delta} \left(\sum_{\tau \in V} \|T^\lambda g_\tau\|_{L^p(w_{B_{K^2}})}^p \right)^{1/p}.$$

Here, the sums are over all caps τ for which $\measuredangle(G^\lambda(\bar{x},\tau),V) \leqslant K^{-1}$, where $\bar{x}$ is the centre of B_{K^2}, and $w_{B_{K^2}}$ is a rapidly decaying weight of the form of that defined in (8.4).

This theorem is a variable coefficient generalisation of a decoupling inequality due to Bourgain [7]. It can be established by adapting the argument of [7] using many of the techniques employed in the current article: see also [2].([32])

Summing together the contributions from the various spatial balls B_{K^2}, it remains to estimate the decoupled contributions $\|T^\lambda f_\tau\|_{L^p(B_R)}$. Since each f_τ has small support, after rescaling one obtains favourable estimates for $\|T^\lambda f_\tau\|_{L^p(B(0,R))}$ by invoking the induction hypothesis. This is made precise by the following lemma.

LEMMA 11.6. (Parabolic rescaling) *Let $1 \leqslant R \leqslant \lambda$, and suppose that f is supported on a ball of radius ϱ^{-1}, where $1 \leqslant \varrho \leqslant R$. For all $p \geqslant 2$ and $\delta > 0$, one has*

$$\|T^\lambda f\|_{L^p(B(0,R))} \lesssim_\delta Q_p(R) R^\delta \varrho^{2n/p-(n-1)} \|f\|_{L^p(B^{n-1})}.$$

The proof of the parabolic rescaling lemma is based on the changes of variables previously encountered in §4.2. For extension operators the argument is simple, consisting of an affine change of variables. In the variable coefficient case, some significant additional complications arise; the details are therefore postponed until the following subsection.

Having introduced the main tools, the proof of Proposition 11.1 easily follows.

Proof of Proposition 11.1. It suffices to demonstrate the linear estimate for p satisfying the additional condition

$$2\frac{2n-k+2}{2n-k} < p; \tag{11.5}$$

the result for the remaining value of p then follows immediately by Hölder's inequality.

Let $\varepsilon > 0$ be given. By hypothesis,

$$\sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap B(0,R) \neq \varnothing}} \min_{V_1,\ldots,V_A} \max_{\tau \notin V_a} \int_{B_{K^2}} |T^\lambda f_\tau|^p \leqslant C(K,\varepsilon) R^{p\varepsilon/2} \|f\|_{L^p(B^{n-1})}^p,$$

([32]) It is remarked that, since the decoupling estimate is applied at a small spatial scale $K^2 \ll \lambda^{1/2}$, one can avoid the use of the full statement of Theorem 11.5 by appealing to an approximation argument. If one argues in this way, then only Theorem 11.5 for extension operators associated with elliptic-type hypersurfaces is required.

where $V_1, ..., V_A$ are $(k-1)$-planes and the notation $\tau \notin V_a$ signifies that

$$\angle(G^\lambda(\bar{x}, \tau), V_a) > K^{-1},$$

with $\bar{x}$ being the centre of the corresponding K^2-ball B_{K^2}.

For each B_{K^2} fix a choice of $V_1, ..., V_A$ which achieves the minimum above. Then one may write

$$\int_{B_{K^2}} |T^\lambda f|^p \lesssim K^{O(1)} \max_{\tau \notin V_a} \int_{B_{K^2}} |T^\lambda f_\tau|^p + \sum_{a=1}^A \int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p.$$

The first term can be estimated using the hypothesised k-broad estimate; in particular,

$$\int_{B(0,R)} |T^\lambda f|^p \lesssim K^{O(1)} C(K, \varepsilon) R^{p\varepsilon/2} \|f\|_{L^p(B^{n-1})}^p + \sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap B(0,R) \neq \varnothing}} \sum_{a=1}^A \int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p.$$

It remains to bound the narrow term, where the contributions come from caps whose directions make a small angle with one of planes V_a. By Theorem 11.5, for any $\delta' > 0$ one has

$$\int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p \lesssim_{\delta'} K^{(k-2)(p/2-1)+\delta'} \sum_{\tau \in V_a} \int_{\mathbb{R}^n} |T^\lambda f_\tau|^p w_{B_{K^2}}$$

for each $1 \leqslant a \leqslant A$. Thus, summing over the a and all the relevant balls B_{K^2}, one concludes that

$$\sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap B(0,R) \neq \varnothing}} \sum_{a=1}^A \int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p \lesssim_{\delta'} K^{(k-2)(p/2-1)+\delta'} \sum_{\tau: K^{-1}\text{-cap}} \int_{B(0,2R)} |T^\lambda f_\tau|^p.$$

Since each f_τ is supported on a K^{-1}-cap, the summands appearing in the right-hand expression are amenable to parabolic rescaling. In particular, letting $\delta > 0$ be a small number chosen to satisfy the requirements of the forthcoming argument, Lemma 11.6 implies that

$$\int_{B(0,2R)} |T^\lambda f_\tau|^p \lesssim_\delta Q_p(R)^p R^\delta K^{2n-(n-1)p} \|f_\tau\|_{L^p(B^{n-1})}^p.$$

Defining

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

and combining these estimates,

$$\int_{B(0,R)} |T^\lambda f|^p \leqslant (K^{O(1)} C(K, \varepsilon) R^{p\varepsilon/2} + C_{\delta,\delta'} Q_p(R)^p R^\delta K^{-e(k,p)+\delta'}), \|f\|_{L^p(B^{n-1})}^p$$

and so, by definition,

$$Q_p(R)^p \leqslant K^{O(1)}C(K,\varepsilon)R^{p\varepsilon/2} + C_{\delta,\delta'}Q_p(R)^p R^\delta K^{-e(k,p)+\delta'}.$$

Since p satisfies (11.5), it follows that $e(k,p) > 0$, and one may choose $\delta' = \frac{1}{2}e(k,p)$, so that the K exponent in the right-hand term is negative. Thus, if $K = K_0 R^{2\delta/e(k,p)}$ for a sufficiently large constant K_0, depending only on ε, δ, p and n, it follows that

$$Q_p(R)^p \leqslant K^{O(1)}C(K_0 R^{2\delta/e(k,p)}, \varepsilon)R^{p\varepsilon/2} + \tfrac{1}{2}Q_p(R)^p.$$

Recall that, by hypothesis, the constant $C(K,\varepsilon)$ arising from the k-broad estimate grows at most polynomially in K. Consequently, one may choose δ to be small enough (depending only on admissible parameters) so that $Q_p(R) \lesssim_\varepsilon R^\varepsilon$, as required. $\square$

As mentioned above, this argument can be adapted to study the case of general Hörmander-type operators (with potentially mixed signature) to prove Proposition 11.2. The induction quantity $Q_p(R)$ is defined as before, but now the supremum is taken over the larger class of all Hörmander-type operators T^λ which are in a suitably reduced form. The proof of the parabolic rescaling lemma then extends to this setting, mutatis mutandis. The key differences arise in the decoupling inequality. In particular, Theorem 11.5 does not hold at the required level of generality. To see why this is so, consider the example of the extension operator E associated with (a compact piece of) the hyperbolic paraboloid given by the graph of $h(\omega) := \omega_1\omega_2$. If $V := \{x \in \mathbb{R}^3 : x_1 = 0\}$ and G is the relevant Gauss map, then

$$S_\omega := \{\omega \in B^2 : G(\omega) \in V\} = \{\omega \in B^2 : \omega_2 = 0\}.$$

Thus, the Eg_τ for $\tau \in V$ are (distributionally) Fourier supported in a neighbourhood of the ξ_1-axis (which is, in particular, a curve of everywhere zero curvature); here the notation $\tau \in V$ is used to denote that $\angle(G(\tau), V) \leqslant K^{-1}$, consistent with the non-standard notion of containment used in §1.5 and Theorem 11.5. As is well known, in the absence of curvature, no non-trivial decoupling estimates are possible.[33]

The following simple result acts as a substitute for Theorem 11.5.

LEMMA 11.7. (Bourgain–Guth [9]) *Suppose that T^λ is a Hörmander-type operator with reduced phase. If $V \subseteq \mathbb{R}^n$ is an m-dimensional linear subspace, then for all $p \geqslant 2$ and $\delta > 0$ one has*

$$\left\| \sum_{\tau \in V} T^\lambda g_\tau \right\|_{L^p(B_{K^2})} \lesssim_\delta K^{(m-1)(1-2/p)+\delta} \left(\sum_{\tau \in V} \|T^\lambda g_\tau\|^p_{L^p(w_{B_{K^2}})} \right)^{1/p}.$$

---

[33] It is remarked that non-trivial ℓ^p-decoupling estimates are known to hold for the *full* hyperbolic paraboloid; see [8]. The problem here arises because one is forced to consider decoupling along the lower-dimensional submanifold $S_\omega \times \{0\}$.

Here, the sums are over all caps τ for which $\measuredangle(G^\lambda(\bar{x}, \tau), V) \leqslant K^{-1}$, where $\bar{x}$ is the centre of B_{K^2}.

This lemma provides much weaker estimates than those guaranteed by the ℓ^p-decoupling theorem in the positive-definite case: here, the K exponent is larger by a factor of 2 than that appearing in Theorem 11.5. The proof is implicitly contained in the proof of Theorem 4 in [9, §5].

To prove Proposition 11.2, one proceeds as in the proof of Proposition 11.1, first decomposing $B(0, R)$ into balls of radius K^2. For each such ball, the broad term is bounded using the hypothesised k-broad estimate, whilst the narrow term is bounded by Lemma 11.7 together with the induction hypothesis (via parabolic rescaling). The larger exponent incurred by Lemma 11.7 propagates through the argument until one arrives at the estimate

$$Q_p(R)^p \leqslant K^{O(1)} C(K, \varepsilon) R^{p\varepsilon/2} + C_{\delta, \delta'} Q_p(R)^p R^\delta K^{(k-2)(p-2)+2n-(n-1)p+\delta'}.$$

In order to close the induction, once again one must ensure that the K exponent is negative. By choosing δ' appropriately, this is possible if p satisfies the stronger hypothesis $p > 2(n-k+2)/(n-k+1)$, which is precisely the condition featured in the statement of Proposition 11.2.

11.2. Proof of Lemma 11.6

It remains to establish the parabolic rescaling lemma, which is achieved by adapting arguments implicit in [9, §5]. As mentioned in the previous section, some additional complications arise in the case of Hörmander operators (as opposed to the extension case), and the proof of the parabolic rescaling is slightly involved.

It will be useful to work with the following discrete reformulation of the main estimate for the operator T^λ.

LEMMA 11.8. *If $\mathcal{D}$ is a maximal R^{-1}-separated discrete subset of Ω, then*

$$\left\| \sum_{\omega_\theta \in \mathcal{D}} e^{2\pi i \phi^\lambda(\cdot; \omega_\theta)} F(\omega_\theta) \right\|_{L^p(B(0,R))} \lesssim Q_p(R) R^{(n-1)/p'} \|F\|_{\ell^p(\mathcal{D})} \qquad (11.6)$$

for all $F: \mathcal{D} \to \mathbb{C}$.

Proof. Fix $\psi \in C_c^\infty(\mathbb{R}^{n-1})$ supported on $B(0, 2)$ which satisfies $0 \leqslant \psi \leqslant 1$ and $\psi(\omega) = 1$ for all $\omega \in B^{n-1}$, and for each $\omega_\theta \in \mathcal{D}$ define $\psi_\theta(\omega) := \psi(10R(\omega - \omega_\theta))$. Thus, for all $x \in$

$B(0, R)$, the exponential sum appearing in the left-hand side of (11.6) can be expressed as a constant multiple of

$$R^{n-1} \int_{\mathbb{R}^{n-1}} e^{2\pi i \phi^\lambda(x;\omega)} a^\lambda(x;\omega) \left[\tilde{\psi}\left(\frac{x}{R}\right) \sum_{\omega_\theta \in \mathcal{D}} e^{-2\pi i \lambda \Omega_\theta(x/\lambda;\omega)} F(\omega_\theta) \psi_\theta(\omega) \right] d\omega,$$

where $\Omega_\theta(x;\omega) := \phi(x;\omega) - \phi(x;\omega_\theta)$, $\tilde{\psi} \in C_c^\infty(\mathbb{R}^n)$ is a function of n variables which enjoys properties similar to those of ψ and a^λ is a suitable choice of amplitude. Since

$$\sup_{\omega \in \text{supp } \psi_\theta} |\partial_x^\beta \Omega_\theta(x;\omega)| \lesssim_\beta R^{-1}|x| \quad \text{for all } \beta \in \mathbb{N}_0^n \text{ and } x \in X,$$

one may safely remove the $\lambda \Omega_\theta(x/\lambda;\omega)$ term from the phase. More precisely, by expanding $\tilde{\psi}(x) e^{-2\pi i \lambda \Omega_\theta(Rx/\lambda;\omega)}$ as a Fourier series in the variable x, one can show that

$$\left| \sum_{\omega_\theta \in \mathcal{D}} e^{2\pi i \phi^\lambda(x;\omega_\theta)} F(\omega_\theta) \right| \lesssim R^{n-1} \sum_{k \in \mathbb{Z}^n} (1+|k|)^{-(n+1)} |T^\lambda f_k(x)|,$$

where T^λ is a Hörmander-type operator with phase ϕ^λ and

$$f_k(\omega) := \sum_{\omega_\theta \in \mathcal{D}} F(\omega_\theta) c_{k,\theta}(\omega) \psi_\theta(\omega)$$

for some choice of smooth functions $c_{k,\theta}$ satisfying the uniform bound $\|c_{k,\theta}\|_{L^\infty(B^{n-1})} \lesssim 1$. Thus, by the definition of $Q_p(R)$, it follows that

$$\left\| \sum_{\omega_\theta \in \mathcal{D}} e^{2\pi i \phi^\lambda(\cdot;\omega_\theta)} F(\omega_\theta) \right\|_{L^p(B(0,R))} \lesssim Q_p(R) R^{n-1} \sum_{k \in \mathbb{Z}^n} (1+|k|)^{-(n+1)} \|f_k\|_{L^p(B^{n-1})}$$

and, since the ψ_θ are supported on pairwise disjoint sets,

$$\|f_k\|_{L^p(B^{n-1})} \lesssim R^{-(n-1)/p} \left(\sum_{\omega_\theta \in \mathcal{D}} |F(\omega_\theta)|^p \right)^{1/p},$$

concluding the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Lemma 11.6. Recall that the phase of T^λ is given by

$$\phi^\lambda(x;\omega) := \lambda \phi\left(\frac{x}{\lambda};\omega\right),$$

where

$$\phi(x;\omega) = \langle x', \omega \rangle + x_n h(\omega) + \mathcal{E}(x;\omega). \tag{11.7}$$

Let $B(\bar{\omega}, \varrho^{-1})$ be a ball supporting f, where $\bar{\omega} \in B^{n-1}$. If $\widetilde{T}^{\lambda/\varrho^2}$ denotes the parabolically rescaled operator defined in (4.9), with rescaled phase function

$$\tilde{\phi}(x; \omega) := \langle x', \omega \rangle + x_n \tilde{h}(\omega) + \tilde{\mathcal{E}}(x; \omega), \tag{11.8}$$

then it follows that

$$\|T^\lambda f\|_{L^p(B(0,R))} \lesssim \varrho^{(n+1)/p} \|\widetilde{T}^{\lambda/\varrho^2} \tilde{f}\|_{L^p(\tilde{D}_R)}$$

where now $\tilde{D}_R$ is an ellipse with principal axes parallel to the coordinate axes and dimensions $O(R/\varrho) \times ... \times O(R/\varrho) \times O(R/\varrho^2)$ and $\tilde{f}(\omega) := \varrho^{-(n-1)} f(\bar{\omega} + \varrho^{-1}\omega)$. Since

$$\|\tilde{f}\|_{L^p(B^{n-1})} = \varrho^{-(n-1)+(n-1)/p} \|f\|_{L^p(B^{n-1})},$$

given $\delta > 0$, the problem is to show that

$$\|\widetilde{T}^{\lambda/\varrho^2} \tilde{f}\|_{L^p(\tilde{D}_R)} \lesssim_\delta Q_p(R) R^\delta \|\tilde{f}\|_{L^p(B^{n-1})}.$$

Observe that the phase $\tilde{\phi}$ defined in (11.8) is also positive-definite and of reduced form. To lighten the notation, consider once again a general positive-definite reduced phase ϕ as in (11.7), and let T^λ is a Hörmander-type operator associated with ϕ^λ. It suffices to show that

$$\|T^\lambda f\|_{L^p(D_{\mathbf{R}})} \lesssim_\delta Q_p(R) R^\delta \|f\|_{L^p(B^{n-1})}$$

for all $1 \ll R \leqslant R' \leqslant \lambda$ and $\delta > 0$, where

$$D_{\mathbf{R}} := \left\{ x \in \mathbb{R}^n : \left(\frac{|x'|}{R'}\right)^2 + \left(\frac{|x_n|}{R}\right)^2 \leqslant 1 \right\}$$

is an ellipse. Of course, if $R = R'$, then this inequality is immediate from the definition of $Q_p(R)$.

Cover B^{n-1} by a collection of essentially disjoint R^{-1}-caps θ, and decompose f as $f = \sum_\theta f_\theta$. Define

$$T_\theta^\lambda f(x) := e^{-2\pi i \phi^\lambda(x; \omega_\theta)} T^\lambda f(x),$$

so that

$$T^\lambda f(x) = \sum_{\theta: R^{-1}\text{-cap}} e^{2\pi i \phi^\lambda(x; \omega_\theta)} T_\theta^\lambda f_\theta(x).$$

Fix $\delta > 0$ to be sufficiently small for the purposes of the forthcoming argument. Each f_θ is supported on an R^{-1}-ball and is therefore, of course, supported on an $R^{-1+\delta}$-ball. Since $(R^{-1+\delta})^{-1} \leqslant \lambda^{1-\delta}$, one may argue as in the proof of Lemma 5.8 to deduce that

$$T_\theta^\lambda f_\theta(x) = T_\theta^\lambda f_\theta * \eta_{R^{1-\delta}}(x) + \text{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}$$

for some choice of smooth, rapidly decreasing function η such that $|\eta|$ admits a smooth, rapidly decreasing majorant $\zeta \colon \mathbb{R}^n \to [0,\infty)$ which is locally constant at scale 1. In particular, it follows that

$$\zeta_{R^{1-\delta}}(x) \lesssim R^\delta \zeta_{R^{1-\delta}}(y) \quad \text{if } |x-y| \lesssim R. \tag{11.9}$$

Cover $D_{\mathbf{R}}$ by finitely-overlapping R-balls, and let B_R be some member of this cover. Combining the above observations, if $\bar{x}$ denotes the centre of B_R and $z \in B(0,R)$, then

$$|T^\lambda f(\bar{x}+z)| \lesssim R^\delta \int_{\mathbb{R}^n} \bigg| \sum_{\theta : R^{-1}\text{-cap}} e^{2\pi i \tilde{\phi}^\lambda(z;\omega_\theta)} e^{2\pi i \phi^\lambda(\bar{x};\omega_\theta)} T^\lambda_\theta f_\theta(y) \bigg| \zeta_{R^{1-\delta}}(\bar{x}-y)\,dy,$$

where

$$\tilde{\phi}^\lambda(z;\omega_\theta) := \phi^\lambda(\bar{x}+z;\omega_\theta) - \phi^\lambda(\bar{x};\omega_\theta).$$

Taking the L^p-norm in z, it follows from Minkowski's inequality that $\|T^\lambda f\|_{L^p(B_R)}$ is dominated by

$$R^\delta \int_{\mathbb{R}^n} \bigg\| \sum_{\theta : R^{-1}\text{-cap}} e^{2\pi i \tilde{\phi}^\lambda(\,\cdot\,;\omega_\theta)} e^{2\pi i \phi^\lambda(\bar{x};\omega_\theta)} T^\lambda_\theta f_\theta(y) \bigg\|_{L^p(B(0,R))} \zeta_{R^{1-\delta}}(\bar{x}-y)\,dy.$$

By Lemma 11.8, the L^p-norm appearing in the above integrand is bounded by a constant multiple of

$$Q_p(R) R^{(n-1)/p'} \bigg(\sum_{\theta : R^{-1}\text{-cap}} |T^\lambda_\theta f_\theta(y)|^p \bigg)^{1/p}.$$

Applying Hölder's inequality and the locally-constant property (11.9), one deduces that

$$\|T^\lambda f\|_{L^p(B_R)} \lesssim Q_p(R) R^{(n-1)/p'+O(\delta)} \bigg(\int_{\mathbb{R}^n} \sum_{\theta : R^{-1}\text{-cap}} |T^\lambda f_\theta(\bar{x}+z-y)|^p \zeta_{R^{1-\delta}}(y)\,dy \bigg)^{1/p}$$

for all $z \in B(0,R)$. By raising both sides of this estimate to the pth power, averaging in z and summing over all balls B_R in the covering, it follows that $\|T^\lambda f\|_{L^p(D_{\mathbf{R}})}$ is dominated by

$$Q_p(R) R^{(n-1)/p'-n/p+O(\delta)} \bigg(\int_{\mathbb{R}^n} \sum_{\theta : R^{-1}\text{-cap}} \|T^\lambda f_\theta\|^p_{L^p(D_{\mathbf{R}}-y)} \zeta_{R^{1-\delta}}(y)\,dy \bigg)^{1/p}.$$

Observe that, by Hörmander's theorem (Lemma 5.6) and Hölder's inequality, one has

$$\|T^\lambda f_\theta\|_{L^2(D_{\mathbf{R}}-y)} \lesssim R^{-(n-1)(1/2-1/p)+1/2} \|f_\theta\|_{L^p(B^{n-1})}.$$

On the other hand, the trivial estimate

$$\|T^\lambda f_\theta\|_{L^\infty(D_{\mathbf{R}}-y)} \lesssim R^{-(n-1)/p'} \|f_\theta\|_{L^p(B^{n-1})}$$

holds, simply due to Hölder's inequality. Combining the above,

$$\|T^\lambda f_\theta\|_{L^p(D_{\mathbf{R}}-y)} \lesssim R^{-(n-1)/p'+n/p} \|f_\theta\|_{L^p(B^{n-1})}.$$

The desired inequality is now immediate. $\square$

12. An ε-removal lemma

The λ^ε-loss in the linear estimates of Theorems 1.1 and 1.2 can be removed away from the endpoint by an appeal to an ε-removal lemma of the type introduced in [31] (see also [9], [30]). The precise form of the required lemma does not appear in the literature, but it can be deduced by a minor modification of an argument from [31]. For completeness, the details are given presently.

Suppose that T^λ is a Hörmander-type operator with associated phase function ϕ^λ (note that here no additional positive-definite assumption is assumed). Let $\bar{p} \geqslant 2$ and suppose for all $\varepsilon > 0$ the estimate

$$\|T^\lambda f\|_{L^p(B_R)} \lesssim_{\varepsilon,\phi,a} R^\varepsilon \|f\|_{L^p(B^{n-1})} \tag{12.1}$$

holds for all $p \geqslant \bar{p}$, all R-balls B_R for $1 \leqslant R \leqslant \lambda$, and any choice of amplitude function. Under this hypothesis, one wishes to show that the global estimate

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_{\phi,a} \|f\|_{L^p(B^{n-1})} \tag{12.2}$$

is valid for all $p > \bar{p}$.

Definition 12.1. (Tao [31]) Let $R \geqslant 1$. A collection $\{B(x_j, R)\}_{j=1}^N$ of R-balls in $\mathbb{R}^n$ is *sparse* if the centres $\{x_1, ..., x_N\}$ are $(RN)^{\bar{C}}$-separated. Here $\bar{C} \geqslant 1$ is a fixed constant, chosen large enough for the purposes of the proof.

Following [31], the first step towards establishing (12.2) is to reduce the problem to proving estimates for T^λ over sparse families of balls.

LEMMA 12.2. *To prove* (12.2) *for all* $p > \bar{p}$, *it suffices to show that for all* $\varepsilon > 0$ *the estimate*

$$\|T^\lambda f\|_{L^{\bar{p}}(S)} \lesssim_{\varepsilon,\phi,a} R^\varepsilon \|f\|_{L^{\bar{p}}(B^{n-1})} \tag{12.3}$$

holds whenever $R \geqslant 1$ *and* $S \subseteq \mathbb{R}^n$ *is a union of* R-*balls belonging to a sparse collection, for any choice of amplitude function.*

The key step in the proof of Lemma 12.2 is the following covering lemma.

LEMMA 12.3. (Tao [30], [31]) *Suppose that* $E \subseteq \mathbb{R}^n$ *is a finite union of* 1-*cubes and* $N \geqslant 1$. *Define the radii* R_j *inductively by*

$$R_0 := 1 \quad and \quad R_j := R_{j-1}^{\bar{C}} |E|^{\bar{C}} \ \ for \ 1 \leqslant j \leqslant N-1.$$

Then, for each $0 \leqslant j \leqslant N-1$, *there exists a family of sparse collections* $(\mathcal{B}_{j,\alpha})_{\alpha \in A_j}$ *of balls of radius* R_j *such that the index sets* A_k *have cardinality* $O(|E|^{1/N})$ *and*

$$E \subseteq \bigcup_{j=0}^{N-1} \bigcup_{\alpha \in A_j} S_{j,\alpha},$$

where $S_{j,\alpha}$ *is the union of all the balls belonging to the family* $\mathcal{B}_{j,\alpha}$.

Proof of Lemma 12.2. Let $E \subseteq \mathbb{R}^n$ be a finite union of 1-cubes. For $N \geqslant 1$, the covering lemma together with the hypothesis (12.3) imply that

$$\|T^\lambda f\|_{L^{\bar{p}}(E)} \lesssim_{\varepsilon,\phi,a} N|E|^{1/N+\varepsilon\bar{C}^N} \|f\|_{L^{\bar{p}}(B^{n-1})}.$$

Choosing $N \sim \log(1/\varepsilon)$, it follows that

$$\|T^\lambda f\|_{L^{\bar{p}}(E)} \lesssim_{\varepsilon,\phi,a} |E|^{\bar{C}/\log(1/\varepsilon)} \|f\|_{L^{\bar{p}}(B^{n-1})}.$$

It will be convenient to work with the dual operator

$$T^* g(\omega) := \int_{\mathbb{R}^n} e^{-2\pi i \phi^\lambda(x;\omega)} a^\lambda(x;\omega) g(x)\, dx,$$

so that the above estimate can be reformulated as

$$\|T^* g\|_{L^{\bar{p}'}(B^{n-1})} \lesssim_{\varepsilon,\phi,a} |E|^{\bar{C}/\log(1/\varepsilon)} \|g\|_{L^{\bar{p}'}(E)}, \tag{12.4}$$

for g supported on the set E.

Fix $p > \bar{p}$ and $\tau \in \left[-\frac{1}{2}, \frac{1}{2}\right]^n$. Suppose that $g \in L^{p'}(\mathbb{R}^n)$ satisfies $\|g\|_{L^{p'}(\mathbb{R}^n)} = 1$ and is constant on the mesh of 1-cubes centred on points of the lattice $\tau + \mathbb{Z}^n$. Form a level set decomposition of g by writing $g = \sum_{k \in \mathbb{Z}} g_k$, where $g_k := g\chi_{E_k}$ for

$$E_k := \{x \in \mathbb{R}^n : 2^{-k} \leqslant |g(x)| < 2^{-k+1}\}.$$

Chebyshev's inequality implies that $|E_k| \leqslant 2^{kp'}$ for all $k \in \mathbb{Z}$. Furthermore, each set E_k is a union of 1-cubes and therefore, if $E_k \neq \varnothing$, then $|E_k| \geqslant 1$. Combining these observations, one deduces that $E_k = \varnothing$ for all $k < 0$. Since g_k is supported on E_k, one may apply (12.4) to conclude that

$$\|T^* g_k\|_{L^{\bar{p}'}(B^{n-1})} \lesssim_{\varepsilon,\phi,a} |E_k|^{\bar{C}/\log(1/\varepsilon)} \|g_k\|_{L^{\bar{p}'}(\mathbb{R}^n)}. \tag{12.5}$$

Using a simple base-times-height estimate, the right-hand side of (12.5) can be bounded by (a constant multiple of)

$$2^{-k}|E_k|^{\bar{C}/\log(1/\varepsilon)+1/\bar{p}'} \lesssim 2^{-k(1-\bar{C}p'/\log(1/\varepsilon)-p'/\bar{p}')}.$$

Since $p' < \bar{p}'$, by choosing ε sufficiently small, one can ensure that the right-hand exponent is negative, and therefore

$$\|T^* g\|_{L^{p'}(B^{n-1})} \lesssim \|T^* g\|_{L^{\bar{p}'}(B^{n-1})} \leqslant \sum_{k \geqslant 0} \|T^* g_k\|_{L^{\bar{p}'}(B^{n-1})} \lesssim_{\phi,a} 1 = \|g\|_{L^{p'}(\mathbb{R}^n)}.$$

This establishes the dual of the desired estimate (12.2) under the additional hypothesis that the function g is constant on 1-cubes.

It remains to remove the condition that g is constant on 1-cubes. The key observation is that this special case of (12.2) implies the discrete inequality

$$\left\| \sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma+\tau; \cdot)} a^\lambda(\sigma+\tau; \cdot) G(\sigma) \right\|_{L^{p'}(B^{n-1})} \lesssim_{\phi,a} \|G\|_{\ell^{p'}(\mathbb{Z}^n)} \tag{12.6}$$

for all $G \in \ell^{p'}(\mathbb{Z}^n)$ and $\tau \in \left[-\frac{1}{2}, \frac{1}{2}\right]^n$. Indeed, once (12.6) is established, taking $g \in L^{p'}(\mathbb{R}^n)$ belonging to a suitable a-priori class and applying Minkowski's inequality, one deduces that

$$\|T^* g\|_{L^{p'}(B^{n-1})} \leqslant \int_{[-1/2,1/2]^n} \left\| \sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma+\tau; \cdot)} a^\lambda(\sigma+\tau; \cdot) g(\sigma+\tau) \right\|_{L^{p'}(B^{n-1})} d\tau.$$

Combining this with (12.6) and Hölder's inequality yields (12.4).

Thus, the problem is now reduced to proving (12.6). Fix $G \in \ell^{p'}(\mathbb{Z}^n)$ and define

$$\tilde{g}(x) := \sum_{\sigma \in \mathbb{Z}^n} G(\sigma) \chi(x - \sigma - \tau),$$

where χ is the characteristic function of $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$. Since $\tilde{g}$ is constant on 1-cubes, one is free to apply (12.4) to this function. In particular, let $\widetilde{T}^*$ be the dual of a Hörmander-type operator with phase ϕ^λ and amplitude $\tilde{a}^\lambda$, where

$$\tilde{a}(x; \omega) := \left(\prod_{j=1}^n \frac{\sin \pi (\partial_{x_j} \phi)(x; \omega)}{\pi (\partial_{x_j} \phi)(x; \omega)} \right)^{-1} a_0(x; \omega),$$

for a_0 a smooth amplitude which is supported on $X \times \Omega$ and satisfies $a_0(x; \omega) = 1$ for $(x; \omega) \in \operatorname{supp} a$. By the usual reductions (see §4), one may assume from the outset that $|(\partial_{x_j} \phi)(x; \omega)| \leqslant \frac{1}{2}$ for $(x; \omega) \in X \times \Omega$ and $1 \leqslant j \leqslant n$, and hence $\tilde{a}$ is a well-defined, smooth function. Thus, the estimate

$$\|\widetilde{T}^* \tilde{g}\|_{L^{p'}(B^{n-1})} \lesssim_{\phi,a} \|\tilde{g}\|_{L^{p'}(\mathbb{R}^n)}$$

holds, which can be rewritten as

$$\left\| \sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma+\tau; \cdot)} (A_\lambda)^\lambda(\sigma+\tau; \cdot) G(\sigma) \right\|_{L^{p'}(B^{n-1})} \lesssim_{\phi,a} \|G\|_{\ell^{p'}(\mathbb{Z}^n)}, \tag{12.7}$$

where

$$A_\lambda(x; \omega) := \int_{[-1/2,1/2]^n} e^{-2\pi i \lambda(\phi(x+y/\lambda; \omega) - \phi(x; \omega))} \tilde{a}\left(x + \frac{y}{\lambda}; \omega\right) dy.$$

Note that (12.7) is almost the desired expression (12.6), except for the disparity between the amplitude functions. To deal with this slight technicality, observe that, since

$$\lim_{\lambda \to \infty} A_\lambda(x; \omega) = a_0(x; \omega) \quad \text{uniformly,}$$

one may assume that λ is sufficiently large so that $|A_\lambda(x; \omega)| \gtrsim 1$ for all $(x; \omega) \in \operatorname{supp} a$. Furthermore, by applying the mean value theorem to the phase,

$$\|\partial_x^\alpha A_\lambda\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})} \lesssim_{\alpha, \phi, a} 1 \quad \text{for all } \alpha \in \mathbb{N}_0^n,$$

the important observation here being that the derivatives are independent of λ. Thus, the expression appearing in the norm on the left-hand side of (12.6) is given by

$$\sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma+\tau; \omega)} (A_\lambda)^\lambda(\sigma+\tau; \omega) G(\sigma)(\varrho_\lambda)^\lambda(\sigma+\tau; \omega),$$

where the ratio $\varrho_\lambda(x; \omega) := a(x; \omega) A_\lambda(x; \omega)^{-1}$ satisfies

$$\|\partial_x^\alpha \varrho_\lambda\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})} \lesssim_{\alpha, \phi, a} 1 \quad \text{for all } \alpha \in \mathbb{N}_0^n.$$

Taking a Fourier series expansion of ϱ_λ in the x variable and using repeated integration-by-parts to estimate the Fourier coefficients, it follows that

$$\varrho_\lambda(x; \omega) = \sum_{k \in \mathbb{Z}^n} (1+|k|)^{-(n+1)} c_{\lambda, k}(\omega) e^{2\pi i \langle x, k \rangle},$$

where the $c_{\lambda, k}$ are bounded functions, uniformly in λ and k (they do, however, depend on n, ϕ and a). One may therefore bound the left-hand side of (12.6) by a $(1+|k|)^{-(n+1)}$-weighted sum of the left-hand side of (12.7) applied to modulated versions of G. Estimating each summand using (12.7) and summing in k concludes the proof. $\qquad \square$

Given the above reduction, it remains to establish the estimates for T^λ over sparse collections of R-balls.

LEMMA 12.4. *Under the hypothesis* (12.1), *if* $p \geqslant \bar{p}$, *then the estimate*

$$\|T^\lambda f\|_{L^p(S)} \lesssim_{\varepsilon, \phi, a} R^\varepsilon \|f\|_{L^p(B^{n-1})}$$

holds for all $\varepsilon > 0$ *whenever* $S \subseteq \mathbb{R}^n$ *is a union of* R-balls belonging to a sparse collection.

Proof. The proof uses a crude form of wave packet analysis and has much in common with the arguments described in §5. Let $\{B(x_j, R)\}_{j=1}^N$ be the sparse collection of balls whose union is the set S. Clearly it suffices to assume that $R \ll \lambda$ and that all the $B(x_k, R)$

intersect the x-support of a^λ. Furthermore, letting $c_{\mathrm{diam}} > 0$ be a small constant chosen to satisfy the requirements of the forthcoming argument, by applying a partition of unity one may assume that $\operatorname{diam} X < c_{\mathrm{diam}}$, and so

$$\frac{|x_{j_1} - x_{j_2}|}{\lambda} \lesssim c_{\mathrm{diam}} \quad \text{for all } 1 \leqslant j_1, j_2 \leqslant N. \tag{12.8}$$

Fix $\eta \in C^\infty(\mathbb{R}^{n-1})$ satisfying $0 \leqslant \eta \leqslant 1$, $\operatorname{supp} \eta \in B^{n-1}$ and $\eta(z) = 1$ for all $z \in B\left(0, \frac{1}{2}\right)$. For $R_1 := CNR$, where $C \geqslant 1$ is a large constant, define $\eta_{R_1}(z) := \eta(z/R_1)$. Further, let $\psi \in C_c^\infty(\mathbb{R}^{n-1})$ satisfy $0 \leqslant \psi \leqslant 1$, $\operatorname{supp} \psi \subset \Omega$ and $\psi(\omega) = 1$ for ω belonging to the ω-support of a^λ. Fix $1 \leqslant j \leqslant N$ and write

$$e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f = P_j f + (e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f - P_j f) =: P_j f + f_{j,\infty},$$

where $P_j f := \hat{\eta}_{R_1} * [e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f]$. If one defines

$$\operatorname{Err}(x) := \int_{\mathbb{R}^{n-1}} e^{2\pi i (\phi^\lambda(x; \omega) - \phi^\lambda(x_j; \omega))} a^\lambda(x; \omega) f_{j,\infty}(\omega) \, d\omega,$$

then it follows that

$$T^\lambda f(x) = T^\lambda [e^{-2\pi i \phi^\lambda(x_j; \cdot)} P_j f](x) + \operatorname{Err}(x).$$

For $x \in B(x_j, R)$, the term $\operatorname{Err}(x)$ is negligible. Indeed, by Plancherel's theorem,

$$\operatorname{Err}(x) = \int_{\mathbb{R}^{n-1}} \overline{\breve{G}_x(z)} \cdot (1 - \eta_{R_1}(z)) [e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f]^\frown(z) \, dz,$$

where

$$\breve{G}_x(z) = \int_{\mathbb{R}^{n-1}} e^{2\pi i (\langle z, \omega \rangle - \phi^\lambda(x; \omega) + \phi^\lambda(x_j; \omega))} a^\lambda(x; \omega) \, d\omega.$$

Taking the ω-derivatives of the phase of $\breve{G}_x(z)$, one obtains

$$z - \lambda \left(\partial_\omega \phi\left(\frac{x}{\lambda}; \omega\right) - \partial_\omega \phi\left(\frac{x_j}{\lambda}; \omega\right) \right) = z + O(R),$$

$$-\lambda \left(\partial_\omega^\alpha \phi\left(\frac{x}{\lambda}; \omega\right) - \partial_\omega^\alpha \phi\left(\frac{x_j}{\lambda}; \omega\right) \right) = O(R) \qquad \text{for } |\alpha| \geqslant 2.$$

Thus, if z belongs to the support of $1 - \eta_{R_1}$, then integration-by-parts (see Lemma A.1) shows that $G_x(z)$ is rapidly decaying in R_1, and therefore

$$|\operatorname{Err}(x)| \leqslant \operatorname{RapDec}(R_1) \|f\|_{L^p(B^{n-1})}.$$

It remains to bound the contributions arising from the frequency localised pieces. By applying the estimate for T^λ with R^ε-loss over each ball $B(x_j, R)$, one obtains

$$\|T^\lambda f\|_{L^p(S)} \leqslant \bigg(\sum_{j=1}^N \|T^\lambda[e^{-2\pi i \phi^\lambda(x_j; \cdot)} P_j f]\|_{L^p(B(x_j, R))}^p \bigg)^{1/p} + \mathrm{RapDec}(R_1) \|f\|_{L^p(B^{n-1})}$$

$$\lesssim_{\varepsilon, \phi, a} R^\varepsilon \bigg(\sum_{j=1}^N \|P_j f\|_{L^p(B^{n-1})}^p \bigg)^{1/p} + \|f\|_{L^p(B^{n-1})}.$$

Thus, it now suffices to show that

$$\bigg(\sum_{j=1}^N \|P_j f\|_{L^p(\mathbb{R}^{n-1})}^p \bigg)^{1/p} \lesssim \|f\|_{L^p(B^{n-1})}.$$

This estimate follows via interpolation between the endpoint cases $p=2$ and $p=\infty$, which are established presently. The $p=\infty$ case is a trivial consequence of Young's inequality, and so it suffices to consider $p=2$. By duality, the desired inequality is equivalent to

$$\bigg\| \sum_{j=1}^N e^{-2\pi i \phi^\lambda(x_j; \cdot)} \psi \cdot [\hat\eta_{R_1} * g_j] \bigg\|_{L^2(\mathbb{R}^{n-1})} \lesssim \bigg(\sum_{j=1}^N \|g_j\|_{L^2(B^{n-1})}^2 \bigg)^{1/2}. \tag{12.9}$$

By squaring the left-hand side of (12.9), one obtains

$$\sum_{j_1, j_2=1}^N \int_{\mathbb{R}^{n-1}} \overline{G_{j_1, j_2}(\omega)} \hat\eta_{R_1} * g_{j_1}(\omega) \overline{\hat\eta_{R_1} * g_{j_2}(\omega)} \, d\omega,$$

where

$$G_{j_1, j_2}(\omega) := e^{2\pi i (\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega))} \psi(\omega)^2.$$

By Plancherel's theorem, each summand of the above expression can be written as

$$\int_{\mathbb{R}^{n-1}} \overline{\widecheck{G}_{j_1, j_2}(z)} (\eta_{R_1} \check g_{j_1}) * (\eta_{R_1} \check g_{j_2})^\sim(z) \, dz; \tag{12.10}$$

here, $(\eta_{R_1} \check g_{j_2})^\sim(z) := \overline{(\eta_{R_1} \check g_{j_2})(-z)}$. Note that the integrand in (12.10) is supported on a ball of radius $O(R_1)$ about the origin.

Fix $1 \leqslant j_1, j_2 \leqslant N$ with $j_1 \neq j_2$, let $z \in \mathbb{R}^{n-1}$ with $|z| \lesssim R_1 < |x_{j_2} - x_{j_1}|$, and consider

$$\widecheck{G}_{j_1, j_2}(z) = \int_{\mathbb{R}^{n-1}} e^{2\pi i (\langle z, \omega \rangle + \phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega))} \psi(\omega)^2 \, d\omega.$$

This oscillatory integral can be bounded by a simple stationary phase analysis. For $\alpha \in \mathbb{N}^{n-1}$ with $|\alpha| \leqslant 2$, consider the function

$$\partial_\omega^\alpha [\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega)] = \partial_\omega^\alpha \langle \partial_x^\lambda \phi(x_{j_1}; \omega), x_{j_2} - x_{j_1} \rangle + O(c_{\mathrm{diam}} |x_{j_2} - x_{j_1}|),$$

where the remainder term has been estimated using (12.8).

Let $c_{\mathrm{crit}} > 0$ be another small constant, chosen to satisfy the requirements of the forthcoming argument, and $\omega_0 \in \Omega$. Suppose that

$$\left| \pm \frac{x_{j_2} - x_{j_1}}{|x_{j_2} - x_{j_1}|} - G^\lambda(x_{j_1}; \omega_0) \right| \geqslant c_{\mathrm{crit}}, \tag{12.11}$$

where the estimate is interpreted as holding for both choices of sign. Condition (H1) on the phase implies that, for each $\omega_0 \in \Omega$, the vector $G^\lambda(x; \omega_0)$ spans the kernel of $\partial^2_{\omega x} \phi^\lambda(x; \omega_0)$. Consequently, in view of (12.11), one has

$$\left| \partial_\omega [\langle \partial_x \phi^\lambda(x_{j_1}; \omega), x_{j_2} - x_{j_1} \rangle]|_{\omega = \omega_0} \right| \gtrsim |x_{j_2} - x_{j_1}|,$$

and therefore

$$\left| \partial_\omega [\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega)]|_{\omega = \omega_0} \right| \gtrsim |x_{j_2} - x_{j_1}|,$$

provided c_{diam} is sufficiently small. On the other hand, if (12.11) fails, then

$$\partial^\alpha_\omega \left\langle \partial_x \phi^\lambda(x_{j_1}; \omega), \frac{x_{j_2} - x_{j_1}}{|x_{j_2} - x_{j_1}|} \right\rangle \bigg|_{\omega = \omega_0} = \partial^\alpha_\omega \langle \partial_x \phi^\lambda(x_{j_1}; \omega), G^\lambda(x; \omega_0) \rangle|_{\omega = \omega_0} + O(c_{\mathrm{crit}}).$$

If c_{crit} and c_{diam} are both chosen to be sufficiently small, then condition (H2) implies that

$$\left| \det \partial^2_{\omega\omega} [\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega)]|_{\omega = \omega_0} \right| \gtrsim |x_{j_2} - x_{j_1}|^{n-1}.$$

Thus, any critical point of the phase must be (quantitatively) non-degenerate, and one may apply higher-dimensional versions of van der Corput's lemma (see, for instance, [28, Chapter VIII, Proposition 6]) to estimate the oscillatory integral. In particular,

$$|\breve{G}_{j_1, j_2}(z)| \lesssim |x_{j_2} - x_{j_1}|^{-(n-1)/2} \lesssim R_1^{-\bar{C}/2},$$

so that the absolute value of (12.10) is bounded by

$$R_1^{-\bar{C}/2} \|(\eta_{R_1} \breve{g}_{j_1}) * (\eta_{R_1} \breve{g}_{j_2})^\sim\|_{L^1(\mathbb{R}^{n-1})} \lesssim R_1^{-\bar{C}/2} \prod_{i=1}^2 \|\eta_{R_1} \breve{g}_{j_i}\|_{L^1(\mathbb{R}^{n-1})}$$

$$\lesssim R_1^{-\bar{C}/2 + n - 1} \prod_{i=1}^2 \|g_{j_i}\|_{L^2(\mathbb{R}^{n-1})}.$$

Since there are only $O(N^2)$ choice of indices j_1 and j_2, one may invoke the trivial estimate

$$\prod_{i=1}^2 \|g_{j_i}\|_{L^2(\mathbb{R}^{n-1})} \lesssim \sum_{j=1}^N \|g_j\|^2_{L^2(B^{n-1})},$$

and then sum all the contributions from all pairs (j_1, j_2) to bound the off-diagonal terms arising from the left-hand side of (12.9). On the other hand, the diagonal terms provide a favourable contribution of

$$\left(\sum_{j=1}^{N} \|\hat{\eta}_{R_1} * g_j\|_{L^2(B^{n-1})}^2 \right)^{1/2} \lesssim \left(\sum_{j=1}^{N} \|g_j\|_{L^2(B^{n-1})}^2 \right)^{1/2}.$$

Combining these observations concludes the proof of (12.9), and thereby establishes the lemma. $\square$

Appendix A. The integration-by-parts argument

In this appendix, further details of the integration-by-parts argument frequently used in the paper are presented.

LEMMA A.1. *Let* $\phi \in C^\infty(\mathbb{R}^n)$ *be real valued and* $a \in C^\infty(\mathbb{R}^n)$ *be supported in* B^n. *Suppose that, for some* $\lambda, M \geqslant 1$, $N \in \mathbb{N}$ *and all* $z \in \operatorname{supp} a$, *these functions satisfy the following conditions*:

(i) $|\partial_z \phi(z)| \geqslant \lambda$;

(ii) $|\partial_z^\alpha \phi(z)| \leqslant M |\partial_z \phi(z)|$ *for all* $\alpha \in \mathbb{N}_0^n$ *with* $2 \leqslant |\alpha| \leqslant N$;

(iii) $|\partial_z^\alpha a(z)| \leqslant M^{|\alpha|}$ *for all* $\alpha \in \mathbb{N}_0^n$ *with* $|\alpha| \leqslant N$.

Then,

$$\left| \int_{\mathbb{R}^n} e^{i\phi(z)} a(z) \, dz \right| \lesssim_N M^N \lambda^{-N}.$$

The lemma is a standard application of integration-by-parts and the Leibniz rule. Nevertheless, the details of the proof are provided for completeness.

Proof. Define $Q \colon \mathbb{R}^n \to \mathbb{R}$ by $Q(z) := |\partial_z \phi(z)|^2$ and consider the mutually adjoint[34] differential operators

$$Du := \frac{\langle \partial_z u, \partial_z \phi \rangle}{iQ} \quad \text{and} \quad D^* u := i \sum_{k=1}^{n} \partial_{z_k}[(\partial_{z_k} \phi) Q^{-1} u].$$

Note that D fixes the function $e^{i\phi}$ and, consequently,

$$\int_{\mathbb{R}^n} e^{i\phi(z)} a(z) \, dz = \int_{\mathbb{R}^n} [D^N e^{i\phi(z)}] a(z) \, dz = \int_{\mathbb{R}^n} e^{i\phi(z)} (D^*)^N a(z) \, dz.$$

Thus, it suffices to show that

$$|(D^*)^N a(z)| \lesssim_N M^N \lambda^{-N}.$$

---

[34] In the sense that $\int_\Omega (Du)v = \int_\Omega u(D^*v)$, whenever at least one of the functions $u, v \in C^\infty(\Omega)$ has compact support.

It is useful to work with the more general statement

$$|\partial_z^\alpha (D^*)^\mu a(z)| \lesssim_{N,\alpha} M^{\mu+|\alpha|}\lambda^{-\mu} \quad \text{for all } \mu \in \mathbb{N}_0 \text{ and } \alpha \in \mathbb{N}_0^n \text{ satisfying } \mu+|\alpha| \leqslant N,$$

which is amenable to induction on μ. The base case $\mu=0$ follows directly from hypothesis (iii). The inductive step is established by appropriate application of the Leibniz rule and the hypothesised bounds for ϕ.

Fix $0 \leqslant \mu \leqslant N-1$ and $\alpha \in \mathbb{N}_0^n$ such that $\mu+1+|\alpha| \leqslant N$. Denoting by e_k the standard coordinate vectors for $k=1,...,n$, it follows by the definition of D^* and the Leibniz rule that

$$\partial_z^\alpha (D^*)^{\mu+1} a = i \sum_{k=1}^n \sum_{\beta \leqslant \alpha+e_k} \binom{\alpha+e_k}{\beta} \partial_z^{\alpha-\beta+e_k}[(\partial_{z_k}\phi)Q^{-1}]\partial_z^\beta (D^*)^\mu a,$$

where, for every fixed k, the second sum is over all multi-indices $\beta \in \mathbb{N}_0^n$ satisfying $\beta_j \leqslant \alpha_j + \delta_{jk}$ for $1 \leqslant j \leqslant n$. For each such multi-index, $\mu+|\beta| \leqslant \mu+1+|\alpha| \leqslant N$, and therefore the induction hypothesis yields

$$|\partial_z^\beta (D^*)^\mu a(z)| \lesssim_N M^{\mu+|\beta|}\lambda^{-\mu}. \tag{A.1}$$

On the other hand, condition (ii) together with the Leibniz rule implies that

$$|\partial_z^\gamma Q^{-1}(z)| \lesssim_N M^{|\gamma|}|Q(z)|^{-1} \quad \text{for all } \gamma \in \mathbb{N}_0^n \text{ with } |\gamma| \leqslant N.$$

Thus, again using (ii) and the Leibniz rule,

$$|\partial_z^\gamma [(\partial_{z_k}\phi(z))Q^{-1}(z)]| \lesssim_N M^{|\gamma|}|\partial_z\phi(z)|^{-1} \leqslant M^{|\gamma|}\lambda^{-1},$$

where the last step is by (i). Applying the above estimate with $\gamma=\alpha-\beta+e_k$ and combining this with (A.1), one deduces that

$$|\partial_z^\alpha (D^*)^{k+1} a(z)| \lesssim_N \sum_{k=1}^n \sum_{\beta \leqslant \alpha+e_k} M^{|\alpha|-|\beta|+1}\lambda^{-1}M^{\mu+|\beta|}\lambda^{-\mu} \lesssim_N M^{\mu+1+|\alpha|}\lambda^{-(\mu+1)},$$

which closes the induction. $\qquad\square$

References

[1] BAGBY, T., BOS, L. & LEVENBERG, N., Multivariate simultaneous approximation. *Constr. Approx.*, 18 (2002), 569–577.

[2] BELTRAN, D., HICKMAN, J. & SOGGE, C. D., Variable coefficient Wolff-type inequalities and sharp local smoothing estimates for wave equations on manifolds. To appear in *Anal. PDE*.

[3] BENNETT, J., Aspects of multilinear harmonic analysis related to transversality, in *Harmonic Analysis and Partial Differential Equations*, Contemp. Math., 612, pp. 1–28. Amer. Math. Soc., Providence, RI, 2014.

[4] BENNETT, J., CARBERY, A. & TAO, T., On the multilinear restriction and Kakeya conjectures. *Acta Math.*, 196 (2006), 261–302.

[5] BOURGAIN, J., L^p-estimates for oscillatory integrals in several variables. *Geom. Funct. Anal.*, 1 (1991), 321–374.

[6] — Some new estimates on oscillatory integrals, in *Essays on Fourier Analysis in Honor of Elias M. Stein* (Princeton, NJ, 1991), Princeton Math. Ser., 42, pp. 83–112. Princeton Univ. Press, Princeton, NJ, 1995.

[7] — Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces. *Israel J. Math.*, 193 (2013), 441–458.

[8] BOURGAIN, J. & DEMETER, C., Decouplings for curves and hypersurfaces with nonzero Gaussian curvature. *J. Anal. Math.*, 133 (2017), 279–311.

[9] BOURGAIN, J. & GUTH, L., Bounds on oscillatory integral operators based on multilinear estimates. *Geom. Funct. Anal.*, 21 (2011), 1239–1295.

[10] CARLESON, L. & SJÖLIN, P., Oscillatory integrals and a multiplier problem for the disc. *Studia Math.*, 44 (1972), 287–299.

[11] CHEN, X., KAYAL, N. & WIGDERSON, A., Partial derivatives in arithmetic complexity and beyond. *Found. Trends Theor. Comput. Sci.*, 6 (2010), 1–138 (2011).

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

[13] GUTH, L., A restriction estimate using polynomial partitioning. *J. Amer. Math. Soc.*, 29 (2016), 371–413.

[14] — Restriction estimates using polynomial partitioning II. *Acta Math.*, 221 (2018), 81–142.

[15] GUTH, L. & KATZ, N. H., On the Erdős distinct distances problem in the plane. *Ann. of Math.*, 181 (2015), 155–190.

[16] HICKMAN, J. & ROGERS, K. M., Improved Fourier restriction estimates in higher dimensions. *Camb. J. Math.*, 7 (2019), 219–282.

[17] HÖRMANDER, L., Oscillatory integrals and multipliers on FL^p. *Ark. Mat.*, 11 (1973), 1–11.

[18] LEE, S., Improved bounds for Bochner–Riesz and maximal Bochner–Riesz operators. *Duke Math. J.*, 122 (2004), 205–232.

[19] — Linear and bilinear estimates for oscillatory integral operators related to restriction to hypersurfaces. *J. Funct. Anal.*, 241 (2006), 56–98.

[20] MINICOZZI, W. P. I & SOGGE, C. D., Negative results for Nikodym maximal functions and related oscillatory integrals in curved space. *Math. Res. Lett.*, 4 (1997), 221–237.

[21] MOCKENHAUPT, G., SEEGER, A. & SOGGE, C. D., Local smoothing of Fourier integral operators and Carleson-Sjölin estimates. *J. Amer. Math. Soc.*, 6 (1993), 65–130.

[22] SOGGE, C. D., Oscillatory integrals and spherical harmonics. *Duke Math. J.*, 53 (1986), 43–65.

[23] — On the convergence of Riesz means on compact manifolds. *Ann. of Math.*, 126 (1987), 439–447.

[24] — Concerning the L^p norm of spectral clusters for second-order elliptic operators on compact manifolds. *J. Funct. Anal.*, 77 (1988), 123–138.

[25] — *Fourier Integrals in Classical Analysis*, second edition. Cambridge Tracts in Mathematics, 210. Cambridge Univ. Press, Cambridge, 2017.

[26] STEIN, E. M., *Singular Integrals and Differentiability Properties of Functions*. Princeton Mathematical Series, 30. Princeton Univ. Press, Princeton, NJ, 1970.

[27] — Oscillatory integrals in Fourier analysis, in *Beijing Lectures in Harmonic Analysis* (Beijing, 1984), Ann. of Math. Stud., 112, pp. 307–355. Princeton Univ. Press, Princeton, NJ, 1986.

[28] — *Harmonic Analysis*: *Real-Variable Methods*, *Orthogonality*, *and Oscillatory Integrals*. Princeton Mathematical Series, 43. Princeton Univ. Press, Princeton, NJ, 1993.

[29] STONE, A. H. & TUKEY, J. W., Generalized "sandwich" theorems. *Duke Math. J.*, 9 (1942), 356–359.

[30] TAO, T., The weak-type endpoint Bochner–Riesz conjecture and related topics. *Indiana Univ. Math. J.*, 47 (1998), 1097–1124.

[31] — The Bochner–Riesz conjecture implies the restriction conjecture. *Duke Math. J.*, 96 (1999), 363–375.

[32] TAO, T., VARGAS, A. & VEGA, L., A bilinear approach to the restriction and Kakeya conjectures. *J. Amer. Math. Soc.*, 11 (1998), 967–1000.

[33] WANG, H., A restriction estimate in $\mathbb{R}^3$ using brooms. Preprint, 2018. arXiv:1802.04312[math.CA].

[34] WARNER, F. W., *Foundations of Differentiable Manifolds and Lie Groups*. Graduate Texts in Mathematics, 94. Springer, New York–Heidelberg, 1983.

[35] WISEWELL, L., Kakeya sets of curves. *Geom. Funct. Anal.*, 15 (2005), 1319–1362.

[36] WOLFF, T. H., *Lectures on Harmonic Analysis*. University Lecture Series, 29. Amer. Math. Soc., Providence, RI, 2003.

[37] WONGKEW, R., Volumes of tubular neighbourhoods of real algebraic varieties. *Pacific J. Math.*, 159 (1993), 177–184.

[38] ZAHL, J., On the Wolff circular maximal function. *Illinois J. Math.*, 56 (2012), 1281–1295.

[39] ZHANG, R., Polynomials with dense zero sets and discrete models of the Kakeya conjecture and the Furstenberg set problem. *Selecta Math.*, 23 (2017), 275–292.

LARRY GUTH
Department of Mathematics
Massachusetts Institute of Technology
182 Memorial Drive
Cambridge, MA, 02139
U.S.A.
lguth@math.mit.edu

MARINA ILIOPOULOU
Department of Mathematics
University of California, Berkeley
Berkeley, CA, 94720-3840
U.S.A.
m.iliopoulou@kent.ac.uk

Current address:
School of Mathematics,
Statistics and Actuarial Science
Sibson Building, Room 262
University of Kent
Parkwood Road
Canterbury CT2 7FS
U.K.

JONATHAN HICKMAN
Eckhart hall Room 414
Department of Mathematics
University of Chicago
5734 S. University Avenue
Chicago, IL, 60637
U.S.A.
jonathan.hickman@ed.ac.uk

Current address:
School of Mathematics
University of Edinburgh
James Clerk Maxwell Building
Peter Guthrie Tait Road
Edinburgh EH9 3FD
U.K.