IMPLICITLY OSCILLATORY MULTILINEAR INTEGRALS
FIRST ENCOUNTERS WITH THE FIFTH KIND
PADOVA MINICOURSE LECTURE NOTES

MICHAEL CHRIST

Contents

1. Introduction

An archetypal (bilinear) oscillatory integral inequality states that

$$\Big| \iint_{\mathbb{R}^d\times\mathbb{R}^d} f(x)\, g(y)\, e^{i\lambda\phi(x,y)}\, \eta(x,y)\, dx\, dy \Big| \le C|\lambda|^{-\gamma} \|f\|_{L^2} \|g\|_{L^2} \tag{1.1}$$

where $\lambda \in \mathbb{R}$ is a large parameter, the real-valued phase function ϕ is smooth and is nondegenerate in a natural sense (to be discussed below), f, g are arbitrary L^2 functions, η is a smooth compactly supported cutoff function, and $\gamma > 0$ and $C < \infty$ depend on ϕ but not on f, g, λ. Its main features are

(1) The decaying factor $|\lambda|^{-\gamma}$,
(2) The absence of any smoothness hypothesis on the measurable factors f, g,
(3) The interplay between the structure of ϕ and the product structure of $f(x)\, g(y)$.

Date: $\boxed{\text{Draft of June 20, 2022}}$.

Research supported by NSF grant DMS-1901413.

Where ϕ is nonconstant, $e^{i\lambda\phi}$ oscillates rapidly when λ is large, creating cancellation that potentially results in smallness of the integral.

Implicitly oscillatory integrals, in contrast, involve no overtly oscillatory factor $e^{i\lambda\phi}$. Instead, the measurable factors f_j are themselves assumed to be oscillatory, but in a non-explicit and unstructured manner. A typical multilinear form of this type takes the form

$$\int_{\mathbb{R}^D} \prod_{j\in J} (f_j \circ \varphi_j)(x)\,\eta(x)\,dx \tag{1.2}$$

where J is a finite index set, $\varphi_j : \mathbb{R}^D \to \mathbb{R}^d$ are smooth submersions from $\mathbb{R}^D$ to a lower-dimensional space $\mathbb{R}^d$, and the functions $f_j : \mathbb{R}^d \to \mathbb{C}$ are merely measurable. The desired upper bound is expressed in terms of strictly negative order $H^{-\sigma}$ Sobolev norms of these functions, rather than L^2 or L^p norms. If $f_k = O(1)$ in L^2 norm, and if f_k is rapidly oscillatory, then $\|f_k\|_{H^{-\sigma}}$ is small, so such an upper bound would guarantee corresponding smallness of the integral (1.2). The presence of more factors $f_j \circ \varphi_j$ than the dimension D of the ambient space of integration is a second essential feature distinguishing nontrivial instances of (1.2) from (1.1).

These integrals arise in connection with multilinear maximal functions in harmonic analysis, with the theory of weak limits, and with Ramsey theory. There are connections with web geometry and with sublevel set inequalities, which are a key element of their analysis.

These lectures will begin with a discussion of some classical results concerning bilinear forms with explicit oscillatory factors and associated sublevel set inequalities. Multilinear (explicitly) oscillatory integrals will then be discussed. Finally, inequalities for implicitly oscillatory integrals will be introduced. Proofs will be outlined, with some key steps presented in detail and some key ingredients of the proofs developed. The connections listed above will be discussed briefly.

The course is based on recent work of the speaker and co-authors, and is an introduction to more recent (and more technically complicated) work. Familiarity with measure and integration, basic Fourier analysis, multivariable calculus, and other miscellaneous topics is assumed.

2. Oscillatory integrals of the first kind

In the terminology of Stein [40], oscillatory integrals of the first kind are integral expressions of the form

$$I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\,\eta(x)\,dx \tag{2.1}$$

where $\lambda \in \mathbb{R}$ is a large parameter, $\eta \in C_0^\infty$ is a smooth[1] compactly supported cutoff function, and the phase function ϕ is smooth and real-valued. Note the absence of any unknown measurable factors f, in contrast to (1.1) and (1.2).

Variants of this formulation arise: η and/or ϕ might have only some finite degree of differentiability; the integral might be over an interval in $\mathbb{R}^1$ with no cutoff function η; et cetera. In these lectures, I'm interested in smooth ϕ and η; ϕ will often be assumed to be real analytic, and sometimes a polynomial.

In these lectures, we are concerned primarily with upper bounds asserting that such expressions are small when $|\lambda|$ is large. We are less concerned with analysis of the optimal rate of decay as $|\lambda| \to \infty$.

[1] "Smooth" means C^∞, unless otberwise indicated.

Proposition 2.1. *If $\nabla\phi$ vanishes nowhere on the support of η then*

$$I(\lambda) = O(|\lambda|^{-N}) \; \forall\, N < \infty. \tag{2.2}$$

The fact that the Fourier coefficients of smooth functions tend rapidly to 0 is a special case;

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

equals $I(\lambda)$ with $\lambda = |\xi|$ and $\phi(x) = x \cdot v$ where $v = \xi/|\xi|$.

Proof of Proposition 2.1. In the one-dimensional case,

$$I(\lambda) = \int \frac{d}{dx} e^{i\lambda\phi(x)}\, (i\lambda\phi'(x))^{-1}\, \eta(x)\, dx = i\lambda^{-1} \int e^{i\lambda\phi(x)} \frac{d}{dx}(\phi'(x)^{-1}\eta(x))\, dx = O(|\lambda|^{-1})$$

and in the same way, N integrations by parts give a bound $O(|\lambda|^{-N})$. The higher-dimensional case reduces to $d = 1$ by a partition of unity and Fubini's theorem. □

There are variants in which ϕ is assumed to be real analytic and exponential decay with respect to λ is shown, using deformation of the contour of integration in $\mathbb{C}^d$.

Zelditch [48] has pointed out that it is unknown, for $d \geq 3$, whether the converse holds: if ϕ is (say) real analytic, if η does not vanish identically, and if $I(\lambda) = O(|\lambda|^{-N})$ for every N, must ϕ have no critical points in the interior of the support of η?

For ϕ with critical points, the most basic result concerns the case of nondegenerate critical points, meaning that the Hessian matrix $D^2\phi$ is nonsingular at each critical point. In that situation critical points are isolated. If $\bar{x}$ is the only critical point of ϕ in the support of η then

$$I(\lambda) = c^d \lambda^{-d/2} |\det(D^2\phi(\bar{x}))|^{-1/2} e^{i\lambda\phi(\bar{x})}\eta(\bar{x}) + O(|\lambda|^{-(d+2)/2})$$

as $|\lambda| \to \infty$ for a certain explicit constant c. There is a full asymptotic expansion in terms of powers $\lambda^{-(d+2k)/2}$.

If ϕ is a polynomial with finitely many critical points then $I(\lambda)$ has an asymptotic expansion, as $\lambda \to +\infty$, in powers of λ and $\log(\lambda)$. The case $d = 1$ is elementary; if $n \geq 2$, if $\phi^{(k)}(\bar{x}) = 0$ for all $1 \leq k < n$, and if η is supported in a neighborhood of $\bar{x}$ in which ϕ has no other critical points, then

$$|I(\lambda)| = O(|\lambda|^{-1/n} \quad \text{as } |\lambda| \to \infty.$$

The case $d = 2$ has been thoroughly analyzed by Varchenko [46], Karpushkin [30], and others; the formulation of the results is quite a bit more intricate. The theory for higher dimensions is less satisfactory.

Van der Corput's lemma provides an important upper bound in the one-dimensional case.

Proposition 2.2. *Let ϕ be real-valued. Let $a \leq b \in \mathbb{R}$.*

Suppose that $\phi \in C^1$ satisfies $|\phi'(x)| \geq r > 0$ for every $x \in [a, b]$. Suppose moreover that ϕ' is monotone on $[a, b]$. Then

$$|\int_a^b e^{i\phi(x)}\, dx| \leq r^{-1}. \tag{2.3}$$

Let $k \ge 2$, and let $\phi \in C^k([a,b])$. If the k-th derivative of ϕ satisfies $|\phi^{(k)}(x)| \ge r$ for every $x \in [a,b]$ then

$$|\int_a^b e^{i\phi(x)}\,dx| \le C_k r^{-1/k}. \tag{2.4}$$

These bounds hold with constants C_k independent of a,b.

Proof. Consider first the case $k=1$. Assume temporarily that $\phi \in C^2$. Integrate by parts to get a constant times

$$\int_a^b e^{i\phi(x)} \frac{\phi''(x)}{\phi'(x)^2}\,dx$$

plus two boundary terms, proportional to $\phi'(a)^{-1}$ and $\phi'(b)^{-1}$, respectively. The boundary terms are OK. Assuming without no loss of generality that the monotonic function ϕ' is nondecreasing, the main term has absolute value

$$\le \int_a^b |\frac{\phi''(x)}{\phi'(x)^2}|\,dx = \int_a^b \frac{\phi''(x)}{\phi'(x)^2}\,dx = \phi'(a)^{-1} - \phi'(b)^{-1} \le r^{-1}.$$

The hypothesis that $\phi \in C^2$ can be easily removed by a approximation argument.

The case of $k \ge 2$ is proved by induction on k. Let $\rho > 0$. We have $|(\phi^{(k-1)})'| \ge r$, so the set of $x \in [a,b]$ at which $|\phi^{(k-1)}| \le \rho$ is an interval J of length $\le r^{-1}\rho$. Therefore $|\int_J e^{i\phi}| \le r^{-1}\rho$. The set $[a,b]\setminus J$ is a union of two or fewer intervals $\tilde J$. On each, $|\phi^{(k-1}| \ge \rho$, so the induction hypothesis gives $|\int_{\tilde J} e^{i\phi}| = O(\rho^{-1/(k-1)})$. In all, we find that

$$|\int_{[a,b]} e^{i\phi}| \lesssim (\rho/r) + (\rho^{-1/(k-1)}).$$

Choosing ρ to optimize this bound establishes the lemma.

We leave it to the audience to check that this proof does work for $k=2$. □

The following simpler bounds are closely related to van der Corput's lemma.

Proposition 2.3. *Let ϕ be real-valued. Let $a \le b \in \mathbb{R}$.*

Let $\phi \in C^1$. If $|\phi'(x)| \ge r > 0$ for every $x \in [a,b]$ Then for every $\varepsilon > 0$,

$$|\{x \in [a,b] : |\phi(x)| \le \varepsilon\}| \le \varepsilon r^{-1}. \tag{2.5}$$

Let $k \ge 2$, and let $\phi \in C^k([a,b])$. If the k-th derivative of ϕ satisfies $|\phi^{(k)}(x)| \ge r$ for every $x \in [a,b]$ then

$$|\{x \in [a,b] : |\phi(x)| \le \varepsilon\}| \le C_k \varepsilon^{1/k} r^{-1/k}. \tag{2.6}$$

The first bound is obvious. The second follows from the same induction argument as was used in the proof of Proposition 2.2. □

The next result will be used in proofs to follow.

Proposition 2.4. *Let ϕ be real-valued and C^∞. Let η be smooth and compactly supported. Suppose that there exists a multi-index α such that $\partial_x^\alpha \phi$ vanishes nowhere on the support of η. Then there exist $C < \infty$ and $\tau > 0$ such that for every $\lambda > 0$,*

$$|\int_{\mathbb{R}^d} e^{i\lambda\phi}\,\eta| \le C\lambda^{-\tau}. \tag{2.7}$$

Proof. For $d = 1$ and $|\alpha| = 1$ this is a far weaker conclusion than that of Proposition 2.1. For $d = 1$ and $|\alpha| \ge 2$, it is an immediate consequence of Proposition 2.2.

Let $d > 1$. Let $\bar{x}$ belong to the support of η. There exists a vector v such that $(v \cdot \nabla)^{|\alpha|}\phi(\bar{x}) \ne 0$. By introducing a partition of unity, we may assume that this holds at every point of the support of η. Matters are thus reduced to the one-dimensional case. □

Lemma 2.5. *Let $N, d \ge 1$. There exist $\tau > 0$ and $A, \sigma < \infty$ with the following property. Let $Q \subset \mathbb{R}^d$ be the closed unit cube. Let $f : Q \to \mathbb{C}$ be a C^{N+1} function, Define*

$$\delta = \min_{x \in Q} \sum_{0 \le |\alpha| \le N} \left| \frac{\partial^\alpha f}{\partial x^\alpha}(x) \right| \tag{2.8}$$

and $B = 1 + \sum_{|\alpha|=N+1} \|\partial^\alpha f\|_{C^0(Q)}$. For each $\varepsilon > 0$ let $S(f, \varepsilon) = \{x \in Q : |f(x)| \le \varepsilon\}$. Then

$$|S(f, \varepsilon)| \le C\varepsilon^\tau \delta^{-\sigma} B^A. \tag{2.9}$$

The complication, in comparison to the van der Corput-type bound, is the need for a factor B^A that takes into account possible large values of partial derivatives of order $N+1$. Such a factor is needed, as the next example shows.

Example. Consider the one-parameter family of functions

$$f_\varepsilon(x) = \varepsilon \sin(\varepsilon^{-1} x)$$

for $x \in [0, 1]$ and $\varepsilon \in (0, 1]$. These satisfy $\sum_{n=0}^{2} |f_\varepsilon^{(n)}(x)| \ge c > 0$, uniformly for all x and all $\varepsilon \in (0, 1]$, but despite this uniform lower bound, $\{x \in [0, 1] : |f_\varepsilon(x)| > \varepsilon/2\}$ has measure bounded below by a positive constant, uniformly in ε. □

Proof of Lemma 2.5. For the proof, partition Q into subcubes, each of sidelength r with r defined by $Br^{N+1} = c_0\varepsilon$. If the constant c_0 is chosen to be sufficiently small then on each subcube, f differs from a Taylor polynomial of degree N by at most ε. Thus matters reduce to the previously known situation in which f is equal to such a polynomial. One must sum the resulting bound over the cubes, resulting in a loss of a power of r. □

Lemma 2.6. *Let f be real-valued and satisfy the hypotheses of Lemma 2.5, with the sole change that only derivatives of f of strictly positive orders are taken into account in the definition of δ:*

$$\delta = \min_{x \in Q} \sum_{1 \le |\alpha| \le N} \left| \frac{\partial^\alpha f}{\partial x^\alpha}(x) \right|. \tag{2.10}$$

Then there exist $\tau > 0$ and $A, C, \sigma < \infty$ such that for all $\mathbb{R} \ni \lambda \ge 0$,

$$\Big| \int_Q e^{i\lambda f} \Big| \le C \min\big(1, \lambda^{-\tau}\delta^{-\sigma}B^A\big). \tag{2.11}$$

The proof follows the same strategy as that of Lemma 2.5: Decompose into subcubes of sidelength satisfying $B\lambda r^{N+1} = 1$; on each subcube, make an affine change of variables that converts it to an interval of length 1, and observe that on each, f can then be replaced by a polynomial of degree N, plus a remainder h such that $e^{i\lambda h}$ is slowly varying. Details are left to the reader. □

The next result is a well-known theorem of Łojasiewicz.

Lemma 2.7. *Let $\Omega \subset \mathbb{R}^d$ be a compact convex set and let $U \supset \Omega$ be an neighborhood of Ω. Let $f : \Omega \to \mathbb{R}$ be C^ω. Let $\Sigma = \{x \in U : f(x) = 0\}$. If f does not vanish identically then there exist $C < \infty$ and $\tau > 0$ such that for every $\varepsilon > 0$,*

$$\big|\{x \in \Omega : \text{distance}(x, \Sigma) < \varepsilon\}\big| \le C\varepsilon^\tau. \tag{2.12}$$

This can be proved using the Weierstrass Preparation Theorem. □

The next lemma was introduced into this subject by Bourgain [4].

Lemma 2.8. *Let $\phi = \phi(x,y)$ be real-valued and C^ω in a neighborhood of a compact convex set $\Omega \times \Omega' \subset \mathbb{R}^d \times \mathbb{R}^{d'}$. Let $\bar y \in \Omega'$. There exists $N < \infty$ such that*

$$\sup_{y \in \Omega'} \sum_{0 \le |\alpha| \le N} \big|\partial_y^\alpha \phi(x,y)\big|^2 \le C \sum_{0 \le |\alpha| \le N} \big|\partial_y^\alpha \phi(x,\bar y)\big|^2 \quad \forall\,(x,y) \in \Omega \times \Omega'. \tag{2.13}$$

We will not discuss the proof of Lemma 2.8. A good reference for a proof is [42].

The right-hand of (2.13) is a real analytic function of x. This can be useful in deducing properties of the supremum on the left-hand side. For instance:

Lemma 2.9. *Let $\phi(x,y)$ be as in the preceding lemma. Suppose that for any x, the function $y \mapsto \phi_x(y) = \phi(x,y)$ does not vanish identically. Then there exist $C < \infty$ and $\tau > 0$ such that for every $x \in \Omega$,*

$$\big|\{y \in \Omega' : |\phi_x(y)| < \varepsilon\}\big| \le C\varepsilon^\tau \quad \forall \varepsilon > 0.$$

Lemma 2.9 follows from the combination of Lemma 2.5 with Lemma 2.8. The uniformity of this inequality with respect to the parameter x will be useful below.

3. A connection

Oscillatory bounds directly imply sublevel bounds. Suppose we wish to establish an upper bound for $|\{x \in B : |\phi(x)| < \varepsilon\}|$ for a real-valued function ϕ. Assume that ϕ is defined and satisfies an oscillatory inequality $\int e^{i\lambda\phi}\zeta = O(|\lambda|^{-\gamma})$ for some cutoff function ζ that is positive in a neighborhood of B.

Choose another auxiliary function $h \in C_0^\infty(\mathbb{R})$ satisfying $h \equiv 1$ in $[-1,1]$. It suffices to bound $\int h(\phi(x)/\varepsilon)\,\zeta(x)\,dx$.

Writing h as the inverse Fourier transform of $\widehat{h}$ gives

$$\int h(\phi(x)/\varepsilon)\,\zeta(x)\,dx = c\int_{\mathbb{R}} \widehat{h(\xi)} \int e^{i\xi\lambda\phi(x)}\,\zeta(x)\,dx\,d\xi$$

with $\lambda = \varepsilon^{-1}$. By hypothesis, the inner integral is $O((1+\lambda|\xi|)^{-\gamma})$ and therefore

$$\int h(\phi(x)/\varepsilon)\,\zeta(x)\,dx = O\big(\int_{\mathbb{R}} (1+|\xi|)^{-2}\cdot(1+\lambda|\xi|)^{-\gamma}\,d\xi\big).$$

The factor $O((1+\lambda|\xi|)^{-\gamma})$ yields an improvement unless $\lambda|\xi|$ is roughly comparable to 1; the set of all]ξ satisfying that constraint has Lebesgue measure $O(\lambda^{-d})$. An elementary calculation gives a bound $O(\varepsilon^\tau)$ for the integral on the right-hand side, for some $\tau = \tau(\gamma,d) > 0$.

Is there a converse? Do sublevel set bounds imply corresponding oscillatory integral bounds? One hint that such a converse may be less direct lies in the observation that oscillatory integral bounds imply a stronger type of sublevel set inequality! To see this, define

$$\|x\|_{\mathbb{R}/\mathbb{Z}} = \text{distance}(x, \mathbb{Z})$$

for $x \in \mathbb{R}$. Consider the "mod $\mathbb{Z}$" sublevel set

$$S(\lambda, \varepsilon) = \{x \in B : \|\lambda\phi(x)\|_{\mathbb{R}/\mathbb{Z}} \le \varepsilon\}.$$

Then assuming the same oscillatory integral inequality for ϕ as above,

$$|S(\lambda, \varepsilon)| = O(\varepsilon + \lambda^{-\tau}\varepsilon^{\tau})$$

for some $\tau(\gamma, d) > 0$.

For the proof, construct a nonnegative auxiliary function $h_\varepsilon : \mathbb{R} \to [0, \infty)$ that is 1-periodic and satisfies $h_\varepsilon(x) = 1$ whenever $\|x\|_{\mathbb{R}/\mathbb{Z}} \le \varepsilon$, whose Fourier coefficients satisfy[2]

$$|\widehat{h_\varepsilon}(n)| = O(\varepsilon \cdot (1 + \varepsilon|n|)^{-2}) \ \ \forall\, n \in \mathbb{Z}.$$

Apply the same reasoning as above, considering $\int h_\varepsilon(\lambda\phi(x))\, \zeta(x)\, dx$ and expanding $h_\varepsilon(\lambda\phi(x)) = \sum_n \widehat{h_\varepsilon}(n)\, e^{2\pi i\lambda\phi(x)}$. The term with $n = 0$ gives rise to a term independent of λ and proportional to ε. The remaining terms gives rise to a remainder that is $O(\varepsilon^{\tau}\lambda^{-\tau})$ upon summation over $n \ne 0$.

4. Oscillatory integrals of the second kind

In the terminology of Stein [40], oscillatory integrals of the second kind are linear operators

$$T_\lambda(f)(x) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x,y)}\, f(y)\, \eta(x, y)\, dy. \tag{4.1}$$

As before, λ is a large real parameter (which we henceforth assume to be positive; replacing ϕ by $-\phi$ reduces matters to positive λ), the phase function ϕ is smooth and real-valued but now has domain $\mathbb{R}^d \times \mathbb{R}^d$, and η is smooth with compact support in the product space. The new element is f, which is Lebesgue measurable. The goal is to obtain upper bounds under mild restrictions on f, with no assumption of any smoothness or quantitative continuity.[3] Most frequently, one regards T_λ is a bounded linear operator between two Lebesgue $X = L^p$ and $Y = L^q$, and one seeks a bound of the form $\|T_\lambda f\|_X \le C\lambda^{-\gamma}\|f\|_Y$.

The same terminology refers to $\mathbb{C}$-valued bilinear forms

$$\mathcal{T}_\lambda(f, g) = \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{i\lambda\phi(x,y)}\, f(y)\, g(x)\, \eta(x, y)\, dx\, dy. \tag{4.2}$$

At least two generalizations are quite natural. First, there is the case in which x, y lie in Euclidean spaces of unequal dimensions. This situation arises in Fourier restriction inequalities, in the study of dispersive wave equations, and in the analysis of Bochner-Riesz Fourier multiplier operators. This situation has been quite extensively studied, but is not the topic of these lectures. Second, one may consider multilinear scalar-valued forms

$$\mathcal{T}_\lambda(\mathbf{f}) = \mathcal{T}_{\phi,\lambda}(\mathbf{f}) = \int_{(\mathbb{R}^d)^J} e^{i\lambda\phi(x)} \prod_j f_j(x_j)\, \eta(x)\, dx \tag{4.3}$$

where J is a finite index set, $x = (x_j \in \mathbb{R}^d : j \in J)$, and $\mathbf{f} = (f_j : j \in J)$, and $\eta \in C_0^\infty$ is a cutoff function whose exact identity is unimportant, on whose support hypotheses are to be imposed.

The most basic result is due to Hörmander.

---

[2]Fix a smooth compactly supported function η that is $\equiv 1$ in $[-1, 1]$. Define $h_\varepsilon(x) = \eta(\varepsilon x)$ for all $x \in [-\frac{1}{2}, \frac{1}{2}]$, and extend h_ε to be periodic with period 1.

[3]If some smoothness is assumed, then f should be regarded as a factor of the cutoff function η, so that one is dealing with an oscillatory integral of the first kind with a perhaps less than ideal cutoff function.

Proposition 4.1. *Suppose that the mixed Hessian matrix*

$$D^2_{x,y}\phi(x,y) = \Big(\frac{\partial^2 \phi}{\partial x\,\partial y}\Big)$$

is nonsingular at every point in the support of η. Then

$$|\mathcal{T}_\lambda(f,g)| \le C\lambda^{-d/2}\|f\|_{L^2}\|g\|_{L^2} \tag{4.4}$$

for every $f,g \in L^2(\mathbb{R}^d)$. The constant C depends on d,ϕ,η but not on f,g,λ.

The most basic example is $\phi(x,y) = x\cdot y$. By choosing $\eta \equiv 1$ in some neighborhood of the origin, substituting $x = \lambda^{1/2}\tilde{x}$ and $y = \lambda^{1/2}\tilde{y}$, then letting $\lambda \to \infty$, one concludes that

$$\Big|\int_{\mathbb{R}^d\times\mathbb{R}^d} e^{ix\cdot y}\, f(x)\, g(y)\, dx\, dy\Big| \le C\|f\|_{L^2}\|g\|_{L^2}$$

for all $f,g \in L^2(\mathbb{R}^d)$, for a certain universal constant[4] C. Equivalently, by duality, the Fourier transform is bounded from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$.

Proof of Proposition 4.1. By a partition of unity we may reduce to the case in which η is supported in a small neighborhood of a point $(\bar{x},\bar{y})$.

Use the so-called TT^* argument: Write

$$\|T_\lambda(f)\|^2_{L^2} = \langle T^*_\lambda T_\lambda f,\, f\rangle \le \|T^*_\lambda T_\lambda f\|_{L^2}\|f\|_{L^2}$$

to reduce matters to controlling the operator norm of the composition $T^*_\lambda \circ T_\lambda$.

Compute

$$T^*_\lambda T_\lambda(f)(x) = \int K_\lambda(x,x')f(x')\,dx'$$

with

$$K_\lambda(x,x') = \int_{\mathbb{R}^d} e^{i\lambda[\phi(x',y)-\phi(x,y)]}\zeta(x,x',y)\,dy$$

for a certain[5] $\zeta \in C^\infty_0(\mathbb{R}^{3d})$. This is an oscillatory integral of the *first* kind, depending on the parameter $(x,x') \in \mathbb{R}^{d+d}$. The assumption that the mixed Hessian matrix $D^2_{x,y}\phi$ of ϕ is nonsingular implies that

$$|\nabla_y(\phi(x,y)-\phi(x',y))| \ge c|x-x'|$$

and a moderately careful[6] inspection of the proof of Proposition 2.1 yields an upper bound

$$|K_\lambda(x,x')| \le C_N\cdot(\lambda|x-x'|)^{-N}\ \forall\, N < \infty.$$

There is also a trivial upper bound

$$|K_\lambda(x,x')| \le C < \infty,$$

uniformly in (x,x').

[4] It is possible to obtain the optimal constant by carrying this argument out with a careful choice of cutoff function, but our focus in these lectures is not directed towards optimal constants.

[5] The subscript 0 always indicates compact support.

[6] Here's the issue: For each pair (x,x'), one wants to find a constant vector v satisfying $|v\cdot\nabla_y(\phi(x,y)-\phi(x',y))| \ge c|x-x'|$. This inequality should hold for all y in a neighborhood of a basepoint y_0. Then one can rotate coordinates so that $v\cdot\nabla = \frac{\partial}{\partial y_1}$ and integrate by parts with respect to y_1. To get the desired bound, it's essential that this neighborhood be independent of (x,x'). This is a delicate point, because no lower bound is available for $|x-x'|$.

We have

$$\langle T_\lambda^* T_\lambda f,\, f\rangle = \int K_\lambda(x,x)\, f(x)\, \overline{f(x')}\, dx\, dx' \le C \int (1+\lambda|x-x'|)^{-d-1}\, |f(x)|\, |f(x')|\, dx\, dx'.$$

We have reduced matters to an operator defined by integration against a nonnegative kernel function K_λ, and moreover, a completely explicit one.

I claim that

$$\Big| \int_{\mathbb{R}^{d+d}} (1+\lambda|x-x'|)^{-d-1}\, f(x)\, g(y)\, dx\, dy\Big| \le C_d \lambda^{-d} \|f\|_{L^2} \|g\|_{L^2} \quad \forall\, f, g \in L^2(\mathbb{R}^d).$$

This follows from the general inequality

$$\Big| \iint f(x)\, g(y)\, \mathcal{K}(x,y)\, dx\, dy\Big| \le A \|f\|_2 \|g\|_2 \tag{4.5}$$

with

$$A = \max\big(\sup_x \int |\mathcal{K}(x,y)|\, dy + \sup_y \int |\mathcal{K}(x,y)|\, dx \big). \tag{4.6}$$

This inequality is valid for any $\mathcal{K}$. Apply it with $\mathcal{K}(x,y) = (1+\lambda|x-y|)^{-d-1}$, for which the quantity A in (4.6) is $A = c\lambda^{-d}$. □

(4.5) follows from

$$\Big| \iint f(x)\, g(y)\, \mathcal{K}(x,y)\, dx\, dy\Big| \le A \min\big(\|f\|_1 \|g\|_\infty,\, \|f\|_\infty \|g\|_1 \big) \tag{4.7}$$

by the Riesz-Thorin interpolation theorem. (4.7) is immediate from the definition of A. □

If the determinant $\det(D^2_{x,y}\phi)$ of the mixed Hessian of ϕ vanishes at some point, but does not vanish identically, then one may hope for a bound with a lesser power of λ^{-1} — analogous to the discussion of van der Corput's lemma for $\int_a^b e^{i\phi}$ — of the form

$$|\mathcal{T}_\lambda(f,g)| \le C\lambda^{-\gamma} \|f\|_{L^p} \|g\|_{L^q},$$

obtaining an *optimal* power of λ, as $\lambda \to +\infty$, in terms of p, q, d and an appropriate quantitative description of the degree to which the determinant vanishes at a point of degeneracy. This theory is quite successful for $d = 2$, and less complete in higher dimensions. Again, optimal constants, and optimal exponents, are not the focus of these lectures; we are headed towards *terra incognita* in which virtually no optimal exponents are known, and the appropriate goal is to show that inequalities hold with *some* positive exponents.

There is a kind of symmetry, or gauge invariance, in this theory. If two phase functions $\phi, \tilde\phi$ are related by

$$\tilde\phi(x,y) = \phi(x,y) + \psi_1(x) + \psi_2(y) \tag{4.8}$$

for some functions ϕ_j, then the forms $\mathcal{T}_{\phi,\lambda}, \mathcal{T}_{\tilde\phi,\lambda}$ satisfy exactly the same inequalities in terms of $\|f\|_{L^p} \|g\|_{L^q}$, since

$$\mathcal{T}_{\tilde\phi,\lambda}(f,g) = \mathcal{T}_{\phi,\lambda}(\tilde f, \tilde g)$$

with $\tilde f = e^{i\lambda\psi_1} f$ and $\tilde g = e^{i\lambda\psi_2} g$; these have the same Lebesgue norms as f, g, respectively. Thus (4.8) defines an equivalence relation on phase functions, with two phase functions in the same equivalence class being genuinely equivalent for our purposes. The mixed Hessian is natural from this perspective, since $D^2_{x,y}\tilde\phi \equiv D^2_{x,y}\phi$ for equivalent phase functions.

An example that does not satisfy the Hörmander hypothesis is

$$\iint_{\mathbb{R}\times\mathbb{R}} e^{i\lambda(x-y)^3}\, f(x)\, g(y)\, \eta(x,y)\, dx\, dy.$$

This form satisfies an upper bound of the form $O(\lambda^{-1/3}\|f\|_{L^2}\|g\|_{L^2})$. The exponent $\frac{1}{3}$ is intermediate between $\frac{d}{2}=\frac{1}{2}$ and the trivial exponent 0. The Hessian, $D^2_{x,y}\phi = -6(x-y)$, vanishes on the diagonal but does not vanish identically. Assume that $\eta(0,0)\neq 0$. Choosing each of f,g to be the indicator function of an interval of length $c\lambda^{-1/3}$ centered at 0 with $c=(\pi/4)^{1/3}$, the real part of $\phi(x,y)$ is $\ge 2^{-1/2}$ whenever $f(x)g(y)\neq 0$ and it follows that $\mathcal{T}_{\phi,\lambda}(f,g)$ has order of magnitude $\lambda^{-1/3}\|f\|_2\|g\|_2$.

We next discuss a version of Proposition 4.1 that allows for some degeneracy, and that will be used in a key proof below. In all results concerning real analytic phases, we will always assume that the domain of the phase is connected, without stating this assumption explicitly. Thus an assumption that a related quantity does not vanish identically, means that it does not vanish identically on any open set.

Proposition 4.2. *Consider a bilinear form*

$$\mathcal{T}_\lambda(f,g) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} e^{i\lambda\phi(x,y)}\, f(x)\, g(y)\, \eta(x,y)\, dx\, dy$$

with real analytic real-valued phase function ϕ. Suppose that the mixed Hessian matrix $D^2_{x,y}\phi$ of ϕ does not vanish identically. There exist $C<\infty$ and $\gamma>0$ satisfying

$$|\mathcal{T}_\lambda(f,g)| \le C(1+\lambda)^{-\gamma}\|f\|_{L^2}\|g\|_{L^2}\ \forall\, f,g\in L^2(\mathbb{R}^d).$$

Proof. In proving a result of this type, we may always assume that (x,y) varies over a Cartesian product domain, by covering a small neighborhood of the support of η by a finite union of small products of cubes, introducing a finite smooth partition of unity subordinate to this cover, and regarding the auxiliary partition functions as additional factors of the cutoff function η.

Following the proof of Proposition 4.1, we apply Cauchy-Schwarz to obtain

$$|\mathcal{T}_\lambda(f,g)|^2 \le C\|g\|_{L^2}^2 \Big|\iint K_\lambda(x,x')\, f(x)\, \overline{f}(x')\, dx\, dx'\Big|$$

with

$$K_\lambda(x,x') = \int e^{i\lambda\psi(x,x',y)}\, \zeta(x,x',y)\, dy$$

where

$$\psi(x,x',y) = \phi(x,y) - \phi(x',y).$$

Fix any point $\bar y$. By hypothesis, $\frac{\partial\psi}{\partial y}(x,x',y)$ does not vanish identically. Therefore there exists N such that $\sum_{1\le|\alpha|\le N}|\partial_y^\alpha\psi(x,x',y)|$ is uniformly comparable to $\sum_{1\le|\alpha|\le N}|\partial_y^\alpha\psi(x,x',\bar y)|$. Equivalently, $(\sum_{1\le|\alpha|\le N}|\partial_y^\alpha\psi(x,x',y)|)^2$ is uniformly comparable to the real analytic function $\sum_{1\le|\alpha|\le N}\partial_y^\alpha\psi(x,x',\bar y)^2$.

Therefore by Lemma 2.6, there exist $\tau>0$ and $C<\infty$ such that for each (x,x'),

$$\Big|\int e^{i\lambda\psi(x,x',y)}\, \zeta(x,x',y)\, dy\Big| \le C(1+\lambda^2\theta(x,x'))^{-\tau}$$

where

$$\theta(x,x') = \sum_{1\le|\alpha|\le N}|\partial_y^\alpha\psi(x,x',\bar y)|^2.$$

The function θ is real analytic[7] and does not vanish identically. The next lemma therefore suffices to complete the proof. □

Lemma 4.3. *Let θ be a real analytic function in a bounded connected domain Ω that does not vanish identically. For any $\tau > 0$ there exist $\gamma > 0$ and $C < \infty$ such that for all $\lambda > 0$,*

$$\iint_\Omega (1 + \lambda|\theta(x,y)|)^{-\tau} |f(x)|\,|g(y)|\,dx\,dy \le C\lambda^{-\gamma} \|f\|_{L^2}\|g\|_{L^2} \ \forall\, f,g \in L^2.$$

Proof. This is an almost immediate consequence of the uniform sublevel set inequality of Lemma 2.9.

For any x, the function $(y,x') \mapsto \psi(x,x',y) = \phi(x,y) - \phi(x',y)$ does not vanish identically. If it did, then $\phi(x',y)$ would be independent of x' for every y, contradicting the hypothesis that the mixed Hessian of ϕ does not vanish identically. Therefore there exists no x for which $\theta(x,x')$ vanishes identically as a function of x'.

Now let $\mathcal{K}(x,y) = (1 + \lambda\theta(x,y))^{-\tau}$. By Lemma 2.9, there exist $C < \infty$ and $\rho > 0$ such that $S_x(\delta) = \{y : \theta(x,y) < \delta\}$ satisfies $|S_x(\delta)| \le C\delta^\rho$ uniformly in x, δ. Choose $\delta = \lambda^{-1/2}$. Then

$$\int_{S_x(\delta)} \mathcal{K}(x,y)\,dy \le |S_x(\delta)| \le C\lambda^{-\rho/2}$$

while

$$\int_{y \notin S_x(\delta)} \mathcal{K}(x,y)\,dy \le C \int_{y \notin S_x(\delta)} \lambda^{-\tau/2}\,dy = O(\lambda^{-\tau/2}).$$

Therefore $\sup_x \int \mathcal{K}(x,y)\,dy = O(\lambda^{-c})$ for a certain $c > 0$. Since $\mathcal{K}(x,y) \equiv \mathcal{K}(y,x)$, the same integral bound holds with the roles of the two variables interchanged. From (4.5) we conclude that

$$\int \mathcal{K}(x,y)\,|f(x)|\,|f(y)|\,dx\,dy = O(\lambda^{-c})\|f\|_{L^2}\|g\|_{L^2}.$$

□

We have proved a bound for an oscillatory integral by reducing matters to a sublevel set inequality. This process is a theme that runs through these lectures. □

5. The third kind

This is a catch-all term (invented for these lectures) which refers primarily to oscillatory integrals of the second kind that involve singular kernels, or involve suprema over some parameters. Some of these are of enormous importance, but they are not the focus of these lectures. I'll review a few canonical examples, and then move on.[8]

Example. The most classical example is the (sequence of) Dirichlet kernels. The partial sums

$$S_N(f)(x) = \sum_{n=-N}^{N} \widehat{f}(n)\, e^{inx}$$

of the Fourier series of a periodic function of one variable are represented by convolution $S_N(f) = (2\pi)^{-1} \int_{\mathbb{R}/2\pi\mathbb{Z}} f(y)\, D_N(x-y)\,dy$ with

$$D_N(x) = \frac{\sin((N+\frac12)x)}{\sin(\frac12 x)}.$$

[7] This analyticity is the benefit of replacing a supremum over y by evaluation at a single point $\bar{y}$.

[8] In the oral lectures, I'll say even less about the third kind in order to save time.

D_N is highly oscillatory for large N, but there is an additional singularity at $x = 0$ due to the vanishing denominator $\sin(x/2)$. Riesz's fundamental theorem about L^p norm convergence of Fourier series can be formulated as the inequality

$$|\iint D_N(x-y)\, f(y)\, g(x)\, dy\, dx| \le C\|f\|_{L^p}\|g\|_{L^q}|$$

for all dual pairs of exponents $p, q \in (1, \infty)$.

Example. The maximal operator that arises in Carleson's theorem on almost everywhere convergence of Fourier series (formulated in the nonperiodic version) is

$$S^* f(x) = \sup_{\lambda \in \mathbb{R}} \Big| \int_{\mathbb{R}} f(y)\, (x-y)^{-1}\, e^{i\lambda y}\, dy \Big|.$$

Here one has a singular factor $(x-y)^{-1}$, an oscillatory factor $e^{i\lambda y}$, and a supremum over all parameters λ. The Carleson-Hunt theorem states that

$$\|S^* f\|_{L^p(\mathbb{R})} \le C_p \|f\|_{L^p(\mathbb{R})}$$

for every $p \in (1, \infty)$. From this it follows in a few lines that $S_N(f) \to f$ almost everywhere on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ as $N \to \infty$, for all $f \in L^p(\mathbb{T})$.

Example. The initial value problem for the time-dependent Schrödinger equation is

$$iu_t = \Delta_x u \quad \text{with } u(0, x) = f(x)$$

where $(t, x) \in (0, \infty) \times \mathbb{R}^d$, the initial datum $f : \mathbb{R}^d \to \mathbb{C}$ is given, and u is the unknown. Up to constant factors which I have not checked, there is an explicit formula for the solution as an oscillatory integral of the second kind:

$$u(t, x) = c_1 \int_{\mathbb{R}^d} t^{-d/2} f(y)\, e^{ic_2(x-y)^2/t}\, dy$$

for certain explicit nonzero constants c_1, c_2. In studying convergence of the solution $u(t, \cdot)$ to the initial datum f, one is led to the maximal operator

$$\mathcal{M}(f)(x) = \sup_{t>0} |u(t, x)|.$$

This is an oscillatory integral operator of the third kind. Carleson asked for which exponents s it is true that $u(t, x) \to f(x)$ for almost every $x \in \mathbb{R}^d$, for every f in the Sobolev space $W^{2,s}(\mathbb{R}^d)$ of functions having s derivatives[9] in L^2. This problem has been solved in the last few years — a major advance — but again, is not the focus of these lectures.

Example. Damped oscillatory integrals (of the second kind)

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} e^{i\lambda\phi(x,y)}\, f(x)\, g(y)\, a(x, y)\, \eta(x, y)\, dx\, dy$$

incorporate a smooth factor a which vanishes where the determinant of the mixed Hessian $D^2_{x,y}\phi(x, y)$ vanishes, thus compensating for deficiency of the phase ϕ. A goal is to find sharp hypotheses on the pair (ϕ, a) which guarantee a bound $O(\lambda^{-d/2}\|f\|_{L^2}\|g\|_{L^2})$.

[9]That is, $\int |\widehat{f}(\xi)|^2(1 + |\xi|^2)^s\, d\xi < \infty$.

6. The fourth kind

In this fourth category I place multilinear analogues of oscillatory integrals of the second kind. Recall that oscillatory integrals of the second kind can be regarded either as linear operators, or as bilinear scalar-valued forms. In the higher-order multilinear context, I will stick with the scalar-valued formulation, considering forms

$$\mathcal{T}_\lambda(\mathbf{f}) = \int_{(\mathbb{R}^d)^n} e^{i\lambda\psi(\mathbf{x})} \prod_{j=1}^{n} f_j(x_j)\, \eta(\mathbf{x})\, d\mathbf{x} \tag{6.1}$$

with $n \geq 3$ and $\mathbf{x} = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n$. Here $\mathbf{f} = (f_1, \ldots, f_n)$, $\eta \in C_0^\infty$ is a compactly supported smooth cutoff function, ϕ is a smooth real-valued phase function, and λ is a large positive real parameter. The functions f_j are assumed to be measurable and sufficiently bounded to ensure integrability, but no smoothness of any type is assumed of these.

For trilinear forms

$$\mathcal{T}_\lambda(f_1, f_2, f_3) = \int_{[0,1]^3} e^{i\lambda\psi(x_1,x_2,x_3)} \prod_{j=1}^{3} f_j(x_j)\, d\mathbf{x} \tag{6.2}$$

with ψ real analytic, an inequality

$$|\mathcal{T}_\lambda(\mathbf{f})| \leq C\lambda^{-\gamma} \prod_j \|f_j\|_2$$

follows at once by applying the bilinear theory to

$$\int_{[0,1]^2} e^{i\lambda\psi(x_1,x_2,x_3)} \prod_{j=1}^{2} f_j(x_j)\, dx_1\, dx_2$$

with x_3 fixed, then integrating the bound so obtained with respect to x_3. The same can be done with the roles of the indices $1, 2, 3$ permuted arbitrarily. Thus either ψ is equivalent to 0 in the sense that it can be decomposed as $\psi = \sum_{j=1}^{3} \psi_j(x_j)$, in which case no bound at all holds with a negative power of λ, or an upper bound of the form $C\lambda^{-\gamma} \prod_j \|f_j\|_{L^2}$ follows from the bilinear theory. So long as one does not pursue the questions of optimal, or near-optimal, values of the exponent γ, that answers our basic question. For bounds for (6.2) that do not follow from the bilinear theory, see [37] and [15].

We will therefore focus on more singular versions

$$\mathcal{T}_\lambda(\mathbf{f}) = \int_{\mathbb{R}^D} e^{i\lambda\psi(x)} \prod_{j\in J} (f_j \circ \varphi_j)(x)\, \eta(x)\, dx \tag{6.3}$$

with J a finite index set of cardinality $|J| \geq 3$ and with $\varphi_j : \mathbb{R}^D \to \mathbb{R}^d$ smooth submersions. Such a form may also be regarded (modulo technicalities) as an integral over a submanifold $\Lambda \subset (\mathbb{R}^d)^J$ of dimension D:

$$\int_\Lambda e^{i\lambda\Psi(\mathbf{x})} \prod_{j\in J} f_j(x_j)\, \eta(\mathbf{x})\, d\sigma(\mathbf{x})$$

with redefined phase Ψ, with cutoff function η, and with Λ having positive codimension. The latter perspective emphasizes two governing features. First, only parameters $(x_j : j \in J)$ in a lower-dimensional set directly interact, in contrast to our discussion of forms of the second kind, in which all pairs (x, y) in the Cartesian product $\mathbb{R}^d \times \mathbb{R}^d$ of the domains of f, g appear. Second, there is an interaction between the phase Ψ and the submanifold Λ.

We'll sometimes drop the cutoff function η, and instead will integrate over a ball, or a rectangle. This formulation is essentially equivalent for our purpose.

Perhaps the simplest examples are

$$\mathcal{T}_\lambda(f_1, f_2, f_3) = \int_{B\subset\mathbb{R}^2} e^{i\lambda(x-y)^n}\, f_1(x)\, f_2(y)\, f_3(x+y)\, dx\, dy \tag{6.4}$$

with $f_j : \mathbb{R}^1 \to \mathbb{C}$ Lebesgue measurable. We are potentially interested in upper bounds of the form

$$|\mathcal{T}_\lambda(\mathbf{f})| \le C\lambda^{-\gamma} \prod_{j=1}^{3} \|f_j\|_{L^{p_j}} \tag{6.5}$$

with $\gamma > 0$. What is known about such bounds for these simplest examples? Relatively little. First, for $n = 1$, there is no decay, that is, (6.5) does not hold for any $\gamma > 0$, not even with the strongest norm, L^∞, placed on each factor f_j. Indeed,

$$e^{i\lambda(x-y)^1}\, f_1(x)\, f_2(y)\, f_3(x+y) = \tilde f_1(x)\, \tilde f_2(y)\, f_3(x+y)$$

with

$$\tilde f_1(x) = e^{i\lambda x} f_1(x) \ \text{ and } \ \tilde f_2(y) = e^{-i\lambda y} f_2(y),$$

completely eliminating the oscillatory factor without changing L^∞ norms.

The same holds for $n = 2$. Now

$$(x-y)^2 = 2x^2 + 2y^2 - (x+y)^2; \tag{6.6}$$

thus $e^{i\lambda(x-y)^2}$ is expressed as a product of three factors, each of which can be incorporated into one of the three factors f_j.

We come to the case $n \ge 3$.

Proposition 6.1. *For each $\mathbb{N} \ni n \ge 3$ there exists $\gamma = \gamma(n) > 0$ for which the forms $\mathcal{T}_\lambda$ defined in* (6.5) *satisfy*

$$|\mathcal{T}_\lambda(\mathbf{f})| \le C\lambda^{-\gamma} \prod_{j=1}^{3} \|f_j\|_{L^\infty}. \tag{6.7}$$

This is proved below. One has a bound with $\gamma = 0$ whenever two exponents p_j are equal to 1 and the third is equal to ∞. Thus by (multilinear) interpolation[10] one obtains from the endpoint case $\mathbf{p} = (\infty,\infty,\infty)$ an inequality with $\gamma = \gamma(\mathbf{p}) > 0$ for any $\mathbf{p}$ with each $p_j > 1$ and $\sum_j p_j^{-1} < 2$. Thus so long as one does not seek optimal decay exponents γ, one might as well restriction to the case where each function f_j is measured via the L^∞ norm in studying inequalities of the type in Proposition 6.1.

In the simplest case $n = 3$, the forms (6.4) have been been studied in a series of papers, including [20], [23], [27], [47], [34], [24].

With L^∞ norms on the right-hand side, (6.5) does not hold for any $\gamma > \frac{1}{3}$; see the next paragraph. It holds for $\gamma = \frac{1}{4}$, with respect to $L^p \times L^p \times L^p$ for a certain $p < \infty$. But it remains entirely open for $\frac{1}{4} < \gamma \le \frac{1}{3}$. An optimal exponent γ (indeed, this exponent is $\frac{1}{4}$) is known for certain finite $\mathbf{p}$, but it remains unknown whether the stronger L^∞ hypothesis results in faster decay.

[10]The proof of the classical Riesz-Thorin complex interpolation theorem applies, almost verbatim.

To see that (6.5) cannot hold for any $\gamma > \frac{1}{3}$ when $n = 3$, consider any smooth, compactly supported f_1, f_2, and choose f_3 so that $f_3(x+y) \equiv 1$ where $f_1(x)f_2(y) \neq 0$. Change variables $(x,y) \mapsto (u,v) = (x-y, x+y)$ to obtain

$$\frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{i\lambda v^3}\,\eta(\tfrac{1}{2}(u+v), \tfrac{1}{2}(v-u)\,du\,dv.$$

The inner integral takes the form

$$c\lambda^{-1/3}\eta(\tfrac{1}{2}v, \tfrac{1}{2}v)) + O(\lambda^{-2/3})$$

so we obtain

$$c\lambda^{-1/3}\int_{\mathbb{R}} \eta(v/2, v/2)\,dv + O(\lambda^{-1/3}).$$

If $\eta \geq 0$ does not vanish identically then the coefficient of the leading term is nonzero.

We have learned three lessons. Firstly, oscillation can be an illusion; one must beware of additive relations such as (6.6), which can reveal that a phase function is degenerate in the sense that it can be rewritten as a multilinear oscillatory form in an equivalent way with no oscillatory factor at all. Secondly, it need not be immediately apparent whether a particular phase function is degenerate in this sense; consider the cases $n = 2$ versus $n = 3$ above. Thirdly, even in the very simplest examples, optimal exponents are unknown in the multilinear (that is, $|J| \geq 3$) theory.

We now prove that $|\mathcal{T}_\lambda(\mathbf{f})| \leq C\lambda^{-\gamma}\prod_{j=1}^3 \|f_j\|_\infty$ for $n \geq 3$, for some $\gamma = \gamma_n > 0$.

Proof of Proposition 6.1. By Cauchy-Schwarz,

$$|\mathcal{T}_\lambda(\mathbf{f})|^2 \leq C\|f_1\|_2^2 \iiint e^{i\lambda[(x-y')^n - (x-y)^n]} f_2(y')\overline{f_2}(y) f_3(x+y')\overline{f_3}(x+y)\zeta(x,y,y')\,dx\,dy\,dy'$$

for a certain $\zeta \in C_0^\infty(\mathbb{R}^3)$. Substituting $(x,y',y) \mapsto (z,y,u)$ with $y' = y+u$ and $z = x+y$ allows us to rewrite the triple integral on the right-hand side as

$$\int_{\mathbb{R}} \Big(\iint_{\mathbb{R}^2} e^{i\lambda P_u(y,z)} F_2^u(y) F_3^u(z)\tilde{\zeta}(y,z,u)\,dy\,dz\Big)\,du$$

with $\tilde{\zeta}$ smooth and compactly supported, with

$$\begin{aligned} F_2^u(y) &= f_2(y')\overline{f_2}(y) = f_2(y+u)\overline{f_2}(y) \\ F_3^u(z) &= f_3(x+y')\overline{f_3}(x+y) = f_3(z+u)\overline{f_3}(z) \\ P_u(y,z) &= (x-y')^n - (x-y)^n = (z-2y-u)^n - (z-2y)^n. \end{aligned}$$

The inner integral is an oscillatory integral of the second kind. The Hessian matrix is the 1×1 matrix with entry

$$\frac{\partial^2}{\partial y\,\partial z} P_u(y,z) = -2n(n-1)[(z-2y-u)^{n-2} - (z-2y)^{n-2}].$$

This vanishes identically for $n = 1$ and for $n = 2$, but for $n \geq 3$ one has

$$\frac{\partial^{n-3}}{\partial z^{n-3}}\frac{\partial^2}{\partial y\,\partial z} P_u(y,z) = 2n!\,u.$$

By Proposition 4.2 (more precisely, by its proof) there exist $C < \infty$ and $\tau > 0$ such that for every $u \in \mathbb{R}$,

$$\Big|\iint_{\mathbb{R}^2} e^{i\lambda P_u(y,z)} F_2^u(y) F_3^u(z)\tilde{\zeta}(y,z,u)\,dy\,dz\Big| \leq C(1+\lambda|u|)^{-\tau}\|F_1^u\|_{L^2}\|F_2^u\|_{L^2}.$$

To complete the proof, it suffices to observe that $\|F_j^u\|_{L^2} \le C\|f_j\|_{L^\infty}^2$ and

$$\int_B (1+\lambda|u|)^{-\tau}\,du \le C_{\tau,B}\min(\lambda^{-\tau},\,\lambda^{-1})$$

for any ball $B \subset \mathbb{R}^1$. □

7. More about the fourth kind

Li, Tao, Thiele, and the author [9] investigated multilinear oscillatory integrals in the singular case $|J|d > D$ in 2005. Their theory was developed for the special case of polynomial phases ϕ. Much of the theory was extended to real analytic phases by Greenblatt [26], but we maintain the restriction to polynomial phases in these notes. We do this in preparation for the discussion below of oscillatory integrals of the fifth kind — which is the main focus of these lectures.

Consider

$$\mathcal{T}_\lambda(\mathbf{f}) = \int_{\mathbb{R}^D} e^{i\lambda P(x)} \prod_{j\in J} (f_j \circ L_j)(x)\,\eta(x)\,dx \tag{7.1}$$

where J is a finite index set, $L_j : \mathbb{R}^D \to \mathbb{R}^d$ are linear and surjective, $\eta \in C_0^\infty$ is a smooth cutoff function, and $P : \mathbb{R}^D \to \mathbb{R}^1$ is a real-valued polynomial.

We seek upper bounds of the form

$$|\mathcal{T}_\lambda(\mathbf{f})| \le C\lambda^{-\gamma} \prod_{j\in J} \|f_j\|_{L^\infty}, \tag{7.2}$$

with the unimportant constant $C < \infty$ permitted to depend on η. As noted above, combining such a bound with more trivial bounds without decaying factors via complex interpolation leads to a wider family of bounds with $\lambda^{-\gamma(\mathbf{p})} \prod_{j\in J} \|f_j\|_{L^{p_j}}$ on the right-hand side.

We are potentially also interested in weaker bounds

$$|\mathcal{T}_\lambda(\mathbf{f})| \le C\theta(\lambda) \prod_{j\in J} \|f_j\|_{L^\infty} \tag{7.3}$$

where $\theta(\lambda) \to 0$ as $\lambda \to \infty$, in those situations in which (7.2) is either untrue, or unknown.

Definition 7.1. *P is degenerate (relative to the family $\{L_j : j \in J\}$) if there exist polynomials Q_j such that*

$$P \equiv \sum_{j\in J} (Q_j \circ L_j). \tag{7.4}$$

P is of course said to be nondegenerate (relative to $\{L_j : j \in J\}$ if it is not degenerate.

It is shown in [9] that P is degenerate in this sense, if and only if it is degenerate in the alternative sense that there exist *distributions* that satisfy $P \equiv \sum_{j\in J}(h_j \circ L_j)$.

As was observed above, if P is degenerate then even the weaker bound (7.3) cannot hold.

Proposition 7.1. [9] *P is nondegenerate if and only if there exists a constant-coefficient differential operator $\mathcal{L}$ that satisfies*

$$\begin{cases} \mathcal{L}(P) \ne 0 \\ \mathcal{L}(g \circ L_j) \equiv 0 \ \forall\, g \in C^\infty,\ \forall\, j \in J. \end{cases} \tag{7.5}$$

We say that $\mathcal{L}$ witnesses the nondegeneracy of P. Thus Proposition 7.1 says that nondegeneracy is always witnessed by some constant-coefficient partial differential operator.

Obviously the existence of such an $\mathcal{L}$ implies the degeneracy of P, but the converse also holds. This provides a useful criterion for nondegeneracy.

Example 7.2. *Let $D = 3$, $d = 1$, and $P(x_1, x_2, x_3) = x_3^2$. Let v_j be nonzero vectors in the light cone in $\mathbb{R}^3$; thus*

$$v_{3,j}^2 = v_{1,j}^2 + v_{2,j}^2.$$

Let $L_j(x) = x \cdot v_j$. The d'Alembertian operator

$$\Box = \frac{\partial^2}{\partial x_3^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2}$$

witnesses the nondegeneracy of P.

In this example, the index set J can be chosen to have arbitrarily large cardinality, though the phase P has degree 2. The method used above in the proof of Proposition 6.1 does not apply here unless $|J| \le 2$.

Question 7.3. *Consider any $D, d, P, J, \{L_j\}$. If P is nondegenerate relative to $\{L_j\}$, does the inequality* (7.2) *hold for some C, γ? Does at least the weaker inequality* (7.3) *hold for some function θ that tends to zero?*

We have already seen, in the discussion of oscillatory integrals of the first kind, that there are closely allied inequalities concerning sublevel sets. To the above data $D, d, P, J, \{L_j\}$, any $\varepsilon > 0$, any ball $B \subset \mathbb{R}^D$, and any measurable functions $f_j : \mathbb{R}^d \to \mathbb{R}$ associate the sublevel sets

$$S(\mathbf{f}, \varepsilon) = \{x \in B : \big|P(x) - \sum_{j\in J} (f_j \circ L_j)(x)\big| < \varepsilon\}. \tag{7.6}$$

The sublevel analogues of (7.2) and (7.3) are respectively

$$\big|S(\mathbf{f}, \varepsilon)\big| \le C\varepsilon^\gamma \tag{7.7}$$

$$\big|S(\mathbf{f}, \varepsilon)\big| \le \theta(\varepsilon) \quad \text{for some } \theta \tag{7.8}$$

with θ denoting an auxiliary function that depends only on $D, d, P, J, \{L_j\}, B$ and satisfies $\lim_{\varepsilon\to 0} \theta(\varepsilon) = 0$.

Question 7.4. *Consider any $D, d, P, J, \{L_j\}, B$. If P is nondegenerate relative to $\{L_j\}$, does the inequality* (7.7) *hold for some C, γ? Does at least the weaker inequality* (7.8) *hold for some function θ that tends to zero?*

An affirmative answer to either part of Question 7.3 implies an affirmative answer to the corresponding part of Question 7.4, by the reasoning that we have already discussed in §3. Indeed, fix an auxiliary Schwartz function $\varphi \ge 0$ satisfying $\varphi(0) > 0$. Inequality (7.7) is completely equivalent to

$$\int_B \varphi(\varepsilon^{-1}(P - \sum_j (f_j \circ L_j))) = O(\varepsilon^\gamma).$$

the integral equals

$$\int_{\mathbb{R}} \widehat{\varphi}(\xi) \Big(\int_B e^{i\varepsilon^{-1}\xi(P(x) - \sum_j (f_j(L_j(x))} \, dx \Big) \, d\xi.$$

The inner integral can be rewritten as

$$\int_B e^{i\lambda P(x)} \prod_j (g_j \circ L_j)(x)\, dx$$

with $\lambda = \varepsilon^{-1}\xi$ and with $g_j = e^{if_j} = O(1)$ in L^∞ norm. If (7.2) holds then (slurring over a technical point, the presence of a factor of $\mathbf{1}_B$ rather than of a smooth cutoff function η) this inner integral is $O(1+\varepsilon^{-1}|\xi|)^{-\gamma}$. The inequality (7.7) follows.

In the same way, (7.8) follows from (7.3).

One of the two main results of [9] answers Question 7.3 satisfactorily in the codimension one case $d = D-1$.

Theorem 7.2. [9] *Consider any $D, d, P, J, \{v_j\}$ with $d = D-1$. Let $L_j(x) = \langle x, v_j\rangle$. If P is nondegenerate then* (7.2) *holds for some $C < \infty$ and $\gamma > 0$.*

In this situation, the degree of P must be $\ge |J|$ for nondegeneracy to hold (provided that L_i is never a scalar multiple of L_j unless $i = j$). The method of proof is the same as that of Proposition 6.1, with a sequence of applications of Cauchy-Schwarz eliminating factors $f_j \circ L_j$ one by one.

At the opposite extreme, when D is arbitrary and $d = 1$, Question 7.3 remains largely open. The following very partial result is the second main theorem of [9].

When $d = 1$, we write each L_j in the form

$$L_j(x) = x \cdot v_j$$

for a unique vector $v_j \in \mathbb{R}^D$.

Theorem 7.3. [9] *Consider any $D, d, P, J, \{v_j\}$ with $d = 1$. Suppose that for any $J' \subset J$ of cardinality $\le D$, $\{v_j : j \in J'\}$ is linearly independent. If P is nondegenerate, and if $|J| < 2D$, then* (7.2) *holds.*

Even in the example (7.2) above, it remains an open question whether (7.2) holds for $|J| = 6 = 2D$.

In § 9 I outline a proof of Theorem 7.3 by a method different from the (significantly simpler!) original proof of [9]. This discussion introduces a technical device (local Fourier decomposition), a theme (decomposition into structured and "pseudorandom" components), and a second theme (reduction of an oscillatory inequality to a related sublevel set inequality) that will all be relied upon in our discussion of oscillatory integrals of the fifth kind.

8. A naive attack

We discuss an unsuccessful attack on Theorem 7.3. This may help to motivate the successful, but more complicated, proof in §9.

Assuming as we may that the support of η has small diameter, expand each f_j in Fourier series

$$f_j(y) = \sum_{n\in\mathbb{Z}} \widehat{f_j}(n)\, e^{2\pi i n y}$$

and insert each of these representations to obtain

$$\mathcal{T}_\lambda(\mathbf{f}) = \sum_{n_1,n_2,\dots,n_{|J|}} \prod_{k\in J} \widehat{f}_k(n_k) \int e^{i\lambda P(x)} e^{2\pi i \sum_j n_j v_j\cdot x}\, \eta(x)\, dx.$$

Denoting the integral by $I_\lambda(\mathbf{n})$ with $\mathbf{n} = (n_j : j \in J)$, this gives

$$|\mathcal{T}_\lambda(\mathbf{f})| \le \sum_{\mathbf{n}\in\mathbb{Z}^J} |I_\lambda(\mathbf{n})| \prod_{k\in J} |\widehat{f}_k(n_k)|. \tag{8.1}$$

$I_\lambda(\mathbf{n})$ is an oscillatory integral of the first kind, depending on a parameter $\mathbf{n}$. The theory of oscillatory integrals of the first kind gives the promising upper bound

$$|I_\lambda(\mathbf{n})| \le C\lambda^{-\gamma} \text{ uniformly in } \mathbf{n}.$$

On the other hand, the multiple infinite series

$$\sum_{\mathbf{n}\in\mathbb{Z}^J} \prod_{k\in J} |\widehat{f}_k(n_k)|$$

converges if and only if every f_k has Fourier coefficients in ℓ^1 (or if some factor vanishes identically) — far too strong a condition to be useful since in typical situations, $|I_\lambda(\mathbf{n})| \ge c\lambda^{-\gamma}$ uniformly for all $(\lambda, \mathbf{n})$ satisfying $|\mathbf{n}| \le c\lambda$. Indeed, the net frequency $\sum_j n_j\xi_j v_j \in \mathbb{R}^D$ in the integral defining $I_\lambda(\mathbf{n})$ is then small relative to $\lambda|\nabla P|$, so no additional cancellation can be expected.

The Fourier coefficients belong to ℓ^2, but no better, even if $f_j \in L^\infty$. Suppose that each f_j satisfies $|\widehat{f}_j(n_j)| \ge c\lambda^{-1/2}$ for all $|n_j| \le c\lambda$, as is consist with Parseval's relation.

$$\sum_{|\mathbf{n}|\le c\lambda} \prod_{k\in J} |\widehat{f}_k(n_k)| \gtrsim \lambda^D \lambda^{-D/2} = \lambda^{D/2}.$$

Thus in order for this analysis to succeed, we would need $\gamma > D/2$. That's too much; not a single example is known of a phase with at least one critical point satisfying a bound with such a large exponent.[11] This method of proof seems to break down irrevocably. Nonetheless, a modification will succeed. See §9.

9. Analysis of certain trilinear (explicitly) oscillatory integrals

Set $r = \frac{3}{4}$. Any value in $(\frac{1}{2}, 1)$ will work, and $3/4$ is chosen only for the sake of concreteness. I will most often write r, rather than the value $3/4$, in an attempt to improve the readability of the formulas and to remind the reader of where various quantities originate. Let $J = \{1, 2, \dots, 2D-1\}$.

Assume $f_j \in L^\infty$ with $\|f_j\|_\infty \le 1$ for each $j \in J$; we seek a bound $\mathcal{T}_\lambda(\mathbf{f}) = O(\lambda^{-\gamma})$. Partition $\mathbb{R}^1$ into intervals I_m of lengths $|I_m| = \lambda^{-r}$. Decompose each f_j in local Fourier series of the type

$$f_j(y) = \sum_{m\in\mathbb{Z}} \eta_m(y) \sum_{n\in\mathbb{Z}} a_{j,m}(n) e^{\pi i\lambda^r ny}$$

where each η_m is supported in the interval I_m^* concentric with I_m satisfying $|I_m^*| = 3|I_m|$ and satisfies natural bounds;

$$\|\frac{d^k}{dx^k}\eta_m\|_{C^0} \le C_k\lambda^{rk} \ \forall\, k \ge 0,$$

and

$$\sum_n |a_{j,m}(n)|^2 = O(1).$$

The latter bound comes from Parseval's theorem, an affine change of variables for each m, and the assumption that $\|f_j\|_{L^\infty} = O(1)$.

[11] In this connection, recall the question of Zelditch mentioned above.

Repeating the above analysis leads to an upper bound for $|\mathcal{T}_\lambda(\mathbf{f})|$ of the form

$$\sum_{\mathbf{m}}\sum_{\mathbf{n}}\prod_{k\in J}|a_{k,m_k}(n_k)|\cdot|I_\lambda(\mathbf{n},\mathbf{m})| \tag{9.1}$$

with

$$I_\lambda(\mathbf{n},\mathbf{m})=\int_{\mathbb{R}^D}e^{i\lambda P(x)}e^{i\pi\lambda^r\sum_{j\in J}n_jv_j\cdot x}\,\zeta_{\mathbf{m}}(x)\,dx \tag{9.2}$$

with cutoff functions $\zeta_{\mathbf{m}}(x)=\eta(x)\prod_j\eta_{m_j}(x\cdot v_j)$.

For most $\mathbf{m}$, $\zeta_{\mathbf{m}}$ will vanish identically; $\eta_{m_j}(x\cdot v_j)=0$ unless $x\cdot v_j$ lies in the small interval I_{m_j}. If $\zeta_{\mathbf{m}}$ does not vanish identically, we say that the tuple of indices $\mathbf{m}$ is interacting. For each interacting tuple, choose $\bar x_{\mathbf{m}}$ in the support of $\zeta_{\mathbf{m}}$. For any x in the support of $\zeta_{\mathbf{m}}$,

$$|\lambda\nabla P(x)-\lambda\nabla P(\bar x)|=O(\lambda\cdot\lambda^{-3/4})=O(\lambda^{1/4})\ll\lambda^{3/4}=\lambda^r.$$

Thus (provided that λ is large) if

$$|\lambda\nabla P(\bar x_{\mathbf{m}})+\pi\lambda^r\sum_j n_jv_j|\ge\lambda^r \tag{9.3}$$

then

$$|\lambda\nabla P(x_{\mathbf{m}})+\pi\lambda^r\sum_j n_jv_j|\ge\tfrac12|\lambda\nabla P(\bar x_{\mathbf{m}})+\pi\lambda^r\sum_j n_jv_j| \tag{9.4}$$

for every point x in the support of $\zeta_{\mathbf{m}}$. We have a trivial bound

$$|I_\lambda(\mathbf{n},\mathbf{m})|\le C\lambda^{-Dr}.$$

From repeated integrations by parts there follows

$$\begin{aligned}|I_\lambda(\mathbf{n},\mathbf{m})|&\le C_N\lambda^{-Dr}(1+\lambda^{-r}|\lambda\nabla P(\bar x_{\mathbf{m}})+\pi\lambda^r\sum_j n_jv_j|)^{-N}\\&=C_N\lambda^{-Dr}(1+|\lambda^{1-r}\nabla P(\bar x_{\mathbf{m}})+\pi\sum_j n_jv_j|)^{-N}\end{aligned} \tag{9.5}$$

for every $N<\infty$.

We say that an interacting index $\mathbf{m}$ is *nonstationary* if

$$|\lambda^{1-r}\nabla P(\bar x_{\mathbf{m}})+\pi\sum_j n_jv_j|\ge\lambda^\rho. \tag{9.6}$$

We have shown that if $\mathbf{m}$ is nonstationary then

$$|I_\lambda(\mathbf{n},\mathbf{m})|\le C_N\lambda^{-N}\ \forall\,N<\infty. \tag{9.7}$$

Now return to (9.1) and consider the sum on the right-hand side. Each noninteracting index $\mathbf{m}$ contributes 0 since then $I_\lambda(\mathbf{n},\mathbf{m})=0$ for every $\mathbf{n}$. For nonstationary interacting $\mathbf{m}$ we have the very favorable bound $|I_\lambda(\mathbf{n},\mathbf{m})|=O(\lambda^{-N})$ for every N, but this is not sufficient since (as we have seen in the unsuccessful analysis above!) $\sum_{\mathbf{n}}\prod_{k\in J}|a_{k,m_k}(n_k)|$ is far from being absolutely convergent given the only available information, which is that $a_{k,m}\in\ell^2$ uniformly in k,m.

The key is that (9.5) contains additional information, beyond (9.7), that is favorable for typical $\mathbf{n}$.

Lemma 9.1. *Let $\{v_j : 1 \le j \le 2D-1\}$ be a collection of vectors in $\mathbb{R}^D$. Suppose that any D of these are linearly independent. Then there exist $C < \infty$ and $\tau \in (0,1)$ such that for any sequences $a_k \in \ell^2$, for any $b \in \mathbb{C}$,*

$$\sum_{\mathbf{n}\in\mathbb{Z}^{2D-1}} (1 + \big|b - \sum_{k=1}^{2D-1} n_k v_k\big|)^{-2D} \prod_{k=1}^{2D-1} |a_k(n_k)| \le C \prod_k \|a_k\|_{\ell^2}^{1-\tau} \prod_k \|a_k\|_{\ell^\infty}^{\tau}. \tag{9.8}$$

The particular exponent $2D$ is of no particular significance; this lemma holds provided a sufficiently large exponent is written in its place, and we will invoke the lemma in a situation where (9.5) provides any exponent that we desire.

Having a bound with positive powers of ℓ^∞ norms of Fourier coefficients on the right-hand side will be decisive later in the proof. If the number of indices j is $2D$, rather than being strictly smaller than $2D$, then such a bound would not be true; this is the point at which the assumption that $|J| < 2D$ is exploited in an essential way.

Proof. It suffices to bound the left-hand side by

$$C \prod_{k\in S} \|a_k\|_{\ell^1} \prod_{k\notin S} \|a_k\|_{\ell^\infty}$$

for any subset $S \subset \{1, 2, \dots, 2D-1\}$ of cardinality D. For then complex interpolation gives the indicated bound, with $\tau = D/(2D-1)$.

Since the hypotheses and conclusions are invariant under permutations of the indices $j \in \{1, 2, \dots, 2D-1\}$, it suffices to prove this for $S = \{1, 2, \dots, D\}$. Now

$$\sum_{\mathbf{n}\in\mathbb{Z}^{2D-1}} (1 + \big|b - \sum_{k=1}^{2D-1} n_k v_k\big|)^{-2D} \prod_{k=1}^{2D-1} |a_k(n_k)|$$

$$\le \prod_{k>D} \|a_k\|_{\ell^\infty} \sum_{(n_1,\dots,n_D)\in\mathbb{Z}^D} \prod_{j\le D} |a_{j,n_j}| \sum_{(n_{D+1},\dots,n_{2D-1})\in\mathbb{Z}^{D-1}} (1+\big|\beta(b, n_1, \dots, n_D) - \sum_{k=D}^{2D-1} n_k v_k\big|)^{-D}$$

where $\beta(b, n_1, \dots, n_D) = b - \sum_{j\le D} n_j v_j$, though this formula need not concern us; all that we need is that β is independent of $(n_{D+1}, \dots, n_{2D-1})$.

The inner sum is bounded, uniformly in all $\beta(b, n_1, \dots)$, since $\{v_k : D \le k \le 2D-1\}$ is linearly independent. (The verification is left as an elementary exercise for the reader.) Therefore the right-hand side is majorized by

$$C \prod_{k>D} \|a_k\|_{\ell^\infty} \sum_{(n_1,\dots,n_D)\in\mathbb{Z}^D} \prod_{j\le D} |a_{j,n_j}| = C \prod_{k\notin S} \|a_k\|_{\ell^\infty} \prod_{j\in S} \|a_j\|_{\ell^1},$$

as claimed. □

Returning to the proof of Theorem 7.3, let $\sigma > 0$ be a small exponent to be chosen below and split each f_j as $f_j = f_j^{\diamondsuit} + f_j^{\heartsuit}$ so that the local Fourier coefficients $a_{j,m}^{\heartsuit}, a_{j,m}^{\diamondsuit}$ satisfy

$$a_{j,m}^{\heartsuit}(n) = \begin{cases} a_{j,m}(n) & \text{if } |a_{j,m}(n)| \le \lambda^{-\sigma} \\ 0 & \text{otherwise.} \end{cases} \tag{9.9}$$

Thus $\|a_{j,m}^{\diamondsuit}\|_{\ell^2} = O(1)$ and for each j, m, $a_{j,m}^{\diamondsuit}(n) = 0$ for all but $O(\lambda^{2\sigma})$ indices n. The latter is an immediate consequence of the bound $\|a_{j,m}^{\diamondsuit}\|_{\ell^2} \le \|a_{j,m}\|_{\ell^2} = O(1)$.

$\mathcal{T}_\lambda(\mathbf{f})$ is thus expressed as a sum of 2^{2D-1} terms $\mathcal{T}_\lambda(\mathbf{g})$ where each g_j is equal either to $f_j^{\heartsuit}$, or to $f_j^{\diamondsuit}$. The term with each $g_j = f_j^{\diamondsuit}$ will be discussed below. For each of the other terms, at least one g_l equals $f_l^\heartsuit$ and hence all of its local Fourier coefficent series satisfy $\|a^\heartsuit_{j,m}\|_{\ell^\infty} = O(\lambda^{-\sigma})$. It follows from the above analysis that

$$|\mathcal{T}_\lambda(\mathbf{g})| = O((\lambda^{-\sigma})^\tau) \tag{9.10}$$

for a certain exponent $\tau > 0$. For any $\sigma > 0$, this is a bound of the desired type.

It remains to bound $|\mathcal{T}_\lambda(f_1^\diamondsuit, \ldots, f_{2D-1}^\diamondsuit)|$. Each $f_j^\diamondsuit$ can be written as a sum of $O(\lambda^{2\sigma})$ local Fourier monomial functions F, each of which takes the form

$$F_j(y) = \sum_m b_{j,m} e^{i\pi\lambda^r k_{j,m} y} \eta_m(y) \tag{9.11}$$

with a Fourier series with a single term $e^{i\pi\lambda^r k y}$ for each m and with coefficients $b_{j,m} = O(1)$. By multilinearity of $\mathcal{T}_\lambda$, $\mathcal{T}_\lambda(f_1^\diamondsuit, \ldots, f_{2D-1}^\diamondsuit)$ is a sum of $O(\lambda^{4D\sigma})$ terms $\mathcal{T}_\lambda(F_1, \ldots, F_{2D-1})$ with each F_j taking the form (9.12). If we can show that there exist $C < \infty$ and $\tau > 0$ such that

$$|\mathcal{T}_\lambda(F_1, \ldots, F_{2D-1})| \le C\lambda^{-\tau} \tag{9.12}$$

for any tuple of functions of the special form (9.11), then in all we will have

$$|\mathcal{T}_\lambda(f_1^\diamondsuit, \ldots, f_{2D-1}^\diamondsuit)| \le C\lambda^{-\tau+4D\sigma}, \tag{9.13}$$

and choosing $\sigma < \tau/4D$ will complete the proof.

We have arrived at a crucial juncture in the proof. To prepare for the proof of (9.12), change notation, defining

$$g_j(y) = -\pi\lambda^{r-1} k_{j,m} \quad \text{for } y \in I_m \text{ for each } j, m. \tag{9.14}$$

Thus we view the frequencies $k_{j,m}$ as functions.

Now return to the full expansion of $\mathcal{T}_\lambda(F_1, \ldots,)$, which simplifies because there is no longer a sum over $\mathbf{n}$. Since $bj, m = O(1)$,

$$|\mathcal{T}_\lambda(F_1, \ldots, F_{2D-1})| \le C \sum_{\mathbf{m}} |I_\lambda(\mathbf{m})|$$

with

$$I_\lambda(\mathbf{m}) = \int_{\mathbb{R}^D} e^{i\lambda P(x)} e^{i\pi\lambda^r \sum_{j\in J} n_j v_j \cdot x} \zeta_{\mathbf{m}}(x)\, dx.$$

We need only sum over stationary interacting tuples $\mathbf{m}$. These satisfy

$$|\lambda^{1-r}\nabla P(\bar{x}_{\mathbf{m}}) - \lambda \sum_j g_j(L_j(\bar{x}_{\mathbf{m}}))v_j| \le \lambda^\rho$$

and therefore

$$|\nabla P(x) - \sum_j (g_j \circ L_j)(x) v_j| \le C\lambda^{r-1+\rho} + C\lambda^{-r}$$

for every x in the support of $\zeta_{\mathbf{m}}$. Because $r = \frac{3}{4}$, $\lambda^{r-1+\rho} + \lambda^{-r} = O(\lambda^{-1/8})$ provided ρ is chosen to be $< \frac{1}{8}$. Thus for every stationary interacting $\mathbf{m}$,

$$|\nabla P(x) - \sum_j (g_j \circ L_j)(x) v_j| \le C\lambda^{-1/8} \quad \text{for every } x \text{ in the support of } \zeta_{\mathbf{m}}. \tag{9.15}$$

The precise exponents are of no importance; various choices have been made to arrange that a strictly negative power of λ arises on the right-hand side of this last inequality.

The sum of the contributions of all noninteracting $\mathbf{m}$ to $\mathcal{T}_\lambda(F_1, \ldots, F_{2D-1})$ is zero. There are at most $O(\lambda^{(2D-1)r})$ tuples $\mathbf{m}$ in all, and each nonstationary $\mathbf{m}$ contributes $O(\lambda^{-M})$)

for any $M < \infty$, so the sum of the contributions of all nonstationary interacting $\mathbf{m}$ u is $O(\lambda^{-N})$ for every $N < \infty$.

Each stationary interacting $\mathbf{m}$ contributes at most $C\lambda^{-Dr}$. $\zeta_{\mathbf{m}}$ is supported on a ball of radius $O(\lambda^{-r})$ in the ambient space $\mathbb{R}^D$. *I will proceed as if these balls were pairwise disjoint.* One can very easily reduce matters to that case by partitioning the indices m_j into congruence classes modulo some fixed integer p and reducing to the situation in which for each j, only indices in a single congruence class, depending on j, arise.

Under this pretence, the sum of the contributions of all nonstationary interacting $\mathbf{m}$ is majorized by a constant multiple of the sum of the measures of these balls. Thus after a slight redefinition of the functions g_j, the sum of these contributions is majorized by a constant multiple of the measure of the sublevel set

$$S(\mathbf{g},\varepsilon) = \big\{x \in B : \big|\nabla P(x) - \sum_{j=1}^{2D-1} (g_j \circ L_j)(x)\, v_j\big| \le \varepsilon\big\} \tag{9.16}$$

where $\varepsilon = \lambda^{-1/8}$.

We have seen above that an upper bound for the oscillatory form $\mathcal{T}_\lambda$ would imply an upper bound for the measures of sublevel sets of the form $\{x \in B : |P(x) - \sum_j (h_j \circ L_j)(x)| < \delta\}$. We have now succeeded in conversely controlling the oscillatory form by the measure of an associated sublevel set — but it is a different sublevel set. Both ∇P and v_j are $\mathbb{R}^D$–valued; this is a sublevel set associated to an $\mathbb{R}^D$–valued function, rather than to a scalar function. Having a vector-valued function in the definition of the sublevel set will turn out to be a decisive advantage.

Lemma 9.2. *Let D, v_j, L_j be as above. Let $P : \mathbb{R}^D \to \mathbb{R}^1$ be nondegenerate relative to $\{L_j : 1 \le j \le 2D-1\}$. There exist $C < \infty$ and $\gamma > 0$ such that for any measurable functions g_j and any $\varepsilon > 0$, the sublevel sets $S(\mathbf{g},\varepsilon)$ defined in* (9.16) *satisfies*

$$|S(\mathbf{g},\varepsilon)| \le C\varepsilon^\gamma. \tag{9.17}$$

Proof. By a linear change of variables we can reduce matters to the case in which for each $j \le D$, $v_j = e_j$, the j-th unit coordinate vector in $\mathbb{R}^D$. Thus $L_j(x) = x_j$ for $j \le D$.

The $(2D-1) \times D$ matrix whose columns are the vectors $v_j \in \mathbb{R}^D$ has rank D, and moreover, any D of its columns are linearly independent, by assumption. By elementary row operations, $D-1$ of the D inequalities defining membership in $S(\mathbf{g},\varepsilon)$ can be used to eliminate the $D-1$ quantities $g_{D+1}, \ldots, g_{2D-1}$, leaving a single scalar inequality of the form

$$\big|Q(x) - \sum_{j=1}^{D} \alpha_j g_j(x_j)\big| = O(\varepsilon) \ \forall\, x \in S(\mathbf{g},\varepsilon)$$

for some vector $w \in \mathbb{R}^D$ and for some constants α_j, with Q equal to the polynomial $Q = (w\cdot\nabla)P$ for some $w \in \mathbb{R}^D$. Moreover, the linear independence hypothesis ensures that every α_j is nonzero. (Exercise!)

There exists $x' = (x'_2, \ldots, x'_D)$ such that $|\{x_1 : (x_1; x') \in S(\mathbf{g},\varepsilon)\}| \gtrsim |S(\mathbf{g},\varepsilon)|$. Write the inequality as

$$g_1(x_1) = \alpha_1^{-1} Q(x_1, x') + \sum_{2 \le j \le D} \alpha_1^{-1}\alpha_j g_j(x'_j) + O(\varepsilon)$$

whenever $(x_1; x') \in S(\mathbf{g},\varepsilon)$. Thus we conclude that there exist a polynomial Q_j of degree less than or equal to the degree of ∇P, and a subset $S' \subset S(\mathbf{g},\varepsilon)$ satisfying

$$|S'| \gtrsim |S(\mathbf{g},\varepsilon)|^2,$$

such that $g_1(x_1) = Q_1(x_1) + O(\varepsilon)$ whenever $x \in S'$.

Replacing $S(\mathbf{g}, \varepsilon)$ by S' and repeating this argument with the roles of the indices $j = 1, 2$ interchanged yields the same conclusion for $j = 2$, and this can be repeated D times. Moreover, the initial setup was invariant under permutation of all $2D - 1$ indices j, so with $D - 1$ more iterations this process can be applied to all indices j. The conclusion is that there exist a measurable set $S^* \subset S(\mathbf{g}, \varepsilon)$ satisfying $|S^*| \gtrsim |S(\mathbf{g}, \varepsilon)|^c$ with $c = 2^{2D-1}$ and polynomials Q_j of degrees not exceeding the degree of ∇P such that

$$\big|\nabla P(x) - \sum_{j=1}^{2D-1} (Q_j \circ L_j)(x)\, v_j\big| = O(\varepsilon) \ \forall\, x \in S^*.$$

Equivalently,

$$R(x) = \sum_{k=1}^{D} \big(e_k \cdot \nabla P(x) - \sum_{j=1}^{2D-1} (Q_j \circ L_j)(x)\, v_j\big)^2$$

satisfies

$$R(x) = O(\varepsilon^2) \ \forall\, x \in S^*.$$

R is a polynomial, whose degree is at most $\mathcal{D}$, twice the degree of P.

R does not vanish identically. If it did, then $\nabla P \equiv \sum_{j=1}^{2D-1} (Q_j \circ L_j) v_j$ for some polynomials $Q_j : \mathbb{R}^1 \to \mathbb{R}$. Let p_j be an antiderivative of Q_j and let $h = P - \sum_j (p_j \circ L_j)$. Then $\nabla h = \nabla P - \sum_j (Q_j \circ L_j) v_j = 0$, so h is constant. Therefore upon adding h to p_1 we obtain $P = \sum_j (p_j \circ L_j)$, contradicting the nondegeneracy hypothesis — which had not been used in the proof until this point!

The quantity $\|R\|$ is not only nonzero; it is bounded below by a strictly positive constant, uniformly for all choices of the Q_j of degrees $\le \mathcal{D}$. For ∇P does not belong to the finite-dimensional vector space of all $\mathbb{R}^D$-valued polynomials $\sum_j (Q_j \circ L_j) v_j$ with Q_j of degree $\le \mathcal{D}$, so the distance (with respect to any fixed norm) from ∇P to that subspace is strictly positive. $\|R\|$ is comparable to the square of that distance.

Therefore by Lemma 2.9 concerning sublevel sets, there exist $C < \infty$ and $\tau > 0$, which depend only on D, B, and the degree of P, such that for any $\delta > 0$

$$\big|\big\{x \in B : R(x) \le \delta \|R\|\big\}\big| \le C\delta^\tau \tag{9.18}$$

where $\|\cdot\|$ denotes any fixed norm on the space of polynomials of degree at most $\mathcal{D}$. Thus

$$\big|\big\{x \in B : R(x) \le C\varepsilon^2\big\}\big| \le C'\varepsilon^{2\tau} \|R\|^{-\tau}. \tag{9.19}$$

This completes the proof of Lemma 9.2, and with it, the proof of Theorem 7.3. □

10. On the rationally commensurate case

Definition 10.1. $\{L_j : j \in J\}$ *is* rationally *commensurate if there exist choices of a basis for* $\mathbb{R}^D$*, and for each* j *a* j*-dependent basis for the target space* $\mathbb{R}^d$*, with respect to which each* L_j *is represented by a* $D \times d$ *matrix with rational entries.*

Perhaps the simplest example that is not rationally commensurate, is as follows.

Example 10.2. *The family of four mappings* $\mathbb{R}^2 \to \mathbb{R}^1$ *given by* $(x, y) \mapsto x$*;* $\mapsto y$*;* $\mapsto x + y$*; and* $\mapsto x - \sqrt{2}y$ *is not rationally commensurate.*

Continue to write $L_j(x) = x \cdot v_j$.

Theorem 10.1. [14] *Consider any $D, d, P, J, \{v_j : j \in J\}$ with $d = 1$. Suppose that for any $J' \subset J$ of cardinality $\le D$, $\{v_j : j \in J'\}$ is linearly independent. If a quadratic polynomial P is nondegenerate, and if $|J| = 2D$, then there exists a function θ satisfying $\theta(\lambda) \to 0$ as $\lambda \to \infty$ such that $|\mathcal{T}_\lambda(\mathbf{f})| \le \theta(\lambda) \prod_j \|f_j\|_{L^\infty}$ for all $\mathbf{f}$ and all $\lambda \ge 0$.*

Note the assumption that P is quadratic. Theorem 10.1 was proved only in the last few days, and I have not yet verified details of the extension to polynomials P of arbitrary degrees.

Theorem 10.1 is based on a connection with combinatorics; indeed, an extension of Szemerédi's theorem due to Furstenberg and Katznelson is one of several central elements in its proof. We will merely present some key ideas and ingredients of the analysis.

$f|_S$ denotes the restriction of a function f to a set S.

Definition 10.3. *Let $P : \mathbb{R}^D \to \mathbb{R}$ be a polynomial. P is said to be* nondegenerate with a finite witness*, relative to $\mathcal{L}$, if there exists a finite set $S \subset \mathbb{R}^d$ such that the restriction $P|_S$ of P to S does not belong to the span of the set of all functions $(f_j \circ L_j)|_S$.*

Equivalently, there should exist a finite set $S \subset \mathbb{R}^d$ and scalars c_s such that

$$\begin{cases} \sum_{x \in S} c_x P(x) \ne 0 \\ \sum_{x \in S} c_x (f_j \circ L_j)(x) = 0 \ \forall \mathbf{f}. \end{cases} \tag{10.1}$$

The usefulness of discrete characterizations of nondegeneracy in the context of oscillatory integral theory was recognized and exploited in [6]. In that spirit, from the Szemerédi–type theorem of Furstenberg and Katznelson [22] we will deduce:

Proposition 10.2. *Suppose that a real-valued polynomial P is nondegenerate with a finite witness, with respect to $\mathcal{L} = \{L_j\}$. Then there exists a function Θ satisfying*

$$\lim_{\varepsilon \to 0^+} \Theta(\varepsilon) = 0 \tag{10.2}$$

such that for every $\varepsilon > 0$ and any measurable functions f_j,

$$|S(\mathbf{f}, \varepsilon)| \le \Theta(\varepsilon). \tag{10.3}$$

Finitely witnessed nondegeneracy obviously implies nondegeneracy.

Theorem 10.3. *Let $\mathcal{L}$ be rationally commensurate, let a polynomial $P : \mathbb{R}^D \to \mathbb{R}$ be nondegenerate relative to $\mathcal{L}$. Then P is nondegenerate with a finite witness relative to $\mathcal{L}$.*

Proposition 10.2 and Theorem 10.3 together yield a sublevel set inequality, albeit not one of power law $O(\varepsilon^\gamma)$ type, for arbitrarily large finite index sets J, in the rationally commensurate case.

Theorem 10.4. *Let a polynomial P be nondegenerate with respect to a rationally commensurate collection $\mathcal{L}$. Then there exists a function Θ satisfying*

$$\lim_{\varepsilon \to 0^+} \Theta(\varepsilon) = 0 \tag{10.4}$$

such that for every $\varepsilon > 0$ and all measurable functions f_j,

$$|S(\mathbf{f}, \varepsilon)| \le \Theta(\varepsilon). \tag{10.5}$$

The property of being rationally commensurate is not necessary in order for (10.5) to hold, in general; see Theorems 7.2 and 7.3, above.

Denote by $\#(A)$ the cardinality of a set A. Furstenberg and Katznelson [22] have shown that for any finite set $S \subset \mathbb{Z}^d$ there exists a positive function θ, satisfying $\theta(N) \to 0$ as $N \to \infty$, such that for any set $A \subset \{1, 2, \cdots, N\}^d$, either there exist $0 \neq n \in \mathbb{Z}$ and $x \in \mathbb{Z}^d$ such that $x + nS \subset A$, or $\#(A) \le \theta(N) N^d$. The next result, a continuuum analogue of this theorem, is an almost immediate consequence.

Proposition 10.5. *Let $B \subset \mathbb{R}^d$ be a bounded region, and let $S \subset \mathbb{R}^d$ be a nonempty finite set. There exists a positive function Θ satisfying $\lim_{r \to 0^+} \Theta(r) = 0$, depending only on S and on B, with the following property: For any Lebesgue measurable set $E \subset B$ and any $r > 0$, either (i) there exist $x \in B$ and $t \ge r$ such that $x + tS \subset E$, or (ii) $|E| \le \Theta(r)$.*

Outline of the proof of Theorem 10.1:

(1) A purely algebraic argument proves that nondegeneracy implies finitely witnessed nondegeneracy, in the rationally commensurate case.
(2) Theorem 10.4 has an analogue for vector-valued functions f_j and polynomials P.
(3) A Cauchy-Schwarz/T^*T–type argument converts the multilinear form T_λ in $2D$ general functions $f_j(y)$ to a form in $2D - 1$ functions $F_j(y,t) = f_j(y+t)\overline{f_j}(y)$ indexed by $\tilde{J} \subset J$, in a new ambient space of dimension $2D - 1$.
(4) Because P is quadratic rather than of higher degree, the local Fourier series expansions of §9 turn out not to be needed. Expanding each $F_j(y,t)$ in Fourier series $\sum_n a_j(t,n) e^{2\pi i n y}$, implementing a $\diamondsuit/\heartsuit$–type decomposition, integrating by parts to reduce matters to the contribution of stationary $\mathbf{n} = (n_j : j \in \tilde{J})$, and reducing from sums $\sum_n a_j(t,n) e^{2\pi i n y}$ with $O(\lambda^{2\sigma})$ nonzero coefficients $a_j(t,\cdot)$ for each t to monomials $a_j(t,n(t)) e^{2\pi i n(t) y}$, the problem is reduced to a vector-valued sublevel sublevel set inequality for $2D - 1$ $\mathbb{R}^D$–valued functions in the auxiliary space $v_{2D}^\perp$.
(5) Moreover, the new vector-valued polynomial that appears in the definition of this sublevel set is nondegenerate,[12] and of course the structure is rationally commensurate. Therefore Theorem 10.4 applies, completing the proof.

11. Introduction to implicitly oscillatory integrals

By an implicitly oscillatory integral I mean a multilinear form, acting on a tuple $\mathbf{f}$ of L^∞ functions, of the type

$$\mathcal{T}(\mathbf{f}) = \int_{\mathbb{R}^D} \prod_{j \in J} (f_j \circ \varphi_j)\, \eta. \tag{11.1}$$

Here $\eta \in C_0^\infty$ is a compactly supported smooth auxiliary function. The principal ingredient is a tuple $\mathbf{\Phi} = (\varphi_j : j \in J)$ of C^ω surjective mappings φ_j from a neighborhood of the support of η to some lower-dimensional target space $\mathbb{R}^d$.

In this formulation, there is no explicitly oscillatory factor $e^{i\lambda\phi}$. We have also dropped the assumption that the mappings L_j are linear. The assumption of real analyticity simplifies aspects of the theory; a C^ω function in a connected open set either vanishes identically, or vanishes only to finite order at any point, while for C^∞ functions there is an intermediate possibility.

[12]Example: $Q(y) = y^2$ satisfies $Q(y+t) - Q(y) = 2ty + t^2$, which is a linear function of y for each t, but as a function of t remains quadratic. This is a source of nondegeneracy for the sublevel set problem here.

We seek inequalities asserting significant cancellation in the integral if f_j are highly oscillatory. A natural formulation is phrased in terms of negative order Sobolev norms through inequalities of the form

$$|\mathcal{T}(\mathbf{f})| \le C \prod_j \|f_j\|_{W^{p,s}} \tag{11.2}$$

where $W^{p,s}$ is the space of functions having s derivatives in L^p, and s *is strictly negative.*

Having such an inequality is equivalent to the following property: There exists $\gamma > 0$ such that for any functions $f_j \in L^\infty$, for any $\lambda \ge 1$, if there exists $k \in J$ such that $\widehat{f_k}(\xi) = 0$ whenever $|\xi| \le \lambda$, then

$$|\mathcal{T}(\mathbf{f})| \le C\lambda^{-\gamma} \prod_j \|f_j\|_{L^\infty}. \tag{11.3}$$

The inequality is essentially invariant under real analytic diffeomorphisms of both the ambient space $\mathbb{R}^D$ and the target spaces $\mathbb{R}^d$, with each target space allowed its own diffeomorphism. Indeed, let $\widehat{f}$ be supported where $|\xi| \ge \lambda$. Let h be a diffeomorphism of $\mathbb{R}^d$, and let $\zeta \in C_0^\infty(\mathbb{R}^d)$. Let $g = (f_k \circ h) \cdot \zeta$. Then g need not satisfy the same Fourier support restriction, even if h is the identity. However, this is misleading; g effectively satisfies the support restriction since it can be decomposed as $g = g_0 + g_1$ with $\|g_0\|_\infty = O(\|g\|_\infty)$, $\widehat{g_0}$ supported where $|\xi| \ge \lambda^{1/2}$, and $\|g_1\|_\infty = O(\lambda^{-N})$ for all $N < \infty$. (Verification is left as an exercise for the audience.)

A simple example is $f_k(y) = e^{i\lambda v \cdot y}$ for any nonzero v. More generally, $h : \mathbb{R} \to \mathbb{R}$ be a C^2 function with nowhere vanishing derivative, and choose $\zeta \in C_0^\infty(\mathbb{R}^1)$ so that $\zeta \circ \varphi_k \equiv 1$ on the support of η. Set $f_k(y) = e^{i\lambda\psi(y)}\,\zeta(y)$. However, general L^∞ functions f_k with the Fourier support restriction are far less structured than the oscillatory factors $e^{i\lambda\phi}$ with which we have been dealing.

The hypothesis on f_k implies that f_k is highly oscillatory, in the particular sense that $\int_I f_k = O(\lambda^{-1})$ uniformly for all bounded intervals $I \subset \mathbb{R}$.

We restrict attention to the nontrivial case, in which $D < |J|d$.

Consider the case in which every mapping $\varphi_j = L_j$ is linear. Suppose for simplicity that $\int \eta \ne 0$, though this is inessential. Let λ be a large scalar parameter. For each j, consider some $w_j \in \mathbb{R}^d$. Choose each $f_j(y) = e^{i\lambda w_j \cdot y}$. Then

$$\mathcal{T}(\mathbf{f}) = \int e^{i\lambda x \cdot w}\,\eta(x)\,dx$$

where

$$w = \sum_j L_j^*(w_j).$$

If $w = 0$ then the integral is independent of λ (and is nonzero). On the other hand, the hypothesis on the Fourier transform of f_k is satisfied provided that $w_k \ne 0$. Thus for linear mappings $\varphi_j = L_j$, the inequality (11.3) cannot hold, unless the mapping $(w_1, \dots, w_{|J|}) \mapsto \sum_j L_j^*(w_j)$ from $(\mathbb{R}^d)^J$ to $\mathbb{R}^D$ is injective. This injectivity cannot hold unless $D \ge |J|d$, the trivial case.

Up to bounded factors, the inequalities in question are invariant under real analytic changes of variables in $\mathbb{R}^D$ (that is, from a neighborhood of the support of η to another domain in $\mathbb{R}^D$) and in the $|J|$ target spaces $\mathbb{R}^d$. Thus the product of C^ω diffeomorphism groups is, loosely speaking, a symmetry group for this problem.

Definition 11.1. *Let φ_j be real analytic surjections from a nonempty open set in $\mathbb{R}^D$ to $\mathbb{R}^d$. $\{\varphi_j : j \in J\}$ is linearizable if there exist C^ω diffeomorphisms h for $\mathbb{R}^D$ and h_j for $\mathbb{R}^d$ such that in some nonempty open subset of $\mathbb{R}^D$, all of the compositions $h_j \circ \varphi_j \circ h^{-1}$ are linear.*

In particular, the inequality (11.3) fails whenever $\{\varphi_j : j \in J\}$ is linearizable.

Question 11.2. *Does the inequality* (11.3) *hold for every family $\{\varphi_j : j \in J\}$ of real analytic surjections that is not linearizable?*

An immediate application is to weak continuity of mappings $\mathbf{f} \mapsto \prod_{j\in J}(f_j \circ \varphi_j)$. Let $p < \infty$ and $s < 0$ be parameters for which the inequality is valid.[13] Let $(\mathbf{f}_\nu : \nu \in \mathbb{N})$ be a sequence of tuples $\mathbf{f}_\nu = (f_{j,\nu} : j \in J)$ of functions in L^p that converges weakly to a limit $\mathbf{f}$, in the sense that $\int_{\mathbb{R}^d} f_{j,\nu}\, h \to \int_{\mathbb{R}^d} f_j\, h$ for every continuous compactly supported test function h, for each $j \in J$. Then $limit_{\nu\to\infty}\mathcal{T}(\mathbf{f}_\nu) = \mathcal{T}(\mathbf{f})$. That is,

$$f_{j,\nu} \rightharpoonup f_j \ \forall\, j \Rightarrow \prod_j (f_{j,\nu} \circ \varphi_j) \rightharpoonup \prod_j (f_j \circ \varphi_j).$$

This follows from the $W^{p,s}$ inequality, because the natural inclusion of L^p into $W^{p,s}$ is compact for strictly negative s.

Joly, Métivier, and Rauch [29] proved weak continuity for $D = 2$ and $|J| = 3$, with somewhat weaker hypotheses. Their proof was based on microlocal defect measures and an argument by contradiction, and yielded no quantitative upper bound corresponding to (11.3).

For threefold products, an inequality in terms of negative order Sobolev norms was proved by Bourgain [3] in a particular case, and later by the author [15] in a relatively general C^ω case. More recently, Evans [21] has developed another proof for threefold products. A more streamlined version of the analysis in [15], with certain supplementary hypotheses relaxed, was presented in [19].

12. Some results

We assume henceforth that the target spaces $\mathbb{R}^d$ have dimension $d = 1$, and that the ambient space $\mathbb{R}^D$ has dimension $D = 2$. In general, decreasing D with d and $|J|$ held fixed increases the difficulty of the problem, so $D = 2$ is the most difficult case. On the other hand, increasing $|J|$ increases difficulty. $|J| = 2$ is the classical bilinear case, called type two in the classification that we are using. Thus $|J| = 3$ is the simplest genuinely multilinear case.

12.1. **The trilinear case.**

In this case, an inequality of the desired type is proved. Let $\varphi_j : \mathbb{R}^2 \to \mathbb{R}^1$ be real analytic mappings. Let $\eta \in C_0^\infty(\mathbb{R}^2)$ be infinitely differentiable and have compact support. Consider a trilinear form

$$\mathcal{T}(\mathbf{f}) = \int_{\mathbb{R}^2} \prod_{j=0}^{2} (f_j \circ \varphi_j)(x)\, \eta(x)\, dx \tag{12.1}$$

acting on ordered triples $\mathbf{f} = (f_0, f_1, f_2)$ of functions $f_j : \mathbb{R}^1 \to \mathbb{C}$.

[13] We exclude $p = \infty$ solely in order to avoid irrelevant discussion about possible definitions of $W^{\infty,s}$ for negative s. The discussion extends to $p = \infty$ with any reasonable definition.

Theorem 12.1. *Let U be a connected neighborhood of the support of η. Let $\varphi_j : U \to \mathbb{R}^1$ be real analytic. Assume that for any $i \neq j \in \{0,1,2\}$, $\det(\nabla\varphi_i, \nabla\varphi_j)$ does not vanish identically in any nonempty open set. Assume that for any nonempty connected open subset $U' \subset U$, for any $\mathbf{g} \in C^\omega(\Phi(U'))$ that satisfies $\sum_{j=0}^2 (g_j \circ \varphi_j) \equiv 0$ in U', each g_j is constant in $\varphi_j(U')$.*

Then there exist $p < \infty$, $\sigma < 0$, and $C < \infty$ such that for all Lebesgue measurable functions $\mathbf{f} = (f_0, f_1, f_2) \in (L^p \times L^p \times L^p)(\mathbb{R}^1)$, the integral defining $\mathcal{T}(\mathbf{f})$ converges absolutely and

$$|\mathcal{T}(\mathbf{f})| \leq C \prod_{j=0}^{2} \|f_j\|_{W^{p,\sigma}}. \tag{12.2}$$

The example $\varphi_j(x,t) = x + t^j$ for $j \in J = \{0,1,2\}$ satisfies all hypotheses of Theorem 12.1.

This was first proved in [15] under certain auxiliary hypotheses, for instance, a transversality hypothesis that $\nabla\varphi_i, \nabla\varphi_j$ are everywhere linearly independent for all $i \neq j$, rather than merely being linearly independent on the complement of an analytic variety of positive codimension. The general case was proved in [19] using ingredients from [15], [8], and [16].

12.2. **A small digression.** Before discussing the trilinear case, we pause to discuss a particularly simple-looking question about multilinear integral forms in which increasing the number of factors f_j changes the situation dramatically.

Consider firstly forms

$$T(f_1, f_2)(x) = \int_{-1}^{1} f_1(x+y)\, f_2(x-y)\, dy.$$

If $f_1, f_2 \in L^1$ then $T(f_1, f_2)$ is finite almost everywhere; indeed, $T(f_1, f_2) \in L^1$ and if f_1, f_2 are nonnegative then $\|T(f_1, f_2)\|_{L^1}$ is a positive constant multiple of $\|f_1\|_1 \|f_2\|_1$. This statement is optimal in various ways; in particular, for any $p < 1$ there exist $f_j \in L^p$ such that $T(f_1, f_2) = \infty$ on a set of positive measure.[14]

Now consider trilinear forms

$$T(\mathbf{f}) = T(f_1, f_2, f_3)(x) = \int_{-1}^{1} f_1(x+y)\, f_2(x-y)\, f_3(x+\alpha y)\, dy$$

with $\alpha \notin \{0, 1, -1\}$. Assume that each function f_j is nonnegative. If $p \geq \frac{3}{2}$ and each $f_j \in L^p$ then $T(\mathbf{f})$ is finite almost everywhere; indeed, $T(\mathbf{f}) \in L^1$. This can be proved, for instance, by showing that $\|T(\mathbf{f})\|_1 \leq C\|f_i\|_1 \|f_j\|_1 \|f_k\|_\infty$ for every permutation (i, j, k) of $(1, 2, 3)$, and invoking multilinear interpolation.

Can the threshold $p = \frac{3}{2}$ for almost everywhere finiteness of the integral be improved? This question was investigated in [10], where two things were proved: Firstly, if $\alpha \in \mathbb{Q}$ then Yes, there exist $p < \frac{3}{2}$ and $q > 0$ such that

$$\|T(\mathbf{f})\|_{L^q(I)} \leq C_I \prod_{j=1}^{3} \|f_j\|_p$$

for every bounded interval $I \subset \mathbb{R}$. Secondly, if $\alpha \notin \mathbb{Q}$ then No; for any $p < \frac{3}{2}$ there exist $f_j \in L^p$ such that $T(\mathbf{f}) = \infty$ at every point of a set of positive Lebesgue measure.

[14] This is trivial; consider building blocks $f_j = \delta^{-1/p}\mathbf{1}_{[} - \delta, \delta]$ and form appropriate infinite linear combinations of these.

The proof of the former result relies on a variant and generalization of an inequality of Katz and Tao [31] and related to work of Bourgain [5] on the Kakeya problem, descending in part from work of Gowers [25] on Szemerédi's theorem. The proof of the latter relies on a lattice construction based on a Diophantine approximation property of irrational numbers.

These results are not directly relevant to the main topics of these lectures. They do illustrate that increasing the number of factors can have a dramatic impact, at least in one context.

12.3. **The quadrilinear case.** Our main result for the quadrilinear case is the following theorem. Its main hypothesis is that there exists an open set $U \subset \mathbb{R}^2$ containing the support of the cutoff function η such that for any connected open subset $U' \subset U$, for any functions $g_j \in C^\omega(\varphi_j(U'))$, if $\sum_{j=1}^4 (g_j \circ \varphi_j) \equiv 0$ in U' then each g_j is constant in $\varphi_j(U')$. As we have seen in the trilinear case and in the discussion of forms of the fourth kind, such a hypothesis is a necessary condition for there to be a valid bound in terms of strictly negative order Sobolev norms.

There are also certain auxiliary hypotheses. One auxiliary hypothesis is that for any $i \neq j \in J$, $\{\nabla\varphi_i(x), \nabla\varphi_j(x)\}$ is linearly indpendent at every x in a neighborhood of the support of η. Another similar transversality hypothesis is also imposed. A second auxiliary hypothesis is phrased in terms of auxiliary nowhere vanishing C^ω vector fields V_j that satisfy $V_j(\varphi_j) \equiv 0$. This second hypothesis is that for any permutation (i, j, k, l) of the indices in J, for any exponent $\tau \in \mathbb{R}$, the function

$$\frac{V_l\varphi_i \cdot |V_k\varphi_i|^\tau}{V_l\varphi_j \cdot |V_k\varphi_j|^\tau}$$

does not take the form $\frac{h_i \circ \varphi_i}{h_j \circ \varphi_j}$ for any C^ω functions h_i, h_j in any nonempty open subset of $\mathbb{R}^2$. The third auxiliary hypothesis is of the same type as the main hypothesis, but applies to a related problem involving three functions F_j which arises in the proof. It is that for any $k \in J$, on any nonempty connected open set $\omega \subset \mathbb{R}^2$, any C^ω solution $(F_j : j \neq k)$ of

$$\sum_{j \neq k} (F_j \circ \varphi_j) \cdot V_k(\varphi_j) \cdot \nabla\varphi_j = 0$$

vanishes identically.

Theorem 12.2. *Let $\mathbf{\Phi}$ satisfy the main hypothesis and the auxiliary hypotheses. For any $p > 2$ and any $\eta \in C_0^\infty(B)$, there exist $\sigma < 0$ and $C < \infty$ such that for every four-tuple $\mathbf{f} = (f_1, \dots, f_4)$ of functions in $L^p(\mathbb{R}^1)$,*

$$|\mathcal{T}(\mathbf{f})| \leq C \prod_{j \in \mathbf{4}} \|f_j\|_{W^{p,\sigma}}. \tag{12.3}$$

A special case of the corresponding result with two factors f_j is a restatement of this well-known fact: Let $t \mapsto \gamma(t)$ be a real analytic mapping from $\mathbb{R}^1$ to $\mathbb{R}^2$. Let $\eta \in C^\infty(\mathbb{R})$ be compactly supported. Define a measure μ in $\mathbb{R}^2$ by $\int g\, d\mu = \int_{\mathbb{R}} g(\gamma(t))\, \eta(t)\, dt$. If the range of γ is not contained in any affine subspace of $\mathbb{R}^2$ then there exists $\delta > 0$ such that the Fourier transform of μ satisfies $|\mu(\xi)| = O(|\xi|^{-\delta})$ as $|\xi| \to \infty$.

12.4. **A variant with functions of two variables.** A particular variant, in which the ambient space of integration is $\mathbb{R}^3$ and each f_j has domain $\mathbb{R}^2$ rather than $\mathbb{R}^1$, was studied in [8]. Let ζ be a smooth function with compact support in $\mathbb{R}^2 \times (\mathbb{R}^1 \setminus \{0\})$. Let $f_j : \mathbb{R}^2 \to \mathbb{C}$. Consider the trilinear form

$$\Lambda(f_1, f_2, f_3) = \int_{\mathbb{R}^3} f_1(x+t, y)\, f_2(x, y+t^2)\, f_3(x,y)\, \zeta(x,y,t)\, dx\, dy\, dt. \tag{12.4}$$

Theorem 12.3. *There exist constants $C > 0$ and $\sigma > 0$ so that for all test functions f_1, f_2, f_3,*

$$|\Lambda(f_1, f_2, f_3)| \le C \|f_1\|_{H^{(-\sigma,0)}} \|f_2\|_{H^{(0,-\sigma)}} \|f_3\|_{L^\infty}, \tag{12.5}$$

where the constant C only depends on ζ, and

$$\|f\|^2_{H^{(a,b)}} = \int_{\mathbb{R}^2} |\widehat{f}(\xi_1, \xi_2)|^2 \, (1 + |\xi_1|^2)^{\frac{a}{2}} \, (1 + |\xi_2|^2)^{\frac{b}{2}} \, d\xi_1 \, d\xi_2.$$

The form Λ is invariant under multiplication of $f_1(x,y)$ by $h_1(y)$, multiplication of $f_2(x,y)$ by $h_2(x)$, and division of $f_3(x,y)$ by $h_1(y)h_2(x)$. Thus no bound in terms of conventional Sobolev norms $\|f_j\|_{H^{-\sigma}(\mathbb{R}^2)}$ for $j = 1, 2$ is valid, and likewise there is no possible bound in terms of any negative order norm of f_3.

The inequality (12.5) is equivalent to the estimate

$$|\Lambda(f_1, f_2, f_3)| \le C\lambda^{-\sigma} \|f_1\|_{L^2} \|f_2\|_{L^2} \|f_3\|_{L^\infty}$$

valid for all $\lambda \ge 1$ under the assumption that $\widehat{f_j}$ is supported where $|\xi_j| \asymp \lambda$ for at least one index $j = 1, 2$.

13. Consequences

13.1. **Weak continuity of products.** For any open set $S \subset \mathbb{R}^2$, we write "$g_\nu \rightharpoonup g$ in S" to mean that $\int g_\nu \, h \to \int g \, h$ as $\nu \to \infty$, for every C^∞ function h supported in S.

Corollary 13.1. *Let $\mathbf{\Phi}$ satisfy the hypotheses of Theorem 12.2. Let $p > 2$. Let $(\mathbf{f}_\nu : \nu \in \mathbb{N})$ be a sequence of 4-tuples of Lebesgue measurable functions. Suppose that $\|f_{\nu,j}\|_{L^p}$ is bounded, uniformly in ν, j. Suppose that for each $j \in \mathbf{4}$, $f_{\nu,j} \rightharpoonup f_j$ in $\varphi_j(B)$ as $\nu \to \infty$. Then*

$$\prod_{j \in \mathbf{4}} (f_{\nu,j} \circ \varphi_j) \rightharpoonup \prod_{j \in \mathbf{4}} (f_j \circ \varphi_j) \quad \text{in } B. \tag{13.1}$$

13.2. **Ramsey theory.** Another application is a combinatorial result on the existence of patterns in subsets of $[0,1]^2$ having positive Lebesgue measure.

Theorem 13.2. *Let $\varepsilon \in (0, \frac{1}{2})$ and $E \subset [0,1]^2$ a measurable set of Lebesgue measure at least ε. Then there exist*

$$(x,y), (x+t, y), (x, y+t^2) \in E$$

with $t > \exp(-\exp(\varepsilon^{-C}))$ for some constant $C > 0$ not depending on E or ε.

It is the quantitative lower bound on t that is significant; without any lower bound, the conclusion follows directly from the Lebesgue density theorem. A result with worse ε-dependence can also be deduced from a general theorem of Bergelson and Leibman [1].

Theorem 13.2 recovers the following quantitative nonlinear Roth theorem of Bourgain [3]: for every measurable $E \subset [0,1]$ with $|E| \ge \varepsilon$ there exist $t > \exp(-\exp(\varepsilon^{-c}))$ and x with x, $x+t$, $x+t^2$ all in E. It suffices to apply Theorem 13.2 to the set $\widetilde{E} = \{(x,y) \in [0,1]^2 : x - y \in E\}$.

13.3. **Maximal functions.** Consider the bilinear maximal operator

$$M(f_1, f_2)(x, y) = \sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |f_1(x + t, y) f_2(x, y + t^2)|\, dt. \tag{13.2}$$

Theorem 13.3. *For every $p, q \in (1, \infty)$, $r \in [1, \infty)$ with $p^{-1} + q^{-1} = r^{-1}$ there exists $C \in (0, \infty)$ such that for all test functions f_1, f_2,*

$$\|M(f_1, f_2)\|_{L^r} \le C\|f_1\|_{L^p}\|f_2\|_{L^q}.$$

Note that for $r > 1$, the claim in Theorem 13.3 follows immediately from Hölder's inequality and the Hardy-Littlewood maximal theorem (in fact, this argument also gives $L^p \times L^{p'} \to L^{1,\infty}$ bounds for all $p \in (1, \infty)$).

Here is a variant, acting on functions of a single real variable. Let $f_j : \mathbb{R}^1 \to \mathbb{C}$, and for each $x \in \mathbb{R}^1$ define

$$\mathcal{M}(f_1, f_2)(x) = \sup_{r\in 2^{\mathbb{Z}}} |\int_{S^1} f_1(x + y_1)\, f_2(x + y_2)\, d\sigma(y)|$$

where σ denotes arc length measure on the unit circle $S^1 \subset \mathbb{R}^2$. This can alternatively be written as

$$\mathcal{M}(f_1, f_2)(x) = \sup_{k\in\mathbb{Z}} |\int_0 2\pi f_1(x + 2^k \cos(\theta))\, f_2(x + 2^k \sin(\theta))\, d\theta|.$$

Theorem 13.4. *Let $I_0 \subset \mathbb{R}$ be a nonempty open interval. Let $\varphi : I_0 \to \mathbb{R}^2$ be a real analytic mapping that satisfies the three hypotheses. Let $\eta : I_0 \to [0, \infty)$ be infinitely differentiable and have compact support in I_0. For any $p_1, p_2 > 1$ there exists $C < \infty$ such that*

$$\|\mathcal{M}(f_1, f_2)\|_{L^q(\mathbb{R}^1)} \le C\|f_1\|_{L^{p_1}(\mathbb{R}^1)}\|f_2\|_{L^{p_2}(\mathbb{R}^1)} \tag{13.3}$$

where $q = q(p_1, p_2)$.

The range of exponents q for which the conclusion holds for some (p_1, p_2) extends below $q = 1$, indeed, to all $q > \frac{1}{2}$. Lacey has shown that if $\gamma_j(t) = c_j t$ for distinct constants c_j then $\mathcal{M}$ satisfies the indicated inequalities in the range $q(p_1, p_2) > \frac{2}{3}$.

A well known result for maximal linear operators in the same spirit as Theorem 13.4 states that if $k < d$, if $\gamma : \mathbb{R}^k \to \mathbb{R}^d$ is real analytic in a neighborhood of a compact set K, and if the range of γ is not contained in any affine subspace of $\mathbb{R}^d$, then the maximal function $\sup_{r\in 2^{\mathbb{Z}}} \int_K |f(x+r\gamma(t))|\, \eta(t)\, dt$ is bounded on $L^p(\mathbb{R}^d)$. See [43] for an introduction to this circle of ideas. A prototypical example is

$$\mathcal{M}^0 f(x) = \sup_{r\in 2^{\mathbb{Z}}} \Big| \int_{S^{d-1}} f(x + ry)\, d\sigma(y) \Big|$$

where $x \in \mathbb{R}^d$, $f : \mathbb{R}^d \to \mathbb{C}$, and σ is surface measure on the unit sphere $S^{d-1} \subset \mathbb{R}^d$. A key fact is that if $\eta \in C^\infty$ is supported in K, then the Fourier transform of the measure defined by $d\mu(t) = \eta(\gamma(t))\, dt$ satisfies $\widehat{\mu(\xi)} = O(|\xi|^{-\delta})$ as $|\xi| \to \infty$, for some $\delta > 0$. A central element of our analysis, Theorem 12.1, states that under natural hypotheses, a bilinear analogue of this Fourier transform decay property holds.

We do not know whether $\mathcal{M}$ map $L^1 \times L^p$ for $p > 1$, or even to weak L^1 for $p = 1$. Even for maximal linear operators such as $\mathcal{M}^0$, it remains an open question whether weak type $(1, 1)$ inequalities hold.

For those exponents $(p_1, p_2) \in (1, \infty)^2$ satisfying $q > 1$, that is, $p_1^{-1} + p_2^{-1} < 1$, Theorem 13.4 is an easy of consequence of linear one-dimensional Calderón-Zygmund theory.

Corollary 13.5. *For $r \in \mathbb{R}$ define*

$$B_r(f_1, f_2)(x) = \int_0^{2\pi} |f_1(x + r\cos(\theta)) \cdot f_2(x + r\sin(\theta))|\, d\theta.$$

Let $q = q(p_1, p_2) \in (0, \infty)$ satisfy $q^{-1} = p_1^{-1} + p_2^{-1}$ for $(p_1, p_2) \in [1, \infty]^2$. For each $(p_1, p_2) \in (1, \infty]^2$ there exists $C < \infty$ such that

$$\|\sup_{r \in 2^{\mathbb{Z}}} |B_r(f_1, f_2)|\|_{L^q(\mathbb{R}^1)} \le C \prod_{j=1}^{2} \|f_j\|_{L^{p_j}(\mathbb{R}^1)}. \tag{13.4}$$

The quantity $J(\theta)$ defined above is equal to $\cos(\theta)+\sin(\theta)$ in this special case. It vanishes at $\theta = \frac{3\pi}{4}$ and at $-\frac{\pi}{4}$. These points play distinguished roles in our analysis.

There have been several earlier works concerning maximal operators associated to bilinear forms

$$B_r(f_1, f_2)(x) = \int_{S^{2d-1}} |f_1(x + ry_1)\, f_2(x + ry_2)|\, d\sigma(y_1, y_2) \tag{13.5}$$

with $(y_1, y_2) \in S^{2d-1} \subset \mathbb{R}^d \times \mathbb{R}^d$ and with σ denoting surface measure on S^{2d-1} for $d \ge 1$, with suprema taken over $r \in (0, \infty)$ and/or $r \in 2^{\mathbb{Z}}$. The lowest-dimensional case $d = 2$ is the most singular case, and had remained open until [19]. (The case $r \in \mathbb{R}^+$ turns out to be quite a bit easier than the case $r \in 2^{\mathbb{Z}}$; it can be treated using only routine theory of the Hardy-Littlewood maximal function in $\mathbb{R}^1$ together with Hölder's inequality.)

Work on various variants, including multilinear expressions of general degree d with integration over S^{d-1} is in progress.

13.4. **A bilinear singular integral operator.** Consider the bilinear singular integral operator

$$T(f_1, f_2)(x, y) = \text{p.v.} \int_{\mathbb{R}} f_1(x + t, y) f_2(x, y + t^2)\, \frac{dt}{t}, \tag{13.6}$$

defined *a priori* for test functions $f_1, f_2 : \mathbb{R}^2 \to \mathbb{C}$. We obtain the following result.

Theorem 13.6. *Let $p, q \in (1, \infty)$, $r \in [1, 2)$ satisfy $p^{-1} + q^{-1} = r^{-1}$. Then T extends to a bounded operator $L^p \times L^q \to L^r$.*

The operator (13.6) is a variant of the bilinear Hilbert transform, analyzed by Lacey and Thiele [32], [33] in important work. The operator T is also a variant of the *triangular Hilbert transform*

$$(f_1, f_2) \mapsto \text{p.v.} \int_{\mathbb{R}} f_1(x + t, y) f_2(x, y + t)\, \frac{dt}{t}. \tag{13.7}$$

Determining whether the triangular Hilbert transform satisfies any Lebesgue norm bounds remains a significant open problem.

Theorem 13.6 unifies two previously known inequalities. First, it implies L^p bounds for a variant of the bilinear Hilbert transform with curvature,

$$(f_1, f_2) \mapsto \text{p.v.} \int f_1(x + t) f_2(x + t^2)\, \frac{dt}{t}. \tag{13.8}$$

The $L^2 \times L^2 \to L^1$ bound for this operator was first proved by Li [34]. Second, Theorem 13.6 yields L^p bounds for the operator

$$f \mapsto \sup_{N \in \mathbb{R}} \Big| \text{p.v.} \int_{\mathbb{R}} f(x - t) e^{iNt^2}\, \frac{dt}{t} \Big|, \tag{13.9}$$

due to Stein (see also Stein and Wainger [44]). This is a variant of Carleson's maximal operator, with t^2 in place of t^1 in the phase.

14. Sublevel set inequalities

Let $B \subset \mathbb{R}^2$ be a ball of positive radius, and let $\varphi_j : B \to \mathbb{R}^1$ be real analytic for $j \in \{1, 2, 3\}$. Suppose that $\nabla\varphi_j$ are pairwise linearly independent at each point in B. Let $0 \le \eta \in C^\infty(B)$.

14.1. **The constant-coefficient case.** The functional equation $f(x)+g(y)+h(x+y) = 0$, has been widely studied. Its solutions are the ordered triples $(f(x), g(y), h(x + y)) = (ax + c_1, ay + c_2, a(x + y) - c_1 - c_2)$ with a, c_1, c_2 all constant, and no others. Approximate solutions, in a certain sense, have been studied in [11], [12]. We consider here the more general functional equation

$$\sum_{j=1}^{3}(f_j \circ \varphi_j) = 0 \text{ almost everywhere} \tag{14.1}$$

where the mappings φ_j need not be linear, and the functions f_j are real-valued. We discuss related sublevel sets

$$S(\mathbf{f}, \varepsilon) = \{\mathbf{x} \in B : |\sum_{j=1}^{3}(f_j \circ \varphi)(\mathbf{x})| \le \varepsilon\} \tag{14.2}$$

associated to ordered triples $\mathbf{f}$ of scalar-valued functions.

The inequality (14.2) differs from corresponding inequalities studied and exploited in various proofs above in two ways: it is homogeneous rather than inhomogeneous, and it is a single scalar inequality, rather than a system of two scalar inequalities.

Theorem 12.1 has the following implication concerning the nonexistence of nontrivial solutions of (14.1).

Corollary 14.1. *Let $B \subset \mathbb{R}^2$ be a closed ball of positive, finite radius. For $j \in \{1, 2, 3\}$ let $\varphi_j \in C^\omega$ map a neighborhood of B to $\mathbb{R}$, and suppose that $\nabla\varphi_j$ are pairwise linearly independent at each point of B. Suppose that the curvature of the web defined by $(\varphi_j : j \in \{1, 2, 3\})$ does not vanish identically on B. Let $\mathbf{f}$ be an ordered triple of Lebesgue measurable real-valued functions. Suppose that for each index j and each $t \in \mathbb{R}$,*

$$|\{x : f_j(x) = t\}| = 0. \tag{14.3}$$

If $\mathbf{f}$ is a solution of the functional equation (14.1) then each function f_j is constant.

In particular, all C^ω solutions $\mathbf{f}$ of (14.1) are constants. Indeed, one of the three component functions f_j must fail to satisfy the hypothesis (14.3), and hence must be constant. It follows immediately from the functional equation (14.1) that the other two component functions are also constant. □

A more quantitative statement is as follows.

Corollary 14.2. *Let $B \subset \mathbb{R}^2$ be a closed ball of positive, finite radius. For $j \in \{1, 2, 3\}$ let $\varphi_j \in C^\omega$ map a neighborhood of B to $\mathbb{R}$, and suppose that $\nabla\varphi_j$ are pairwise linearly independent at each point of B. Suppose that the curvature of the web defined by $(\varphi_j : j \in \{1, 2, 3\})$ does not vanish identically on B. There exist $\delta > 0$ and $C < \infty$ such that for* *any ordered triple* $\mathbf{f}$ *of Lebesgue measurable real-valued functions and any* $r \in (0,\infty)$*, the sublevel set* $S(\mathbf{f},r)$ *satisfies*

$$|S(\mathbf{f},r)| \le C \sup_{t\in\mathbb{R}} \big|\{x \in \varphi_j(B) : |f_j(x)-t| \le r\}\big|^{\delta} \tag{14.4}$$

for each $j \in \{1,2,3\}$*.*

We have discussed how oscillatory inequalities imply sublevel set inequalities by considering

$$\int_{\mathbb{R}^2} h(\varepsilon^{-1}\sum_j (f_j\circ\varphi_j)\,\eta,$$

rewriting this as a superposition

$$\int_{\mathbb{R}} \widehat{h}(\xi/\varepsilon) \int_{\mathbb{R}^2} \prod_j (F_{j,\xi/\varepsilon}\circ\varphi_j)(x)\,\eta(x)\,dx\,d\xi$$

of oscillatory integrals with $F_{j,t}(y) = e^{itf_j(y)}$, and invoking an oscillatory bound for the inner integral for each ξ. In order to prove the above two corollaries by this strategy, one must be able to control the Sobolev norms of e^{itf_j} that arise when Theorem 12.1 is invoked; and one must take into account the obvious fact that the sublevel set can be large if the functions f_j are nearly constant on large sets. The last step relies on the next lemma, which asserts (for large parameters A) that if f_j is suitably far from being constant, then $F_{j,t}$ has suitably small negative order Sobolev norm for most large parameters t.

Lemma 14.3. *Let* $\sigma < 0$*. Let* $I \subset \mathbb{R}$ *be a bounded interval. Then there exists* $C < \infty$ *such that for any real-valued function* $f \in L^2(\mathbb{R})$ *supported in a fixed bounded set, for any* $A \in (0,\infty)$*,*

$$\int_{\lambda\le A} \|\mathbf{1}_I e^{i\lambda f}\|_{H^\sigma}^2\,d\lambda \le CA \sup_{t\in\mathbb{R}} \big|\{x\in I : |f(x)-t| \le A^{-1}\}\big|^{|\sigma|}. \tag{14.5}$$

Proof. It suffices to treat the case $A = 1$, since the substitution $\lambda = A\tau$ reduces the general case to this one.

Let h be a nonnegative Schwartz function satisfying $h(y) \ge 1$ for all $y \in [-1,1]$, with $\widehat{h}$ supported in $[-1,1]$.

$$\begin{aligned}
\int_{\lambda\le 1} \|\mathbf{1}_I e^{i\lambda f}\|_{H^\sigma}^2\,d\lambda &\le \int h(\lambda)\|\mathbf{1}_I e^{i\lambda f}\|_{H^\sigma}^2\,d\lambda\\
&= \int h(\lambda)\int_{\mathbb{R}} \Big|\int e^{i\lambda f(x)}e^{-ix\xi}\,\mathbf{1}_I(x)\,dx\Big|^2 (1+\xi^2)^\sigma\,d\xi\,d\lambda\\
&= \int_{\mathbb{R}} h(\lambda)\int_{\mathbb{R}}\iint_{I\times I} e^{i\lambda[f(x)-f(y)]}e^{-i(x-y)\xi}\,dx\,dy(1+\xi^2)^\sigma\,d\xi\,d\lambda\\
&= \iint_{I\times I}\Big(\int_{\mathbb{R}} e^{-i(x-y)\xi}(1+\xi^2)^\sigma\,d\xi\Big)A\widehat{h}(A(f(y)-f(x)))\,dx\,dy\\
&\le CA\iint_{I\times I} |x-y|^{-1-\sigma}\,\big|\widehat{h}(A(f(y)-f(x)))\big|\,dx\,dy.
\end{aligned}$$

Since $\sigma < 0$, this is majorized by

$$CA \iint_{I^2} |x-y|^{-1+|\sigma|} \mathbf{1}_{|f(x)-f(y)| \le A^{-1}}(x,y)\,dx\,dy$$
$$\le CA \sup_{y\in I} \int_I |x-y|^{-1+|\sigma|} \mathbf{1}_{|f(x)-f(y)| \le A^{-1}}(x)\,dx$$
$$\le CA \sup_t \big|\{x \in I : |f(x)-t| \le A^{-1}\}\big|^{|\sigma|}.$$

□

14.2. **Variable coefficients.** The next result is concerned with general sublevel set inequalities with variable coefficients. It is one of the main ingredients of the proof of Theorem 12.2.

Let φ_j be real analytic mappings, and let a_j be C^ω real-valued coefficients. Let $B \subset \mathbb{R}^2$ be a ball. Associated that a tuple $\mathbf{f} = (f_j)$ of real-valued measurable functions are sublevel sets

$$S(\mathbf{f}, \varepsilon) = \big\{x \in B : |\sum_{j=1}^3 a_j (f_j \circ \varphi_j)(x)\big| < \varepsilon\big\}. \tag{14.6}$$

Sublevel sets of this type, associated to *triples* $(\varphi_1, \varphi_2, \varphi_3)$ of mappings, arise naturally in the analysis of *quadrilinear* oscillatory integrals $\int_{\mathbb{R}^2} \prod_{j=1}^4 (f_j \circ \varphi_j)\,\eta$. Indeed, consider this situation: Fix $r = \frac{3}{4}$. Let λ be an arbitrarily large positive parameter. Partition $\mathbb{R}$ into intervals I_m of length λ^{-r}. Assume that each factor f_j takes the special form

$$f_j(y) = \sum_m a_{j,m} e^{i\lambda g_j(m) y}\,\zeta_m(y)$$

with $a_{j,m} = O(1)$. After dealing with various remainder terms that are relatively easily shown to be $O(\lambda^{-c})$, we require an upper bound for $|\mathcal{T}(\mathbf{f})$ of the form

$$\sum_{\mathbf{m}} \Big| \int e^{i\lambda \sum_j g_j(m_j) \nabla\varphi_j(\bar{x}_{\mathbf{m}})\cdot x} \zeta_{\mathbf{m}}(x) \Big|.$$

Isolating the contributions of those indices $\mathbf{m}$ that are interacting and stationary leads to the need for a sublevel type inequality. But for which function? It is for a vector-valued function

$$\sum_{j=1}^4 (g_j \circ \varphi_j)(x)\, \nabla\varphi_j(x),$$

taking values in $\mathbb{R}^2$, with variable (vector-valued) coefficients $\nabla\varphi_j$. Thus we have two linear functions of three scalar-valued unknown functions g_j. One of these two linear functions can be used to eliminate g_4, leaving a scalar-valued sublevel set problem in three unknown functions g_1, g_2, g_3, with variable coefficients that are expressible in terms of $\{\nabla\varphi_j : 1 \le j \le 4\}$.

The proof of Theorem 12.2 has two main parts. The first is a reduction — based on ideas that have been explained above in simpler contexts — to the case of functions of the special form discussed in the preceding paragraph. This reduces matters to a three term sublevel set inequality, with variable coefficients. Such an inequality does not follow (so far as this author is aware) from any known trilinear oscillatory inequality. The second main part of the proof of Theorem 12.2 is consequently a proof of the next theorem.

Theorem 14.4. *Let $a_j, \varphi_j \in C^\omega(\tilde{B})$. Assume that the coefficients a_j vanish nowhere in B, and that for each $i \neq j \in \{1,2,3\}$, $\nabla\varphi_i(x)$ and $\nabla\varphi_j(x)$ are linearly independent at each $x \in B$.*

Assume that for any nonempty open set $U \subset \tilde{B}$, if $f_j : \varphi_j(U) \to \mathbb{R}$ are real analytic and $\sum_{j=1}^3 a_j \cdot (f_j \circ \varphi_j) \equiv 0$ in U then each $f_j \equiv 0$ in $\varphi_j(U)$.

Assume that for each permutation (i,j,k) of $(1,2,3)$, for any $\tau \in \mathbb{R}$ and any nonempty open set $U \subset \tilde{B}$, the function $\frac{a_i|V_k\varphi_i|^\tau}{a_j|V_k\varphi_j|^\tau}$ cannot be expressed in U as $\frac{h_i\circ\varphi_i}{h_j\circ\varphi_j}$ for any real analytic functions h_i, h_j.

There exist $C < \infty$ and $\tau > 0$ with the following property. Let $\varepsilon > 0$ be arbitrary. For any ordered triple $\mathbf{f}$ of Lebesgue measurable functions satisfying

$$|f_3(y)| \geq 1 \quad \textit{for every } y \in \varphi_3(B),$$

$$|S(\mathbf{f}, \varepsilon)| \leq C\varepsilon^\tau. \tag{14.7}$$

14.3. **Sublevel set inequalities with arbitrarily many summands.** In this section we consider sublevel sets $S(\mathbf{f}, \varepsilon)$ for sums $\sum_{j\in J} a_j(x)\,(f_j \circ \varphi_j)(x)$ with an arbitrary number $|J|$ of terms, in the special case in which all mappings φ_j are *linear*. Such a situation is specified by a datum

$$\mathcal{D} = (J, \mathbf{\Phi}, \mathbf{a}) = (J, \{\varphi_j : j \in J\}, \{a_j : j \in J\}).$$

The following result applies to sublevel set inequalities with arbitrarily many summands.

Theorem 14.5. *Let $B \subset \mathbb{R}^2$ be a closed ball of positive, finite radius. Let $\tilde{B} \subset \mathbb{R}^2$ be an open neighborhood of B. Let $|J| \geq 1$. For each $j \in \{1, 2, \ldots, |J|\}$ let $\varphi_j : \mathbb{R}^2 \to \mathbb{R}^1$ be a surjective linear mapping, and let $a_j : \tilde{B} \to \mathbb{C}$ be real analytic. Let $\mathcal{D} = (|J|, \{\varphi_j\}, \{a_j\})$.*

Suppose that for any two distinct indices $i \neq j \in J$, $\nabla\varphi_i, \nabla\varphi_j \in \mathbb{R}^2$ are linearly independent. Suppose that none of the coefficients a_j vanish identically in B. Suppose that for any nonempty open set $U \subset \tilde{B}$, and for $\mathcal{D}$ as well as for any datum $\mathcal{D}^*$ associated to $\mathcal{D}$, any real analytic solution $\mathbf{f}$ of the linear equation associated to $\mathcal{D}^*$ vanishes identically in $\mathbf{\Phi}(U)$.

Then there exist $C < \infty$ and $\tau > 0$ such that for any Lebesgue measurable $\mathbf{f}$ and any $\varepsilon > 0$,

$$|\{x \in S(\mathbf{f}, \varepsilon) : \sum_j |f_j \circ \varphi_j(x)| \geq 1\}| \leq C\varepsilon^\tau. \tag{14.8}$$

The proof of this theorem combines a simplification of the proof of Theorem 14.4 with a recursion. The complexity of this recursion increases quite rapidly as the number of summands increases; indeed, the number n_k of recursions needed to reach $|J| = k$ is roughly $2^{n_{k-1}}$; it takes perhaps on the order of $2^{100{,}000}$ to reach $k = 8$. Each recursive step introduces an associated datum, and each associated datum requires its own auxiliary hypotheses. Thus the number of hypotheses quickly becomes rather large. This provides an additional motivation for removal of as many auxiliary hypotheses as possible.

14.4. **Reformulation of the hypothesis.** We establish an equivalent reformulation of the main hypothesis for the three term sublevel set inequality. This provides a straightforward extension of that hypothesis to C^∞ data $(\mathbf{a}, \mathbf{\Phi})$. To any $\mathbf{f} = (f_1, f_2, f_3)$ associate the function

$$G_{\mathbf{f}}(x) = \sum_{j=1}^{3} a_j(x)(f_j \circ \varphi_j)(x). \tag{14.9}$$

Proposition 14.6. *Let $\mathbf{a}, \mathbf{\Phi}$ be real analytic in a neighborhood $\tilde{B}$ of a closed ball $B \subset \mathbb{R}^2$. Assume that none of the mappings φ_j are constant on B, and that there exists no pair of distinct indices $i \neq j \in \{1,2,3\}$ for which $\nabla\varphi_i$ and $\nabla\varphi_j$ are everywhere linearly dependent. Then the following are equivalent.*

(1) *$(\mathbf{a}, \mathbf{\Phi})$ satisfies the main hypothesis.*
(2) *For each point $\bar{x} \in B$ at which $(\mathbf{a}, \mathbf{\Phi})$ is nondegenerate, there exists a positive integer M such that for any $\mathbf{f} \in C^M$ defined in a neighborhood of $\Phi(\bar{x})$, if $G_{\mathbf{f}}$ vanishes to order M at $\bar{x}$ then every f_j vanishes to order M at $\varphi_j(\bar{x})$.*
(3) *There exist $M \in \mathbb{N}$, $C < \infty$, and $\tau > 0$ and a real analytic variety $\Sigma \subset \tilde{B}$ of positive codimension such that for every $\bar{x} \in B$ and every tuple of functions $\mathbf{f} \in C^M$ defined in a neighborhood of $\Phi(\bar{x})$,*

$$\text{distance}(\bar{x}, \Sigma)^\tau \sum_{j=1}^{3} \sum_{0 \leq k \leq M} \left| \left(\frac{d^k}{dy^k} f_j\right)(\varphi_j(\bar{x}))\right| \leq C \sum_{0 \leq |\alpha| \leq M} \left| \frac{\partial^\alpha}{\partial x^\alpha} G_{\mathbf{f}}(\bar{x}) \right|. \tag{14.10}$$

The proof developed below has subsequently been applied in Proposition 7.1 of [19] to prove a corresponding result for inequalities involving two indices j, rather than three.

Condition (3) of the Proposition directly implies (2), which directly implies (1). Conversely, (1) implies (3) by the next lemma and Łojasiewicz's theorem. We say that a function vanishes to order N at a point if the function, and all of its partial derivatives of orders less than or equal to N vanish at that point.

15. $\flat/\sharp$ DECOMPOSITION

A basic strategy for analyzing $\int_{\mathbb{R}^2} \prod_{j=1}^{4} (f_j \circ \varphi_j)\, \eta$, is to apply the Cauchy-Schwarz inequality to eliminate one of the four functions f_j. We have seen, in the analysis above of $\boxed{?}$, how (at least in the case in which φ_j are linear) this Cauchy-Schwarz argument brings in functions $\mathcal{D}_s f_k$ defined by

Definition 15.1. *For $x, s \in \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{C}$,*

$$\mathcal{D}_s f(x) = f(x+s)\overline{f(x)}. \tag{15.1}$$

If f is highly oscillatory, then $\mathcal{D}_s f$ need not be so; if for instance $f(x) = e^{i\lambda x}$ then $\mathcal{D}_s f$ is a constant function of x for each s. On the other hand, if $f(x) = e^{i\lambda x^2}$ with $\lambda \in \mathbb{R}$ and $|\lambda|$ large, then $\mathcal{D}_s f(x)$ takes the form $c(s)e^{i2s\lambda x}$ with $|c(s)| \equiv 1$. The next lemma is based on a simple characterization of those functions for which $\mathcal{D}_s f$ is not highly oscillatory.

Lemma 15.1. [8] *Let $\delta > 0$ and $R \geq 1$. For any $f \in L^2(\mathbb{R})$ there exists a decomposition $f = f_\sharp + f_\flat$ satisfying $\|f_\sharp\|_{L^2} + \|f_\flat\|_{L^2} \lesssim \|f\|_{L^2}$, with the following supplementary properties.*

The summand $f_\sharp$ admits a decomposition

$$f_\sharp(x) = \sum_{n=1}^{M} h_n(x) e^{i\alpha_n x} \tag{15.2}$$

with each $\alpha_n \in \mathbb{R}$, and with each h_n a smooth function satisfying

$$\begin{cases} \|\partial^N h_n\|_\infty \leq C_N R^N \|f\|_\infty \ \forall\, N \geq 0, \\ \|h_n\|_{L^2} \lesssim \|f\|_{L^2}, \\ \widehat{h_n} \text{ is supported in } [-R, R], \\ M \lesssim R^\delta. \end{cases} \tag{15.3}$$

Moreover, the support of $\widehat{f_\sharp}$ is contained in a CR–neighborhood of the support of $\widehat{f}$. Finally, if $f \in L^\infty$ then

$$\|f_\sharp\|_\infty \lesssim M^{1/2}\|f\|_\infty. \tag{15.4}$$

The summand $f_\flat$ satisfies

$$\int_{\mathbb{R}} \int_{|\xi|\le R} |\widehat{\mathcal{D}_s f_\flat}(\xi)|^2 \, d\xi \, ds \lesssim R^{-\delta}\|f\|_{L^2}^4. \tag{15.5}$$

All implicit constants are independent of R, f.

Thus for most s, the function $x \mapsto \mathcal{D}_s f_\flat(x)$ either has small norm, or is highly oscillatory. The conclusion (15.4) is not stated explicitly in [8], but is an immediate consequence of the construction given there.

Lemma 15.1 is proved by repeatedly applying the following simpler decomposition.

Lemma 15.2. *Let $f \in L^2(\mathbb{R}^d)$, $\rho \in (0,1)$ and $R > 0$. Suppose that*

$$\int_{\mathbb{R}^d} \int_{|\xi|\le R} |\widehat{\mathcal{D}_s f}|^2(\xi) \, d\xi \, ds \ge \rho\|f\|_{L^2}^4.$$

Then there exists an orthogonal decomposition $f = g + h$ with $\widehat{g}$ supported in some ball of radius R, $g \perp h$, and $\|g\|_{L^2} \ge \frac{1}{2}\rho^{1/2}\|f\|_{L^2}$.

Proof.

$$\widehat{\mathcal{D}_s f}(\xi) = \int_{\mathbb{R}^d} e^{2\pi i s\cdot(\xi+\xi')}\widehat{f}(\xi+\xi')\overline{\widehat{f}}(\xi')d\xi'$$

and hence

$$|\widehat{\mathcal{D}_s f}(\xi)|^2 = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{2\pi i s\cdot(\xi'-\xi'')}\widehat{f}(\xi+\xi')\overline{\widehat{f}}(\xi')\overline{\widehat{f}}(\xi+\xi'')\widehat{f}(\xi'')\, d\xi'\, d\xi''.$$

Therefore

$$\int_{\mathbb{R}^d} |\widehat{\mathcal{D}_s f}(\xi)|^2 \, ds = \int_{\mathbb{R}^d} |\widehat{f}(\xi+\xi')|^2 |\widehat{f}(\xi')|^2 \, d\xi'$$

and finally

$$\begin{aligned}\int_{\mathbb{R}^d}\int_{|\xi|\le R} |\widehat{\mathcal{D}_s f}(\xi)|^2 \, d\xi\, ds &= \iint_{|\xi-\xi'|\le R} |\widehat{f}(\xi)|^2|\widehat{f}(\xi')|^2 \, d\xi\, d\xi' \\ &\le \|f\|_{L^2}^2 \sup_B \int_B |\widehat{f}|^2\end{aligned}$$

where the supremum is over all balls B of radius R in $\mathbb{R}^d$. Choose B to essentially realize this supremum. The desired decomposition is obtained by defining $g \in L^2$ via $\widehat{g} = \mathbf{1}_B \widehat{f}$ and $h = f - g$. □

16. Proof of the theorem on patterns

Theorem 13.2 is a quantitative Ramsey-type theorem, asserting existence of patterns $(x, x+t, x+t^2)$ in $E \times E \times E$ for measurable sets $E \subset [0,1]$ of arbitrarily small positive Lebesgue measure. The theorem was proved by Bourgain via an energy increment argument (see below) that relies on a particular case of the trilinear oscillatory inequality of Theorem 12.1. In the discussion below, we review that proof with no essential innovations.

Let $\varphi \in C_0^\infty(\mathbb{R}^1)$ be an auxiliary function that satisfies $\varphi \ge 0$, is supported in $[-\frac12, \frac12]$, and satisfies $\varphi(x) > 0$ for all $x \in [-\frac14, \frac14]$. Define $\varphi_r(x) = r\varphi(rx)$ for $r \in (0,\infty)$ and $P_r f = f * \varphi_r$. Define also the translates $\tilde\varphi(x) = \varphi(x-1)$ and $\tilde\varphi_R(x) = \varphi_R(x - R^{-1})$.

Let $E \subset [0,1]$ be a measurable set, and let $f = \mathbf{1}_E$ be its indicator function. Let $R < R' < R''$ be large positive numbers.

The integral

$$\mathcal{I}(R') = \int_{[0,1]} \int_{R't \in [\frac12, \frac32]} \mathbf{1}_E(x)\, \mathbf{1}_E(x+t)\, \mathbf{1}_E(x+t^2)\, dt\, dx$$

counts the number of (x,t) satisfying $R't \in [\frac12, \frac32]$ such that $x, x+t, x+t^2$ all lie in E.

The main step in the proof is the following energy increment lemma, whose thrust is roughly that either $\mathcal{I}(R')$ is appropriately large, or $\int_{R \lesssim |\xi| \lesssim R'} |\widehat{\mathbf{1}_E}(\xi)|^2\, d\xi$ must be reasonably large.

Lemma 16.1. *There exist constants $A, B, C < \infty$ and $c > 0$ such that for any measurable set $E \subset [0,1]$,*

$$c|E|^3 \le R'\mathcal{I}(R') + C\|P_{R''}f - P_R f\|_{L^2} + C\|P_{R'}f - P_R f\|_{L^2} \tag{16.1}$$

provided that $R < R' < R''$ satisfy

$$R'' \ge A|E|^{-B} \text{ and } R' \ge AR|E|^{-3}. \tag{16.2}$$

To deduce Theorem 13.2 from Lemma 16.1, construct sequences of large positive parameters $R_n, R_n', R_n'' = R_{n+1}$ satisfying $R_1 \ge A|E|^{-B}$ and $R_n' \ge AR_n|E|^{-3}$ for every n. Let $D < \infty$ be a large constant to be chosen below, and choose N to be the smallest integer greater than $D|E|^{-6}$.

There are two cases. Firstly, if there exists $n \in [1,N]$ for which

$$\|P_{R_n''}f - P_{R_n}f\|_{L^2} + \|P_{R_n'}f - P_{R_n}f\|_{L^2} \le \tfrac12 C^{-1}c|E|^3 \tag{16.3}$$

then $\mathcal{I}(R_n') \ge c'(R_n')^{-1}|E|^3 > 0$ by (16.1). Thus there exists $t \ge \frac12 (R_n')^{-1}$ such that $x, x+t, x+t^2$ all belong to E, as was to be shown.

If there does not exist such an n then

$$\sum_{n=1}^N \|P_{R_n''}f - P_{R_n}f\|_{L^2}^2 \ge cN|E|^6$$

or the same holds for $\sum_{n=1}^N \|P_{R_n'}f - P_{R_n}f\|_{L^2}^2$. In the former case, $|E|^{1/2} = \|\widehat{\mathbf{1}_E}\|_{L^2}$ is forced to be very large, since

$$\sum_{n=1}^N \|P_{R_n''}f - P_{R_n f}\|_{L^2}^2 = \int |\widehat{f}(\xi)|^2 \sum_{n=1}^N |\widehat\varphi(R_{n+1}\xi) - \widehat\varphi(R_n\xi)|^2\, d\xi$$

and it is well known, and easily verified, that

$$\sum_{n=1}^\infty |\widehat\varphi(R_{n+1}\xi) - \widehat\varphi(R_n\xi)|^2 < \infty$$

uniformly for all $\xi \in \mathbb{R}$, provided merely that $A \ge 2$ so that $R_{n+1} \ge 2R_n$ for every n.

Consequently

$$\sum_{n=1}^\infty \|P_{R_n''}f - P_{R_n f}\|_{L^2}^2 < \infty.$$

Therefore $N = O(|E|^{-6})$, which is a contradiction if D is chosen to be sufficiently large. The same reasoning applies in the latter case. Thus there must exist $n \le N \le D|E|^{-6}$ for which (16.3) holds. The proof of Theorem 13.2 is complete. □

Proof of Lemma 16.1. In this discussion I will slur over the distinction between integrating over $[0,1]^2$, and integrating over $\mathbb{R}^2$ with respect to a smooth compactly supported cutoff function.

$$R'\mathcal{I}(R') \ge c\int_{[0,1]^2} f(x)\, f(x+t)\, f(x+t^2)\, \tilde\varphi_{R'}(t)\, dt\, dx.$$

In this integral, split the third factor f as $f = (f - P_{R''}f) + P_{R''}f$. By Theorem 12.1, the contribution of the first term, $f - P_{R''}f$, is $O((R'')^{-\sigma})$ for a certain $\sigma > 0$.[15] Thus

$$\begin{aligned} R'\mathcal{I}(R') &\ge c\int_{[0,1]^2} f(x)\, f(x+t)\, (P_{R''}f)(x+t^2)\, \tilde\varphi_{R'}(t)\, dt\, dx - O((R'')^{-\sigma}) \\ &\ge c\int_{[0,1]^2} f(x)\, f(x+t)\, (P_{R}f)(x+t^2)\, \tilde\varphi_{R'}(t)\, dt\, dx - O((R'')^{-\sigma}) - O(\|P_{R''}f - P_R f\|_{L^2}). \end{aligned}$$

To obtain the last line, we have replaced R'' by R in the integral, and have used the fact that $\|\tilde\varphi_{R'}\|_{L^1} = O(1)$ uniformly in R'. The function P_Rf is Lipschitz with Lipschitz bound $O(R)$, and t^2 is $O((R')^{-2})$ in the support of $\tilde\varphi_{R'}(t)$, so $P_Rf(x+t^2) = P_Rf(x) + O(R(R')^{-2})$ on the support of the integrand. Thus

$$\begin{aligned} &\int_{[0,1]^2} f(x)\, f(x+t)\, (P_Rf)(x+t^2)\, \tilde\varphi_{R'}(t)\, dt\, dx \\ &\qquad\qquad \ge \int_{[0,1]^2} f(x)\, f(x+t)\, (P_Rf)(x)\, \tilde\varphi_{R'}(t)\, dt\, dx - O(R(R')^{-2}). \end{aligned}$$

The integral on the right-hand side is equal to

$$\int_{[0,1]} f(x)\, (P_{R'}f)(x+R'^{-1})\, (P_Rf)(x)\, dx$$

because $\tilde\varphi_r(y) = \varphi_r(y - r^{-1})$. The final step here is to replace $(P_{R'}f)(x+R'^{-1})$ by $P_{R'}f(x)$, up to an acceptable remainder term. The last integral is

$$\ge \int_{[0,1]} f(x)\, (P_Rf)(x+R'^{-1})\, (P_Rf)(x)\, dx - O(\|P_Rf - P_{R'}f\|_{L^2}).$$

P_Rf is Lipschitz, with Lipschitz norm $O(R)$, so $|P_Rf(x+R'^{-1}) - P_Rf(x)| = O(R/R')$. Combining all of these steps, we have shown that

$$R'\mathcal{I}(R') \ge c\int_{[0,1]} f \cdot P_Rf \cdot P_R - O\big(\|P_Rf - P_{R'}f\|_{L^2} + \|P_Rf - P_{R''}f\|_{L^2}\big)$$

minus a remainder term that is

$$O((R'')^{-\sigma}) + O(RR'^{-2}) + O(R/R').$$

[15] The factor $\tilde\varphi_{R'}$ does not appear in Theorem 12.1. It can be absorbed by writing $\tilde\varphi_{R'}(t) = \int a(\xi)e^{it\xi}\, d\xi$ with $a \in L^1$ uniformly in R'. Each $e^{it\xi}$ can be rewritten as $e^{-ix\xi}e^{i(x+t)\xi}$, and the first factor can be absorbed into $f(x)$ and the second into $f(x+t)$, resulting for each ξ in a form to which Theorem 12.1 can be directly applied.

We require that $R'' \ge A|E|^{3/\sigma}$ and $R' \ge AR|E|^{-3}$ for a large constant A. Then

$$\text{(16.4)}\quad R'\mathcal{I}(R') \ge c\int_{[0,1]} f\cdot P_R f\cdot P_R \\ - O\big(\|P_R f - P_{R'}f\|_{L^2} + \|P_R f - P_{R''}f\|_{L^2}\big) - a|E|^3$$

where $a \to 0$ as $A \to \infty$. To complete the proof of Lemma 16.1, it suffices to have a strictly positive lower bound for the integral that remains. The next lemma provides such a bound. □

Lemma 16.2. *There exists a constant $c_0 > 0$ such that for every Lebesgue measurable set $E \subset [0,1]$, $f = \mathbf{1}_E$ satisfies*

$$\int_{[0,1]} f\cdot P_R f\cdot P_{R'} f \ge c_0|E|^3$$

uniformly for all $R, R' \ge 2$.

One can regard this integral as representing the probability that $x, x+s, x+t$ all belong to E if x is chosen uniformly at random in $[0,1]$. s uniformly at random satisfying $0 \le s \lesssim 1/R$, and t uniformly at random satisfying $0 \le t \lesssim 1/R'$. The straightforward proof will be given below.

Inserting the conclusion of Lemma 16.2 into (16.4) gives the conclusion of Lemma 16.1 provided that A is chosen sufficiently large to ensure that $a < c_0$. □

A systematic accounting of the parameters — specifically, of the rate of growth of N as a function of $|E|$ — yields the bound $t > \exp(-\exp(\varepsilon^{-C}))$ in the statement of Theorem 13.2. This accounting is left to the eager reader. □

Proof of Lemma 16.2. Assume that $R \le R'$. Tile $\mathbb{R}$ by intervals I of length $1/4R'$, and by intervals of some length between $1/8R$ and $1/4R$, such that each interval J of the greater length is a union of certain intervals I of the smaller length.

For each interval I, $P_{R'}f(x) \ge c'|I|^{-1}\int_I f$ for every $x \in I$, while likewise $P_R f(x) \ge |J|^{-1}\int_J$ for every $x \in J$. Therefore

$$\begin{aligned}\int_{[0,1]} f\cdot P_R f\cdot P_{R'} f &\ge c\sum_J \int_J \sum_{I\subset J}\int_I \big[f\cdot |I|^{-1}\textstyle\int_I f\ \cdot |J|^{-1}\int_J f\big]\\ &= \sum_J |J|^{-1}\textstyle\int_J f\cdot \displaystyle\int_J \sum_{I\subset J} |I|^{-1}(\textstyle\int_I f)^2\\ &\ge \sum_J |J|^{-1}\textstyle\int_J f \displaystyle\int_J (\sum_{I\subset J}\textstyle\int_I f)^2\cdot |J|^{-1}\\ &= \sum_J |J|^{-2}\int_J f(\int_J f)^2\\ &\ge (\sum_J \int_J f)^3\end{aligned}$$

using Hölder's inequality to obtain the third and sixth lines. □

17. On the proof of the three term sublevel set inequality

17.1. **A model argument, and some strategy.** Let $\varphi_j(x,t) = x + t^j$ for $(x,t) \in \mathbb{R}^2$. Let $f_j \in C^2$. We sketch a proof that if $\sum_{j=0}^2 f_j(x+t^j) \equiv 0$ in a connected set Ω, then each f_j is constant in $\varphi_j(\Omega)$. This proof is based on differentiation of the equation.

Differentiate with respect to t to get

$$\sum_{j=1}^{2} jt^{j-1} f_j'(x+t^j) \equiv 0$$

in Ω. Apply $\partial_t - \partial_x$ to get

$$(4t^2 - 2t) f_2''(x+t^2) = 2f_2'(x+t^2)$$

in Ω. This forces $f_2'' \equiv f_2' \equiv 00$ in $\varphi_2(\Omega)$, so f_2 is constant. Thus $f_0(x) + f_1(x+t) = C$, some constant. By freezing x we find that f_1 is locally constant, hence globally constant, in $\varphi_1(\Omega)$. Then the says directly that f_0 is constant. □

The proof of Theorem 14.4 is partly modeled on this argument. Given $\mathbf{f}$ satisfying $|\sum_{j=1}^3 a_j \cdot (f_j \circ \varphi_j)| < \varepsilon$ and $\sum_j |f_j \circ \varphi_j| \ge 1$ at every point of a set $S \subset \mathbb{R}^2$, we aim to prove that there exist differentiable functions g_j and $S' \subset S$ satisfying $|S'| \gtrsim |S|^C$ such that $g_j = f_j + O(\varepsilon)$ on $\varphi_j(S')$, and moreover

$$\big|\nabla \sum_{j=1}^{3} a_j (g_j \circ \varphi_j)\big| = O(\varepsilon) \text{ on } S'.$$

Since the gradient has two components, this is a system of two inequalities. One inequality can be used to eliminate g_3. The result is a single inequality in two unknown functions g_1, g_2, and their derivatives.

The presence of the derivatives is a complication, but having only two mappings $\varphi_j : \mathbb{R}^2 \to \mathbb{R}^1$ in play is an enormous simplification, as we have seen above in the proof of Lemma 9.2. The variable coefficients a_j turn out to be an advantage at this stage.

Thus much of the analysis is a proof of a regularity theorem, in which all functions are well-defined only up to $O(\varepsilon)$ additive corrections. There are three main steps. In the first, a preliminary approximation of f_j by g_j is accomplished, with the domain of each f_j tiled by intervals of length ε and with functions $\tilde{g}_j$ constructed, each of which is constant on each interval in the tiling. In the second step, the domain of each $\tilde{g}_j$ is tiled by intervals of length $\delta = \varepsilon^{3/4}$, and improved approximators g_j are constructed, each of which is affine on each tile with $|\nabla g_j| = O(1)$ and with $|\tilde{g}_j - g_j| = O(\delta^{1+\kappa})$ for a certain exponent $\kappa > 0$. This holds on $S' \subset S$, which satisfies $|S'| \gtrsim |S|^C$. This is sufficient information to lead to the desired bound for $\nabla \sum_j a_j (g_j \circ \varphi_j)$. In this way, a more tractable sublevel set problem is constructed from the given one.

17.2. **Elementary combinatorial inequalities.** I will often write $\mathbb{R}^k$ for $k = 1, 2$ when a subset of $\mathbb{R}^k$ is intended.

For each index j, let V_j be a nowhere vanishing vector field with C^ω coefficients that satisfies $V_j(\varphi_j) \equiv 0$. $\mathbb{R}^2$ is foliated by the one-parameter family of C^ω curves $\{x \in \mathbb{R}^2 : \varphi_j(x) = y\}$, parametrized by $y \in \mathbb{R}^1$. (More precisely, by $y \in \varphi_j(B) \subset \mathbb{R}^1$. For each $x \in \mathbb{R}^2$, ehe mapping $\mathbb{R}^1 \ni t \mapsto e^{tV_j}(x)$ parametrizes this curve, with $y = \varphi_j(x)$. $\{(x,x') \in \mathbb{R}^2 \times \mathbb{R}^2 : \varphi_j(x) = \varphi_j(x')\}$ is thus naturally parametrized by $(x,s) \in \mathbb{R}^2 \times \mathbb{R}$ via $x' = e^{sV_j}(x)$.

The next lemma is a widely used combinatorial principle, adapted to the continuum setting, with sizes of sets measured by Lebesgue measure(s) rather than by cardinality.

Lemma 17.1. *For any $\delta > 0$ and any $S \subset \mathbb{R}^2$,*

$$|\{(x,s) : x \in S \text{ and } e^{sV_j}(x) \in S\}| \geq c\delta|S|^2.$$

Proof. Choose coordinates in which $\varphi_j(x_1, x_2) \equiv x_1$. Then the measure of the set in question is equal to

$$\iiint_{|x_2 - x_2'| \leq \delta} \mathbf{1}_S(x_1, x_2)\mathbf{1}_S(x_1, x_2')\,dx_2\,dx_2'\,dx_1.$$

By Cauchy-Schwarz, for $\delta = 1$,

$$\begin{aligned}
|S| &= \iint \mathbf{1}_S(x_1, x_2)\,dx_1\,dx_2 \\
&\lesssim \Big(\int \big(\int \mathbf{1}_S(x_1, x_2)\,dx_2\big)^2\,dx_1\Big)^{1/2} \\
&= \Big(\iiint \mathbf{1}_S(x_1, x_2)\mathbf{1}_S(x_1, x_2')\,dx_2\,dx_2'\,dx_1\Big)^{1/2},
\end{aligned}$$

which is the bound claimed for $\delta = 1$.

For the general case, let $S_n = \{(x_1, x_2) \in S : n\delta \leq x_2 < (n+1)\delta\}$, apply the case $\delta = 1$ to obtain $|S_n| \lesssim |S_n'|^{1/2}$, and apply Cauchy-Schwarz to obtain

$$|S| = \sum_n |S_n| \leq \sum_n |S_n'|^{1/2} \lesssim \delta^{-1/2}\big(\sum_n |S_n'|\big)^{1/2} = \delta^{-1/2}|S'|.$$

□

Variants of this lemma arise repeatedly in the analysis, and are proved in the same way.

Lemma 17.2. *For any measurable $S \subset \mathbb{R}^2$ there exists a measurable set $S' \subset S$ satisfying $|S'| \gtrsim |S|^2$ such that every $x \in S'$,*

$$|\{x' \in S' : \varphi_j(x') = \varphi_j(x)\}| \gtrsim |S|.$$

In the conclusion, $E = \{x' \in S' : \varphi_j(x') = \varphi_j(x)\}$ is identified with a subset of $\mathbb{R}^1$ via the inverse of the exponential map $t \mapsto e^{tV_j}(x)$, and $|E|$ denotes the one-dimensional Lebesgue measure of this subset of $\mathbb{R}^1$.

This lemma is a direct consequence of Fubini's theorem. Details are left to the reader. □

Throughout the discussion, this type of situation will arise, with a subset E of some Euclidean space naturally identified with a subset of a lower-dimensional space, and the notation $|E|$ will denote this lower-dimensional Lebesgue measure, without any explicit notatioal indication.

Repeated application of Lemma reflemma:doublex gives this conclusion:

Lemma 17.3. *There exist measurable sets $T \subset \mathbb{R}^3$ and $T^* \subset \mathbb{R}^4$, and a point $\bar{x}$, such that*

$$\begin{cases}
|T| \gtrsim |S|^8 \\
T^*| \gtrsim |S|^{16} \\
\forall\, t \in T,\ \ |\{t_4 : (t_1, t_2, t_3, t_4) \in T^*\}| \gtrsim |S|^8, \\
e^{t_3V_1}e^{t_2V_2}e^{t_1V_1}\bar{x} \in S \quad \forall\,(t_1, t_2, t_3) \in T \\
e^{t_4V_2}e^{t_3V_1}e^{t_2V_2}e^{t_1V_1}\bar{x} \in S \quad \forall\,(t_1, t_2, t_3, t_4) \in T^*
\end{cases} \tag{17.1}$$

17.3. **A supplementary hypothesis.** We aim for an upper bound for the Lebesgue measure of

$$\mathbf{S} = \{x \in S(\mathbf{f}, \varepsilon) : |(f_3 \circ \varphi_3)(x)| \ge 1\}.$$

No nontrivial upper bound is available if each f_j is allowed to be small on a large set, so the lower bound $|f_3 \circ \varphi_3)(x)| \ge 1$ is imposed.

Matters reduce to the case in which a supplementary hypothesis

$$|(f_j \circ \varphi_j)(x)| \le 2 \ \forall\, x \in \mathbf{S} \tag{17.2}$$

is also given. Indeed, partition $\mathbf{S}$ as a pairwise disjoint union $\mathbf{S} = \bigcup_{n \ge 0} \bigcup_{k=1}^{3} \mathbf{S}_{n,k}$ with the property that

$$|(f_k \circ \varphi_k)(x)| \ge 2^n \ \forall\, x \in \mathbf{S}_{n,k}$$
$$|(f_j \circ \varphi_j)(x)| < 2^{n+1} \ \forall\, x \in \mathbf{S}_{n,k}.$$

The supplementary hypothesis can be applied for $\mathbf{S}_{n,k}$ upon replacing each f_j by $2^{-n} f_j$ and replacing ε by $2^{-n}\varepsilon$; the formulation of Theorem 14.4 is invariant under permutation of the indices $j \in \{1,2,3\}$, so having a lower bound for $|f_k|$ is functionally equivalent to having a lower bound for $|f_3|$. We find that if the conclusion of Theorem 14.4 can be established under the supplementary hypothesis, it follows that

$$|\mathbf{S}_{n,k}| \lesssim (2^{-n}\varepsilon)^{\tau} \ \forall\, n, k.$$

By summing over n, k we obtain the desired conclusion in general, without any supplementary hypothesis. □

We will assume henceforth that (17.2) is satisfied.

17.4. **Reduction to a sublevel problem with only two unknown functions.** Let $S = S(\mathbf{f}, \varepsilon)$. Form

$$S' = \{(x, s) : \ x \in S,\ e^{sV_3} x \in S, \text{ and } |s| \le \varepsilon\}.$$

By Lemma 17.1, $|S'| \gtrsim \varepsilon |S|^2$. Form

$$\tilde{S} = \{(x, s) : \ x \in S,\ e^{\varepsilon t V_3} x \in S, \text{ and } |t| \le 1\}.$$

Then $|\tilde{S}'| \gtrsim |S|^2$. We aim for a lower bound $|\tilde{S}| \gtrsim \varepsilon^c$, which implies a corresponding bound for $|S|$ with exponent $c/2$.

By dividing through by the nowhere vanishing coefficient a_3 in the definition of $S(\mathbf{f}, \varepsilon)$, we may reduce matters to the case in which $a_3 \equiv 1$. Consider any $(x, t) \in \tilde{S}$ and write $(x, x') = (x, e^{\varepsilon t V_3} x) \in S^2$. Subtraction gives

$$\sum_{j=1}^{2} \big[a_j(x)(f_j \circ \varphi_j)(x) - a_j(x')(f_j \circ \varphi_j)(x') \big] = O(\varepsilon).$$

Now $|x - x'| = O(\varepsilon)$, so $|a_j(x) - a_j(x')| = O(\varepsilon)$, so this can be simplified to

$$\sum_{j=1}^{2} a_j(x) \big[(f_j \circ \varphi_j)(e^{t\varepsilon V_3} x) - (f_j \circ \varphi_j)(x) \big] = O(\varepsilon).$$

Define mappings $\psi_{j,\varepsilon} : \mathbb{R}^3 \to \mathbb{R}^2$ and functions $g_j : \mathbb{R}^2 \to \mathbb{C}$ by

$$\begin{cases} \psi_j^{\varepsilon}(x, t) = (\varphi_j(x), \varepsilon^{-1} \big[\varphi_j(e^{\varepsilon t W_3} x) - \varphi_j(x) \big]) \\ g_j(y, s) = f_j(y + \varepsilon s) - f_j(y). \end{cases} \tag{17.3}$$

Thus

$$(f_j \circ \varphi_j)(e^{\varepsilon t V_3} x) - (f_j \circ \varphi_j)(x) = (g_j \circ \psi_j^\varepsilon)(x,t)$$

and consequently

$$\sum_{j=1}^{2} a_j(x)(g_j \circ \psi_j)(x,t) = O(\varepsilon) \ \forall \, (x,t) \in \tilde{S}.$$

We have arrived at a sublevel set problem in a three-dimensional ambient space, with two unknown functions g_j, each of which is a function of two real variables. Note that no *lower* bound on $|g_j|$ is given or deducible. But our goal is to show that $|g_j|$ are small on suitably large sets, so a proof by contradiction would naturally provide us with such a lower bound.

As $\varepsilon \to 0$, the mappings ψ_j^ε converge in the C^∞ topology to mappings ψ_j. The vector fields W_j^ε can be chosen so that they also converge to limiting vector fields W_j.

17.5. **A two term sublevel set inequality.** This subsection is concerned with two term sublevel set inequalities, for functions of two variables, in a three-dimensional ambient space. The next proposition is a general result that treats a slight simplification of the situation that arose in §17.4. Its proof contains the main elements of a proof for the elaboration needed in that context.

The setup for the next result is as follows. Let $B \subset \mathbb{R}^3$ be a ball, let $\psi_j : B \to \mathbb{R}^2$ be C^ω submersions and let a_j be nowhere vanishing C^ω functions. To any Lebesgue measurable functions $g_1, g_2 : \mathbb{R}^2 \to \mathbb{C}$ associate the sublevel sets

$$S(\mathbf{g}, \varepsilon) = \{x \in B : \Big| \sum_{j=1}^{2} b_j(x)\, (g_j \circ \psi_j)(x) \Big| < \varepsilon\}. \tag{17.4}$$

Let W_j be nowhere vanishing C^ω vector fields in a neighborhood of B that satisfy $W_j(\psi_j) \equiv 0$.

Proposition 17.4. *Suppose that W_1, W_2 satisfy the bracket condition at each point of $\tilde{B} \times (\mathbb{R} \setminus \{0\})$. Suppose also that in every nonempty open subset of a neighborhood of B, any C^ω solution $\mathbf{g}^* = (g_1^*, g_2^*)$ of the equation*

$$\sum_{j=1}^{2} b_j \cdot (g_j^* \circ \psi_j) = 0$$

vanishes identically. Then there exist $C < \infty$ and $\tau > 0$ such that for any $\varepsilon > 0$ and any pair of measurable functions g_j,

$$\big|\{x \in S(\mathbf{g}, \varepsilon) : \max_j |(g_j \circ \psi_j)(x)| \ge 1\}\big| \le C\varepsilon^\tau. \tag{17.5}$$

The constants C, ε_0 depend on B, ψ_j, σ but neither on S nor on the functions g_j.

We next prove Proposition 17.4. The next two lemmas are used in that proof. They are essentially exercises in differential calculus, the implicit function theorem, and aspects of real analytic functions.

Lemma 17.5. *If $\Phi = (\varphi_1, \varphi_2, \varphi_3)$ is not linearizable then the limiting vector fields W_1, W_2 satisfy the bracket condition at every point of $B \times (\mathbb{R} \setminus \{0\})$.*

Since $W_j^\varepsilon \to W_j$ in the C^∞ topology as $\varepsilon \to 0$, the same conclusion holds for $\{W_1^\varepsilon, W_2^\varepsilon\}$ for every sufficiently small $\varepsilon > 0$. Moreover, the bracket condition holds uniformly in ε, uniformly on every compact subset of $B \times (\mathbb{R} \setminus \{0\})$, in the natural sense.

For $x \in \mathbb{R}^3$ define $\Theta_x : \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\Theta_x(t_1, t_2, t_3) = e^{t_3 W_1} e^{t_2 W_2} e^{t_1 W_1} x.$$

Let $J_x(t) = J_x(t_1, t_2, t_3)$ be the Jacobian determinant of Θ_x with respect to t. We will often suppress the subscript x and write simply $\Theta(t)$.

Lemma 17.6. *Let ψ_1, ψ_2, W_1, W_2 be as in Proposition 17.4. Assume that $\{W_1, W_2\}$ satisfies the bracket condition. Then $J_x(t)$ does not vanish identically as a function of $(x, t) \in B \times \mathbb{R}^3$.*

The hypotheses, and therefore the conclusion, are invariant under permutation of the indices $j = 1, 2$.

Lemma 17.7. *Let $S \subset \mathbb{R}^3$ be Lebesgue measurable. There exist $\bar{x} \in B$ and a set $T \subset \mathbb{R}^3$ such that*

$$\begin{cases} |T| \gtrsim |S|^A \\ \Theta(t) \in S \ \forall\, t \in T \end{cases} \tag{17.6}$$

The proof is a repeated application of the proof of Lemma 17.2.

Define $F : \mathbb{R}^3 \to \mathbb{C}$ as follows. Write the relation as

$$(f_1 \circ \psi_1)(x) = a(x)\,(f_2 \circ \psi_2)(x) + O(\varepsilon) \quad \forall\, x \in S$$

where $a \in C^\omega$ vanishes nowhere.

$$\begin{cases} F(0,0,0) = (f_1 \circ \psi_1)(\bar{x}) \\ F(t_1, 0, 0) = F(0,0,0) \\ F(t_1, t_2, 0) = \dfrac{a(\Theta_{\bar{x}}((t_1, t_2, 0)))}{a(\Theta_{\bar{x}}((t_1, 0, 0)))} (f_1 \circ \psi_1)(\Theta_{\bar{x}}((t_1, 0, 0)) \\ F(t_1, t_2, t_3) = F(t_1, t_2, 0). \end{cases} \tag{17.7}$$

The function F is manifestly real analytic; the successive lines of (17.7) define analytic functions whose domains are respectively subsets of $\mathbb{R}^0$, $\mathbb{R}^1$, $\mathbb{R}^2$, and finally $\mathbb{R}^3$.

I claim that

$$F(t) = (f_1 \circ (\psi_1 \circ \Theta))(t) + O(\varepsilon)\ \forall\, t \in T. \tag{17.8}$$

Consequently, since $t \in T \Rightarrow \Theta(t) \in S$, it is also true that

$$F(t) = a(\Theta(t))(f_2 \circ (\psi_2 \circ \Theta))(t) + O(\varepsilon) \quad \forall\, t \in T. \tag{17.9}$$

Proof of (17.8). The relation (17.8) holds for $t = (0,0,0)$ by the first clause in the definition since $\Theta(0,0,0) = \bar{x}$. It consequently holds for all $(t_1, 0, 0) \in T$ since $\Theta(t_1, 0, 0) \in S$ and

$$\psi_1(\Theta(t_1, 0, 0)) = \psi_1(e^{tW_1}\bar{x}) = \psi_1(\bar{x}).$$

Therefore

$$(f_1 \circ \psi_1)(\Theta(t_1, 0, 0)) = (f_1 \circ \psi_1)(\bar{x}) = (f_1 \circ \psi_1)(e^{t_1 W_1}\bar{x}) = (f_1 \circ \psi_1)(\Theta(t_1, 0, 0)).$$

For $t = (t_1, t_2, 0) \in T$,

$$\begin{aligned}(f_1 \circ \psi_1)(\Theta(t)) &= a(\Theta(t))(f_2 \circ \psi_2)(\Theta(t)) + O(\varepsilon) \\ &= a(\Theta(t))(f_2 \circ \psi_2)(\Theta(t_1, 0, 0)) + O(\varepsilon) \\ &= \frac{a(\Theta(t))}{a(\Theta(t_1, 0, 0))} \cdot a(\Theta(t_1, 0, 0))(f_2 \circ \psi_2)(\Theta(t_1, 0, 0)) + O(\varepsilon) \\ &= \frac{a(\Theta(t))}{a(\Theta(t_1, 0, 0))}(f_1 \circ \psi_1)(\Theta(t_1, 0, 0)) + O(\varepsilon) \\ &= F(t) + O(\varepsilon).\end{aligned}$$

The first equality holds because $\Theta(t) \in S$, the second because $\psi_2(\Theta(t_1, 0, 0)) = \psi_2(\Theta(t_1, t_2, 0))$, and the third because $\Theta(t_1, 0, 0) \in S$, while the fourth is the definition of $F(t_1, t_2, 0)$. Finally, for $t = (t_1, t_2, t_3) \in T$ with $t_3 \neq 0$,

$$F(t) = F(t_1, t_2, 0) = (f_1 \circ \psi_1)(\Theta(t_1, t_2, 0)) + O(\varepsilon) = (f_1 \circ \psi_1)(\Theta(t)) + O(\varepsilon).$$

$\square$

Define $\tilde{\psi}_j = \psi_j \circ \Theta$, and $\tilde{a} = a \circ \Theta$. There exist C^ω vector fields $\tilde{W}_j$ whose push-forwards with respect to Θ satisfy $\Theta_*(\tilde{W}_j) = JW_j$, where J is the Jacobian determinant of Θ with respect to t. $\tilde{W}_j$ factors through $\tilde{\psi}_j$, and conversely, if $h \in C^1$ satisfies $\tilde{W}_j(h) \equiv 0$ then h factors through $\tilde{\psi}_j$.

By (17.8) and (17.9),

$$\begin{cases} F(e^{s_1 \tilde{W}_1} t) = F(t) + O(\varepsilon) \\ \tilde{a}(e^{s_2 \tilde{W}_2} t)^{-1} F(e^{s_2 \tilde{W}_2} t) = \tilde{a}(t)^{-1} F(t) + O(\varepsilon) \end{cases}$$

whenever t, $e^{s_1 \tilde{W}_1} t$, and $e^{s \tilde{W}_2} t$ all belong to T. The set of all (t, s_1, s_2) satisfying this condition has five-dimensional Lebesgue measure $\gtrsim |T|^C$ for some $C < \infty$.[16]

The function F is completely determined by ψ_1, ψ_2 and a single datum $(\bar{x}, \alpha)$ with $\alpha = f_1(\psi_1(\bar{x}))$. This datum ranges over a compact subset of $\mathbb{R}^3 \times (\mathbb{C} \setminus \{0\})$ since it is assumed that $|f_j| = O(1)$ and that $\max_j |(f_j \circ \psi_j)(\bar{x})| \geq 1$. Denote by $\mathcal{F}$ the set of all such functions $F_{(\bar{x}, \alpha)}$.

I claim that if some function H that is C^ω in some open subset of $\mathbb{R}^3$ satisfies the exact relations

$$\begin{cases} H(e^{s_1 \tilde{W}_1} t) = H(t) \\ \tilde{a}(e^{s_2 \tilde{W}_2} t)^{-1} H(e^{s_2 \tilde{W}_2} t) = \tilde{a}(t)^{-1} H(t) \end{cases}$$

for all (t, s_1, s_2) in some nonempty open set in $\mathbb{R}^5$, then H vanishes identically. Indeed, there exists a nonempty open subset on which Θ is a bijection. By defining $f_1(y) = H(t)$ with $t \in \tilde{\psi}_1 - 1(\{y\})$ and similarly $f_2(y) = \tilde{a}(t)^{-1} H(t)$ with $t \in \tilde{\psi}_2^{-1}(\{y\})$ we obtain a real analytic exact solution of the equation $(f_1 \circ \psi_1) = a \cdot (f_2 \circ \psi_2)$ in a nonempty open subset of $\mathbb{R}^3$. By hypothesis, $f_j \circ \psi_j$ both vanish identically in that open set. Therefore H vanishes identically in a nonempty open subset of its domain. This reasoning applies in any connected component of the domain of H. Therefore H vanishes identically. $\square$

The family $\mathcal{F}$ of all functions F that are defined by (17.7) from data (f_1, f_2) that are $O(1)$ is parametrized by $(\bar{x}, \alpha)$ with $\alpha = (f_1 \circ \psi_1)(\bar{x})$. The parameter $(\bar{x}, \alpha)$ ranges over a

[16] Justification is needed here, but these lectures do not include full details.

compact subset of $\mathbb{R}^3 \times (\mathbb{C} \setminus \{0\})$. The mapping $(\bar{x}, \alpha, t) \mapsto F(t) = F_{(\bar{x},\alpha)}(t)$ is real analytic. Moreover, F does not vanish identically, since $F(\bar{x}) = \alpha \neq 0$.

Consider the family $\mathcal{G}$ consisting of all functions

$$G_{\bar{x},\alpha}(t, s_1, s_2) = |F(e^{s_1 \tilde{W}_1} t) - F(t)|^2 + |\tilde{a}(e^{s_2 \tilde{W}_2} t)^{-1} F(e^{s_2 \tilde{W}_2} t) - \tilde{a}(t)^{-1} F(t)|^2$$

such that $F \in \mathcal{F}$. This is a family of real analytic functions of (t, s_1, s_2) that depends real analytically on $(\bar{x}, \alpha)$. We have shown, two paragraphs above, that no function $G_{\bar{x},\alpha} \in \mathcal{G}$ vanishes identically. Therefore by Lemma 2.9 there exist C, τ such that

$$\big|\big\{(t, s_1, s_2) : |G_{\bar{x},\alpha}(t, s_1, s_2)| < \delta\big\}\big| \leq C\delta^\tau$$

uniformly for all $(\bar{x}, \alpha)$ and all $\delta \in (0, 1]$.

Choose $\delta = C\varepsilon$ for a large constant C. By our construction, the measure of the set of all such (t, s_1, s_2) is $\geq c|S|^C$ for some $c > 0$ and $C < \infty$. Thus $|T| \leq C\varepsilon^c$, and consequently the same holds for the measure of S, for some (other) constants $c, C \in \mathbb{R}^+$. $\square$

References

[1] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. 9 (1996), 725–753.

[2] W. Blaschke and G. Bol, *Geometrie der Gewebe*, Springer-Verlag, 1938.

[3] J. Bourgain, *A nonlinear version of Roth's theorem for sets of positive density in the real line*, J. Analyse Math. 50 (1988), 169–181

[4] ______, *A remark on the maximal function associated to an analytic vector field*, Analysis at Urbana, Vol. I (Urbana, IL, 1986-1987), 111–132, London Math. Soc. Lecture Note Ser., 137, Cambridge Univ. Press, Cambridge, 1989.

[5] ______, *On the dimension of Kakeya sets and related maximal inequalities*, Geom. Funct. Anal. 9 (1999), 256–282.

[6] A. Carbery, M. Christ, and J. Wright, *Multidimensional van der Corput and sublevel set estimates*, J. Amer. Math. Soc. 12 (1999), no. 4, 981–1015

[7] A. Carbery and J. Wright, *What is van der Corput's lemma in higher dimensions?* Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000). Publ. Mat. 2002, Vol. Extra, 13–26.

[8] M. Christ, P. Durcik, and J. Roos, *Trilinear smoothing inequalities and a variant of the triangular Hilbert transform*, Adv. Math. 390 (2021), Paper No. 107863

[9] M. Christ, X. Li, T. Tao, and C. Thiele, *On multilinear oscillatory integrals, nonsingular and singular*, Duke Math. J. 130 (2005), no. 2, 321–351

[10] M. Christ, *On certain elementary trilinear operators*, Math. Research Letters 8 (2001), 43-56.

[11] ______, *Near-extremizers of Young's inequality for Euclidean groups*, Rev. Mat. Iberoam. 35 (2019), no. 7, 1925–1972

[12] ______, *Young's inequality sharpened*, Geometric aspects of harmonic analysis, 261–298, Springer INdAM Ser., 45, Springer, Cham, [2021]

[13] ______, *Bounds for multilinear sublevel sets via Szemerédi's theorem*, preprint, arXiv:1107.2350 math.CA

[14] ______, *Multilinear bounds under finitely witnessed nondegeneracy*, in preparation

[15] ______, *On trilinear oscillatory integral inequalities and related topics*, preprint, arXiv:2007.12753

[16] ______, *On implicitly oscillatory quadrilinear integrals*, preprint

[17] ______, *A three term sublevel set inequality*, preprint

[18] M. Christ and D. Oliveira e Silva, *On trilinear oscillatory integrals*, Rev. Mat. Iberoam. 30 (2014), no. 2, 667–684.

[19] M. Christ and Z. Zhou, *A class of singular bilinear maximal functions*, preprint, arXiv:2203.16725

[20] D. Dong, D. Maldague, and D. Villano, *Special cases of power decay in multilinear oscillatory integrals*, preprint, arXiv:1904.05428 math.CA

[21] L. C. Evans, in preparation

[22] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. 34 (1978), 275–291.

[23] M. Gilula, P. T. Gressman, and L. Xiao, *Higher decay inequalities for multilinear oscillatory integrals*, Math. Res. Lett. 25 (2018), no. 3, 819–842.
[24] M. Gilula and K. O'Neill, *Oscillatory Loomis-Whitney and projections of sublevel sets*, arXiv:1903.12300, math.CA
[25] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. 8 (1998), no. 3, 529–551.
[26] M. Greenblatt, *Simply nondegenerate multilinear oscillatory integral operators with smooth phase*, Math. Res. Lett. 15 (2008), no. 4, 653–660.
[27] P. T. Gressman and L. Xiao, *Maximal decay inequalities for trilinear oscillatory integrals of convolution type*, J. Funct. Anal. 271 (2016), no. 12, 3695–3726.
[28] L. Hörmander, *Oscillatory integrals and multipliers on FL^p*, Ark. Mat. 11, (1973), 1–11.
[29] J. L. Joly, G. Métivier, and J. Rauch, *Trilinear compensated compactness and nonlinear geometric optics*, Ann. of Math. (2) 142 (1995), no. 1, 121–169
[30] V. N. Karpushkin, *A theorem concerning uniform estimates of oscillatory integrals when the phase is a function of two variables*, J. Soviet Math. 35 (1986), 2809–2826. MR1680177
[31] N. Katz and T. Tao, *Bounds on arithmetic projections, and applications to the Kakeya conjecture*, Math. Res. Lett. 6 (1999), 625–630.
[32] M. Lacey and C. Thiele, *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$*, Ann. of Math. (2) 146 (1997), no. 3, 693–724.
[33] ———, *On Calderón's conjecture*, Ann. of Math. (2) 149 (1999), no. 2, 475–496.
[34] X. Li, *Bilinear Hilbert transforms along curves I: The monomial case*, Anal. PDE 6 (2013), no. 1, 197–220.
[35] A. Nagel, E. M. Stein, and S. Wainger, *Balls and metrics defined by vector fields. I. Basic properties*, Acta Math. 155 (1985), no. 1-2, 103-147
[36] D. H. Phong and E. M. Stein, *The Newton polyhedron and oscillatory integral operators*, Acta Math. 179 (1997), no. 1, 105–152.
[37] D. H. Phong, E. M. Stein, and J. Sturm, *Multilinear level set operators, oscillatory integral operators, and Newton polyhedra*, Math. Ann. 319 (2001), no. 3, 573–596.
[38] ———, *On the growth and stability of real-analytic functions*, Amer. J. Math. 121 (1999), no. 3, 519–554
[39] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 104–109.
[40] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.
[41] ———, *Problems in harmonic analysis related to oscillatory integrals and curvature*, A plenary address presented at the International Congress of Mathematicians held in Berkeley, California, August 1986. Introduced by Ray A. Kunze. ICM Series. American Mathematical Society, Providence, RI, 1988
[42] E. M. Stein and B. Street, *Multi-parameter singular Radon transforms III: Real analytic surfaces*, Adv. Math. 229 (2012), no. 4, 2210–2238
[43] E. M. Stein and S. Wainger, *Problems in harmonic analysis related to curvature*, Bull. Amer. Math. Soc. 84 (1978), no. 6, 1239–1295
[44] E. M. Stein and S. Wainger, *Oscillatory integrals related to Carleson's theorem*, Math. Res. Lett. 8 (2001), no. 5-6, 789–800.
[45] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics, 105. Cambridge University Press, Cambridge, 2006
[46] A. Varchenko, *Newton polyhedra and estimations of oscillatory integrals*, Functional Anal. Appl. 10 (1976), 175–196. MR0422257
[47] L. Xiao, *Sharp estimates for trilinear oscillatory integrals and an algorithm of two-dimensional resolution of singularities*, Rev. Mat. Iberoam. 33 (2017), no. 1, 67–116.
[48] M. Zworski, personal communication

Michael Christ, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA
E-mail address: mchrist@berkeley.edu