BOUNDS FOR MULTILINEAR SUBLEVEL SETS

MICHAEL CHRIST

1. Introduction

Let $\ell_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ be surjective linear transformations, let $P : \mathbb{R}^d \to \mathbb{R}$ be a real-valued polynomial, and let $\eta \in C_0^1(\mathbb{R}^d)$ be a compactly supported, continuously differentiable cutoff function. For $\lambda \in \mathbb{R}$ define the multilinear oscillatory integral forms

$$\mathcal{I}_\lambda(f_1, \cdots, f_n) = \int_{\mathbb{R}^d} e^{i\lambda P(y)} \prod_{j=1}^{n} f_j \circ \ell_j(y)\eta(y)\, dy. \tag{1.1}$$

Under what conditions do there exist $\delta > 0$ and $C < \infty$ such that for all $f_j \in L^\infty(\mathbb{R}^{d_j})$,

$$|\mathcal{I}_\lambda(f_1, \cdots, f_n)| \le C|\lambda|^{-\delta} \prod_j \|f_j\|_{L^\infty} \text{ for all } \lambda \in \mathbb{R}? \tag{1.2}$$

Under what conditions does there exist a function ρ satisfying $\rho(\lambda) \to 0$ as $|\lambda| \to \infty$ such that for all functions $f_j \in L^\infty$,

$$|\mathcal{I}_\lambda(f_1, \cdots, f_n)| \le \rho(\lambda) \prod_j \|f_j\|_{L^\infty} \text{ for all } \lambda \in \mathbb{R}? \tag{1.3}$$

The question in the form (1.2) was posed by Li, Tao, Thiele, and this author in [5]. Nonoscillatory inequalities of the form $\int \prod_j |f_j \circ \ell_j| \lesssim \prod_j \|f_j\|_{L^{p_j}}$ have been studied in [1],[2].

Oscillatory integral inequalities of this type have been extensively studied for $n = 2$, where one is dealing with bilinear forms $\langle T_\lambda(f_1), f_2\rangle$. The associated linear operators T_λ are commonly known as oscillatory integrals of the second type, and a simple necessary and sufficient condition for (1.2) to hold (with some unspecified exponent) is known [7]. There is an extensive literature dealing with more specific inequalities involving L^p norms, in which one seeks the optimal exponent δ as a function of exponents p.

For $n \ge 3$, however, there arises a class of singular oscillatory integrals which have no direct analogues in the bilinear case. These singular cases arise when $d < \sum_j d_j$. Generic ordered n-tuples of points $(x_1, \cdots, x_n) \in$

Date: October 28, 2008.
This research was supported by NSF grant DMS-040126.

$\times_j \mathbb{R}^{d_j}$ then do not contribute to the integral $\mathcal{I}_\lambda$, which may alternatively be expressed as

$$\mathcal{I}_\lambda(f_1, \cdots, f_j) = \int_\Sigma e^{i\lambda Q(x)} \prod_j f_j(x_j) \tilde{\eta}(x)\, d\sigma(x) \tag{1.4}$$

for a certain linear subspace $\Sigma \subset \times_j \mathbb{R}^{d_j}$ of positive codimension. Here $\tilde{\eta}$ is smooth and has compact support, and σ is Lebesgue measure on Σ. Apparently singular situations are always reducible to nonsingular ones in a certain sense in the bilinear case, but this is not so for $n \geq 3$, in general. See §6 below for discussion of this point.

To date little is known about the general singular multilinear case, but there are indications that, as was emphasized for certain related bilinear problems in [3], (1.2),(1.3) are linked to combinatorial issues. More refined questions about the optimal exponent δ in (1.2), if any, and about inequalities with $\prod_j \|f_j\|_{L^\infty}$ replaced by $\prod_j \|f_j\|_{L^{p_j}}$, are premature.

An obvious necessary condition [5] for (1.3) is that P should be *nondegenerate*, relative to $\{\ell_j\}$, in the sense that P cannot be expressed in the form $P = \sum_{j=1}^n p_j \circ \ell_j$ for any measurable functions p_j; this is equivalent [5] to there being no such representation in which p_j are polynomials of degrees not exceeding the degree of P. For the bilinear case $n = 2$, nondegeneracy of P is indeed sufficient for (1.2). The main results of [5] asserted that for $n \geq 3$, nondegeneracy of P implies (1.2), under certain rather restrictive supplementary hypotheses. In particular, this holds when all $d_j = d - 1$, and it holds when all $d_j = 1$ provided that $n < 2d$. No example is known to us in which a nondegenerate polynomial has been shown not to satisfy (1.2), let alone (1.3), but the vast majority of cases remain open.

In the present paper we do not answer these basic questions in any cases; rather, we study a class of weaker inequalities (2.2) which would be implied by (1.3). We establish these inequalities for all nondegenerate polynomials satisfying a natural rationality hypothesis, whereas only quite restricted classes of polynomials were treated in [5]. A second main result sheds additional light on the meaning of nondegeneracy, by establishing its equivalence, under the rationality hypothesis, with a formally stronger property, which we call finitely witnessed nondegeneracy. This furnishes an essential link with additive combinatorics. With this equivalence in hand, the remainder of the proof is a nearly direct application of a generalization of Szemerédi's theorem due to Furstenberg and Katznelson [6].

2. Results

Let $\{\ell_j : \mathbb{R}^d \to \mathbb{R}^{d_j}\}$ be a finite collection of surjective linear mappings. For any Lebesgue measurable functions g_j which are finite almost everywhere, for any $\varepsilon > 0$, and for any compact subset $B \subset \mathbb{R}^d$ consider the *sublevel sets*

$$E_\varepsilon(P, g_1, \cdots, g_n) = \{y \in B : |P(y) - \sum_{j=1}^{n} g_j(\ell_j(y))| < \varepsilon\}. \tag{2.1}$$

If a real-valued measurable function P satisfies (1.2), then there is an upper bound for the measures of these sublevel sets, of the form

$$\big|E_\varepsilon(P, g_1, \cdots, g_n)\big| \le C\varepsilon^\delta \textit{ uniformly for all measurable functions } g_j. \tag{2.2}$$

If instead P satisfies (1.3), then there is a corresponding weakened version of (2.2) in which the right-hand side is replaced by a function of ε which tends to zero as $\varepsilon \to 0$. Because of this connection with multilinear operators, we call sets E_ε of the form (2.1) multilinear sublevel sets.

(2.2) can be deduced from (1.2). To do so, fix a nonnegative cutoff function $h \in C_0^\infty(\mathbb{R})$ satisfying $h(t) = 1$ whenever $|t| \le 1$. Fix also $0 \le \zeta \in C_0^\infty(\mathbb{R}^m)$ such that $\zeta \equiv 1$ on B. Then

$$\begin{aligned}\big|\{x \in B : |P(x) - \sum_j g_j(\ell_j(x))| < \varepsilon\}\big| &\le \int h[\varepsilon^{-1}(P - \sum_j g_j \circ \ell_j)(x)]\, \zeta(x)\, dx \\ &= (2\pi)^{-1}\varepsilon \int_{\mathbb{R}} \widehat{h}(\varepsilon\lambda) \int e^{i\lambda(P(x) - \sum_j g_j(\ell_j(x)))}\zeta(x)\, dx\, d\lambda.\end{aligned}$$

Applying (1.2) to the inner integral and continuing in a straightforward way leads to the sublevel set bounds. (1.3) leads in the same way to a corresponding variant of (2.2).

Our discussion relies on a different notion of degeneracy than that defined above. $f|_S$ will denote the restriction of a function f to a set S.

Definition 2.1. Let $d, d_1, \cdots, d_n$ be arbitrary positive integers. Let $P : \mathbb{R}^d \to \mathbb{C}$ be a polynomial, and let $\ell_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ be linear transformations for $1 \le j \le n$. P is said to be *nondegenerate with a finite witness*, relative to $\{\ell_j\}$, if there exists a finite set $S \subset \mathbb{R}^d$ such that $P|_S$ does not belong to the span of the set of all functions $(f_j \circ \ell_j)|_S$.

The union is taken over all indices j and all functions f_j before the span is formed. An equivalent formulation is that there exist a finite set $S \subset \mathbb{R}^d$ and scalars c_s such that

$$\sum_{s \in S} c_s P(s) \ne 0 \tag{2.3}$$

but

$$\sum_{s \in S} c_s f_j(\ell_j(s)) = 0 \text{ for all indices } j \text{ and all functions } f_j. \tag{2.4}$$

For generic finite sets S the mapping $\ell_1(s)|_S$ is injective (unless $\ell_1 \equiv 0$), whence no such scalars can exist.

The usefulness of discrete characterizations of nondegeneracy in the context of oscillatory integral theory was recognized and exploited in [3].

From the theorem of Furstenberg and Katznelson we will deduce:

Proposition 2.1. *Suppose that a real-valued polynomial P is nondegenerate with a finite witness, with respect to a finite collection of surjective linear transformations $\ell_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$. Then there exists a function Θ satisfying*

$$\lim_{\varepsilon \to 0^+} \Theta(\varepsilon) = 0 \tag{2.5}$$

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

$$|E_\varepsilon(P, f_1, \cdots, f_n)| \le \Theta(\varepsilon). \tag{2.6}$$

In discussing the relation between these two notions of nondegeneracy, we will employ the following auxiliary concept. A related, though distinct, concept was shown in [4] to be natural in the context of a different question about multilinear operators.

Definition 2.2. Let $\ell_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ be finitely many linear transformations. The collection $\{\ell_j\}$ is said to be *rationally commensurate* if there exist invertible $\mathbb{R}$-linear transformations $A : \mathbb{R}^d \to \mathbb{R}^d$ and $A_j : \mathbb{R}^{d_j} \to \mathbb{R}^{d_j}$ such that with respect to the standard bases of $\mathbb{R}^d$ and of $\mathbb{R}^{d_j}$, the linear transformations $\tilde{\ell}_j = A_j^{-1} \circ \ell_j \circ A$ are all represented by matrices with integer entries.

Remark 2.1. It is easy to see that in the rationally commensurate case, if P fails to be nondegenerate with a finite witness, then there can be no sublevel set bound of the form (2.6),(2.5). Indeed, we may change variables to arrange that each ℓ_j maps $\mathbb{Z}^d$ to $\mathbb{Z}^{d_j}$. Let a bounded set $B \subset \mathbb{R}^d$ and $\varepsilon > 0$ be given, and choose $r = r(\varepsilon) > 0$ so that $|P(x) - P(y)| < \varepsilon$ whenever $|x - y| < r/2$. Fix $\rho > 0$ such that $|\ell_j(z)| \le \rho |z|$ for all j and all $z \in \mathbb{R}^d$. Consider the lattice $r\mathbb{Z}^d = \{rn : n \in \mathbb{Z}^d\}$. For each index j, $\ell_j(r\mathbb{Z}^d) \subset \mathbb{R}^{d_j}$ is again a lattice. Let $\mathcal{L}_r = r\mathbb{Z}^d \cap B$.

By assumption, there exist functions f_j such that $P(y) = \sum_j f_j(\ell_j(y))$ for all $y \in \mathcal{L}_r$. Since $\ell_j(\mathcal{L}_r) \subset r\mathbb{Z}^{d_j}$, there exists a constant $c_0 > 0$, independent of r, such that for any $z \ne z' \in \mathcal{L}_r$ and each index j, either $|\ell_j(z) - \ell_j(z')| \ge c_0 r$, or $\ell_j(z) = \ell_j(z')$.

Only the values of f_j on $\ell_j(\mathcal{L}_r)$ come into play. Redefine f_j so as to be constant on the ball $B(y, cr) \subset \mathbb{R}^{d_j}$ of radius cr centered at each point $y \in \ell_j(\mathcal{L}_r)$, where c is a positive constant, independent of r, sufficiently small to ensure that these balls are pairwise disjoint for distinct values of y.

The identity $P(y) = \sum_j f_j(\ell_j(y))$ still holds at each point of $\mathcal{L}_r$ for these modified functions f_j. Moreover, if $x \in \mathbb{R}^d$ and $|x - y| \le c' r$ for some $y \in \mathcal{L}_r$, where c' is another sufficiently small positive constant independent of r, then $|P(x) - \sum_j f_j \circ \ell_j(x)| \le \varepsilon + |P(y) - \sum_j f_j \circ \ell_j(x)|$. By construction, $f_j \circ \ell_j(x) = f_j \circ \ell_j(y)$. Thus $|P(x) - \sum_j f_j \circ \ell_j(x)| < \varepsilon$ whenever the distance from x to $\mathcal{L}_r$ is $< c' r$. The measure of the set of all such points $x \in B$ does not tend to zero as $\varepsilon \to 0$, contradicting (2.6). $\square$

Discrete nondegeneracy clearly implies nondegeneracy. Although we do not know whether the converse holds in general, it is true in the rational case.

Theorem 2.2. *Let $P : \mathbb{R}^d \to \mathbb{C}$ be a polynomial, and let $\ell_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ be a finite collection of surjective linear transformations. If $\{\ell_j\}$ is rationally commensurate, and if P is nondegenerate relative to $\{\ell_j\}$, then P is nondegenerate with a finite witness relative to $\{\ell_j\}$.*

Most of the work in this paper is devoted to proving this purely algebraic fact. Theorem 2.2 implies Remark 2.1 in a stronger form, for degeneracy of P means that there exist functions f_j for which $P - \sum_j f_j \circ \ell_j \equiv 0$, and then $E_\varepsilon = B$ for all $\varepsilon > 0$.

Proposition 2.1 and Theorem 2.2 together yield our main result.

Theorem 2.3. *Let a polynomial P be nondegenerate with respect to a finite rationally commensurate collection of surjective linear transformations. Then there exists a function Θ satisfying*

$$\lim_{\varepsilon \to 0^+} \Theta(\varepsilon) = 0 \tag{2.7}$$

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

$$|E_\varepsilon(P, f_1, \cdots, f_n)| \le \Theta(\varepsilon). \tag{2.8}$$

It bears emphasis that oscillatory integral bounds of the type (1.2) which imply this conclusion were proved in [5] in several cases, without any hypothesis of rational commensurability. Natural questions which remain are whether the commensurability hypothesis is superfluous, and whether Θ may always be taken to be of power law form $C\varepsilon^\delta$. The proof here certainly does not give power law bounds, since it relies on a result of Szemerédi type.

The proof of Theorem 2.3 is sufficiently robust to yield also the following variant.

Theorem 2.4. *Let P be a C^∞ real-valued function defined in a neighborhood of $x_0 \in \mathbb{R}^d$. Let $\{\ell_j\}$ be a rationally commensurate finite collection of surjective linear transformations $\ell_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$. Suppose that some Taylor polynomial for P at x_0 is nondegenerate with respect to $\{\ell_j\}$. Then there exist a neighborhood U of x_0 and a function Θ satisfying $\lim_{\varepsilon\to 0^+} \Theta(\varepsilon) = 0$ such that for every $\varepsilon > 0$ and all measurable functions f_j,*

$$\big|\{x \in U : |(P - \sum_j f_j \circ \ell_j)(x)| < \varepsilon\}\big| \le \Theta(\varepsilon). \tag{2.9}$$

Another extension concerns periodic sublevel sets, in which $P - \sum_j f_j \circ \ell_j$ is viewed as taking values in $\mathbb{R}/2\pi\mathbb{Z}$, rather than in $\mathbb{R}$. Define

$$\|y\| = \text{distance}\,(y, 2\pi\mathbb{Z})$$

for $y \in \mathbb{R}$.

$$E^\dagger_{\varepsilon,\lambda}(P, f_1, \cdots, f_n) = \{x \in B : \|\lambda P(x) - \sum_j f_j(\ell_j(x))\| < \varepsilon\}. \tag{2.10}$$

Let us assume that the cutoff function η appearing in (1.1) is nonnegative, and write $|E| = \int_E \eta$ for any measurable set $E \subset \mathbb{R}^d$. A uniform bound for multilinear oscillatory integrals of the form $|\mathcal{I}_\lambda(f_1, \cdots, f_n) \le C|\lambda|^{-\delta}$ for some $\delta \in (0,1)$ implies uniform bounds of the form

$$\big||E^\dagger_{\varepsilon,\lambda}| - c_0\varepsilon\big| \le C\varepsilon^\delta |\lambda|^{-\delta}, \tag{2.11}$$

where $c_0 = \int_{\mathbb{R}^d} \eta(x)\,dx$. Similarly a uniform decay bound $|\mathcal{I}_\lambda(f_1, \cdots, f_n)| \le \Theta(\lambda)$, where $\Theta(\lambda) \to 0$ as $|\lambda| \to \infty$, implies uniform bounds

$$\big||E^\dagger_{\varepsilon,\lambda}| - c_0\varepsilon\big| \le \theta(\varepsilon, |\lambda|^{-1}) \tag{2.12}$$

where $\theta(s,t) \to 0$ as $\min(s,t) \to 0^+$. Conversely, such inequalities imply uniform decay bounds for oscillatory integrals. Thus it is natural to seek suitable uniform upper bounds for $|E^\dagger_{\varepsilon,\lambda}|$ for nondegenerate polynomial phases P.

Theorem 2.5. *Suppose that P is nondegenerate, relative to a rationally commensurate set $\{\ell_j\}$ of surjective linear mappings. Let $B \subset \mathbb{R}^d$ be a bounded set. Then there exists a positive function Θ satisfying $\Theta(t) \to 0$ as $t \to 0^+$ such that for all measurable functions f_j and all $|\lambda| \ge 1$,*

$$\Big|\{x \in B : \text{distance}\,(\lambda P(x) - \sum_j f_j(\ell_j(x)), 2\pi\mathbb{Z}) < \varepsilon\}\Big| \le \Theta(\varepsilon). \tag{2.13}$$

3. Proof of Proposition 2.1

Proposition 3.1. *Let $B \subset \mathbb{R}^d$ be a bounded region, and let $S \subset \mathbb{R}^d$ be a finite set. There exists a positive function Θ satisfying $\Theta(r) \to 0$ as $r \to 0^+$, 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) $|E| \le \Theta(r)$.*

Proof. Denote by $\#(A)$ the cardinality of a set A. According to a theorem of Furstenberg and Katznelson [6], 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 $n \in \mathbb{Z}$ and $x \in \mathbb{Z}^d$ such that $x + nS \subset E$, or $\#(A) \le \theta(N)N^d$.

Proposition 3.1 follows rather directly from this result. Under the additional assumption that the set S in the hypothesis is contained in $\mathbb{Z}^d$, the reduction goes as follows: Let N be a large positive integer chosen so that $\frac{1}{2}r \le N^{-1} < r$. Define $\Omega = \{\omega = (\omega_1, \cdots, \omega_d) \in \mathbb{R}^d : 0 \le \omega_j < N^{-1}$ for all $1 \le j \le d\}$, and define $\mathcal{L}_{N,\omega} = N^{-1}\mathbb{Z}^d + \omega = \{N^{-1}n + \omega : n \in \mathbb{Z}^d\}$. Let $E \subset B$, and suppose that conclusion (i) of Proposition 3.1 fails to hold. Decompose $E = \cup_{\omega\in\Omega} E_\omega$ where $E_\omega = E \cap \mathcal{L}_{N,\omega}$. Then for any integer j and point $x \in \mathcal{L}_{N,\omega}$, the set $x + jN^{-1}S$ is not contained in E_ω. Applying the theorem of Furstenberg and Katznelson to $\tilde{E}_\omega = \{Ny : y \in E_\omega\}$ yields the bound $\#(E_\omega) \le N^d\theta(N)$. Consequently

$$|E| = \int_\Omega \#(E_\omega)\,d\omega \le N^{-d}\sup_\omega \#(E_\omega) \le \theta(N), \tag{3.1}$$

establishing Proposition 3.1 under the auxiliary hypothesis.

The general case of Proposition 3.1 follows from a particular case of the theorem of Furstenberg and Katznelson by the following lifting argument. Introduce $\mathbb{R}^M = \mathbb{R}^d \times \mathbb{R}^S$ with coordinates (x,t), where $t = (t_s : s \in S) \in \mathbb{R}^S$. Let $e_s \in \mathbb{R}^S$ be the unit vector corresponding to the s-th coordinate. Define $E^\dagger = E \times \mathbb{R}^S$.

Introduce the shear transformation $T : \mathbb{R}^M \to \mathbb{R}^M$ defined by

$$T(x,t) = (x - \sum_{\sigma\in S} t_\sigma \sigma, t), \tag{3.2}$$

and let $E^\ddagger = T(E^\dagger)$. Then for any $r > 0$, $t \in \mathbb{R}^S$, $s \in S$, and $x \in \mathbb{R}^d$,

$$x + rs \in E \text{ if and only if } T(x,t) + (0, re_s) \in E^\ddagger. \tag{3.3}$$

Indeed, $x + rs \in E$ is equivalent to $(x + rs, t + re_s) \in E^\dagger$. Next

$$\begin{aligned} T(x+rs, t+re_s) &= (x + rs - \sum_{\sigma\in S} t_\sigma\sigma - rs, t + re_s) \\ &= (x - \sum_\sigma t_\sigma\sigma, t + re_s) = T(x,t) + (0, re_s), \end{aligned}$$

whence

$$\begin{aligned} x + rs \in E &\Leftrightarrow (x+rs, t+re_s) \in E^\dagger \\ &\Leftrightarrow T(x+rs, t+re_s) \in E^\ddagger \Leftrightarrow T(x,t) + (0, re_s) \in E^\ddagger. \end{aligned}$$

Let $S^\ddagger = \{(0, e_s) : s \in S\} \subset \mathbb{R}^M$. Suppose now that E satisfies the restriction that $x + rS \subset E$ implies $r \le \varepsilon$. Then $E^\ddagger$ satisfies a corresponding restriction: if $z \in \mathbb{R}^M$ and if $z + rS^\ddagger \subset E^\ddagger$, then $r \le \varepsilon$. Indeed, there exists a unique point (x,t) satisfying $T(x,t) = z$. By (3.3), $z + rS^\ddagger \subset E^\ddagger$ if and only if $x + rS \subset E$.

For almost every $x \in \mathbb{R}^m$, we now have a set $E^* = \{(w,t) \in E^\ddagger : w = x \text{ and } |t| \le 1\}$, contained in a fixed bounded subset of $\mathbb{R}^S$, such that for any $z \in \mathbb{R}^S$, if $z + re_s \in E^*$ for every $s \in S$ then $r \le \varepsilon$. As was shown above, this forces $|E^* \cap B'| \le \Theta(\varepsilon)$ for any fixed bounded set B'. Therefore $|E| \le C\Theta(\varepsilon)$ by Fubini's theorem. □

Proof of Proposition 2.1. Suppose that P is nondegenerate with a finite witness. Fix a finite set S and scalars $\{c_s : s \in S\}$ such that $\sum_{s\in S} c_s F(s) = 0$ whenever F takes the form $\sum_j f_j \circ \ell_j$, but $\sum_{s\in S} c_s P(s) = 1$. Let f_j be arbitrary measurable functions; for convenience we assume that f_j is defined on all of $\mathbb{R}^1$.

Set

$$h(y,r) = \sum_{s\in S} c_s\big(P(y+rs) - \sum_j (f_j\circ\ell_j)(y+rs)\big)$$
$$E_\varepsilon = \{y\in B : |\big(P - \sum_j f_j\circ\ell_j\big)(y)| < \varepsilon\}.$$

Then

$$h(y,r) = \sum_{s\in S} c_s\big(P(y+rs) - \sum_j (f_j\circ\ell_j)(y+rs)\big) \equiv \sum_{s\in S} c_s P(y+rs)$$

is a polynomial function of $(y,r)\in\mathbb{R}^d\times\mathbb{R}$. The set S and coefficients c_s were constructed in part to ensure that this polynomial does not vanish identically. Hence there exist $A<\infty$, $\delta>0$, and $C<\infty$ such for any sufficiently small $\rho>0$, $B\subset\mathbb{R}^d$ may be partitioned into the union of $O(\rho^{-d})$ dyadic cubes Q_j of sidelength ρ, together with a remainder set $B\setminus\cup_j Q_j$, in such a way that (i) $|B\setminus\cup_j Q_j|\le C\rho^\delta$ and (ii) for each j, each $x\in Q_j$, and each $r\in(0,\rho]$, $|h(x,r)|\ge r^A$.

Choose $\rho=\varepsilon^{1/2A}$. If $x\in B$ and $x+rS\subset E_\varepsilon$ then

$$|h(x,r)| \le \sum_{s\in S} c_s\big|\big(P-\sum_j f_j\circ\ell_j\big)(x+rs)\big| \le C\sum_{s\in S}\varepsilon, \tag{3.4}$$

which implies that $r^A\lesssim \#(S)\varepsilon$ if $x\in\cup_j Q_j$ and $r\le\rho$, where $\#(S)$ denotes the cardinality of S. Therefore

$$|E_\varepsilon\cap Q_j| \le |Q_j|\Theta(C\varepsilon^{1/A}/\rho), \tag{3.5}$$

by Proposition 3.1 applied to a dilate of Q_j. Here the C depends on the cardinality of S, which is a constant in this context.

Summing over j yields

$$|E_\varepsilon| \le |B\setminus\cup_j Q_j| + \sum_j |E_\varepsilon\cap Q_j| \le C\varepsilon^{\delta/2A} + |B|\Theta(C\varepsilon^{1/2A}), \tag{3.6}$$

which is a bound of the desired form. □

4. Proof of Theorem 2.2

Even if P is nondegenerate, the restriction of P to a generic finite set S will be degenerate relative to $\{\ell_j\}$. Indeed, if the restriction of some ℓ_i to S is injective, then any function on S takes the form $f_i\circ\ell_i$. Thus S is a more promising candidate to be a witness, if all of the mappings ℓ_j are far from being injective on S. This motivates the use of finite lattices as witnesses; the hypothesis of rational commensurability will ensure a strong failure of injectivity for suitable lattices.

$M\mathbb{Z}^d$ will denote the set of all $(x_1,\dots,x_d)\in\mathbb{Z}^d$ for which each coordinate is divisible by M.

Recall that any finitely generated torsion-free $\mathbb{Z}$-module $\mathcal{M}$ is isomorphic to $\mathbb{Z}^n$ for some unique n; n is called the rank of $\mathcal{M}$. Any submodule of $\mathbb{Z}^n$

is finitely generated and torsion-free. By the *rank* of a homomorphism of $\mathbb{Z}$-modules, we mean the rank of its range; only finitely generated and torsion-free ranges will arise in this paper. Let $\mathcal{M} \subset \mathbb{Z}^n$ be a sub-$\mathbb{Z}$-module of rank r, and choose elements $e_1, \cdots, e_r \in \mathcal{M}$ such that the mapping $(x_1, \cdots, x_r) \mapsto x \cdot e = x_1 e_1 + \cdots + x_r e_r$ defines a bijection of $\mathbb{Z}^r$ onto $\mathcal{M}$. If $q : \mathcal{M} \to \mathbb{C}$ is a polynomial, in the sense that q can be represented as a finite linear combination over $\mathbb{C}$ of the monomials $x \cdot e \mapsto x_1^{\gamma_1} \cdots x_n^{\gamma_n}$ with each exponent γ_j a nonnegative integer, then such a representation is unique.

The analysis will make use of difference operators. For any vector y, D_y denotes the operator $D_y f(x) = f(x+y) - f(x)$. These operators all commute with one another. If L is a linear mapping then

$$D_y(f \circ L) = \big(D_{L(y)}(f)\big) \circ L. \tag{4.1}$$

A version of Leibniz's rule is

$$D_v(fg) = D_v(f) \cdot g + f_v \cdot D_v(g) \text{ where } f_v(x) = f(x+v). \tag{4.2}$$

We will need the following elementary property of polynomials, whose proof is omitted.

Lemma 4.1. *For any d, r there exists $N < \infty$ such that for any polynomial $P : \mathbb{R}^d \to \mathbb{C}$ of degree $\le r$, if $P(x) \equiv 0$ for all $x \in \mathbb{Z}^d$ satisfying $|x| \le N$, then $P(x) = 0$ for all $x \in \mathbb{R}^d$.*

The next lemma describes solutions of certain difference equations.

Lemma 4.2. *Let $(v_j) \subset \mathbb{Z}^n$ be any finite list of nonzero vectors, not necessarily distinct, and let $\mathcal{D}$ be the difference operator $\mathcal{D} = \prod_j D_{v_j}$. Then there exist $C, r < \infty$, a positive integer M, and finitely many $\mathbb{Z}$-module homomorphisms $\ell_\gamma : \mathbb{Z}^n \to \mathbb{Z}^{n_\gamma}$ where $n_\gamma < n$, such that for any sufficiently large $N < \infty$ and any function $f : \mathbb{Z}^n \to \mathbb{C}$ which satisfies $\mathcal{D}(f)(x) = 0$ for all $x \in \mathbb{Z}^n$ satisfying $|x| \le N$, there exists a representation*

$$f(x) = \sum_\gamma q_\gamma(x)(h_\gamma \circ \ell_\gamma)(x) \tag{4.3}$$

valid for all $x \in M\mathbb{Z}^n$ satisfying $|x| \le N - C$, where the $q_\gamma : \mathbb{Z}^n \to \mathbb{C}$ are polynomials of degrees at most r.

Sketch of proof. Proceed by induction the number of factors D_{v_j}. Thus suppose it to be given $\mathcal{D}D_w(f) = 0$ vanishes for all x in the region indicated, where $w \ne 0$ and $\mathcal{D}$ is as above. Applying the induction hypothesis gives a representation $D_w(f)(x) = \sum_\gamma q_\gamma(x)(h_\gamma \circ \ell_\gamma)(x)$ of the above form, for $x \in M\mathbb{Z}^n$ satisfying $|x| \le N - C$.

It is awkward to proceed further, because w need not lie in $M\mathbb{Z}$. However, $D_w\mathcal{D}(f) = \mathcal{D}D_w(f)$ vanishes for $|x| \le N$, from which it follows that $\mathcal{D}D_{Mw}(f) = D_{Mw}\mathcal{D}(f)$ also vanishes for $|x| \le N - C(|w|)$; note that M depends only on $\{v_j\}$, not on f. Thus we may suppose from the outset that $w \in M\mathbb{Z}^n$.

It is straightforward to solve the equation $D_w(F)(x) = \sum_\gamma q_\gamma(x)(h_\gamma \circ \ell_\gamma)(x)$ with a solution F in the desired form (4.3), in the region $M\mathbb{Z}^n \cap \{x : |x| \le N - C\}$, with the initial condition that F vanishes on a suitable submodule of rank $n = 1$ which does not contain w. This equation is solved term-by-term, distinguishing the terms for which $\ell_\gamma(w) = 0$ from those for which $\ell_\gamma(w) \ne 0$.

Finally since $D_w(f - F) \equiv 0$ on an appropriate domain, it must take the form $h \circ \ell$, where ℓ has rank $n - 1$ and $\ell(w) = 0$. □

Lemma 4.3. *Let $P : \mathbb{Z}^d \to \mathbb{C}$ be a polynomial which takes the form*

$$P(x) = \sum_{\alpha \in A} (f_\alpha \circ L_\alpha)(x) \text{ for all } x \in \mathbb{Z}^d \text{ satisfying } |x| \le N. \tag{4.4}$$

Here A is a finite set of indices, $L_\alpha : \mathbb{Z}^d \to \mathbb{Z}^{d_\alpha}$ are $\mathbb{Z}$–linear mappings, and f_α are arbitrary functions. If N is sufficiently large then there exist polynomials p_α and positive integers M, N^* such that

$$P(x) = \sum_\alpha p_\alpha \circ L_\alpha \tag{4.5}$$

for all $x \in M\mathbb{Z}^d$ satisfying $|x| \le N^*$. $N^* \to \infty$ as $N \to \infty$, while M and the degrees of the p_α remain uniformly bounded, provided that the linear mappings L_α and the degree of P remain fixed.

The functions f_α in such a decomposition $P = \sum_\alpha f_\alpha \circ L_\alpha$ are not necessarily polynomials. There are also cases in which they are necessarily polynomials, but are not necessarily unique.

A related result was established in [5]: If a polynomial P admits a decomposition $P(x) = \sum_\alpha f_\alpha \circ L_\alpha$ on $\mathbb{R}^d$, where the f_α are merely distributions, then it admits such a decomposition with those distributions replaced by polynomials. The simple proof given in [5] does not seem to adapt directly to the present discrete setting.

By admissible data we mean the collection of mappings L_α, and the degree of P. It will be important, in both the proof and application of Lemma 4.3, that N^*, M and the degrees of p_α depend only on admissible data. We will say that a polynomial has bounded degree if its degree is bounded above by a quantity which depends only on admissible data. Likewise, by a *large finite submodule* of $\mathbb{Z}^n$ we mean, in the context of Lemma 4.3, the set of all $x = (x_1, \cdots, x_n) \in \mathbb{Z}^n$ such that $|x| \le N^\sharp$ and each coordinate x_j is divisible by some $M^\sharp$, where $N^\sharp \to \infty$ while $M^\sharp$ remains uniformly bounded, as the parameter N given in the hypotheses tends to ∞. $N^\sharp, M^\sharp$ are permitted to depend on admissible data. This is of course an abuse of language, since these "submodules" are not closed under addition.

Conclusion of proof of Theorem 2.2. If P is not nondegenerate with a finite witness, then Lemma 4.3, applied to the polynomials $x \mapsto P(Nx)$, asserts that any open ball $B \subset \mathbb{R}^d$ centered at the origin, for each sufficiently large integer N, $P|_{B \cap N^{-1}\mathbb{Z}^d}$ can be expressed in the form $\sum_j Q_j \circ \ell_j$, where the polynomials Q_j may potentially depend on N, but have uniformly bounded degrees. By Lemma 4.1, applied again to $x \mapsto P(MNx)$ for a certain constant M, this implies that $P - \sum_j Q_j \circ \ell_j$ vanishes identically on $\mathbb{R}^d$. Thus P is degenerate relative to $\{\ell_j\}$. $\square$

In the proof of Lemma 4.3, the quantity M appearing in its conclusion will repeatedly be replaced by a larger multiple of itself throughout an inductive procedure. All of these quantities will be denoted by the same symbol M, with the understanding that M is always bounded above by a quantity which depends only on admissible data.

Proof of Lemma 4.3. Lemma 4.3 will be proved by an inductive scheme which involves more general representations of P. To set this up, suppose that $P : \mathbb{Z}^d \to \mathbb{C}$ is a polynomial which takes the form

$$P(x) = \sum_{\alpha \in A} \sum_j (Q_{\alpha,j} \circ L_\alpha) \cdot (h_{\alpha,j} \circ \ell_{\alpha,j})(x) \tag{4.6}$$

for all $x \in \mathbb{Z}^d$ satisfying $|x| \le N$. Here A is a finite set of indices, j ranges over a finite set of indices for each $\alpha \in A$, $L_\alpha, \ell_{\alpha,j}$ are $\mathbb{Z}$–linear mappings from $\mathbb{Z}^d$ to some $\mathbb{Z}^{n(\alpha)}$ and $\mathbb{Z}^{n(\alpha,j)}$, respectively,

$$\text{nullspace}\,(L_\alpha) \subset \text{nullspace}\,(\ell_{\alpha,j}),$$

$h_{\alpha,j}$ are arbitrary functions, and $Q_{\alpha,j}$ are polynomials with domains $\mathbb{Z}^{n(\alpha,j)}$. Admissible data are now the collection of mappings L_α, ℓ_α, and the degrees of P, Q_α. We wish to conclude that if N is sufficiently large then there exist polynomials $Q^*_{\alpha,k}$, linear mappings $\ell^*_{\alpha,k}$, functions $h^*_{\alpha,k}$, and integers M, N^*, such that

$$P(x) = \sum_{\alpha,k} (Q^*_{\alpha,k} \circ L_\alpha) \cdot (h^*_{\alpha,k} \circ \ell^*_{\alpha,k})(x) \tag{4.7}$$

for all $x \in M\mathbb{Z}^d$ satisfying $|x| \le N^*$ where α ranges over the same index set A as in (4.6), k ranges over a finite index set for each $\alpha \in A$, and nullspace$\,(L_\alpha) \subset$ nullspace$\,(\ell^*_{\alpha,k})$. The index sets over which k ranges need not coincide with those over which j ranges in (4.6), and in practice will be larger. Moreover as $N \to \infty$, $N^* \to \infty$ while M and the degrees of the $Q^*_{\alpha,k}$ remain uniformly bounded, provided that the admissible data remain fixed. Finally, (4.7) is required to be a simpler representation than the given one (4.6), in the sense that either the maximum (over all pairs α, k) of the ranks of the $\ell^*_{\alpha,k}$ is strictly less than the maximum rank of all $\ell_{\alpha,j}$, or that the two maxima are equal and that the number of index pairs (α, k) for which $\ell^*_{\alpha,k}$ has this maximal rank is strictly less than the number of pairs (α, j) for which $\ell_{\alpha,j}$ has maximal rank.

At each stage of the induction, a hypothesis on $\mathbb{Z}^d$ leads to a conclusion only on some $M\mathbb{Z}^d$, but then $M\mathbb{Z}^d$ can be reidentified with $\mathbb{Z}^d$ in the following step. Finitely many induction steps bring us to the situation in which every $\ell_{\alpha,j}$ has rank zero. In that case, each $h_{\alpha,j} \circ \ell_{\alpha,j}$ is a constant. For each α, $\sum_j (Q_{\alpha,j} \circ L_\alpha) \cdot h_{\alpha,j} \circ \ell_{\alpha,j}$ is the composition of a single polynomial with L_α, so (4.6) has the desired form and the proof would be complete. Thus in order to prove the lemma, it suffices to carry out this reduction step.

Supposing that some $\ell_{\alpha,j}$ has nonzero rank, choose some α_0 such that ℓ_{α_0} has maximal rank, Equivalently, nullspace$\,(\ell_{\alpha_0}) \subset \mathbb{Q}^d$ has minimal dimension among all such nullspaces, as vector spaces over $\mathbb{Q}$.

Let $\beta \in A$ be an arbitrary index. If

$$\text{nullspace}\,(\ell_\beta) \neq \text{nullspace}\,(\ell_{\alpha_0}) \tag{4.8}$$

then nullspace$\,(\ell_\beta)$ is not a subset of nullspace$\,(\ell_{\alpha_0})$. For each such β choose some vector $y_\beta \in \mathbb{Z}^d \cap \text{nullspace}\,(\ell_\beta) \setminus \text{nullspace}\,(\ell_{\alpha_0})$. Define

$$\mathcal{D} = \prod_\beta D_{y_\beta} \tag{4.9}$$

where the composition product is taken over all indices β satisfying (4.8). Then

$$\mathcal{D}(f_\beta \circ \ell_\beta) \equiv 0 \text{ on } \mathbb{Z}^d$$

for all β satisfying (4.8), for all functions f_β. Indeed, D_{y_β} annihilates all such functions, and the factors in (4.9) all commute.

Consequently there exists a positive integer a such that

$$\mathcal{D}^a\big((Q_\beta \circ L_\beta) \cdot (f_\beta \circ \ell_\beta)\big) \equiv 0 \tag{4.10}$$

for all β satisfying (4.8); a may be chosen to depend only on the degrees of the given polynomials Q_α. This follows from Leibniz's rule (4.2). For D_v^b annihilates any polynomial of degree strictly less than b, for any vector v, and $Q_\alpha \circ L_\alpha$ is a polynomial on $\mathbb{Z}^d$ whose degree does not exceed that of Q_α.

We may also choose a sufficiently large to ensure that $\mathcal{D}^a(P) \equiv 0$. Therefore

$$\sum_\alpha{}' \mathcal{D}^a\big((Q_\alpha \circ L_\alpha) \cdot (h_\alpha \circ \ell_\alpha\big) \equiv 0, \tag{4.11}$$

where the notation $\sum'_\alpha$ indicates that the sum is taken over all indices α which satisfy

$$\text{nullspace}\,(\ell_\alpha) = \text{nullspace}\,(\ell_{\alpha_0}). \tag{4.12}$$

(4.11) holds at all points $x \in \mathbb{Z}^d$ which satisfy $|x| \le N - C$, where C depends only on the vectors y_β and the exponent a, thus only on permissible quantities; in particular, C is independent of N.

Define $U \subset \mathbb{Z}^d$ to be the nullspace of ℓ_{α_0}. Choose a sub-$\mathbb{Z}$-module $V \subset \mathbb{Z}^d$ which is complementary to U in the sense that U, V are linearly independent over $\mathbb{Q}$, and $U \cup V$ spans $\mathbb{Q}^d$ over $\mathbb{Q}$. Then $U + V$ contains $M\mathbb{Z}^d$, for some positive integer M which depends only on ℓ_{α_0}. We write (u, v) to denote an arbitrary point $(u, v) = u + v \in U + V$.

For each index α satisfying (4.12), $(h_\alpha \circ \ell_\alpha)(u,v)$ depends only on v, hence will be denoted as $(h_\alpha \circ \ell_\alpha)(v)$. The factor $(Q_\alpha \circ L_\alpha)(u,v)$ potentially depends on both variables, but any dependence on v can be incorporated into h_α. Thus for each α satisfying (4.12), we abuse notation by writing

$$(Q_{\alpha,j} \circ L_\alpha)(u,v) \cdot (h_{\alpha,j} \circ \ell_{\alpha,j})(v) = \sum_k (\tilde{Q}_{\alpha,j,k} \circ L_\alpha)(u) \cdot (\tilde{h}_{\alpha,j,k} \circ \ell_{\alpha,j})(v)$$

where k runs over a finite index set, and the $\tilde{Q}_{\alpha,j,k}$ are polynomials. Both the cardinalities of these index sets, and the degrees of these polynomials, are bounded above by quantities which depend only on admissible data. It is to be understood that $L_\alpha(u)$ means $L_\alpha(u,v)$ and $\ell_{\alpha,j}(v)$ means $\ell_{\alpha,j}(u,v)$.

Let K be the maximum degree of all the polynomials $\tilde{Q}_{\alpha,j,k}$. Decompose $\tilde{Q}_{\alpha,j,k} = Q^\dagger_{\alpha,j,k} + R_{\alpha,j,k}$ where $Q^\dagger_{\alpha,j,k}$ is homogeneous of degree K, while the remainders $R_{\alpha,j,k}$ have degrees strictly less than K.

By Leibniz's rule (4.2),

$$0 \equiv {\sum_\alpha}' \sum_{j,k} (Q^\dagger_{\alpha,j,k} \circ L_\alpha)(u) \cdot \mathcal{D}^a\big((\tilde{h}_{\alpha,j,k} \circ \ell_{\alpha,j})(v)\big) + R(u,v) \tag{4.13}$$

for all $(u,v) \in U+V$ satisfying $|(u,v)| \le N-C$, where $C < \infty$ depends only on admissible data, and R can be expressed as a polynomial in u of degree $\le K-1$, whose coefficients are functions of v. Each term $(Q^\dagger_{\alpha,j,k} \circ L_\alpha)(u)$ is a homogeneous polynomial of degree K, evaluated at $L_\alpha(u)$, thus is a homogeneous polynomial of degree K as a function of u.

Since the degrees of all polynomials in play here are bounded uniformly in N, it follows from (4.13) and Lemma 4.1 that

$$0 \equiv {\sum_\alpha}' \sum_{j,k} (Q^\dagger_{\alpha,j,k} \circ L_\alpha)(u) \cdot \mathcal{D}^a\big((\tilde{h}_{\alpha,j,k} \circ \ell_{\alpha,j})(v)\big), \tag{4.14}$$

again for all $(u,v) \in U+V$ satisfying $|(u,v)| \le N-C$.

There are now two cases. In Case 1,

$$\mathcal{D}^a\big((\tilde{h}_{\alpha,j} \circ \ell_\alpha)(v)\big) \equiv 0 \tag{4.15}$$

for all $v \in V$ satisfying $|v| \le N-C$, for each α satisfying (4.12). By (4.1),

$$\mathcal{D}^a\big((\tilde{h}_{\alpha,j} \circ \ell_\alpha)\big) = \big(\mathcal{D}'_\alpha(\tilde{h}_{\alpha,j})\big) \circ \ell_\alpha \text{ where } \mathcal{D}'_\alpha = \prod_\beta D^a_{\ell_\alpha(y_\beta)},$$

with the product taken over all β satisfying (4.8). Since $\ell_\alpha(y_\beta) \ne 0$, Lemma 4.2 asserts that for each such pair of indices α, j, the restriction of $\tilde{h}_{\alpha,j}$ to $\ell_\alpha(V)$ can be decomposed as a finite sum of terms, each of which is the product of a polynomial of bounded degree with a function of the form $h^\sharp \circ \ell^\sharp \circ \ell_\alpha$ for some $\mathbb{Z}$–linear mapping $\ell^\sharp$ whose rank is strictly less than the rank of ℓ_α. This representation holds on the set of all $v \in M_jV$ satisfying $|v| \le cN - C$, where M_j, c, C depend only on admissible data. Any polynomial composed with ℓ_α can be rewritten as a polynomial composed with L_α, since nullspace $(L_\alpha) \subset$ nullspace (ℓ_α) by hypothesis. Since $M\mathbb{Z}^d \subset U + V$ for some positive integer M, we have reduced matters to a situation which satisfies the hypothesis of the induction step on a large finite submodule of $\mathbb{Z}^d$.

Consider next Case 2, in which there exists at least one pair α, j for which (4.15) fails to hold. Choose a pair γ, i and some $v_1 \in V$ such that $(\mathcal{D}'_J(\tilde h_{\gamma,i})) \circ \ell_\gamma(v_1) \neq 0$. Specialize (4.14) to $v = v_1$ and solve for $(Q^\dagger_{\gamma,i} \circ L_\gamma)(u)$ as a $\mathbb{C}$-linear combination of the other $(Q^\dagger_{\alpha,j} \circ L_\alpha)(u)$. The term $(Q^\dagger_{\gamma,i} \circ L_\gamma)(u) \cdot \mathcal{D}^a\big(\tilde h_{\gamma,i} \circ \ell_\gamma(v)\big)$ is thus expressed as a $\mathbb{C}$-linear combination of hybrid terms $(Q^\dagger_{\alpha,j} \circ L_\alpha)(u) \cdot \mathcal{D}^a\big(\tilde h_{\gamma,i} \circ \ell_\gamma(v)\big)$ for all (u, v) in a large finite submodule of $U + V$. However, since $\ell_\alpha|_V$ is injective, each $\tilde h_{\gamma,i} \circ \ell_\gamma(v)$ can be reexpressed in the form $h^\flat_{\alpha,j} \circ \ell_\alpha(v)$, and thus each hybrid term can be reexpressed in the form $(Q^\dagger_{\alpha,j} \circ L_\alpha)(u) \cdot \mathcal{D}^a\big(h^\flat_{\alpha,j} \circ \ell_\alpha(v)\big)$.

The result is that the degree of at least one of the polynomials $\tilde Q_{\alpha,j}$ has been decreased. This process can be iterated until either Case 1 arises, or all $\tilde Q_{\alpha,j}$ have degree zero. In the former event, the proof is complete by induction.

In the latter event, these constants can be absorbed into the factors h_α. Then the sum $\sum'_\alpha (h_\alpha \circ \ell_\alpha)$ appearing in the initial representation of P can be rewritten more simply as a single term $h^\dagger_{\alpha_0} \circ \ell_{\alpha_0}$. The relation (4.11) becomes simply $\mathcal{D}^a\big(h^\dagger_{\alpha_0} \circ \ell_{\alpha_0}\big) \equiv 0$. This again makes Lemma 4.2 applicable, so $h^\dagger_{\alpha_0}$ can be represented, on a large finite submodule of its domain, as a finite sum of products of polynomials of bounded degrees multiplied by functions composed with linear mappings of ranks strictly less than the rank of ℓ_{α_0}, all of which factor through ℓ_{α_0}. Thus matters are again reduced to a prior induction step, completing the proof. $\square$

Remark 4.1. Suppose that $d_j = 1$ for all $1 \le j \le n$, so that $\ell_j : \mathbb{R}^d \to \mathbb{R}^1$, and that kernel$(\ell_i) \neq$ kernel(ℓ_j) whenever $i \neq j$. If $P : \mathbb{R}^d \to \mathbb{R}$ is a polynomial, and if $P = \sum_j f_j \circ \ell_j$ for certain distributions f_j defined in $\mathbb{R}^1$, then necessarily each f_j is a polynomial whose degree is majorized by a quantity depending only on d, n and the degree of P. This fact holds without the hypothesis of rational commensurability; it follows by a variant of the reasoning used in the proof of Theorem 2.2, with difference operators replaced by differential operators $\prod_{k \neq j} (v_k \cdot \nabla)^{a_k}$ where $\ell_k(v_k) = 0$ for all $k \neq j$ but $\ell_j(v_j) \neq 0$, and with the exponents a_k chosen so that the degree of the operator exceeds the degree of P.

This is not true without the restriction $d_j = 1$; 0 can be represented as $f_1(x_2, x_3) + f_2(x_1, x_3) + f_3(x_1, x_2)$ in many ways.

Remark 4.2. There are two places in the analysis where available information has been only partially exploited. Firstly, only translates and dilates of a single finite point configuration were used to establish the sublevel set bounds, while our algebraic discussion showed that translates and dilates of a rather large family of configurations are actually excluded from sublevel sets. Secondly, the identity $\sum_s c_s(P-\sum_j f_j\circ\ell_j)(x+rs) = h(x,r)$ was used merely to obtain an inequality $\sum_s |(P-\sum_j f_j\circ\ell_j)(x+rs)| \ge c|h(x,r)|$.

5. Two extensions

We discuss here the proofs for two extensions, Theorems 2.4 and Theorem 2.5. The former concerns C^∞ phases which are not necessarily polynomials, but satisfy a finite order nondegeneracy condition.

Proof of Theorem 2.4. Let P be a C^∞ real-valued function satisfying the hypothesis, and let a small $\varepsilon > 0$ be given. Fix a large positive integer N, to be specified below. Let $\rho > 0$ be a function of ε, to be determined. Partition a neighborhood of U into cubes Q_k, each of sidelength ρ. For each k let P_k be the Taylor polynomial of degree $N-1$ for P at the center point c_k of Q_k. Then

$$|(P-P_k)(x)| \le C\rho^N \text{ for all } x\in Q_k. \tag{5.1}$$

Define $L_k(y) = c_k + \rho y$, and $\tilde P_k = P_k\circ L_k$. $\ell_j\circ L_k$ is now affine linear rather than linear, but for any function f_j we can write $f_j\circ\ell_j\circ L_k = \tilde f_j\circ\ell_j$ where $\tilde f_j$ is an appropriate translate and dilate of f_j, depending on j,k. If N is sufficiently large then $\tilde P_k$ as a mapping whose domain is the unit cube centered at the origin, is nondegenerate relative to the (affine) linear transformations $\ell_j\circ L_k$. Our hypotheses do not guarantee that these polynomials are nondegenerate uniformly in k in any sense, but as in the proof of Proposition 2.1, if ρ is chosen to be an appropriate positive power of ε then the identity $\sum_{s\in S} c_s(\tilde P_k - \sum_j f_j\circ\ell_j)(x+rs) = h_k(x,r) = h(L_k(x),\rho r)$ can be exploited to obtain a bound for most k, while the sum of the measures of the remaining cubes Q_k is small. $\square$

Theorem 2.5 is a stronger result, in which $P-\sum_j f_j\circ\ell_j$ is regarded as taking values in the quotient space $\mathbb{R}/2\pi\mathbb{Z}$. Sublevel sets are then typically larger, yet turn out to satisfy the same upper bounds.

Outline of proof of Theorem 2.5. This follows from a small modification of the arguments already indicated. If P is nondegenerate then after a change of variables, all the ℓ_j can be represented by matrices with integer entries, and there exist a finite set $S\subset\mathbb{Z}^d$ and coefficients $c_s\in\mathbb{Z}$ satisfying (2.3) and (2.4). Indeed, the construction already given yields a finite witness set $S\subset\mathbb{Z}^d$. For such a set, the vector space of all $(c_s)_{s\in S}\in\mathbb{R}^{|S|}$ satisfying (2.4) is the null space of a certain matrix with integer entries, hence is spanned over $\mathbb{R}$ by elements of $\mathbb{Z}^S$.

By taking the coefficients c_s to be integers and repeating the above reasoning as in the above discussion of sublevel sets, we conclude that $E^\dagger_{\varepsilon,\lambda}$ can contain no finite point configuration $\{x + rS : s \in S\}$ for which the pair (x, r) satisfies

$$\text{distance}\,(\lambda h(x, r), 2\pi\mathbb{Z}) \geq C\varepsilon, \tag{5.2}$$

where $h(x, r) = \sum_{s\in S} c_s P(x + rs)$ is a polynomial in r of positive degree, whose coefficients are polynomial functions of x and whose leading coefficient is independent of x.

The set of all (x, r) satisfying (5.2) is more complicated than the corresponding set in the proof of Proposition 2.1, so some additional preparation is needed before the theorem of Furstenberg and Katznelson can be applied. Fix any bounded set $B \subset \mathbb{R}^d$. Using the fact that $r \mapsto h(x, r)$ is a polynomial whose coefficients are polynomials in x, and whose leading coefficient is independent of x, it follows that if A is chosen to be sufficiently large, then there exist $c, C \in \mathbb{R}^+$ such that for all $\delta \in (0, 1]$ and any $|\lambda| \geq 1$,

$$|\{(x, r) \in B \times (0, 1] : \|\lambda h(x, r)\| \leq \delta r^A\}| \leq C\delta^c.$$

This can be established using the fact that if $\phi : [0, 1] \to \mathbb{R}$ is a C^2 function satisfying $\phi'(t) \geq 1$ and $\phi''(t) \geq 0$ for all $t \in [0, 1]$, then $|\{t \in [0, 1] : \|\phi(t)\| \leq \delta\}| \leq C\delta$ for all $\delta > 0$, for a certain absolute constant C.

Therefore there exist $A < \infty$ and $c > 0$ such that whenever $|\lambda| \geq 1$, for any $N \geq 1$ and any $j \in \{1, 2, \cdots, N\}$,

$$|\{(x, t) \in B \times [\tfrac{1}{2}, 1] : \|\lambda h(x, tj/N)\| \leq N^{-A'-A}\}| \leq CN^{-cA'}.$$

Choose A' so that $N^{1-cA'} \equiv N^{-1}$. By applying Fubini's theorem and taking unions of exceptional sets over all the N parameters $j \in \{1, 2, \cdots, N\}$, we lose a factor of N and hence conclude that for any $|\lambda| \geq 1$, there exists $t \in [\frac{1}{2}, 1]$ such that

$$\|\lambda h(x, tj/N)\| \geq cN^{-A-A'} \text{ for all } j \in \{1, 2, \cdots, N\} \text{ and all } x \in B \setminus \mathcal{E},$$

where the exceptional set $\mathcal{E}$ satisfies

$$|\mathcal{E}| \leq CN^{1-cA'} = CN^{-1}.$$

In combination with (5.2), this permits the theorem of Furstenberg and Katznelson to be applied, in the same spirit as in the proof of Proposition 2.1 above. □

A strong variant, to the effect that there is a uniform sublevel set estimate of this form with $\lambda = 1$ for all polynomials P of bounded degree that are uniformly nondegenerate, follows in the same way. A key point is that the proof of Theorem 2.2 produces a finite witness set S which is independent of P, so long as P has bounded degree.

Remark 5.1. It remains an open question whether the multilinear oscillatory integral inequalities (1.2) or (1.3) hold for all nondegenerate polynomial phases P, without additional hypotheses. In the rationally commensurate case, Theorem 2.5 does rule out certain *strong* counterexamples to (1.3). Such a strong counterexample, for some sequence of values of λ tending to $+\infty$, has each function $f_j(y) = f_{j,\lambda}(y)$ of the form $f_{j,\lambda}(y) = e^{-i\phi_{j,\lambda}(y)}$ for some measurable real-valued phase $\phi_{j,\lambda}$, with the phases satisfying distance $\left(\lambda P(y) - \sum_j \phi_{j,\lambda}(\ell_j(y)), 2\pi\mathbb{Z}\right) < \delta(\lambda)$ for all y outside a set E_λ, where $\delta(\lambda) \to 0$ as $\lambda \to +\infty$ through the given sequence, and the measure $d\mu(y) = \eta(y)\,dy$ satisfies $\mu(E_\lambda) \to 0$ as $\lambda \to +\infty$. In such a situation, the $\mathcal{I}_\lambda(f_1, \cdots, f_n)$ would not tend to zero.

6. Discussion

Bilinear case. The bilinear case has been intensively studied. Recall first the nonsingular situation, in which the mapping $y \mapsto (\ell_1(y), \ell_2(y))$ of $\mathbb{R}^d$ to $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ is a bijection. In this case, $\mathcal{I}_\lambda(f_1, f_2)$ can be written as

$$\int e^{i\lambda P(x,y)} f(x)g(y)\eta(x,y)\,dx\,dy,$$

where P is a real-valued polynomial. A necessary and sufficient for (1.2) is that there exist nonzero multi-indices α, β for which $\partial^{\alpha+\beta}P/\partial x^\alpha \partial y^\beta$ does not vanish identically; equivalently, P is not a sum of one function of x plus another function of y.

Next consider the singular bilinear situation. $\mathcal{I}_\lambda(f, g)$ can always be expressed in the form $\int e^{i\lambda P(x,y,z)} f(x, z)g(y, z)\eta(x, y, z)\,dx\,dy\,dz$, where x, y range over Euclidean spaces of arbitrary dimensions, and z over a space of positive dimension. Such an expression satisfies (1.2) if and only if there exist $\alpha, \beta \neq 0$ such that $\partial^{\alpha+\beta}P/\partial x^\alpha \partial y^\beta$ does not vanish identically as a function of all three variables. For on one hand, if all such mixed partial derivatives do vanish identically, then $P(x, y, z)$ can be decomposed in the form $p(x, z)+q(y, z)$. The resulting factors $e^{i\lambda p}$ and $e^{i\lambda q}$ can be incorporated into f, g respectively, and there is consequently no valid inequality (1.2). On the other hand, if some such mixed partial derivative does not vanish identically, then integration with respect to x, y for fixed z sets up a nonsingular problem for a bilinear form. The result of the preceding paragraph gives a bound

$$C \min\left(1, |Q(z)|^{-1}|\lambda|^{-\delta}\right)\|f(\cdot, z)\|_\infty \|g(\cdot, z)\|_\infty$$

for some exponent $\delta > 0$ and some polynomial Q which does not vanish identically. (1.2) easily follows by integration with respect to z.

The conclusion is that any formally singular bilinear situation can be reduced to nonsingular ones by freezing some of the coordinates, exploiting oscillation, then integrating with respect to the frozen coordinates.

Higher order case. In contrast, the singular multilinear forms of higher order studied in this paper are not in general reducible to nonsingular ones in this way. As an example, define $P:\mathbb{R}^3\to\mathbb{R}$ to be $P(x_1,x_2,x_3)=x_3^2$. Fix a large positive integer N, to be specified below. For $j\in\{1,2,3,\cdots,N\}$ choose nonzero unit vectors $v_j=(v_j^1,v_j^2,v_j^3)\in\mathbb{R}^3$, none of which is a scalar multiple of another, all satisfying

$$(v_j^3)^2=(v_j^1)^2+(v_j^2)^2. \tag{6.1}$$

Define $\ell_j(x)=x\cdot v_j=x_1v_j^1+x_2v_j^2+x_3v_j^3$, and consider

$$\mathcal{I}_\lambda(f_1,f_2,\cdots,f_N)=\int_{\mathbb{R}^3}e^{i\lambda P(x)}\prod_{j=1}^N f_j(\ell_j(x))\,\eta(x)\,dx$$

where $\eta\in C_0^\infty$ is a cutoff function which does not vanish identically. This multilinear operator is singular, since the integral is taken over $\mathbb{R}^3$ but the sum of the dimensions of the target spaces of the mappings ℓ_j is N. The differential operator $L=\frac{\partial^2}{\partial x_3^2}-\frac{\partial^2}{\partial x_1^2}-\frac{\partial^2}{\partial x_2^2}$ annihilates $f_j\circ\ell_j$ for all j, by virtue of the equations (6.1), but does not annihilate P. Therefore P is nondegenerate relative to $\{\ell_j:1\le j\le N\}$.

Consider the restriction of P and the ℓ_j to any two-dimensional affine subspace V. Let w_j be the projection of v_j onto the unique parallel translate of V which contains 0. For $x\in V$, $x\cdot v_j=x\cdot w_j$ plus a constant independent of x. V may identified with $\mathbb{R}^2$, and the integral over V is then expressed as $\int_{\mathbb{R}^2}e^{iQ(y)}\prod_{j=1}^N(\tilde f_j(y\cdot w_j))\,\tilde\eta(y)\,dy$, where the phase Q is a *quadratic* polynomial.

We claim that $\{v_j\}$ can be chosen so that for every affine two-dimensional subspace V of $\mathbb{R}^3$, $P|_V$ is degenerate, relative to $\{\ell_j|_V\}$. Thus there is no inequality of the form (1.2), nor any sublevel set bound of the form (2.8),(2.7), relative to V. Therefore the method of reduction to lower dimension by "slicing" is not applicable.

To establish the claim, observe first that any quadratic polynomial $Q:\mathbb{R}^2\to\mathbb{R}$ is necessarily degenerate, relative to any family of three or more mappings of the form $L_j(y)=y\cdot w_j$ which satisfy the requirement that none of the vectors w_j is a scalar multiple of any of the others. This is seen by permuting the indices j and changing coordinates so that $w_1=(1,0)$, $w_2=(0,1)$, and $w_3=(a,b)$ with both a,b nonvanishing. Then any quadratic polynomial in (y_1,y_2) can be expressed as a linear combination of $\{y_1,y_2,y_1^2,y_2^2,(ay_1+by_2)^2\}$.

It remains only to show that N and $\{v_j:1\le j\le N\}$ can be chosen both to satisfy (6.1) and so that for every two-dimensional subspace $V\subset\mathbb{R}^3$, some subcollection of three of the associated vectors $\{w_j:1\le j\le N\}$ has no element equal to a scalar multiple of any other element. Define

$$v_j=(v_j^1,v_j^2,v_j^3)=2^{-1/2}(\cos(2\pi/N),\sin(2\pi/N),1).$$

If N is sufficiently large then the required property clearly holds, for otherwise one can obtain a contradiction by letting $N\to\infty$ and exploiting the compactness of the Grassmann manifold of all subspaces V.

Connection with Gowers uniformity norms. The first step in the algebraic proof of the existence of finite witness sets is to apply a finite difference operator to $P-\sum_j(f_j\circ\ell_j)$ which annihilates $e^{i\lambda P}$ and every term $f_j\circ\ell_j$ but a single one. This operation has an analytic counterpart for multilinear oscillatory integrals. Write $\mathcal{I}_\lambda(f_1,\cdots,f_m)=\langle T_\lambda(f_1,\cdots,f_{m-1}),\,f_m\rangle$ for certain multilinear operators T_λ. Then $|\mathcal{I}_\lambda(f_1,\cdots)|^2\lesssim\|f_m\|_{L^2}^2\int|T_\lambda(f_1,\cdots,f_{m-1})|^2$, and it suffices to obtain an upper bound for the integral. This leads to the elimination of f_m, the replacement of P by a polynomial of lower degree, and the replacement of each remaining f_j by $f_j(y+\ell_j(v))\overline{f_j(y)}$, with an additional integration with respect to $v\in\text{nullspace}\,(\ell_m)$. Iterate this operation until only f_1 remains, and then if necessary iterate finitely many additional times until P is eliminated.

Suppose for simplicity that all the target spaces $\mathbb{R}^{d_j}$ are one-dimensional. One then obtains a bound of the type (1.3) unless the Gowers uniformity norm $\|f_1\|_{U^k(\mathbb{R}^{d_1})}$ is bounded below by $\eta(\lambda)\|f_1\|_{L^\infty}$ for a certain k, where $\eta(\lambda)\to 0$ very slowly as $|\lambda|\to\infty$. The index k which arises depends both on the degree of P, and on the number of functions f_j. This argument applies for each index j, so by multilinearity, the estimation of $\mathcal{I}_\lambda$ reduces to the case in which none of the functions $f_j/\|f_j\|_{L^\infty}$ has very small uniformity norm.

Thus an appropriate description of functions whose uniformity norms are not small should lead to a proof of (1.3). Inverse theorems for these norms over finite cyclic groups are currently known for all k. Their thrust is that if $\|f\|_{U^k}$ is not small relative to $\|f\|_{L^\infty}$, then f can be decomposed into a controlled sum of functions which resemble e^{iQ} for polynomials Q of bounded degree, plus a small remainder. But (slurring over the distinction between Euclidean spaces for the sake of simplicity) the weak inverse theorems currently known for $k>3$ do not seem to be sufficiently strong quantitatively to yield any bound of the form (1.3) for $\mathcal{I}_\lambda$. One needs "global" results on $\mathbb{Z}/N\mathbb{Z}$, which are uniform in N in a suitable sense; results which yield decompositions valid on subsets whose cardinalities are $o(N)$ are not useful here. Results of this sort are indeed known for $k\le 3$. We hope to show in a forthcoming work how an inverse theorem known for the U^3 norm leads to trilinear oscillatory integral bounds for quadratic polynomial phases.

References

[1] J. Bennett, A. Carbery, M. Christ and T. Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals*, Geom. Funct. Anal. 17 (2008), no. 5, 1343–1415.

[2] ———, *Finite bounds in Hölder-Brascamp-Lieb multilinear inequalities*, to appear, Math. Research Letters. math.CA/0505691

[3] 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.

[4] M. Christ, *On the simplest trilinear operators*, Math. Research Letters 8 (2001), 43-56.

[5] M. Christ, X. Li, T. Tao, and C. Thiele, *On multilinear oscillatory integrals, nonsingular and singular*, Duke Math. J. 130 (2005), 321–351.

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

[7] 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.

Michael Christ, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA

E-mail address: mchrist@math.berkeley.edu