A ROTH-TYPE THEOREM FOR DENSE SUBSETS OF $\mathbb{R}^d$.

BRIAN COOK ÁKOS MAGYAR MALABIKA PRAMANIK

Abstract. Let $1 < p < \infty$, $p \neq 2$. We prove that if $d \geq d_p$ is sufficiently large, and $A \subseteq \mathbb{R}^d$ is a measurable set of positive upper density then there exists a $\lambda_0 = \lambda_0(A)$ such for all $\lambda \geq \lambda_0$ there are $x, y \in \mathbb{R}^d$ such that $\{x, x+y, x+2y\} \subseteq A$ and $|y|_p = \lambda$, where $||y||_p = (\sum_i |y_i|^p)^{1/p}$ is the $l^p(\mathbb{R}^d)$-norm of a point $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$. This means that dense subsets of $\mathbb{R}^d$ contain 3-term progressions of all sufficiently large gaps when the gap size is measured in the l^p-metric. This statement is known to be false in the Euclidean l^2-metric as well as in the l^1 and ℓ^∞-metrics, and one of the goals of this note is to understand this phenomenon.

1. Introduction.

A main objective of Euclidean Ramsey theory is the study of geometric configurations in large but otherwise arbitrary measurable sets. A fundamental and representative result in the field states that for any $d \geq 2$, a set $A \subseteq \mathbb{R}^d$ of positive upper Banach density contains all large distances. i.e., for every sufficiently large $\lambda \geq \lambda_0(A)$ there are points $x, x+y \in A$ such that $||y||_2 = \lambda$. Recall that A has positive upper Banach density if

$$\bar{\delta}(A) := \limsup_{N\to\infty} \sup_{x\in\mathbb{R}^d} \frac{|A \cap (x + [0,N]^d)|}{N^d} > 0.$$

The result was obtained independently, along with various generalizatons, by a number of authors, for example Furstenberg, Katznelson and Weiss [5], Falconer and Marstrand [4], and Bourgain [1].

A natural question to ask is whether similar statements exist that involve a larger point configuration. Indeed such results are well-known in the discrete regime of the integer lattice, under suitable assumptions of largeness on the underlying set. Such results can often be translated to existence of configurations in the Euclidean setting as well. For instance, Roth's theorem [9] in the integers states that a subset of $\mathbb{Z}$ of positive upper density contains a three-term arithmetic progression $\{x, x+y, x+2y\}$ and it easily implies that a measurable set $A \subseteq \mathbb{R}$ of positive upper density contains a three-term progression whose gap size can be arbitrarily large. On the other hand, a simple example given in [1] shows that there is a set $A \subseteq \mathbb{R}^d$ in any dimension $d \geq 1$, such that the gap lengths of all 3-progressions in A do not contain *all sufficiently large* numbers. More precisely, the counterexample provided in [1] is the set A of points $x \in \mathbb{R}^d$ such that $|||x||_2^2 - m| \leq \frac{1}{10}$ for some $m \in \mathbb{N}$. The parallelogram identity $2||y||_2^2 = ||x||_2^2 + ||x+2y||_2^2 - 2||x+y||_2^2$ then dictates that $\left|||y||_2^2 - \frac{\ell}{2}\right| \leq \frac{4}{10}$ (for some $\ell \in \mathbb{N}$) for any progression $\{x, x+y, x+2y\} \subseteq A$.

This example however does not exclude the validity of such a result when the gaps are measured using some other metric on $\mathbb{R}^d$ that does not obey the parallelogram law. The aim of this note is to show that this is indeed the case for the l^p metric $||y||_p := (\sum_{i=1}^d ||y_i||^p)^{1/p}$ for all $1 < p < \infty$, $p \neq 2$, and in this sense, a counterexample as described above is more an exception rather than the rule. A more general result of this type in the finite field setting was given by the first two authors in [2]. Variations of our arguments also work for other metrics given by positive homogeneous

1991 Mathematics Subject Classification. 11D72, 11P32.

The second and third authors are respectively supported by grant ERC-AdG. 321104 and an NSERC Discovery grant.

polynomials of degree at least 4 and certain nondegenerate point configurations. We hope to pursue these extensions elsewhere.

2. Main results.

Theorem 2.1. *Let $1 < p < \infty$, $p \neq 2$. Then there exists a constant $d_p \geq 2$ such that for $d \geq d_p$ the following holds. Any measurable set $A \subseteq \mathbb{R}^d$ of positive upper Banach density contains a three-term arithmetic progression $\{x, x+y, x+2y\} \subseteq A$ with gap $||y||_p = \lambda$ for all sufficiently large $\lambda \geq \lambda(A)$.*

Remarks:

(a) The result is sharp in the range of p. Easy variants of the example in [1] show that Theorem 2.1 and in fact even the two-point results of [5, 1, 4] cannot be true for $p = 1$ and $p = \infty$. Indeed, if $A = \mathbb{Z}^d + \epsilon_0[-1, 1]^d$ for some small $\epsilon_0 > 0$, then on one hand A is of positive upper Banach density. On the other, if $x, x+y \in A$ for some $y \neq 0$, then both $||y||_\infty$ and $||y||_1$ are restricted to lie within distance $O(\epsilon_0)$ from some positive integer.

(b) We believe that the p-dependence of the dimensional threshold d_p stated in the theorem is an artifact of our proof. In our analysis, d_p grows without bound as $p \nearrow \infty$, while other implicit constants involved in the proof blow up near $p = 1$ and $p = 2$. See in particular Proposition 2.2 and Lemma 4.3. It would be of interest to determine whether Theorem 2.1 holds for all $d \geq 2$ for the specified values of p.

2.1. **Overview of proof.** We describe below the main elements of the proof. Details will be provided in subsequent sections.

Our main observation is a stronger finitary version of Theorem 2.1 for bounded measurable sets.

Theorem 2.2. *Let $1 < p < \infty$, $p \neq 2$. Let $d \geq d_p$, $\delta > 0$ and let $N \geq N(\delta)$ be sufficiently large. Then for any measurable set $A \subseteq [0, N]^d$ of measure $|A| \geq \delta\, N^d$ the following holds.*
For any lacunary sequence $1 < \lambda_1 < \ldots < \lambda_J \ll N$ with $\lambda_{j+1} \geq 2\lambda_j$ and $J \geq J(\delta)$, there exists a three term arithmetic progression $\{x, x+y, x+2y\} \subseteq A$ such that $||y||_p = \lambda_j$ for some $1 \leq j \leq J$,

Proof of Theorem 2.1. Theorem 2.2 implies Theorem 2.1. Indeed, assume that Theorem 2.1 does not hold. Then there exists an infinite sequence $\{\lambda_j\}_{j=1}^\infty \subseteq \mathbb{N}$ such that $\lambda_j \neq |y|$ for any j and any y which is the gap of a 3-progression contained in A. Without loss of generality the sequence may be assumed to be lacunary, i.e. $\lambda_{j+1} \geq 2\lambda_j$ for all j. Setting $\delta = \overline{\delta(A)}/2$, fix any sufficiently large $J = J(\delta)$ and any sufficiently large box B_N of size $N = N(\delta, \lambda_J)$ on which the density of A is $|A|/N^d \geq \delta$. By translation invariance we may assume $B_N = [0, N]^d$, contradicting Theorem 2.2. $\square$

2.1.1. *A counting function and its variants.* For the rest of the paper we will fix a finite exponent $p > 1$, $p \neq 2$, and for simplicity of notation write $|y| = ||y||_p$. We start by counting three term arithmetic progressions $P = \{x, x+y, x+2y\}$ contained in A via a positive measure σ_λ supported on the l^p-sphere $S_\lambda = \{y \in \mathbb{R}^d_+;\ |y| = \lambda\}$. We have restricted the y variables to the positive quadrant $\mathbb{R}^d_+$ as in this case $|y|^p = \sum y_i^p$ which is easier to work with. Let $f := \mathbf{1}_A$ be the indicator function of a measurable set $A \subseteq [0, N]^d$. As is standard in enumerating configurations, we introduce the counting function

$$\mathcal{N}_\lambda(f) := \int_{\mathbb{R}^d} \int_{S_\lambda} f(x)f(x+y)f(x+2y)\, d\sigma_\lambda(y)\, dx. \tag{2.1}$$

Clearly if $\mathcal{N}_\lambda(f) > 0$ then A must contain a 3-progression $x, x+y, x+2y$ with $|y| = \lambda$. We will define the measure σ_λ via the oscillatory integral

$$\sigma_\lambda(y) := \lambda^{-d+p}\chi_+(y)\int_{\mathbb{R}} e^{i\,(|y|^p-\lambda^p)t}\,dt, \tag{2.2}$$

where $\chi_+(y)$ is the indicator function of the positive quadrant $\mathbb{R}^d_+$. It is well-known (see [7], Ch.2) that the above oscillatory integral defines an absolutely continuous measure with respect to the surface area measure on S_λ whose density function is $|\nabla Q(y)|^{-1}$ with $Q(y) = |y|^p$. The normalizing factor λ^{-d+p} is inserted to ensure that $\sigma_\lambda(S_\lambda) = \sigma_1(S_1) > 0$, which is independent of λ.

Let ψ be a Schwarz function such that $0 \le \psi \le 1$, $\psi(0) = 1$ and $\widehat{\psi} \ge 0$ is compactly supported. Define the quantity

$$\omega_\lambda(y) := \lambda^{-d+p}\chi_+(y)\int_{\mathbb{R}} e^{it\,(|y|^p-\lambda^p)}\psi(\lambda^p t)\,dt. \tag{2.3}$$

Note that by scaling

$$\omega_\lambda(y) = \lambda^{-d}\omega(y/\lambda) = \lambda^{-d}\,\widehat{\psi}\,(\,|y/\lambda|^p - 1\,), \tag{2.4}$$

hence ω_λ is compactly supported on $B(0; C\lambda)$ with

$$\int \omega_\lambda(y)\,dy = \int \omega(y)\,dy = C_\omega > 0.$$

Also define

$$\mathcal{M}_\lambda(f) := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d_+} f(x)f(x+y)f(x+2y)\,\omega_\lambda(y)\,dy\,dx. \tag{2.5}$$

The first step is to show that this quantity is large.

Proposition 2.1. *Let $0 < \delta \le 1$ and let $A \subseteq [0,N]^d$ be such that $|A| \ge \delta N^d$. Then there exists a constant $c(\delta) > 0$ depending only on δ such that for $0 < \lambda \ll N$,*

$$\mathcal{M}_\lambda(\mathbf{1}_A) \;\ge\; c(\delta)\,N^d. \tag{2.6}$$

As we will see in Section 3, the proof of this proposition is in essence Roth's theorem adapted to the Euclidean setting (see [1], Appendix A, Theorem 3) combined with a "symmetrization trick", executed in Lemma 3.1, to ensure that A can be replaced by a similar set of large density that is symmetric about the coordinate hyperplanes.

Next, we define a variant of $\mathcal{M}_\lambda$ indexed by a small $\varepsilon > 0$ which is a good approximation to $\mathcal{N}_\lambda(f)$. Let

$$\omega^\varepsilon_\lambda(y) := \lambda^{-d+p}\int_{\mathbb{R}} e^{it\,(|y|^p-\lambda^p)}\psi(\varepsilon\lambda^p t)\,dt. \tag{2.7}$$

It is easy to see that

$$\omega^\varepsilon_\lambda(y) = \lambda^{-d}\varepsilon^{-1}\widehat{\psi}\left(\frac{|y/\lambda|^p - 1}{\varepsilon}\right) = \lambda^{-d}\omega^\varepsilon_1(y/\lambda). \tag{2.8}$$

Define

$$\mathcal{M}^\varepsilon_\lambda(f) := \int_{\mathbb{R}^d}\int_{\mathbb{R}^d_+} f(x)f(x+y)f(x+2y)\,\omega^\varepsilon_\lambda(y)\,dx\,dy. \tag{2.9}$$

We will establish the error estimate

Proposition 2.2. *Let $f : [0,N]^d \to [-1,1]$ and let $0 < \varepsilon < 1$. Then there exist constants $\gamma = \gamma_p > 0$ and $C_{p,d} > 0$ both independent of λ such that for $0 < \lambda \ll N$,*

$$|\mathcal{N}_\lambda(f) - \mathcal{M}^\varepsilon_\lambda(f)| \leq C_{p,d}\varepsilon^{\gamma d-1}\, N^d. \tag{2.10}$$

In particular, γ_p is independent of d. The constant $C_{p,d} \nearrow \infty$ as $p \to 1$ or 2, while $\gamma_p \to 0$ as $p \nearrow \infty$.

The proof of Proposition 2.2 is based on two facts that may be of independent interest. The first is an inequality showing that the so-called U^3-uniformity norm of Gowers [6] controls expressions like $\mathcal{N}_\lambda(f)$. Let us recall the definition of the U^3 norm for a compactly supported bounded measurable function g:

$$\|g\|^8_{U^3(\mathbb{R}^d)} = \int_{(x,y)\in\mathbb{R}^d\times\mathbb{R}^{3d}} \Big(\prod_{\nu\in\{0,1\}^3} \bar{g}^\nu(x+\nu_1y_1+\nu_2y_2+\nu_3y_3)\Big)\, dxdy, \tag{2.11}$$

where ν_1,ν_2,ν_3 can take the values 0 or 1, $\bar{g}^\nu = \bar{g}$ if $\nu_1+\nu_2+\nu_3$ is odd and $\bar{g}^\nu = g$ otherwise.

Lemma 2.1. *Let $f : [0,N]^d \to [-1,1]$ and let $0 < \varepsilon < 1$. Then for $0 < \lambda \ll N$ one has*

$$|\mathcal{N}_\lambda(f) - \mathcal{M}^\varepsilon_\lambda(f)| \lesssim N^d\, \|\chi_+(\sigma - \omega_1^\varepsilon)\|_{U^3}. \tag{2.12}$$

While it is not apriori clear how to define the U^3-norm of the measure σ defined in (2.2), we note that $\omega_1^\varepsilon \to \sigma$ weakly as $\varepsilon \to 0$. To prove (2.12), we will first establish that $\{\omega_1^\eta\}$ is a Cauchy sequence with respect to the U^3-norm and then define $\|\sigma - \omega_1^\epsilon\|_{U^3} := \lim_{\eta\to 0}\|\omega_1^\eta - \omega_1^\varepsilon\|_{U^3}$. We will also show that

Lemma 2.2. *Let $0 < \varepsilon < 1$, $p > 1$, $p \neq 2$. If $d > 8r_p$ with $r_p = \max(p+1, 2p-1)$, then one has*

$$\|\chi_+(\sigma - \omega_1^\varepsilon)\|_{U^3} \leq C_{p,d}\varepsilon^{\frac{d}{8r_p}-1}. \tag{2.13}$$

where the constant $C_{p,d}$ has the same behaviour as described in Proposition 2.2.

Let us note in passing that (2.12) and (2.13) yield (2.11) with $\gamma = 1/(8r_p)$. The proof of Lemma 2.2 uses in an essential way that on $\mathbb{R}^d_+$ the norm is defined by the expression $|y|^p = \sum_i y_i^p$ for some $p > 1$, $p \neq 2$. In particular, the fact that $x_i^p + (x_i+2y_i)^p - 2(x_i+y_i)^p$ does not vanish identically for $p \neq 1, 2$ plays a pivotal role. Estimate (2.13) does not hold for the Euclidean l^2-metric or the l^1-metric.

2.1.2. *Multilinear Calderón-Zygmund singular integral operators.* The final ingredient in the proof of Theorem 2.2 is an estimate given in [8] for certain multilinear operators similar to the bilinear Hilbert transform. In order to describe the form in which we need this estimate, let us fix $0 < \epsilon \ll 1$ and define the constant $c_1(\epsilon)$ as follows,

$$c_1(\varepsilon)\int \omega(y)\, dy = \int \omega_1^\varepsilon(y)\, dy. \tag{2.14}$$

We will see in Lemma 4.1 that $c_1(\varepsilon) \approx 1$, i.e. is bounded by two positive constants depending only on the dimension d. Write $k^\varepsilon(y) := \omega^\varepsilon_\lambda(y) - c_1(\varepsilon)\,\omega_{\lambda_j}(y)$, and

$$\mathcal{E}_\lambda(f) := \mathcal{M}^\varepsilon_\lambda(f) - c_1(\varepsilon)\mathcal{M}_\lambda(f) = \int\int f(x)f(x+y)f(x+2y)\, k^\varepsilon(y)\, dy\, dx, \tag{2.15}$$

so that by (2.14) one has the cancellation property

$$\int k^\varepsilon(y)\, dy = \int (\omega^\varepsilon_\lambda(y) - c_1(\varepsilon)\,\omega_\lambda(y))\, dy = \int (\omega_1^\varepsilon(y) - c_1(\varepsilon)\,\omega(y))\, dy = 0. \tag{2.16}$$

The key estimate concerning the operator $\mathcal{E}_\lambda$ is the following:

Proposition 2.3. *Suppose that $\{\lambda_j : 1 \le j \le J\}$ is a lacunary sequence (finite or infinite) with $\lambda_{j+1} \ge 2\lambda_j$ for all j. Then for any $f : [0,N]^d \to [-1,1]$,*

$$\sum_{j=1}^{J} |\mathcal{E}_{\lambda_j}(f)|^2 \le C_\epsilon N^d ||f||_4^4 \le C_\epsilon N^{2d},$$

where the constant C_ϵ depends only on the quantity ϵ used to define $\mathcal{E}_\lambda$ and is in particular, independent of f and the number J of elements in the lacunary sequence.

We provide details of this result in Section 5.

Proof of Theorem 2.2. Assuming Propositions 2.1, 2.2 and 2.3 for now, the proof proceeds by contradiction. Assume that there exist arbitrarily large N, a measurable set $A \subseteq [0,N]^d$ with $|A| \ge \delta N^d$, and a sequence of non-admissible progression gaps $\lambda_1 < \lambda_2 < \cdots < \lambda_J \ll N$ for some $J \ge J(\delta)$, such that $\mathcal{N}_{\lambda_j}(f) = 0$ for $f = 1_A$. The sequence may be chosen to be lacunary, and J may be assumed to be arbitrarily large as well, by choosing N large enough. Thus, for $1 \le j \le J$,

$$0 = \mathcal{N}_{\lambda_j}(f) = c_1(\epsilon)\mathcal{M}_{\lambda_j}(f) + \left[\mathcal{N}_{\lambda_j}(f) - \mathcal{M}^\epsilon_{\lambda_j}(f)\right] + \mathcal{E}_{\lambda_j}(f)$$

In view of Propositions 2.1 and 2.2, and recalling that $c_1(\epsilon) \approx 1$, we find that for some sufficiently small ϵ depending on p, d and δ, the inequality

$$\frac{c(\delta)}{2}N^d \le \left(c(\delta) - C_{p,d}\epsilon^{\gamma d-1}\right)N^d \le \left|c_1(\epsilon)\mathcal{M}_{\lambda_j}(f) + \left[\mathcal{N}_{\lambda_j}(f) - \mathcal{M}^\epsilon_{\lambda_j}(f)\right]\right| = \left|\mathcal{E}_{\lambda_j}(f)\right|$$

holds for every $1 \le j \le J$. Squaring both sides and summing over all $j \le J$ yields, after an application of Proposition 2.3 with $f = 1_A$,

$$c(\delta)JN^{2d} \le \sum_{j=1}^{J} |\mathcal{E}_{\lambda_j}(f)|^2 \le C_\epsilon N^{2d}.$$

This implies that $J \le C_{p,d,\delta}$, contradicting the hypothesis that J can be chosen arbitrarily large. □

3. The main term

We now set about proving the main propositions leading up to the theorem. In this section we prove Proposition 2.1, via an application of Roth's theorem on compact abelian groups (see [1], Appendix A, Theorem 3). The compact group of interest is of course the d-dimensional torus Π^d.

Our first task is to justify that without loss of generality, the set A can be chosen to be symmetric about the coordinate hyperplanes, that is $A = \tau^i(A)$ where $\tau^i(x_i) = -x_i$, $\tau^i(x_j) = x_j$ $(j \ne i)$, for all $1 \le i \le d$. For a given $t \in \mathbb{R}$ let τ^i_t denote the reflection $\tau^i_t(x_i) = 2t - x_i$, $\tau^i_t(x_j) = x_j$ for $j \ne i$.

Lemma 3.1. *Let $0 < \delta < 1$, $N > 1$ and let $A \subseteq [0,N]^d$ such that $|A| \ge \delta N^d$. Then there exists a point $\alpha = (\alpha_i)_{1 \le i \le d}$ such that $|A'| \ge \delta' N^d$, where $A' = \left[\bigcap_{1\le i\le d} \tau^i_{\alpha_i}(A)\right] \cap A$ and $\delta' \ge 2^{-2d+1}\delta^{2^d} > 0$.*

Proof. By scaling one may assume $A \subseteq [0,1]^d$, $|A| \ge \delta$. We will inductively define a family of sets $A \supset A_1 \supset \ldots \supset A_d$ and finally set $A' = A_d$. Write $f = \mathbf{1}_A$, and estimate, using the Cauchy-Schwarz inequality

$$\iint_{(x_1,t)\in[0,1]^2} \int_{x'\in[0,1]^{d-1}} f(x_1,x')f(2t-x_1,x')\,dt\,dx_1\,dx'$$
$$= \frac{1}{2}\int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} f(x_1,x')\,dx_1\right)^2 dx' \ge \frac{\delta^2}{2}.$$

Since for a fixed t,

$$|A \cap \tau_t^1(A)| = \int_{[0,1]} \int_{[0,1]^{d-1}} f(x_1, x') f(2t - x_1, x') \, dx' \, dx_1,$$

and is vanishing unless $0 \leq t \leq 1$, there must exist some $t = \alpha_1$ such that $|A \cap \tau_{\alpha_1}^1(A)| \geq \delta^2/2$. Set $A_1 := A \cap \tau_{\alpha_1}^1(A)$ and $\delta_1 := \delta^2/2$. Now repeat the procedure inductively in each coordinate for $i = 2, \ldots, d$ to generate the points $\alpha_2, \ldots, \alpha_d$ and the sets $A_2 \supset \ldots \supset A_d$. It is immediate from the construction that $|A_i| \geq \delta_i$ where $\delta_1 = \delta^2/2$ and $\delta_{i+1} = \delta_i^2/2$ and that the set A_i is invariant under the reflections $\tau_j(\alpha_j)$ for all $1 \leq j \leq i$. An easy calculation shows that $|A_d| \geq 2^{-2d+1}\delta^{2^d}$ and satisfies the conclusion of the lemma with $\alpha = (\alpha_1, \ldots, \alpha_d)$. □

Proof of Proposition 2.1. If we define $A'' := A' - \alpha$ with A' and α constructed in the above lemma, then $A'' \subseteq [-N, N]^d$ will be symmetric with respect to all the coordinate hyperplanes and will have density $\delta'' \geq 2^{-3d+1}\delta^{2^d} > 0$. Since $\mathcal{M}_\lambda(1_A) \geq \mathcal{M}_\lambda(1_{A'}) = \mathcal{M}_\lambda(1_{A''})$, we will assume without loss of generality for the remainder of this section that $A \subseteq [-N, N]^d$ is invariant under all reflections $x_i \to -x_i$. Then it is easy to see that for such $f = \mathbf{1}_A$,

$$2^d \mathcal{M}_\lambda(f) = \mathcal{M}'_\lambda(f) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) f(x+y) f(x+2y) \widetilde{\omega}_\lambda(y) \, dx \, dy,$$

where $\widetilde{\omega}_\lambda$ is defined the same way as ω_λ in (2.3), except without the factor $\chi_+(y)$. Thus in proving Proposition 2.1 we may omit the restriction $y \in \mathbb{R}^d_+$.

Fix $\delta \in (0, 1]$ and $\lambda \ll N$ for a sufficiently large N. Consider a real-valued function $f : [-N, N]^d \to [0, 1]$ with $\int f \geq \delta N^d$. Equipartition the cube $[-N, N]^d$ in the natural way into disjoint boxes with sides parallel to the coordinate axes and length $\ell = c\lambda$ by choosing a sufficiently small number $c > 0$ so that N/l is a positive integer. Enumerate the boxes by $\{B_i : 1 \leq i \leq L\}$. Each box B_i may then be identified with its leftmost endpoint x_i, so that $B_i = [0, \ell]^n + x_i$.

We split f into pieces restricted to each box. More precisely, define $g_i(x) : [0, \ell]^d \to [0, 1]$ by $g_i(x) = \mathbf{1}_{[0,l]^n}(x + x_i) f(x + x_i)$, so that $f = \sum_{i=1}^L g_i$ on $[-N, N]^d$. The non-negativity of f implies the bound

$$\mathcal{M}'_\lambda(f) \geq \sum_{i=1}^{L} \iint_{x,y \in \mathbb{R}^d} g_i(x) g_i(x+y) g_i(x+2y) \, \omega_\lambda(y) \, dy \, dx. \tag{3.1}$$

Recall from (2.4) that $\omega_\lambda(y) = \lambda^{-d} \, \widehat{\psi}\,(\,|y/\lambda|^p - 1\,)$; hence we may choose ψ such that $\widehat{\psi}\,(\,|y/\lambda|^p - 1\,) \geq 1/10$ for $y \in [-\ell, \ell]^d$. Then (3.1) yields

$$\mathcal{M}'_\lambda(f) \geq \frac{\lambda^{-d}}{10} \sum_{i=1}^{L} \iint_{x,y \in \mathbb{R}^d} g_i(x) g_i(x+y) g_i(x+2y) \, dy \, dx. \tag{3.2}$$

Now identify Π^d with the cube $[-\frac{1}{2}, \frac{1}{2}]^d$. After a change of variable $(x, y) \mapsto 10\ell(x, y)$, each summand on the right hand side of (3.2) may be written as

$$\begin{aligned}(10\ell)^{2d} &\iint_{x,y \in \mathbb{R}^d} g_i(10\ell x) g_i(10\ell(x+y)) g_i(10\ell(x+2y)) \, dy \, dx \\ &= (10\ell)^{2d} \iint_{x,y \in \Pi^d} g_i(10\ell x) g_i(10\ell(x+y)) g_i(10\ell(x+2y)) \, dy \, dx.\end{aligned}$$

Note that the support assumptions on g_i dictate that the integrand is supported on $[-\frac{1}{10}, \frac{1}{10}]^{2d} \subset \Pi^d$, as indicated in the last step. If we also have that

$$\int_{\Pi^d} g_i(10\ell x)\, dx \geq \eta > 0 \qquad \text{for some index } i, \tag{3.3}$$

then Roth's theorem on compact abelian groups would imply that for such an index

$$\iint_{\Pi^d \times \Pi^d} g_i(10\ell x) g_i(10\ell(x+y)) g_i(10\ell(x+2y))\, dx\, dy \;\geq\; c_0(d,\eta),$$

where $c_0(d,\eta) > 0$ is a constant depending only on d and η. We will prove in Lemma 3.2 below that (3.3) holds with $\eta = \delta(10)^{-d}/2$ for at least $\delta L/2$ indices i. Summing over all these indices in (3.2) then leads to the bound

$$\mathcal{M}'_\lambda(f) \geq c_0(d,\eta)\lambda^{-d}\frac{\delta}{2}L(10\ell)^{2d} = c(d,\delta),$$

as claimed. □

Lemma 3.2. *The relation* (3.3) *holds with* $\eta = \frac{\delta(10)^{-d}}{2}$ *for at least* $\delta L/2$ *indices* i.

Proof. This is a simple pigeonholing argument. Let I denote the number of indices i for which the integral inequality in (3.3) holds. After a scaling change of variable, this is the same set of indices i for which $\int_{[0,\ell]^d} g_i(x)\, dx \geq (\delta/2)\ell^d$. By our hypothesis on f,

$$\int f(x)\, dx = \sum_{i=1}^{L} \int g_i(x)\, dx \geq \delta N^d. \tag{3.4}$$

On the other hand, $0 \leq g_i \leq 1$ is supported on $[0,\ell]^d$, so $||g_i||_1 \leq \ell^d$ for trivial reasons. This leads to the estimate

$$\sum_{i=1}^{L} \int g_i(x) \leq (\delta/2)N^d + I\ell^d. \tag{3.5}$$

for each $1 \leq i \leq L$. Combining the lower bound in (3.4) with the upper bound in (3.5) and recalling that $L = N^d/\ell^d$ leads to the claimed statement. □

4. Error estimates.

As indicated in the introduction, our main objective here is to prove Proposition 2.2, which is a direct consequence of Lemma 2.1 and 2.2. Before turning our attention to the proof of these lemmas, let us first make the simple but important observation that $\int \omega_1^\varepsilon(y)\, dy \approx 1$ uniformly for sufficiently small $\varepsilon > 0$. Let $\nu_p := \big|\{y;\ |y|_p \leq 1\}\big|$, then by homogeneity $\big|\{y;\ |y|_p \leq \eta\}\big| = \nu_p\, \eta^d$. This fact was used in the proof of Theorem 2.2 in order to bound from below the main term $c_1(\epsilon)\mathcal{M}_{\lambda_j}(f)$.

Lemma 4.1. *There exist constants* $0 < c_1 < c_2$ *depending only on the function* ψ *and the parameter* d *and* p *such that*

$$c_1 \leq \int_{\mathbb{R}^d} \omega_1^\varepsilon(y)\, dy \leq c_2, \tag{4.1}$$

uniformly for $0 < \varepsilon < \frac{1}{10d}$.

Proof. By the definition of the function ψ we have $|\widehat{\psi}(t)| \geq c$ for $|t| \leq \tau$ for some constants $c > 0$ and $0 < \tau \leq 1$. Then by (2.7) one has

$$\int \omega_1^\varepsilon(y)\, dy \geq \frac{c}{\varepsilon}\,\big|\{y;\ 1-\tau\varepsilon \leq |y|^p \leq 1+\tau\varepsilon\}\big| = \frac{c\nu_p}{\varepsilon}\left((1+\tau\varepsilon)^{d/p} - (1-\tau\varepsilon)^{d/p}\right) \geq c_1,$$

uniformly for $0 < \varepsilon \leq 1/10d$ as $p > 1$. Similarly, as $0 \leq \widehat{\psi} \leq 1$ and $\widehat{\psi}$ is supported on $(-2, 2)$,

$$\int \omega_1^\varepsilon(y)\, dy \leq \frac{1}{\varepsilon} \left|\{y;\ 1 - 2\varepsilon \leq |y|^p \leq 1 + 2\varepsilon\}\right| = \frac{\nu_p}{\varepsilon} \left((1+2\varepsilon)^{d/p} - (1-2\varepsilon)^{d/p}\right) \leq c_2.$$

□

Since $\omega_1^\varepsilon(y)$ is invariant under reflections to the coordinate hyperplanes we have that

$$\int \chi_+(y)\omega_1^\varepsilon(y)\, dy = 2^{-d} \int \omega_1^\varepsilon(y)\, dy,$$

and the above lemma holds for $\chi_+(y)\omega_1^\varepsilon(y)$ as well.

To prove Lemma 2.1 we need the following result.

Lemma 4.2. *Let $f : [-N, N]^d \to [-1, 1]$ and $g : [0, \lambda]^d \to [-1, 1]$ be given functions. Then*

$$\int_{x,y \in \mathbb{R}^d} f(x)f(x + y)f(x + 2y)g(y)\, dx\, dy \ \lesssim\ N^d\, \lambda^{d/2}\, \|g\|_{U^3}, \tag{4.2}$$

where the implicit constant depends only on d.

Proof. The proof involves several changes of variables and successive applications of the Cauchy-Schwarz inequality. Set

$$T = \int_{x,y\in\mathbb{R}^d} f(x)f(x + y)f(x + 2y)g(y)\, dx\, dy.$$

Applying the Cauchy-Schwarz inequality in the x integration to get

$$T^2 \leq N^d \int_x \int_y \int_{y'} f(x + y)f(x + 2y)f(x + y')f(x + 2y')g(y)g(y')\, dx\, dy\, dy'$$

Use the substitution $y' = y + h$ followed by the substitution $x \to x - y$, and define

$$\Delta_h F(x) = F(x + h)\overline{F(x)}$$

for a generic complex valued function F. Then one may write

$$T^2 \leq N^d \int_x \int_y \int_h \Delta_h f(x)\Delta_{2h} f(x + y)\Delta_h g(y)\, dx\, dy\, dh.$$

The integrals in y, h may be restricted to a region with $|y|$, $|h| \lesssim \lambda$ due to the support of g. Then another application of Cauchy-Schwarz in the x and h integration gives

$$T^4 \lesssim N^{3d}\lambda^d \int_{x,h,y,y'} \Delta_{2h} f(x + y)\Delta_{2h} f(x + y')\Delta_h g(y)\Delta_h g(y')\, dx\, dy\, dh\, dy'$$

Again use the substitutions $y' = y + k$ and $x \to x - y$ in turn to get

$$T^4 \lesssim N^{3d}\lambda^d \int_{x,y,k,h} \Delta_{2h} f(x)\Delta_{2h} f(x + k)\Delta_h g(y)\Delta_h g(y + k) dx\, dy\, dh\, dk.$$

One final application of the Cauchy-Schwarz inequality in x and h and k integration gives

$$T^8 \lesssim N^{7d}\lambda^{4d} \int_{x,y,h,k,y'} 1_{[0,N]^d}(x)\Delta_h g(y)\Delta_h g(y')\Delta g(y + k)\Delta_h g(y' + k) dx\, dy\, dy'\, dh\ dk.$$

The x integration may be carried out, and the applying the substitution $y' \to y + l$ gives the final form

$$T^8 \lesssim N^{8d}\lambda^{4d} \int_{y,h,k,l} \Delta_{h,k,l} f(y)\, dy\, dh\, dk\, dl \tag{4.3}$$

where $\Delta_{h,k,l}$ is well defined as the composition of the operators Δ_h, Δ_k, and Δ_l. The integral is easily verified to be $\|g\|_{U^3}^8$, which completes the proof. $\square$

Proof of Lemma 2.1. As indicated in the introduction, the right hand side of (2.12) is to be interpreted as

$$||\chi_+(\sigma - \omega_1^\epsilon)||_{U^3} := \lim_{\eta\to 0} ||\chi_+(\omega_1^\eta - \omega_1^\epsilon)||_{U^3}.$$

Since the integral representation of

$$\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\epsilon(f) = \lim_{\eta\to 0} \mathcal{M}_\lambda^\eta(f) - \mathcal{M}_\lambda^\epsilon(f)$$

is of the form (4.2), we may apply Lemma 4.2 with $g(y) = \chi_+(y)\big(\omega_\lambda^\eta(y) - \omega_\lambda^\epsilon(y)\big)$ to get

$$\big|\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\epsilon(f)\big| \lesssim N^d\lambda^{d/2} \lim_{\eta\to 0} ||\chi_+\big(\omega_\lambda^\eta - \omega_\lambda^\epsilon\big)||_{U^3}.$$

Since ω_λ^ϵ is a rescaled version of ω_1^ϵ, scaling properties of the U^3 norm imply that

$$||\chi_+\big(\omega_\lambda^\eta - \omega_\lambda^\epsilon\big)||_{U^3} = \lambda^{-d}\lambda^{d/2}||\chi_+\big(\omega_1^\eta - \omega_\lambda^\epsilon\big)||_{U^3},$$

which leads to the claimed upper bound. $\square$

Next we turn to the proof of Lemma 2.2. In what follows we assume that λ and $N \gg \lambda$ are fixed, and f is the characteristic function of a set $A \subset [-N,N]^d$ with measure δN^d. First we need an estimate for one-dimensional scalar oscillatory integrals of the following type.

Lemma 4.3. *Let $1 < p < \infty$, $p \neq 2$. For a smooth cut-off function ϕ, let ϕ_+ be its restriction to the positive real numbers and define the integral*

$$I(t) := \int_{y,h,k,l\in\mathbb{R}} \triangle_{h,k,l}\left(\phi_+(y)e^{it|y|^p}\right) dy\, dh\, dk\, dl. \tag{4.4}$$

Then there exists a constant $r = r(p) > 0$ such that

$$|I(t)| \leq C_p|t|^{-\frac{1}{r}}, \quad \textit{for } |t| \geq 1. \tag{4.5}$$

One may take $r(p) = p+1$ for $1 < p < 2$, and $r(p) = 2p-1$ for $p > 2$. The constant C_p is finite in the indicated range of p, and tends to infinity as $p \to 1$ or 2.

Proof. Replacing $y+h$ by a new variable y', the integral $I(t)$ may be rewritten as

$$I(t) = \int |I_{k,l}(t)|^2\, dk\, dl,$$

where

$$I_{k,l}(t) = \int \triangle_{k,l}\phi_+(y)\, e^{it\psi_{k,l}(y)}\, dy,$$

with

$$\triangle_{k,l}\phi_+(y) = \phi_+(y)\phi_+(y+k)\phi_+(y+l)\phi_+(y+k+l),$$

$$\psi_{k,l}(y) = y^p + (y+k+l)^p - (y+k)^p - (y+l)^p.$$

The reason we can write $\psi_{k,l}(y)$ in this form is that y, $y+k$, $y+l$, $y+k+l$ are all positive on the support of $\triangle_{k,l}\phi_+$.

It is clear that $I_{k,l}(\eta)$ is uniformly bounded, hence $I(t)$ receives small contribution from regions where at least one of the integration variables k, l is small. For a small parameter $0 < \eta < 1$ to be chosen later, we may therefore write

$$I(t) = \int_{|k|,|l| \geq \eta} |I_{k,l}(t)|^2 \, dk \, dl + O(\eta).$$

We now estimate the integral $I_{k,l}(t)$ for fixed k, l assuming $|k|, |l| \geq \eta$. Introducing a smooth partition of unity, we have $I_{k,l}(t) = I'_{k,l}(t) + J_{k,l}(t),$ where the domain of integration of $I'_{k,l}(t)$ ranges over those y for which at least one of the quantities $y,\ y+k,\ y+l,\ y+k+l$ is $O(\eta)$. Thus $I'_{k,l}(t) = O(\eta)$.

For $J_{k,l}(t)$ one may write, using Taylor's remainder formula

$$\psi_{k,l}(y) = klp(p-1) \int_{[0,1]^2} (y + uk + sl)^{p-2} \, du \, ds.$$

Therefore its derivative is given by

$$\psi'_{k,l}(y) = klp(p-1)(p-2) \int_{[0,1]^2} (y + uk + sl)^{p-3} \, du \, ds.$$

By our assumptions, we have that

$$\eta \lesssim y + uk + sl \lesssim 1,$$

uniformly for $0 \leq u, s \leq 1$, and also that $|k|, |l| \gtrsim \eta$. Thus for $p > 1$, $p \neq 2$,

$$|\psi'_{k,l}(y)| \gtrsim \eta^{p-1},$$

with an implicit constant independent of k and l. Then, writing $\psi = \psi_{k,l}$ and χ for the amplitude, integration by parts yields

$$\begin{aligned} J_{k,l}(t) &= \int \frac{d}{dy} \left(e^{it\psi(y)} \right) \frac{\chi(y)}{it\psi'(y)} \, dy = -\frac{1}{it} \int e^{it\psi(y)} \frac{d}{dy} \left(\frac{\chi(y)}{\psi'(y)} \right) dy \\ &= \frac{1}{it} \int e^{it\psi(y)} \left(\frac{\chi'}{\psi'} + \frac{\chi \psi''}{\psi'^2} \right) dy. \end{aligned}$$

Here we have used the support properties of χ in the form $||\chi||_\infty =)(1)$ and $||\chi'||_\infty =)(\eta^{-1})$. Therefore

$$|J_{k,l}(t)| \lesssim |t|^{-1} \left(\eta^{-p} + \eta^{-2p+2} \right) \leq |t|^{-1} \eta^{-r'(p)},$$

with $r'(p) = \max(p, 2p-2) > 0$. This implies that $|I_{k,l}(t) = O(\eta) + O(|t|^{-1}\eta^{-r'_p})$, choosing $\eta := |t|^{-\frac{1}{r_p+1}}$ yields

$$|I(t)| \lesssim |t|^{-\frac{1}{r_p}}, \tag{4.6}$$

with $r(p) = r'(p) + 1 = \max(p+1, 2p-1)$ and the lemma follows. □

Proof of Lemma 2.2. Since $\omega_1^\eta - \omega_1^\epsilon$ is compactly supported, say on $[-C, C]^d$, let us fix a smooth cutoff function $\Phi(y) = \phi^{\otimes d}(y)$ where ϕ is a smooth bump function supported on an interval of the form $[-2C, 2C]$ and identically 1 on the middle half of it. Then,

$$\|\chi_+(\omega^\eta - \omega^\varepsilon)\|_{U^3} = \|\Phi_+(y) \int_t (\psi(\eta t)) - \psi(\varepsilon t)) e^{i(|y|^p - 1)t} dt||_{U^3(dy)}, \tag{4.7}$$

with $Phi_+(y) = \chi_+(y)\Phi(y)$

Applying Minkowski's inequality to the right hand side of the last equation (4.7) this is further estimated by

$$\int_t |\psi(\eta t)) - \psi(\varepsilon t)| \, ||\Phi_+(y)e^{it\,|y|^p}||_{U^3(\mathbb{R}^d)} dt. \tag{4.8}$$

Note that as $\Phi_+(y)e^{it\,|y|^p} = \prod_{i=1}^d \phi_+(y_i)e^{it\,y_i^p}$, we have

$$||\Phi_+(y)e^{it\,|y|^p}||_{U^3(\mathbb{R}^d)} \;=\; ||\phi_+(y)e^{it\,y^p}||^d_{U^3(\mathbb{R})},$$

where the one-dimensional integrals

$$||\phi_+(y)e^{it\,y^p}||^8_{U^3(\mathbb{R})} = \int_{y,h,k,l\in\mathbb{R}} \left(\phi_+(y)e^{it|y|^p}\right) dy\,dh\,dk\,dl$$

are estimated in Lemma 4.3. Thus, we have for $|t| \geq 1$

$$||\Phi_+(y)e^{it\,|y|^p}||_{U^3(y)} \;\lesssim\; |t|^{-\frac{d}{8r}},$$

with $r = r(p)$ given in (4.5). Inserting this bound into (4.8), we complete the estimation as follows,

$$\begin{aligned}\int_t |\psi(\eta t)) - \psi(\varepsilon t)| \, |t|^{-\frac{d}{8r}}\, dt &\leq \int \big[|\psi(\eta t)| + |\psi(\epsilon t)|\big] t^{-d/8r}\, dt \\ &\lesssim \eta^{\frac{d}{8r}-1} + \varepsilon^{\frac{d}{8r}-1} \lesssim \epsilon^{\frac{d}{8r}-1}\end{aligned}$$

which is bounded uniformly in η provided that $\eta \ll \epsilon$ and$d > 8r$. □

5. A result from time-frequency analysis.

Here we will prove Proposition 2.3 by using the main result of [8]. The necessary verifications of the hypotheses of [8] will be done subsequently.

Proof of Proposition 2.3. By the Cauchy-Schwarz inequality and support restrictions on f,

$$\begin{aligned}|\mathcal{E}_{\lambda_j}(f)|^2 &\lesssim N^d \iiint f(x+y)f(x+z)f(x+2y)f(x+2z)\, k_j^\varepsilon(y)k_j^\varepsilon(z)\, dydzdx \\ &= N^d \iiint f(x)f(x+z-y)f(x+y)f(x+2z-y)\, k_j^\varepsilon(y)k_j^\varepsilon(z)\, dydzdx.\end{aligned} \tag{5.1}$$

Summing (5.1) for $1 \leq j \leq J$ one has

$$\sum_{j=1}^J |\mathcal{E}_{\lambda_j}(f)|^2 \;\lesssim\; N^d \int f(x)\, \mathcal{K}_J^\varepsilon(f,f,f)(x)\, dx, \tag{5.2}$$

where the trilinear operator $\mathcal{K}_J^\epsilon$ on the right side is given by

$$\mathcal{K}_J^\varepsilon(f_1,f_2,f_3)(x) := \int\int f_1(x+z-y)f_2(x+y)f_3(x+2z-y)K_J^\varepsilon(y,z)\, dy\, dz, \tag{5.3}$$

with integration kernel

$$K_J^\varepsilon(y,z) = \sum_{j=1}^J k_j^\varepsilon(y)k_j^\varepsilon(z), \qquad \text{where} \qquad k_j^\epsilon(y) = \omega_{\lambda_j}(y) - c_1(\epsilon)\omega_{\lambda_j}(y). \tag{5.4}$$

We will justify in Lemma 5.1 below that this is a Calderón-Zygmund kernel, L^p mapping properties of associated operators of the type given in (5.3) are studied by Muscalu, Tao and Thiele [8]. Indeed, by Theorem 1.1 in [8] one has the estimate

$$||\mathcal{K}_J^\varepsilon(f,f,f)||_{L^{4/3}(\mathbb{R}^d)} \;\leq\; C_\varepsilon\, \|f\|^3_{L^4(\mathbb{R}^d)}, \tag{5.5}$$

with a constant $C_\varepsilon > 0$ independent of J and the sequence $\lambda_1 < \ldots < \lambda_J$. Then

$$\begin{aligned}\sum_{j=1}^{J} |\mathcal{E}_{\lambda_j}(f)|^2 &\lesssim N^d \int f(x)\mathcal{K}_J^\varepsilon(f,f,f)(x)\,dx \\ &\leq C_\varepsilon\, N^d\, \|\mathcal{K}_J^\epsilon(f,f,f)\|_{L^{4/3}(\mathbb{R}^d)}\, \|f\|_{L^4(\mathbb{R}^d)} \leq C_\epsilon N^d ||f||_4^4, \end{aligned} \tag{5.6}$$

as claimed. □

Begin by rewriting (5.6) as

$$\mathcal{K}_J^\varepsilon(f_1,f_2,f_3)(x) := \iiint_{(\mathbb{R}^d)^3} e^{ix\cdot(\xi_1+\xi_2+\xi_3)}\widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)m(\xi_1,\xi_2,\xi_3)\,d\xi_1\,d\xi_2\,d\xi_3, \tag{5.7}$$

where

$$\begin{aligned} m(\xi_1,\xi_2,\xi_3) &= \iint_{(\mathbb{R}^d)^2} K_J^\varepsilon(y,z)e^{-iy\cdot(\xi_1-\xi_2+\xi_3)}e^{ix\cdot(\xi_1+2\xi_3)}dy\,dz \\ &= \widehat{K_J^\varepsilon}(-\xi_1+\xi_2-\xi_3,\xi_1+2\xi_3) \end{aligned} \tag{5.8}$$

Set

$$\Gamma = \{(\xi_1,\xi_2,\xi_3,\xi_4) \in (\mathbb{R}^d)^4 : \xi_1+\xi_2+\xi_3+\xi_4 = 0\}, \text{ and} \tag{5.9}$$

$$\Gamma' = \{(\xi_1,\xi_2,\xi_3,\xi_4) \subset \Gamma : \xi_1-\xi_2+\xi_3 = 0,\ \xi_1+2\xi_3 = 0, \tag{5.10}$$

so that $\dim(\Gamma) = 3d$ and $\dim(\Gamma') = d$. By taking the Fourier transform one may re-express $\mathcal{K}_J^\epsilon$ as a multiplier, namely,

$$\widehat{\mathcal{K}_J^\varepsilon}(f_1,f_2,f_3)(-\xi_4) := \int_\Gamma \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)m(\xi_1,\xi_2,\xi_3,\xi_4)\,d\xi_1\,d\xi_2\,d\xi_3, \tag{5.11}$$

identifying the operator $\mathcal{K}_J^\varepsilon$ with the ones studied in [8]. The only difference is that the functions f_i are now defined on $\mathbb{R}^d$ instead of on $\mathbb{R}$; however this does not affect the arguments given there. To clarify, Γ' is a graph over d of the canonical variables, and $\mathcal{K}_J^\epsilon$ generates a 4-linear form, so in the notation of [3], $k = d$ and $n = 4$. Thus the rank of the operator as described in this paper is $m = k/d = 1$, an integer $< 2 = n/2$. As has been pointed out in page 295 of [3], the higher-dimensional adaptation of the main result of [8] in this setting of integral rank is fairly straightforward, provided certain requirements on the multiplier is met. Thus, in order to apply the main result of [8] one needs to establish certain growth and differentiability properties of the multiplier $m(\xi)$. This is the goal of the following lemma.

Lemma 5.1. *Given an integration kernel K_J as in* (5.4)*, with the summands obeying the cancellation condition* (2.16)*, let $m(\xi)$ be the associated multiplier defined in* (5.8)*, where $\xi = (\xi_1,\ldots,\xi_3) \in \mathbb{R}^{3d}$ is a coordinate system identifying the subspace Γ. Then for any multi-index $\alpha \in \mathbb{Z}^{3d}_{\geq 0}$ one has the estimate*

$$|\partial_\xi^\alpha m(\xi)| \leq C_{\alpha,\varepsilon}\,(dist(\xi,\Gamma'))^{-|\alpha|}. \tag{5.12}$$

Here Γ and Γ' are as in (5.9) *and* (5.10).

Remark: The crucial point here is that the constant $C_{\alpha,\varepsilon}$ is independent of J and the lacunary sequence $\lambda_1 < \ldots < \lambda_J$. Also the above estimate is needed just up to some fixed finite order. Once this is established our main result Theorem 2.2 follows as explained at the end of Section 2.

Proof. Since m is essentially $\widehat{K_J^\epsilon}$ composed with a linear transformation, we study the latter function in detail. The relation (5.4) implies that

$$\widehat{K_J^\varepsilon}(\eta,\zeta) = \sum_{j=1}^{J} \widehat{k_j^\varepsilon}(\eta)\widehat{k_j^\varepsilon}(\zeta)$$

and

$$\widehat{k_j^\varepsilon}(\eta) = \int_{y\in\mathbb{R}^d} e^{iy\cdot\eta}\left(\omega_{\lambda_j}^\varepsilon(y) - c_1(\varepsilon)\omega_{\lambda_j}(y)\right)\,dy = \widehat{\omega_1^\varepsilon}(\lambda_j\eta) - c_1(\varepsilon)\widehat{\omega}(\lambda_j\eta).$$

Let us recall the definitions $\omega_1^\varepsilon(y) = \varepsilon^{-1}\widehat{\psi}((|y|^p - 1)/\varepsilon)$, where $\omega(y) = \omega_1^1(y)$ with $\widehat{\psi}$ is a compactly supported smooth function. Therefore for all multi-indices α

$$|\partial_\eta^\alpha \widehat{\omega_1^\varepsilon}(\eta)| \le C_\alpha. \tag{5.13}$$

Integrating by parts k times in the integral expression for $\partial_\eta^\alpha \omega_1^\epsilon$ we also obtain

$$|\partial_\eta^\alpha\, \widehat{\omega_1^\varepsilon}(\eta)| \le C_{\alpha,k}\,\varepsilon^{-|\alpha|}\,(1+|\eta|)^{-k}. \tag{5.14}$$

Thus for any $k\in\mathbb{N}$

$$|\partial_\eta^\alpha\, \widehat{k_j^\varepsilon}(\eta)| \le C_{\alpha,k}\,\varepsilon^{-|\alpha|}\,(1+|\lambda_j\eta|)^{-k}. \tag{5.15}$$

The cancellation property (2.16) gives that $\widehat{k_j^\varepsilon}(0) = 0$, which leads to the additional pointwise estimate

$$|\widehat{k_j^\varepsilon}(\eta)| \le C\,\lambda_j|\eta|. \tag{5.16}$$

This implies that $\widehat{K_J^\varepsilon}(0,\zeta) = \widehat{K_J^\varepsilon}(\eta,0) = 0$ and

$$|\widehat{K_J^\varepsilon}(\eta,\zeta)| \lesssim \sum_{j\le J} \min\left(\lambda_j|\eta|, \frac{1}{\lambda_j|\eta|}\right) \lesssim 1 \tag{5.17}$$

as the sequence $\mu_j := \lambda_j|\eta|$ is lacunary and $\|\widehat{k_j^\varepsilon}\|_\infty \lesssim 1$. This shows that the multiplier $m(\xi)$ defined in (5.11) is bounded. To estimate its partial derivatives we apply (5.15) with $k = |\alpha|+|\beta|+1$ to write

$$\begin{aligned}
|\partial_\eta^\alpha\partial_\zeta^\beta\;\widehat{K_J^\varepsilon}(\eta,\zeta)| &\le C_{\alpha,\beta,\varepsilon}\sum_{j\le J}\lambda_j^{|\alpha|+|\beta|}\,\min\left((1+|\lambda_j\eta|)^{-|\alpha|-|\beta|-1}, (1+|\lambda_j\zeta|)^{-|\alpha|-|\beta|-1}\right)\\
&\le C_{\alpha,\beta,\varepsilon}\,\min\left(|\eta|^{-|\alpha|-|\beta|}, |\zeta|^{-|\alpha|-|\beta|}\right) \le C_{\alpha,\beta,\varepsilon}\,(|\eta|+|\zeta|)^{-|\alpha|-|\beta|}.
\end{aligned}$$

Here we have used the fact that

$$\sum_{j\le J}\mu_j^k(1+\mu_j)^{-k-1} \le \sum_{j:\,\mu_j\le 1}\mu_j + \sum_{j:\,\mu_j\ge 1}\mu_j^{-1} \le C,$$

for the lacunary sequences $\mu_j := \lambda_j|\eta|$ and $\mu_j := \lambda_j|\zeta|$ with $k = |\alpha|+|\beta| \ge 1$. By (5.8) this leads to the estimate

$$|\partial_\xi^\alpha\, m(\xi)| \le C_{\alpha,\varepsilon}\,(|-\xi_1+\xi_2-\xi_3|^{-|\alpha|} + |\xi_1+2\xi_3|^{-|\alpha|}) \le C'_{\alpha,\varepsilon}\,dist(\xi,\Gamma')^{-|\alpha|}.$$

for $\xi = (\xi_1,\xi_2,\xi_3)\in\mathbb{R}^{3d}\simeq\Gamma$ and any multi-index $\alpha\in(\mathbb{Z}_+)^{3d}$. This proves Lemma 5.1. □

References

[1] J. Bourgain, *A Szemerédi type theorem for sets of positive density in* $\mathbb{R}^k$. Israel J. Math., 54(3), (1986), 307-316.

[2] B. Cook, Á Magyar, *On restricted arithmetic progressions over finite fields*, Online J. Anal. Comb. 7 (2012), 1-10

[3] C. Demeter, M. Pramanik, C. Thiele, *Multilinear singular operators with fractional rank*, Pacific J. Math. 246 (2010), no. 2, 293324.

[4] K. Falconer, J. Marstrand, *Plane sets with positive density at infinity contain all large distances*, Bull. London Math. Soc. 18 (1986), no. 5, 471474.

[5] H. Furstenberg, Y. Katznelson, B. Weiss. *Ergodic theory and configurations in sets of positive density*, Mathematics of Ramsey theory. Springer Berlin Heidelberg, (1990) 184-198.
[6] W.T. Gowers. *A new proof of Szemerdi's theorem for arithmetic progressions of length four*, Geometric and Functional Analysis, 8(3), (1998) 529-551.
[7] L. Hörmander *The Analysis of Linear Partial Differential Operators. Vol. I: Distribution Theory and Fourier Analysis*, Springer Berlin-Heidelberg (1983)
[8] C. Muscalu, T. Tao, C. Thiele, *Multi-linear operators given by singular multipliers*, Journal of the AMS, 15.2 (2002), 469-496.
[9] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28, (1953). 104109.
[10] T. Ziegler, *Nilfactors of* $\mathbb{R}^m$*-actions and configurations in sets of positive upper density in* $\mathbb{R}^m$., J. Anal. Math. 99 (2006), 249-266.

Brian Cook
University of Wisconsin, Madison, USA.
Electronic address: bcook@math.wisc.edu

Akos Magyar
University of Athens, Georgia, USA.
Electronic address: magyar@math.uga.edu

Malabika Pramanik
University of British Columbia, Vancouver, Canada.
Electronic address: malabika@math.ubc.ca