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

BRIAN COOK  ÁKOS MAGYAR  MALABIKA PRAMANIK

Abstract. Let $A \subseteq \mathbb{R}^d$ with $d > 8$ be a measurable set of positive upper density. We prove that 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 $|y|_4 = \lambda$, where $|y|_4 = (\sum_i y^4_i)^{1/4}$ is the $l^4$-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 arbitrary large gaps when the gap size is measured by the $l^4$-metric. This is known to be false in the ordinary $l^2$-metric and one of the goals of this note is to understand this phenomenon.

1. Introduction.

Euclidean Ramsey theory studies geometric configurations in large but otherwise arbitrary measurable sets. It was shown by Furstenberg, Katznelson and Weiss [4] and independently by Bourgain [1] that a set $A \subseteq \mathbb{R}^d$ of positive upper Banach density contains all large distances; that is for every sufficiently large $\lambda \geq \lambda_0(A)$ there are points $x,x+y \in A$ such that $|y| = \lambda$. Recall that $A$ has positive upper Banach density if

$$\bar{\delta}(A) := \limsup_{N \to \infty} \sup_{x \in \mathbb{R}^d} \frac{|A \cap (x + [0,N]^d)|}{N^d} > 0.$$\n
Roth’ theorem [6] states that a subset of $\mathbb{Z}$ of positive upper density contains a 3-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 3-term progressions whose gaps can be arbitrary large. However a simple example given in [1] shows that there is a set $A \subseteq \mathbb{R}^d$ in any dimension $d$, such that the gaps of $|y|$ of all 3-term progressions in $A$ do not contain all large numbers. Indeed, take $A$ to be the set of points $x \in \mathbb{R}^d$ such that $|x|^2 - m \leq \frac{1}{10}$ for some $m \in \mathbb{N}$. Then by the parallelogram identity $||y|^2 - l| \leq \frac{4}{10}$ (for some $l \in \mathbb{N}$) for any progression $\{x,x+y,x+2y\} \subseteq A$. This does not exclude the validity of such a result when the gaps are measure via some other metric on $\mathbb{R}^d$, which does not satisfy the parallelogram identity. The aim of this note is to show that this is indeed the case for the $l^4$ metric $|y|_4 = (\sum_{i=1}^d y_i^4)^{1/4}$. It is plausible that our arguments also work for metrics given by positive homogeneous polynomials of degree at least 4 however we do not pursue this here. A more general result of this type in the finite field setting was given by the first two authors in [2].

2. Main results.

We describe below the main elements of the proof the details will be given in subsequent sections. For the rest of the paper we will use $l^4$-norm and the associated metric on $\mathbb{R}^d$, and for simplicity of notation we will write $|y| = (\sum y^4_i)^{1/4}$. Our main result is the following

**Theorem 2.1.** Let $d > 8$ and let $A \subseteq \mathbb{R}^d$ be a measurable set such that $\bar{\delta}(A) > 0$. There exists $\lambda_0 = \lambda_0(A)$ such that for all $\lambda \geq \lambda_0$ there exists $x,y \in \mathbb{R}^d$ such that $\{x,x+y,x+2y\} \subseteq A$ and $|y| = \lambda$.

We will prove the following stronger finitary version for measurable sets $A \subseteq [0,N]^d$.

1991 Mathematics Subject Classification. 11D72, 11P32.
The second author is supported by grant ERC-AdG. 321104.
Theorem 2.2. Let \( d > 8, \delta > 0 \) and let \( N \geq N(\delta) \) be sufficiently large. Then for any measurable set \( A \subseteq [0, N]^d \) of measure \( m(A) \geq \delta N^d \) 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 3-term progression \( \{x, x+y, x+2y\} \subseteq A \) such that \( |y| = \lambda_j \) for some \( 1 \leq j \leq J \),

It is clear that Theorem 2.2 implies Theorem 2.1. Indeed, assume that Theorem 2.1 does not hold. Then set \( \delta := \delta(A)/2 \) there exists an infinite lacunary sequence \( \{\lambda_j\}_{j=1}^{\infty} \subseteq \mathbb{N} \), \( \lambda_{j+1} \geq 2\lambda_j \) such that \( \lambda_j \neq |y| \) for any \( j \) and any \( y \) which is the gap of a 3-term progression contained in \( A \). Fix a sufficiently large \( J = J(\delta) \) and a sufficiently large box \( B_N \) of size \( N = N(\delta, \lambda_J) \) on which the density of \( A \) is \( m(A)/N^d \geq \delta \). By translation invariance we may assume \( B_N = [0, N]^d \) contradicting Theorem 2.2.

We start by counting 3-term progressions \( x, x+y, x+2y \) contained in \( A \) via a positive measure \( \sigma_\lambda \) supported on the \( l^1 \)-sphere \( S_\lambda = \{y \in \mathbb{R}^d; |y| = \lambda\} \). Let \( f := 1_A \) be the indicator function of a measurable set \( A \subseteq [0, N]^d \). Let

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

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

\[
\sigma_\lambda(y) := \lambda^{-d+4} \int_{\mathbb{R}} e^{i(|y|^4 - \lambda^4)t} \, dt. \tag{2.1}
\]

It is well-known [5], Ch.2 that the above oscillatory integral defines an absolute continuous measure with respect to the surface area measure on \( S_\lambda \) whose density function is \( |\nabla Q(y)|^{-1} \) with \( Q(y) = |y|^4 \). The normalizing factor \( \lambda^{-d+4} \) is inserted to have \( \sigma_\lambda(S_\lambda) = \sigma_1(S_1) > 0 \), which is independent of \( \lambda \).

Let \( 0 \leq \psi \leq 1, \psi(0) = 1 \) be a Schwarz function such that its Fourier transform \( \widehat{\psi} \geq 0 \) and is compactly supported. Define the quantity

\[
\omega_\lambda(y) := \lambda^{-d+4} \int_{\mathbb{R}} e^{it(|y|^4 - \lambda^4)} \psi(\lambda^4 t) \, dt.
\]

Note that by scaling

\[
\omega_\lambda(y) = \lambda^{-d} \omega(y/\lambda) = \lambda^{-d} \widehat{\psi}(|y/\lambda|^4 - 1),
\]

hence

\[
\int \omega_\lambda(y) \, dy = \int \omega_1(y) \, dy = C_\omega > 0.
\]

Define

\[
\mathcal{M}_\lambda(f) := \int_{\mathbb{R}^d} \int_{S_\lambda} f(x)f(x+y)f(x+2y)\omega_\lambda(y) \, dy \, dx,
\]

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

Proposition 2.1. Let \( 0 < \delta \leq 1 \) and let \( f : [0, N]^d \to [0, 1] \) be such that \( \int f \geq \delta \).

Then there exists a constant \( c(\delta) > 0 \) depending only on \( \delta \) such that for \( 0 < \lambda \ll N \),

\[
\mathcal{M}_\lambda(f) \geq \frac{1}{c(\delta)} N^d.
\]

Next, we define a quantity corresponding to a small \( \varepsilon > 0 \) which is a good approximation to \( \mathcal{N}_\lambda(f) \). Let

\[
\omega_\lambda^\varepsilon(y) := \lambda^{-d+4} \int_{\mathbb{R}} e^{it(|y|^4 - \lambda^4)} \psi(\varepsilon \lambda^4 t) \, dt.
\]
It is easy to see that
\[
\omega_\lambda^\varepsilon(y) = \lambda^{-d}\varepsilon^{-1}\lambda^{-1}\left(\frac{|y/\lambda|^4 - 1}{\varepsilon}\right) = \lambda^{-d}\omega_1^\varepsilon(y/\lambda).
\]
(2.2)

Define
\[
\mathcal{M}_\lambda^\varepsilon(f) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(x+y)f(x+2y)\omega_\lambda^\varepsilon(y) \, dx \, dy.
\]

We have the error estimate

**Proposition 2.2.** Let \( f : [0,N]^d \to [-1,1] \) and let \( d > 8, \ 0 < \varepsilon < 1 \). Then for \( 0 < \lambda \ll N \) one has
\[
|\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\varepsilon(f)| \lesssim \varepsilon^{d-1} N^d.
\]

Note that right side is independent of \( \lambda \). The proof of Proposition 2.2 is based on two facts. The first is an inequality showing that the so-called \( U^3 \)-uniformity norm of Gowers [3] controls expressions like \( \mathcal{N}_\lambda(f) \). Let is recall its definition for a bounded measurable function \( g : [-N,N]^d \to \mathbb{C} \),
\[
\|g\|_{U^3}^8 = \int_{x,y_1,y_2,y_3} \prod_{\nu_1,\nu_2,\nu_3=0,1} \tilde{g}''(x + \nu_1 y_1 + \nu_2 y_2 + \nu_3 y_3) \, dx \, dy_1 \, dy_2 \, dy_3,
\]
(2.3)

where \( \nu_1,\nu_2,\nu_3 \) can take the values 0 or 1, \( \tilde{g}'' = \tilde{g} \) if \( \nu_1 + \nu_2 + \nu_3 \) is odd and \( \tilde{g}'' = g \) otherwise.

**Lemma 2.1.** Let \( f : [0,N]^d \to [-1,1] \) and let \( d > 8, \ 0 < \varepsilon < 1 \). Then for \( 0 < \lambda \ll N \) one has
\[
|\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\varepsilon(f)| \lesssim N^d \|\sigma - \omega_1^\varepsilon\|_{U^3}.
\]

It is not immediately clear how to define the \( U^3 \)-norm of the measure \( \sigma \) defined in (2.1). Note that \( \omega_1^\varepsilon \to \sigma \) weakly as \( \varepsilon \to 0 \), and we will show that \( \omega_1^\varepsilon \) form a Cauchy sequence with respect to the \( U^3 \)-norm and define \( \|\sigma\|_{U^3} := \lim_{\varepsilon \to 0} \|\omega_1^\varepsilon\|_{U^3} \). In fact we’ll show

**Lemma 2.2.** Let \( d > 8, \ 0 < \varepsilon < 1 \). One has
\[
\|\sigma - \omega_1^\varepsilon\|_{U^3} \lesssim \varepsilon^{d-1}.
\]
(2.4)

The proof uses in an essential way the metric is defined by the degree 4 expression \( |y|^4 = \sum_i y_i^4 \), estimate (2.4) does not hold for the Euclidean metric.

The final ingredient to prove Theorem 2.2 is an estimate given in [7] for certain multilinear operators similar to the bilinear Hilbert transform. To see how it comes into the picture, assume that Theorem 2.2 does not hold for a lacunary sequence \( \lambda_1 < \lambda_2 < \ldots < \lambda_J \ll N \), that is \( \mathcal{N}_{\lambda_j}(f) = 0 \) for \( f = 1_A \) for all \( 1 \leq j \leq J \). For \( 0 < \varepsilon \ll 1 \) define the quantity \( c_1(\varepsilon) \) so that
\[
c_1(\varepsilon) \int \omega(y) \, dy = \int \omega_1^\varepsilon(y) \, dy.
\]
(2.5)

It is easy to see that \( c_1(\varepsilon) \approx 1 \), i.e. is bounded by two positive constants depending only on the dimension \( d \). Then for a sufficiently small \( \varepsilon = \varepsilon(\delta) \) we have by Propositions 2.1 and 2.2
\[
|\mathcal{M}_{\lambda_j}^\varepsilon(f) - c_1(\varepsilon)\mathcal{M}_{\lambda_j}(f)| \gtrsim c(\delta) N^d,
\]
(2.6)

for all \( 1 \leq j \leq J \) as long as \( \varepsilon = \varepsilon(\delta) \) is chosen sufficiently small. Writing \( k_j^\varepsilon(y) := \omega_{\lambda_j}^\varepsilon(y) - c_1(\varepsilon)\omega_{\lambda_j}(y) \), we have that
\[
\mathcal{E}_{\lambda_j}(f) := \mathcal{M}_{\lambda_j}^\varepsilon(f) - c_1(\varepsilon)\mathcal{M}_{\lambda_j}(f) = \int \int f(x)f(x+y)f(x+2y)k_j^\varepsilon(y) \, dy \, dx.
\]
Moreover, by (2.5) one has the cancelation property
\[
\int k_j^\epsilon(y) \, dy = \int (\omega^\epsilon_{\lambda_j} - c_1(\epsilon) \omega_{\lambda_j}) \, dy = \int (\omega^\epsilon_1 - c_1(\epsilon) \omega) \, dy = 0. \tag{2.7}
\]
Thus by the Cauchy-Schwarz inequality
\[
|\mathcal{E}_{\lambda_j}(f)|^2 \lesssim N^d \int \int \int f(x+y) f(x+z) f(x+2y) f(x+2z) k_j^\epsilon(y) k_j^\epsilon(z) \, dy \, dz \, dx
\]
\[
= N^d \int \int \int f(x) f(x+z-y) f(x+y) f(x+2z-y) k_j^\epsilon(y) k_j^\epsilon(z) \, dy \, dz \, dx. \tag{2.8}
\]
Summing (2.8) for \(1 \leq j \leq J\) one has
\[
\sum_{j=1}^J |\mathcal{M}_{\lambda_j}^\epsilon(f) - \mathcal{M}_{\lambda_j}(f)|^2 \lesssim N^d \int f(x) \mathcal{K}_j^\epsilon(f, f, f)(x) \, dx,
\]
where the trilinear operator on the right side is given by
\[
\mathcal{K}_j^\epsilon(f_1, f_2, f_3)(x) := \int \int f_1(x+z-y) f_2(x+y) f_3(x+2z-y) K_j^\epsilon(y, z) \, dy \, dz, \tag{2.9}
\]
with kernel
\[
K_j^\epsilon(y, z) = \sum_{j=1}^J k_j^\epsilon(y) k_j^\epsilon(z).
\]
This is a Calderon-Zygmund kernel, \(L^p\) mapping properties of associated operators of the type given in (2.9) are studied by Muscalu, Tao and Thiele [7]. Indeed, by Theorem 1.1 in [7] one has the estimate
\[
\|\mathcal{K}_j^\epsilon(f, f, f)\|_{L^{4/3}(\mathbb{R}^d)} \leq C_\epsilon \|f\|_{L^4(\mathbb{R}^d)},
\]
with a constant \(C_\epsilon > 0\) independent of \(J\) and the sequence \(\lambda_1 < \ldots < \lambda_J\). Then
\[
\sum_{j=1}^J |\mathcal{M}_{\lambda_j}^\epsilon(f) - \mathcal{M}_{\lambda_j}(f)|^2 \lesssim N^d \int f(x) \mathcal{K}_j^\epsilon(f, f, f)(x) \, dx
\]
\[
\leq C_\epsilon N^d \|f\|_{L^{4/3}(\mathbb{R}^d)} \|f\|_{L^4(\mathbb{R}^d)} \leq C_\epsilon N^{2d}. \tag{2.10}
\]
However by (2.6) we have that the left side of (2.10) is at least \(c(\delta)^2 J N^{2d}\), thus choosing \(J = J(\epsilon, \delta)\) sufficiently large we get a contradiction, and Theorem 2.1 follows.

3. The main term.

In this section we prove Proposition 2.1, which proven by 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 \(\mathbb{T}^d\).

First we fix \(\delta \in (0, 1]\) and \(\lambda \ll N\) for a sufficiently large \(N\). Next consider a given real valued function \(f : [0, N]^d \to [0, 1]\) with \(\int f \geq 3N^d\). Equipartition the cube \([0, N]^d\) in the natural way into disjoint boxes of side length \(l = \epsilon \lambda\) by choosing a sufficiently small number \(c > 0\) so that \(N/l\) is a positive integer. Enumerate the boxes by \((B_i)_{i=1}^N\) in some fashion. Each box \(B_i\) may then be identified with its leftmost endpoint \(x_i\), so that \(B_i = [0, l]^n + x_i\).
Now define \(g_i(x) : [0, l]^d \to [0, 1]\) by \(g_i(x) = 1_{[0, l]^n}(x + x_i)f(x + x_i)\). We have the bound
\[
M_\lambda(f) \geq \sum_{i=1}^L \int_{x \in \mathbb{R}^d} \int_{y \in \mathbb{R}^d} g_i(x)g_i(x + y)g_i(x + 2y) \omega(y) \, dy \, dx. \tag{3.1}
\]

Recall that
\[
\omega(y) = \lambda^{-d} \hat{\psi}(|y/\lambda|^4 - 1),
\]
and hence we may thus assume that \(\psi\) is chosen such that \(\hat{\psi}(|y/\lambda|^4 - 1) \geq 1/10\) for \(y \in [-l, l]^d\). Then (3.1) yields
\[
M_\lambda(f) \geq \lambda^{-d} \sum_{i=1}^L \int_{x \in \mathbb{R}^d} \int_{y \in \mathbb{R}^d} g_i(x)g_i(x + y)g_i(x + 2y) \, dy \, dx. \tag{3.2}
\]

Now identify \(\Pi^d\) with the cube \([-1/2, 1/2]^d\). The summands on the right hand side of (3.2) may be written as
\[
(10l)^{2d} \int_{x \in \mathbb{R}^d} \int_{y \in \mathbb{R}^d} g_i(10lx)g_i(10l(x + y))g_i(10l(x + 2y)) \, dy \, dx
\]
\[
= (10l)^{2d} \int_{x \in \Pi^d} \int_{y \in \Pi^d} g_i(10lx)g_i(10l(x + y))g_i(10l(x + 2y)) \, dy \, dx.
\]

Here we have chosen the constant \(c\) so that the integrand is supported on \(\Pi^d\), which can clearly simultaneously be done. If we also have that \(\int_{\Pi^d} g_i(10lx) \, dx \geq \eta > 0\), then by Roth’s Theorem on Compact Abelian Groups we have
\[
\int_{\Pi^d} \int_{\Pi^d} g_i(10lx)g_i(10l(x + y))g_i(10l(x + 2y)) \, dx \, dy > c(\eta),
\]
where \(c(\eta) > 0\) is a constant depending only on \(\eta\).

To complete the proof of Proposition 2.1 we simply need enough of the \(g_i\) with sufficient density.

**Lemma 3.1.** For at least \((\delta/4)L\) of the \(i = 1, \ldots, L\) we have that \(\int g_i \geq (\delta/4)|B_i|\).

**Proof.** Let \(I\) denote the number of indices \(i\) with \(\int g_i \geq (\delta/4)|B_i|\). We have
\[
\sum_{i=1}^L \int g_i \geq (\delta/2)N^d
\]
and also
\[
\sum_{i=1}^L \int g_i \leq (\delta/4)N^d + I N^d / L,
\]
as \(|B_i| = N^d / L\) for each \(1 \leq i \leq L\). \(\square\)

4. **Error estimates.**

We turn to the proof of Proposition 2.2 and recall this is to follow from Lemma 2.1 and Lemma 2.2. To prove Lemma 2.1 we use Lemma 2.2 in tandem with the following result.

**Lemma 4.1.** Let \(f : [0, N]^d \to [-1, 1]\) and \(g : [0, \lambda]^d \to [-1, 1]\) be given functions. Then one has
\[
\int_{x,y \in \mathbb{R}^d} f(x) f(x + y) f(x + 2y) g(y) \, dx \, dy \lesssim N^d \lambda^{d/2} \|g\|_{L^3}.
\]
Proof. Set
\[ T = \int_{x, y \in \mathbb{R}^d} f(x)f(x + y)f(x + 2y)g(y) \, dx \, dy, \]
and begin by applying the Cauchy-Schwarz inequality in the $x$ integration to get
\[ T^2 \leq N^d \int_x \int_y f(x + y) \int_y f(x + 2y) \int_y f(x + y') f(x + 2y') g(y) g(y') \, dx \, dy \, dy' \]
Use the substitution $y' = y + h$ followed by the substitution $x \to x - y$, and define $\Delta_h F(x) = F(x + h)F(x)$ for a generic function $F$ to rewrite this as
\[ T^2 \leq N^d \int_x \int_y \Delta_h f(x) \Delta_2h f(x + y) \Delta_h g(y) \, dx \, dy \, dh. \]
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
\[ T^4 \lesssim N^3 \lambda^d \int_{x, y, h, y'} \Delta_2h f(x + y) \Delta_2h f(x + y') \Delta_h g(y) \Delta_h g(y') \, dx \, dy \, dh \, dy' \]
Again use the substitutions $y' = y + k$ and $x \to x - y$ in turn to get
\[ T^4 \lesssim N^3 \lambda^d \int_{x, y, k, h} \Delta_2h f(x) \Delta_2h f(x + k) \Delta_h g(y) \Delta_h g(y + k) \, dx \, dy \, dh \, dk. \]
One final application of the Cauchy-Schwarz inequality in $x$ and $h$ and $k$ integration gives
\[ T^8 \lesssim N^7 \lambda^d \int_{x, y, h, k, y'} 1_{[0,N]^3}(x) \Delta_h g(y) \Delta_h g(y') \Delta_h g(y + k) \Delta_h g(y' + k) \, dx \, dy \, dy' \, dh \, dk. \]
The $x$ integration may be carried out, and the applying the substitution $y' \to y + l$ gives the final form
\[ T^8 \lesssim N^8 \lambda^d \int_{y, h, k, l} \Delta_{h, k, l} f(y) \, dy \, dh \, dk \, dl \quad (4.1) \]
where $\Delta_{h, k, l}$ is well defined as the composition of the operators $\Delta_h, \Delta_k,$ and $\Delta_l$. The integral in (4.1) is easily verified to be $\|g\|_{U^3}^8$ as defined in (2.3), and this completes the proof. \qed

Next we turn to the proof of Lemma 2.2. In what follows we assume that $\lambda$ and $N \gg \lambda$ are fixed, and $f$ is the characteristic function of a set $A \subset [0, N]^d$ with measure $\delta N^d$.

Proof. (Lemma 2.2)
Let $\eta \ll \varepsilon$ be a small parameter. We need to estimate
\[ \|\psi^{\eta} - \psi^{\varepsilon}\|_{U^3} \]
independent of $\eta$. By (2.2), we can find a smooth cutoff function $\Phi(y) = \varphi^{\otimes d}(y)$ where $\varphi$ is a smooth bump function supported on an interval of the form $[-C, C]$ for some constant $C > 0$, such that
\begin{align*}
\|\psi^{\eta} - \psi^{\varepsilon}\|_{U^3} &= \|\lambda^{-d+4}\Phi(y/\lambda) \int_t (\psi(\eta\lambda^4 t)) - \psi(\varepsilon\lambda^4 t))e^{i(|y|^4 - \lambda^4 t)}dt\|_{U^3(dy)} \\
&= \|\lambda^{-d}\Phi(y/\lambda) \int_t (\psi(\eta t)) - \psi(\varepsilon t))e^{i(|y/\lambda|^4 - 1)t}dt\|_{U^3(dy)} \\
&= \|\lambda^{-d/2}\Phi(y) \int_t (\psi(\eta t)) - \psi(\varepsilon t))e^{i(|y|^4 - 1)t}dt\|_{U^3(dy)}. \quad (4.2)
\end{align*}
The first equality is from the substitution $\lambda^4 t \to t$. The second follows from the fact that $\|F(y/\lambda)\|_{U^3(dy)} = \|\lambda^{d/2}F(y)\|_{U^3(dy)}$, which is obtained simply by scaling.
Apply Minkowski’s inequality to the right hand side of the last equality in (4.2) to get the upper bound
\[ \lambda^{-d/2} \int \left| \psi(\eta t) - \psi(\varepsilon t) \right| \left\| \Phi(y) e^{it|y|^4} \right\|_{U^3(y)} dt. \]

Note that
\[ \Delta_{h,k,l}(\Phi(y) e^{it|y|^4}) = (\Delta_{h,k,l} \Phi(y)) e^{24it \sum y_i h_i k_i l_i} \]
As \( \Phi(y) = \prod_{i=1}^d \varphi(y_i) \) is a tensor power we may consider then the one dimensional integrals
\[ \int_{y,h,k,l \in \mathbb{R}} (\Delta_{h,k,l} \varphi(y)) e^{24it y h k l} dy dh dl, \]
which is clearly bounded above by
\[ \int_{h,k,l \in \mathbb{R}} \int_{y \in \mathbb{R}} (\Delta_{h,k,l} \varphi(y)) e^{it y h k l} dh dk dl \bigg| dy = \int_{h,k,l \in \mathbb{R}} \left| \Delta_{h,k,l} \varphi(24t h k l) \right|. \]
We also have the bound
\[ \varphi(24thkl) \lesssim (1 + 24|thkl|)^{-2}. \]
Now we partition the \( h, k, l \) integration ranges by dyadic intervals. Then we have the upper bounds
\[ \sum_{i_1, i_2, i_3} (1 + |t|2^{i_1+i_2+i_3})^{-2} 2^{-i} \sum_{i_1, i_2, i_3} (1 + |t|2^m)^{-2} 2^m m^2 \]
\[ \lesssim \sum_{2^m \ll |t|} 2^m m^2 + |t|^{-1} \sum_{2^m \gtrsim |t|} m^2 \lesssim (\log t)^3 |t|^{-1}. \]
To finish the proof we need to bound
\[ \int_{t} \left| \psi(\eta t) - \psi(\varepsilon t) \right| |t|^{-d/8+v} dt. \]
independent of \( \eta \), where \( v \) is an small positive parameter. In fact this is bounded by
\[ \int_{|t|^2 \geq d} t^{-d/8+v} dt \lesssim \varepsilon^{1/9} \]
provided that \( d > 8 \) as \( |\psi(\eta t) - \psi(\varepsilon t)| \leq 1 - \psi(\varepsilon t). \)

The proof of Proposition 2.1 remains. We have
\[ N^\lambda(f) - M^\xi(f) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(x+y)(\sigma_\lambda(y) - \omega_\lambda(y)) dx dy. \]
This may be written as
\[ \lim_{\eta \to 0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(x+y)f(x+2y)(\omega_\lambda^\eta(y) - \omega_\lambda^\xi(y)) dx dy. \]
The function \( \omega_\lambda^\eta - \omega_\lambda^\xi \) is supported on \([-\lambda', \lambda']\) with \( \lambda' \approx \lambda \). Then by Lemma 4.1
\[ \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(x+y)f(x+2y)(\omega_\lambda^\eta(y) - \omega_\lambda^\xi(y)) dx dy \right| \lesssim N^d \lambda^{d/2} \left\| \omega_\lambda^\eta - \omega_\lambda^\xi \right\|_{U^3}. \]
Now by Lemma 2.2 we have
\[ \left\| \omega_\lambda^\eta - \omega_\lambda^\xi \right\|_{U^3} \lesssim \lambda^{-d/2} \]
the proof Proposition 2.1 easily follows.
5. A RESULT FROM TIME-FREQUENCY ANALYSIS.

Here we prove estimate (2.10) by using the main result of [7]. Begin by rewriting (2.9) as

\[ K_\varepsilon^j(f_1, f_2, f_3)(x) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{ix \cdot (\xi_1 + \xi_2 + \xi_3)} \hat{f}_1(\xi_1) \hat{f}_2(\xi_2) \hat{f}_3(\xi_3) m(\xi_1, \xi_2, \xi_3) d\xi_1 d\xi_2 d\xi_3, \]

where

\[ m(\xi_1, \xi_2, \xi_3) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_\varepsilon^j(y, z) e^{-iy \cdot (\xi_1 - \xi_2 + \xi_3)} e^{iz \cdot (\xi_1 + 2\xi_3)} dy dz = \hat{K}_\varepsilon^j(-\xi_1 + \xi_2 - \xi_3, \xi_1 + 2\xi_3) \]

(5.1)

Set \( \Gamma = \{(\xi_1, \xi_2, \xi_3, \xi_4) : \xi_1 + \xi_2 + \xi_3 + \xi_4 = 0\} \) and let \( \Gamma' \subset \Gamma \) be the subspace which satisfies the additional equations \( \xi_1 - \xi_2 + \xi_3 = 0 \) and \( \xi_1 + 2\xi_3 = 0 \). Note that \( dim(\Gamma) = 3d \) and \( dim(\Gamma') = d \).

By taking the Fourier transform one may further write

\[ \hat{K}_\varepsilon^j(f_1, f_2, f_3)(-\xi_4) := \int_{\Gamma} \hat{f}_1(\xi_1) \hat{f}_2(\xi_2) \hat{f}_3(\xi_4) m(\xi_1, \xi_2, \xi_3, \xi_4) d\xi_1 d\xi_2 d\xi_3, \]

(5.2)

identifying the operator \( K_\varepsilon^j \) with the ones studied in [7], the difference being that the functions \( f_i \) are now defined on \( \mathbb{R}^d \) instead of on \( \mathbb{R} \) which however does not affect the arguments given there.

To apply the main result of [7] one needs to establish certain properties of the multiplier \( m(\xi) \), namely the following estimate

**Lemma 5.1.** Let \( m(\xi) = m(\xi_1, \ldots, \xi_4) \) be the multiplier defined in (5.1), with \( \xi = (\xi_1, \ldots, \xi_4) \in \Gamma \). Then for any multi-index \( \alpha = (\alpha_1, \ldots, \alpha_4d) \) one has the estimate

\[ |\partial^{\alpha}_{\xi} m(\xi)| \leq C_{\alpha, \varepsilon} (\text{dist}(\xi, \Gamma'))^{-|\alpha|}. \]

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 as explained at the end of Section 2.

**Proof.** Observe that

\[ \hat{K}_\varepsilon^j(\eta, \zeta) = \sum_{j=1}^{J} \hat{k}_\varepsilon^j(\eta) \hat{k}_\varepsilon^j(\zeta) \]

and

\[ \hat{k}_\varepsilon^j(\eta) = \int_{y \in \mathbb{R}^d} e^{iy \cdot \eta} \left( \omega^\varepsilon_{\lambda_j}(y) - c_1(\varepsilon) \omega_{\lambda_j}(y) \right) dy = \tilde{\omega}_1^\varepsilon(\lambda_j \eta) - c_1(\varepsilon) \tilde{\omega}(\lambda_j \eta). \]

Recalling the definitions \( \omega^\varepsilon_1(y) = \varepsilon^{-1} \tilde{\psi}(|y|^4 - 1)/\varepsilon \) and \( \omega(y) = \omega^1(y) \) with \( \tilde{\psi} \) being a compactly supported smooth function we have for all multi-index \( \alpha \)

\[ |\partial^\alpha_{\eta} \tilde{\omega}_1^\varepsilon(\eta)| \leq C_\alpha, \]

and moreover by integrating by parts \( k \)-times one has

\[ |\partial^\alpha_{\eta} \tilde{\omega}_1^\varepsilon(\eta)| \leq C_{\alpha, k} \varepsilon^{-|\alpha|} (1 + |\eta|)^{-k}, \]

thus for any \( k \in \mathbb{N} \)

\[ |\partial^\alpha_{\eta} \hat{k}_\varepsilon^j(\eta)| \leq C_{\alpha, k} \varepsilon^{-|\alpha|} (1 + |\lambda_j \eta|)^{-k}. \]

(5.3)

By the cancelation property (2.7) we also have that \( \hat{k}_\varepsilon^j(0) = 0 \) and hence

\[ |\hat{k}_\varepsilon^j(\eta)| \leq C \lambda_j |\eta|. \]
This implies that $\hat{K}_j^\varepsilon(0, \zeta) = \hat{K}_j^\varepsilon(\eta, 0) = 0$ and

$$|\hat{K}_j^\varepsilon(\eta, \zeta)| \lesssim \sum_{j \leq J} \min \left( \lambda_j |\eta|, \frac{1}{\lambda_j |\eta|} \right) \lesssim 1$$

as the sequence $\mu_j := \lambda_j |\eta|$ is lacunary and $\|\hat{K}_j^\varepsilon\|_\infty \lesssim 1$. This shows that the multiplier $m(\xi)$ defined in (5.2) is bounded. To estimate its partial derivatives we apply (5.3) with $k = |\alpha| + |\beta| + 1$ to write

$$|\partial_\eta^\alpha \partial_\zeta^\beta \hat{K}_j^\varepsilon(\eta, \zeta)| \leq C_{\alpha, \beta, \varepsilon} \sum_{j \leq J} \lambda_j^{|\alpha|+|\beta|} \min \left( (1 + |\lambda_j \eta|)^{-|\alpha|-|\beta|-1}, (1 + |\lambda_j \zeta|)^{-|\alpha|-|\beta|-1} \right)$$

$$\leq C_{\alpha, \beta, \varepsilon} \min \left( |\eta|^{-|\alpha|-|\beta|}, |\zeta|^{-|\alpha|-|\beta|} \right) \leq C_{\alpha, \beta, \varepsilon} \left( |\eta| + |\zeta| \right)^{-|\alpha|-|\beta|}. \tag{5.4}$$

Here we have used the fact that

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

for the lacunary sequences $\mu_j := \lambda_j |\eta|$ and $\mu_j := \lambda_j |\zeta|$ with $k = |\alpha| + |\beta| \geq 1$.

By (5.1) this yields to the estimate for $\xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 \simeq \Gamma$ and any multi-index $\alpha \in (\mathbb{Z}_+)^{3d}$

$$|\partial^{\alpha} m(\xi)| \leq C_{\alpha, \varepsilon} \left( |\xi_1 + \xi_2 - \xi_3|^{-|\alpha|} + |\xi_1 + 2\xi_3|^{-|\alpha|} \right) \leq C_{\alpha, \varepsilon} \operatorname{dist}(\xi, \Gamma')^{-|\alpha|}.$$

This proves Lemma 5.1. □

References