On Triples in Arithmetic Progression

Jean Bourgain

An Exposition by Ben Green

November 1999

Contents

The following is an exposition of a brilliant paper [1] of Bourgain. In this paper he improves the known bounds in Roth's Theorem on arithmetic progressions (APs) of length 3. Whilst the improvement is not huge (and in fact not big enough to guarantee that the primes contain infinitely many 3-term APs on density grounds alone) the argument is an exceedingly ingenious adaptation of the Hardy-Littlewood "circle method" which is certainly worth knowing about.

Before we can even outline the approach taken by Bourgain in any meaningful way, it is necessary to set up some nomenclature and to introduce a few technical preliminaries.

1 Local Structure of Sets

A very common type of argument in Bourgain's paper is the following. One has finite sets $A, B \subseteq \mathbb{Z}$, and a function $f : \mathbb{Z} \to \mathbb{R}$ for which, say,

$$\sum_{n \in A} f(n) \ \geq \ \eta|A|.$$

One then wishes to conclude that there is some translate $B' = B + m$ for which

$$\sum_{n \in B'} f(n) \ \geq \ (1-\epsilon)\eta|B|$$

for some small ϵ. This section is devoted to an exploration of situations under which such a principle holds. One feels that the principle is doomed to failure unless B is much "smaller" than A (unless B equals A, of course). However one also needs A to "look like B locally" to avoid examples such as $A = \{0, 5, 10, \dots, 5(n-1)\}$, $B = \{0, 1, 2, 3, 4\}$. In such an example the behaviour of f on A gives very little information on the behaviour of f on translates of B.

After some thought the following definition seems natural.

Definition 1 *Let*

$$Q(n) \ = \ |\{m \in A | \, n \in B + m\}| \ = \ |(n - B) \cap A| \, .$$

Then we say that A looks κ-locally like B *if*

$$\sum_n |Q(n) - A(n)|B|| \ \leq \ \kappa|A||B|.$$

To relate this to Bourgain's paper, we note that if α and β are the characteristic measures associated to the sets A and B then A looks κ-locally like B precisely when

$$\|\alpha * \beta - \alpha\|_1 \ \leq \ \kappa.$$

Let us now see how this definition relates to the type of averaging argument discussed above. Let $f : \mathbb{Z} \to \mathbb{R}$ be a function with $\|f\|_\infty \leq 1$, and suppose that A looks κ-locally like B. Then

$$\begin{aligned} \left| \sum_{m \in A} \sum_{n \in B+m} f(n) - |B| \sum_{n \in A} f(n) \right| \ &= \ \left| \sum_n (Q(n) - |B|A(n)) \, f(n) \right| \\ &\leq \ \kappa|A||B|. \end{aligned} \quad (1)$$

Hence we have the following Lemma.

Lemma 2 *Suppose that* $f : \mathbb{Z} \to \mathbb{R}$ *has* $\|f\|_\infty \le 1$*. Suppose that*

$$\left|\sum_{n\in A} f(n)\right| \;\ge\; \eta|A|$$

and that A *looks* $\epsilon\eta$*-locally like* B*. Then*

(i)

$$\sum_{m\in A}\left|\sum_{n\in B+m} f(n)\right| \;\ge\; (1-\epsilon)\eta|A||B|$$

and

(ii) *There is a translate* B' *of* B *with*

$$\left|\sum_{n\in B'} f(n)\right| \ge (1-\epsilon)\eta|B|.$$

In the sequel we shall use Lemma 2 and also inequality (1) directly.

It turns out that in cases that will interest us, the notion of κ-local likeness allows one to give some rather strong information about triples in Arithmetic Progression.

Lemma 3 *Suppose that* $0 \in A$*,* $0 \in B$ *and that* A *is symmetric about 0. Suppose that* A *is* κ*-locally like* $2B$*. Then there are at least* $(1-\kappa)|A||B|$ *triples* $(n_1, n_2, n_3) \in A \times A \times B$ *with* $n_1 + n_2 = 2n_3$*.*

Proof The number N of such triples is

$$\begin{aligned}\sum_{n_2\in A}\sum_{n_1} A(n_1)2B(n_1+n_2) &= \sum_{n_2\in A}\sum_{n_1} A(-n_1)2B(n_2-n_1)\\ &= \sum_{n_2\in A}(A*2B)(n_2).\end{aligned}$$

Now

$$\sum_m |(A*2B)(m) - A(m)|B|| \;\le\; \kappa|A||B|$$

and so

$$\begin{aligned}|N - |A||B|| &= \left|\sum_{n_2\in A}(A*2B)(n_2) - |A||B|\right|\\ &= \left|\sum_{n_2\in A}\{(A*2B)(n_2) - |B|A(n_2)\}\right|\\ &\le \sum_{n_2\in A}|(A*2B)(n_2) - |B|A(n_2)|\\ &\le \kappa|A||B|.\end{aligned}$$

Hence $N \ge (1-\kappa)|A||B|$ as claimed. □.

2 A First Sketch of the Argument

We start by recalling, very briefly, the usual approach taken in proving Roth's Theorem. One takes a set A of density δ in $\mathbb{Z}/N\mathbb{Z}$, and compares the number of length 3 Arithmetic Progressions in A with $\frac{1}{2}\delta^3 N^2$. This is roughly the number of 3-term APs in a random subset of $\mathbb{Z}/N\mathbb{Z}$. The difference D between these two quantities can be expressed using the Fourier Coefficients $\hat{A}(r)$ of A. If D is small then A contains a progression of length 3 becuase it approximates a random set. Otherwise D is large, and we can deduce that some $\hat{A}(r)$ is large for $r \neq 0$. This information in turn allows us to deduce that A has increased density $\delta + c\delta^2$ in some reasonably large Arithmetic Progression P. But P is affinely equivalent to $\{1, \ldots, N\}$, and so we can iterate the argument. However one can only increment the density $O(\delta^{-1})$ times before it becomes greater than 1, which is clearly impossible. Hence if A is large enough then it contains a 3-term AP.

Bourgain's point of departure seems to be the following. Suppose that

$$\hat{A}(r) \;=\; \sum_n A(n)e^{2\pi inr/N}$$

is large. To show that A has increased density in some progression P, one has to somehow get rid of the exponential terms appearing here. In the usual proof of Roth's Theorem this is done by splitting up $\mathbb{Z}/N\mathbb{Z}$ into small progressions on which $e^{2\pi inr/N}$ is roughly constant as n varies. This, however, is rather inefficient – rather a lot of small progressions are required. Suppose instead that one forgets about progressions, and splits $\mathbb{Z}/N\mathbb{Z}$ up into *sets* on which $\|nr/N\|$ is roughly constant. We could easily deduce that A has increased density on one of these sets. Unfortunately however this information is not equivalent to the original hypothesis, since one of the new sets is not affinely equivalent to $\{1, \ldots, N\}$. Hence we have to strengthen the entire hypothesis that we are trying to prove.

The "sets" that we are discussing here are of course just translates of Bohr Neighbourhoods. Hence we shall try to prove something like the following.

Conjecture 4 *Let A be a subset of some Bohr Neighbourhood Λ, such that $|A| = \delta|\Lambda|$. Then for fixed δ and "sufficiently large" Λ, A contains a three-term Arithmetic Progression.*

Since $\mathbb{Z}/N\mathbb{Z}$ is trivially a Bohr Neighbourhood, we might hope that this would imply Roth's Theorem with a better bound.

There are many difficulties to overcome in order to make the above idea work, as we shall discover. These stem principally from three facts.

(i) If Λ is a Bohr Neighbourhood then it is rather difficult to say anything about the number of 3-term APs in Λ. In particular one does not seem to be able to say that it is anywhere near $|\Lambda|^2$. This means that comparing the number of APs in A with the number of APs in a random subset of Λ with density δ does not give strong information.

(ii) If $\Lambda' \subseteq \Lambda$ are Bohr Neighbourhoods then it is not *a priori* at all obvious that Λ looks κ-locally like Λ' for small κ, even when Λ' is much smaller than Λ. Without this being the case, the many averaging arguments used in the proof of Roth's Theorem will not be available.

(iii) Bohr Neighbourhoods are further from being groups than intervals are. In the usual proof of Roth's Theorem it is a short step from $\{1, \ldots, N\}$ to $\mathbb{Z}/N\mathbb{Z}$, a group on which one can do Fourier Analysis. With a Bohr Neighbourhood it is rather unnatural to embed in a cyclic group like this.

Problem (iii) is resolved by slightly changing our definition of what a Bohr Neighbourhood is, and working all the time in the group $\mathbb{Z}$. The Pontryagin Dual of $\mathbb{Z}$ is of course $\mathbb{T}$, the circle group, and so our Fourier Transforms will be defined on this. We contrast this with the situation in Roth's Theorem (at least in Gowers' [3] version of the proof) in which both functions and their Fourier Transforms are defined on $\mathbb{Z}/N\mathbb{Z}$ (where N is a prime).

3 Bohr Neighbourhoods

3.1 Definitions and Elementary Properties

We begin by defining what we mean by a Bohr Neighbourhood from now on.

Definition 5 *Let $\theta = \{\theta_1, \ldots, \theta_d\} \in \mathbb{R}^d$, and let ϵ and M be real numbers with $\epsilon < \frac{1}{2}$. Then we define the* Bohr Neighbourhood $\Lambda_{\theta,\epsilon,M}$ *to be the set of all $n \in \mathbb{Z}$ such that $|n| \leq M$ and $\|n\theta_j\| \leq \epsilon$ for $j = 1, \ldots, d$.*

This is clearly very similar to the "mod N" version of the same name. We take the opportunity to record here some simple facts about Bohr Neighbourhoods which will be useful later.

Lemma 6 *$|\Lambda_{\theta,\epsilon,M}| \geq \epsilon^d M$.*

Proof Let $\mathbb{S}^d$ be the unit torus $\mathbb{R}^d/\mathbb{Z}^d$. Consider the set of all $P_n = (\|n\theta_1\|, \ldots, \|n\theta_d\|) \in \mathbb{S}^d$ for integers $n \in [1, M]$. This has size M, so some ϵ-cube $\mathcal{B}$ of $\mathbb{S}^d$ contains at least $M\epsilon^d$ of the P_i (this "obvious" averaging argument actually requires careful analysis its justification). Let $\mathcal{C}$ be the set of all $n \in [1, M]$ for which $P_n \in \mathcal{B}$. Then there is an injection

$$\phi : \mathcal{C} \to \Lambda_{\theta,\epsilon,M}$$

defined by $\phi(n) = n - n_0$, where $n_0 \in \mathcal{C}$ is arbitrary. $\square$

Lemma 7 *$|\Lambda_{\theta,\epsilon,M}| < 8^{d+1}|\Lambda_{\theta,\frac{\epsilon}{2},\frac{M}{2}}|$*

Proof Divide $\Lambda_{\theta,\epsilon,M}$ into sets A_i such that

(i) $\{(\|n\theta_1\|, \ldots, \|n\theta_d\|) \,|\, n \in A_i\}$ is contained in an $\frac{\epsilon}{2}$-cube in $\mathbb{S}^d$;

(ii) A_i is contained in an interval of length $\frac{M}{2}$.

This can be achieved with 8^{d+1} sets A_i. Each A_i injects to $\Lambda_{\theta,\frac{\epsilon}{2},\frac{M}{2}}$ by sending n to $n - n_0$, where $n_0 \in A_i$ is arbitrary. The result follows. $\square$

Bourgain's Paper has a nice alternative derivation of these results using Fourier Analysis.

In view of the difficulties (i) and (ii) mentioned above, together with Lemma 3, we are clearly going to be interested in finding out when a Bohr Neighbourhood Λ looks κ-locally like another neighbourhood Λ'. We deal with this now.

3.2 Bohr Neighbourhoods and Local Likeness

We will be interested only in the following rather specific version of the question. Let $A = \Lambda_{\theta,\epsilon,M}$ and let $B \subseteq \Lambda_{\theta,\gamma\epsilon,\gamma M}$. When does A look κ-locally like B? A crucial observation is that, because of the structure of Bohr Neighbourhoods, the answer to this question depends on how A behaves near its "edges". To this end (and with hindsight) we make the following definition.

Definition 8 *Fix $\theta \in \mathbb{R}^d$. Then we say that a pair (ϵ, M) is* regular *if*

$$1 - 100d|\gamma| \;\leq\; \frac{|\Lambda_{\theta,(1+\gamma)\epsilon,(1+\gamma)M}|}{|\Lambda_{\theta,\epsilon,M}|} \;\leq\; 1 + 100d|\gamma|$$

whenever $|\gamma| \leq \frac{1}{100d}$.

If (ϵ, M) is regular for some $\theta \in \mathbb{R}^d$ then we also describe the Bohr Neighbourhood $\Lambda_{\theta,\epsilon,M}$ as regular. The reason for making the definition in exactly this way is that, as we shall show later, no pair (ϵ, M) is very far from a regular pair. Before doing that, however, we show how regularity governs questions of local-likeness.

Proposition 9 *Suppose (ϵ, M) is regular. Let $A = \Lambda_{\theta,\epsilon,M}$, and let $B \subseteq \Lambda_{\theta,\gamma\epsilon,\gamma M}$ where $\gamma \leq \frac{1}{200d}$. Then A looks $400d\gamma$-locally like B.*

Proof Let $Q(n) = |\{m \in A | n \in B + m\}|$. Suppose that $n \in B + m$ for some $m \in A$. Then $n \in \Lambda_{\theta,(1+\gamma)\epsilon,(1+\gamma)M}$. Therefore we have that $Q(n) = 0$ if $n \notin \Lambda_{\theta,(1+\gamma)\epsilon,(1+\gamma)M}$. Suppose now that $n \in \Lambda_{\theta,(1-\gamma)\epsilon,(1-\gamma)M}$. Then for all $b \in B$ we have $n - b \in A$ and so $Q(n) = |B|$. For all other values of n, $Q(n)$ lies between 0 and $|B|$. It follows that

$$\begin{aligned}
\sum_n |Q(n) - A(n)|B|| &\leq 2|B| \left|\Lambda_{\theta,(1+\gamma)\epsilon,(1+\gamma)M} \setminus \Lambda_{\theta,(1-\gamma)\epsilon,(1-\gamma)M}\right| \\
&\leq 2|A||B| \cdot ((1 + 100d\gamma) - (1 - 100d\gamma)) \\
&\leq 400d\gamma|A||B|.
\end{aligned}$$

But this is precisely what it means for A to look $400d\gamma$-locally like B. □

In general, local likeness is not preserved under passing to subsets. We leave it to the reader to construct, for any $\kappa > 0$, sets A, B and $C \subseteq B$ for which A looks κ-locally like B, but A does not look $\frac{1}{2}$-locally like C' for any translate C' of C. The proof of Proposition 9 used some of the specific structure of Bohr Neighbourhoods.

The following immediate corollary of Proposition 9 allows us, in view of Lemma 3, to discuss 3-term APs in the context of Bohr Neighbourhoods.

Corollary 10 *Suppose (ϵ, M) is regular and let $\gamma \leq \frac{1}{400d}$. Then $\Lambda_{\theta,\epsilon,M}$ looks $800d\gamma$-locally like $2\Lambda_{\theta,\gamma\epsilon,\gamma M}$.*

Proof Simply apply Proposition 9 and note that $2\Lambda_{\theta,\gamma\epsilon,\gamma M} \subseteq \Lambda_{\theta,2\gamma\epsilon,2\gamma M}$. □

Combining this with Lemma 3 gives

Corollary 11 *Suppose that (ϵ, M) is regular and let $\delta \leq \frac{1}{400d}$. Let $\Lambda = \Lambda_{\theta,\epsilon,M}$ and $\Lambda' = \Lambda_{\theta,\gamma\epsilon,\gamma M}$. Then there are at least $(1 - 800d\gamma)|\Lambda||\Lambda'|$ triples $(n_1, n_2, n_3) \in \Lambda \times \Lambda \times \Lambda'$ with $n_1 + n_2 = 2n_3$.*

3.3 Finding Regular Bohr Neighbourhoods

Of course everything that we proved in the last section is useless until we have said something about regular pairs (ϵ, M) (for a fixed $\theta \in \mathbb{R}^d$). A pair (ϵ, M) is regular if $|\Lambda_{\theta,(1+\gamma)\epsilon,(1+\gamma)M}|$ varies in a controlled manner. Recall Definition 8 for the precise details. Let (ϵ, M) be a not-necessarily-regular pair. We are going to show in Lemma 12 below that there is $\alpha \in [\frac{1}{2}, 1]$ for which $(\alpha\epsilon, \alpha M)$ is regular. This is what we mean by saying that every pair is close to a regular one.

Set $f(\alpha) = |\Lambda_{\theta,\alpha\epsilon,\alpha M}|$. Then $f(\alpha)$ is a non-decreasing function on $[\frac{1}{2}, 1]$ and, by Lemma 7,

$$f(1) \ \leq \ 8^{d+1} f\left(\tfrac{1}{2}\right) \tag{2}$$

. It turns out that these facts alone are enough to prove the following.

Lemma 12 *n Let (ϵ, M) be a pair of positive real numbers with $\epsilon < 1$. Then there is a real number $\alpha \in [\frac{1}{2}, 1]$ for which $(\alpha\epsilon, \alpha M)$ is regular.*

Proof It clearly suffices to show that there is $\alpha \in [\frac{1}{2}, 1]$ such that

$$1 - 100d|\gamma| \ \leq \ \left|\frac{f((1+\gamma)\alpha)}{f(\alpha)}\right| \ \leq \ 1 + 100d|\gamma|$$

for all $|\gamma| \leq \frac{1}{100d}$. Suppose then that this is false. Observe that $\frac{1}{1-x} \geq 1 + x$ when $x \geq 0$; hence for every $\alpha \in [\frac{1}{2}, 1]$ there is $t_\alpha \in [0, \frac{1}{100d}]$ such that

$$\left|\frac{f\left((1+t_\alpha)\alpha\right)}{f\left((1-t_\alpha)\alpha\right)}\right| \ \geq \ 1 + 100dt_\alpha \tag{3}$$

$$\geq \ e^{50dt_\alpha}, \tag{4}$$

the last step following because $1 + x \geq e^{\frac{1}{2}x}$ for $x \leq 1$. At this point we pause to prove a small covering lemma. This can be traced back at least as far as Croft [2] but probably ranks as "folklore".

Lemma 13 *Suppose a finite collection of closed intervals $I_1, \ldots, I_k$ covers $[0, 1]$. Then we can pick a subcollection $I_{i_1}, \ldots, I_{i_m}$ whose members are disjoint except possibly at their endpoints, with total measure at least $\frac{1}{2}$.*

Proof Without loss of generality suppose that the collection $I_1, \ldots, I_k$ is minimal in that if any I_j is removed, the intervals no longer cover $[0, 1]$. It is then easy to see that no point x lies in three of the I_j, because there are two intervals I_r and I_s containing x such that any other I_t containing x lies in $I_r \cup I_s$. But it is then easy to see what the intervals "look like". Suppose that the $I_j = [a_j, b_j]$ with $a_1 \leq a_2 \leq \ldots \leq a_k$. Then

$$a_1 \leq a_2 \leq b_1 \leq a_3 \leq b_2 \leq a_4 \leq \ldots \leq b_{k-1} \leq b_k.$$

It follows that the two collections $I_1 \cup I_3 \cup \ldots$ and $I_2 \cup I_4 \cup \ldots$ contain disjoint intervals. The result is now obvious. □

To apply Lemma 13, recall (3). By compactness we may take a finite set

$$\{\alpha_1, \ldots, \alpha_k\} \subseteq \left[\frac{1}{2} + \frac{1}{100d}, 1 - \frac{1}{100d}\right]$$

such that the intervals $[(1 - t_{\alpha_i})\, \alpha_i, (1 + t_{\alpha_i})\, \alpha_i]$ cover $\left[\frac{1}{2} + \frac{1}{100d}, 1 - \frac{1}{100d}\right]$. Since $t_{\alpha_i} \leq \frac{1}{100d}$, all of these intervals are contained in $[\frac{1}{2}, 1]$. By Lemma 13, we can pick a disjoint subcollection of measure at least $\frac{1}{4}\left(1 - \frac{1}{100d}\right) > \frac{1}{5}$. Letting these intervals correspond to $\{\alpha_1, \ldots, \alpha_l\}$, one has

$$2\sum_{i=1}^{l} \alpha_i t_{\alpha_i} \;>\; \frac{1}{5}$$

and so

$$\sum_{i=1}^{l} t_{\alpha_i} \;>\; \frac{1}{10}.$$

Using this in (3) gives

$$\prod_{i=1}^{l} \left| \frac{f\left((1 + t_{\alpha_i})\alpha_i\right)}{f\left((1 - t_{\alpha_i})\alpha_i\right)} \right| \;\geq\; e^{50d \sum_i t_{\alpha_i}} \;>\; e^{5d}.$$

However the left hand side is at most $\frac{f(1)}{f(\frac{1}{2})}$, and hence by (2) we have

$$8^{d+1} \;>\; e^{5d}.$$

This is a contradiction, and Lemma 12 is established. □

4 A Second Sketch of the Argument

Having dispensed with preliminaries, we can now outline the mode of attack. Let a set A be given, such that A has density δ in some Bohr Neighbourhood $\Lambda = \Lambda_{\theta,\epsilon,M}$. Suppose for a contradiction that A has no 3-term AP. We want to be able to say that A looks significantly different from a random subset of Λ with density δ. Recall, however, that we could not give strong information about APs in Λ itself, and so it is at first sight not possible to proceed along these lines. However suppose that (ϵ, M) is a regular pair, and that $\Lambda' = \Lambda_{\theta,c\epsilon,cM}$ wih c small. Then, by Corollary 11, the number of 3-term APs $(n_1, n_2, n_3) \in \Lambda \times \Lambda \times \Lambda'$ is close to its maximum possible value $|\Lambda||\Lambda'|$. Suppose that $A' = A \cap \Lambda'$ has density δ' in Λ'. Then we can try to use the information that A has no non-trivial 3-term APs to show that $A \times A \times A'$ looks significantly different from a random subset of $\Lambda \times \Lambda \times \Lambda'$ of density $\delta \times \delta \times \delta'$. Of course we would also have to interpret the last (rather nonsensical) statement suitably. The problem with this is that we have very little control over δ'. In fact (since the property of having a 3-term AP is translation invariant) one only needs some translate of A to have density about δ on both Λ and Λ'. However even this might not be achievable. We now explain the way around this, which is one of the less transparent complexities of Bourgain's Paper.

5 Controlling Density on Smaller Bohr Neighbourhoods

Suppose that A has density δ on some $\Lambda = \Lambda_{\theta,\epsilon,M}$, where (ϵ, M) is regular. Suppose that A has no non-trivial 3-term APs. Our aim is then to deduce that A has density $\delta + O(\delta^2)$ on some reasonably large regular Bohr Neighbourhood Λ'. As in the usual proof of Roth's Theorem, this will lead to a proof of Conjecture 4 in the case Λ regular. This in turn will imply the traditional form of Roth's Theorem, because the set $[-N, N]$ is easily seen to be a regular Bohr Neighbourhood. Let $\Lambda_1 = \Lambda_{\theta,c_1\epsilon,c_1M}$ and $\Lambda_2 = \Lambda_{\theta,c_2\epsilon,c_2M}$ be regular, with $1 \gg c_1 \gg c_2$ such that Λ looks κ_1-locally like Λ_1 and Λ_1 looks κ_2-locally like Λ_2. The values of c_1,c_2, κ_1 and κ_2 will be specified later. We show next that either there is already a density increment for some translate of A in either Λ_1 or Λ_2, or else one can find some translate $A' = A + m$ which has density approximately δ on both Λ_1 and Λ_2. In this latter case we really will get useful information by comparing $(A\cap\Lambda_1)\times(A\cap\Lambda_1)\times(A\cap\Lambda_2)$ with something like a random subset of $\Lambda_1 \times \Lambda_1 \times \Lambda_2$ of the appropriate density.

Lemma 14 *Let $\delta_1(m)$ be the density of $A + m$ on Λ_1. Then*

$$\left|\sum_{m\in\Lambda} (\delta_1(m) - \delta)\right| \le \kappa_1|\Lambda|.$$

Proof We have

$$\sum_{m\in\Lambda} (\delta_1(m) - \delta) = \frac{1}{|\Lambda_1|}\left(\sum_{m\in\Lambda}\sum_{n\in\Lambda_1+m} A(n) - |\Lambda_1|\sum_{n\in\Lambda} A(n)\right).$$

Recalling that Λ looks κ_1-locally like Λ_1, the result follows immediately from (1). □

Lemma 15 *Let $\delta_1(m)$, $\delta_2(m)$ be the densities of $A + m$ on Λ_1 and Λ_2 respectively. Then either there is m such that $|\delta-\delta_1(m)| \le 6\kappa_1$ and $|\delta-\delta_2(m)| \le 6\kappa_1$, or there is m such that $\delta_1(m) \ge \delta+\kappa_1$, or else there is m such that $\delta_2(m) \ge \delta + \kappa_1$.*

Proof We note that, since $\Lambda_2 \subseteq \Lambda_1$, it follows from Proposition 9 that Λ looks κ_1-locally like Λ_2. Suppose now that the result is false. Then, for all $m \in \Lambda$, either $|\delta - \delta_1(m)| > 6\kappa_1$ or

$|\delta - \delta_2(m)| > 6\kappa_1$. Without loss of generality assume that $|\delta - \delta_1(m)| > 6\kappa_1$ for at least $\frac{1}{2}|\Lambda|$ values of $m \in \Lambda$, so that

$$\sum_{m \in \Lambda} |\delta - \delta_1(m)| \; > \; 3\kappa_1 |\Lambda|.$$

Since $\delta_1(m) < \delta + \kappa_1$ for all m, we have

$$\sum_{\substack{m \in \Lambda \\ \delta_1(m) > \delta}} |\delta - \delta_1(m)| \; \leq \; \kappa_1 |\Lambda|.$$

Therefore

$$\sum_{\substack{m \in \Lambda \\ \delta_1(m) \leq \delta}} (\delta - \delta_1(m)) \; > \; 2\kappa_1 |\Lambda|,$$

and so

$$\begin{aligned} \left| \sum_{m \in \Lambda} (\delta_1(m) - \delta) \right| \; &\geq \; \sum_{\substack{m \in \Lambda \\ \delta_1(m) \leq \delta}} (\delta - \delta_1(m)) - \sum_{\substack{m \in \Lambda \\ \delta_1(m) > \delta}} |\delta - \delta_1(m)| \\ &> \; \kappa_1 |\Lambda|. \end{aligned}$$

This contradicts Lemma 14. □

We shall require, it turns out, that $\kappa_1 \leq 2^{-17}\delta^2$. What value of c_1 will be required to achieve this? From Proposition 9, we see that any $c_1 \leq d^{-1}2^{-27}\delta^2$ will do. By Lemma 12, we can pick some

$$c_1 \in \left[d^{-1}2^{-28}\delta^2, d^{-1}2^{-27}\delta^2 \right] \tag{5}$$

such that $(c_1\epsilon, c_1 M)$ is regular. We assume from now on that such a c_1 has been chosen. The value of c_2 has yet to be specified, but provided $c_1 \geq c_2$ we have the following conclusion.

Proposition 16 *Let $\Lambda_1 = \Lambda_{\theta, c_1\epsilon, c_1 M}$ and $\Lambda_2 = \Lambda_{\theta, c_2\epsilon, c_2 M}$ be regular, with c_1 satisfying*

$$d^{-1}2^{-28}\delta^2 \; \leq \; c_1 \; \leq \; d^{-1}2^{-27}\delta^2.$$

Let $\delta_i(m)$ denote the density of $A + m$ on Λ_i. Then there is some m for which one of the following is true.

(i) $\delta_1(m) \geq \delta + 2^{-20}\delta^2$;

(ii) $\delta_2(m) \geq \delta + 2^{-20}\delta^2$;

(iii) $\delta_1(m) \geq \delta - 2^{-13}\delta^2$ and $\delta_2(m) \geq \delta - 2^{-13}\delta^2$.

Proof This is an easy consequence of Lemma 15 and Proposition 9. □

Until the very end of the argument we will work under the assumption that (iii) holds. Indeed (i) and (ii) already represent density increments of precisely the type we are aiming for. The properties of A on Λ itself will no longer concern us. Hence, replacing A by some translate $A+m$ if necessary, we assume henceforth that

$$\delta_1 \;\geq\; \delta - 2^{-13}\delta^2 \tag{6}$$

and

$$\delta_2 \;\geq\; \delta - 2^{13}\delta^2, \tag{7}$$

where δ_i is the density of A on $\Lambda_i = \Lambda_{\theta,c_i\epsilon,c_iM}$, c_1 satisfies (5), and $c_2 \leq c_1$. One further piece of notation: we will write $A_1 = A \cap \Lambda_1$ and $A_2 = A \cap \Lambda_2$.

6 The Hardy-Littlewood Method

In this section we actually perform the comparison between $A_1 \times A_1 \times A_2$ and a random subset of $\Lambda_1 \times \Lambda_1 \times \Lambda_2$ of density $\delta_1^2\delta_2$. We acknowledge that this is still a slightly nonsensical statement. Let I_1 be the number of 3-term APs in $A_1 \times A_1 \times A_2$. Assuming that A contains no nontrivial APs, we have that $I_1 = |A_2|$. For a function $f : \mathbb{Z} \to \mathbb{C}$ define the Fourier Transform $\hat{f} : \mathbb{T} \to \mathbb{C}$ by

$$\hat{f}(x) \;=\; \int_0^1 f(x)e^{2\pi inx}\,dx$$

as usual. Then it is not hard to see that

$$I_1 \;=\; \int_0^1 \hat{A}_1(x)^2\hat{A}_2(-2x)\,dx,$$

where we have identified the sets A_1 and A_2 with their characteristic functions. Let

$$I_2 \;=\; \delta_1^2\delta_2\int_0^1 \hat{\Lambda}_1(x)^2\hat{\Lambda}_2(-2x)\,dx.$$

This is supposed to be a guess at the number of 3-term APs in a random subset of $\Lambda_1 \times \Lambda_1 \times \Lambda_2$ of density $\delta_1 \times \delta_1 \times \delta_2$. Indeed it is actually equal to $\delta_1^2\delta_2$ times the number of 3-term APs in $\Lambda_1 \times \Lambda_1 \times \Lambda_2$. Suppose that c_2 is chosen so that Λ_1 looks $\frac{1}{5}$-locally like $2\Lambda_2$. By Corollary 10, a sufficient condition for this is that

$$c_2 \;\leq\; 2^{-13}d^{-1}c_1. \tag{8}$$

By Lemma 12 and (5), we can find c_2 with

$$2^{-42}d^{-2}\delta^2 \;\leq\; c_2 \;\leq\; 2^{-40}d^{-2}\delta^2 \tag{9}$$

such that $\Lambda_2 = \Lambda_{\theta,c_2\epsilon,c_2M}$ is regular and (8) is satisfied.

Now by Lemma 3 or Corollary 11 we have that

$$I_2 \;\geq\; \frac{4}{5}\delta_1^2\delta_2|\Lambda_1||\Lambda_2|. \tag{10}$$

For us to be able to derive useful information from the fact that $I_1 = |A_2|$ (i.e. the fact that A_2 has no non-trivial 3-term AP) we will need, say, that $I_1 \leq \frac{1}{5}\delta_1^2\delta_2|\Lambda_1||\Lambda_2|$. This will certainly be true if

$$\delta_1^2|\Lambda_1| \;\geq\; 5. \tag{11}$$

When we come to iterate our entire argument, obtaining density increments $\delta' = \delta + O(\delta^2)$ in successive Bohr Neighbourhoods, it will be (11) that finally determines the bound that we shall get. Indeed the successive Bohr Neighbourhoods can never get smaller than the bound specified by (11).

Recalling from (5) and (6) that $c_1 \geq d^{-1}2^{-28}\delta^2$ and $\delta_1 \geq \frac{\delta}{2}$, we see using Lemma 6 that (11) will certainly hold if

$$M\epsilon^d \;\geq\; \left(\frac{2^{61}d^2}{\delta^6}\right)^d. \tag{12}$$

Here we have been rather crude in order to make the expression a little neater, but this slackness makes almost no difference to the final bound.

Supposing then that (12) is satisfied, we have that

$$|I_1 - I_2| \;\geq\; \frac{1}{2}\delta_1^2\delta_2|\Lambda_1||\Lambda_2|. \tag{13}$$

The next part of the argument consists of estimating $|I_1 - I_2|$ in a different way, a way which will eventually allow us to obtain the density increment that we have discussed so much. We have

$$\begin{aligned}
|I_1 - I_2| \;&=\; \left|\int_0^1 \hat{A}_1(x)^2\hat{A}_2(-2x)\,dx \;-\; \delta_1^2\delta_2\int_0^1 \hat{\Lambda}_1(x)^2\hat{\Lambda}_2(-2x)\,dx\right| \\
&\leq\; \left|\int_0^1 \hat{A}_1(x)^2\left(\hat{A}_2(-2x) - \delta_2\hat{\Lambda}_2(-2x)\right)\,dx\right| \;+\; \delta_2\left|\int_0^1 \left(\hat{A}_1(x)^2 - \delta_1^2\hat{\Lambda}_1(x)^2\right)\hat{\Lambda}_2(-2x)\,dx\right| \\
&\leq\; \|\hat{A}_2(-2x) - \delta_2\hat{\Lambda}_2(-2x)\|_\infty \int_0^1 \left|\hat{A}_1(x)\right|^2\,dx \\
&\qquad\qquad + \delta_2\int_0^1 \left|\hat{A}_1(x)^2 - \delta_1^2\hat{\Lambda}_1(x)^2\right|\left|\hat{\Lambda}_2(-2x)\right|\,dx.
\end{aligned} \tag{14}$$

By (13), we see that either

$$\|\hat{A}_2(-2x) - \delta_2\hat{\Lambda}_2(-2x)\|_\infty \int_0^1 \left|\hat{A}_1(x)\right|^2\,dx \;\geq\; \frac{1}{4}\delta_1^2\delta_2|\Lambda_1||\Lambda_2| \tag{15}$$

or

$$\int_0^1 \left|\hat{A}_1(x)^2 - \delta_1^2\hat{\Lambda}_1(x)^2\right|\left|\hat{\Lambda}_2(-2x)\right|\,dx \;\geq\; \frac{1}{4}\delta_1^2|\Lambda_1||\Lambda_2|. \tag{16}$$

If (15) holds then, by Parseval's Theorem, one has

$$\|\hat{A}_2 - \delta_2\hat{\Lambda}_2\|_\infty \;\geq\; \frac{1}{4}\delta_1\delta_2|\Lambda_2|. \tag{17}$$

This statement is saying that A_2 looks significantly unlike a random subset of Λ_2 with density δ_2. This case, in fact, is rather similar to the argument followed in the usual proof of Roth's Theorem. We will show that, indeed, working with Bohr Neighbourhoods allows us to get a substantial (i.e. $O(\delta^2)$) density increment on a reasonably large Λ_3. This was our original motivation for working with Bohr Neighbourhoods instead of progressions. The main difficulty comes from the fact that we must also derive a density increment from (16) above. This proves to be much more difficult, and is an unfortunate byproduct of the more complicated analysis that has been necessary in dealing with Bohr Neighbourhoods. We shall obtain a density increment from (15) now: this is what we call the First Case. The rest of the paper will be devoted to obtaining a density increment from (16), which we call the Second Case.

7 Obtaining a Density Increment in the First Case

From (6) and (7) we have $\delta_1 > \frac{1}{2}\delta$ and $\delta_2 > \frac{1}{2}\delta$. Hence, writing (17) out in full one gets

$$\left|\sum_{n\in A_2}(A(n)-\delta_2)\,e^{2\pi inx_0}\right| \;\geq\; \frac{\delta^2}{16}|\Lambda_2| \tag{18}$$

for some $x_0 \in \mathbb{T}$. Suppose that Λ_3 is a regular Bohr Neighbourhood on which $e^{2\pi inx_0}$ is roughly constant. If Λ_2 can be efficiently covered by translates of Λ_3, then we can envisage showing that some translate of A has increased density on Λ_3 much as in the usual proof of Roth's Theorem. Such a Λ_3 is given by $\Lambda_3 = \Lambda_{\theta',\gamma c_2\epsilon,\gamma c_2 M}$ where $\theta' = \theta \cup \{x_0\}$ (so that $\theta' \in \mathbb{R}^{d+1}$) and γ is small, so that Λ_2 looks locally like Λ_3. To this end recall from Proposition 9 that if $\gamma \leq \frac{1}{200d}$ then Λ_2 looks $400d\gamma$-locally like Λ_3. Applying Lemma 2 to (18) gives

$$\sum_{m\in\Lambda_2}\left|\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\,e^{2\pi inx_0}\right| \;\geq\; \left(\frac{\delta^2}{16}-400d\gamma\right)|\Lambda_2||\Lambda_3|$$

and so

$$\sum_{m\in\Lambda_2}\left|\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\right| + \sum_{m\in\Lambda_2}\left|\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\left(e^{2\pi imx_0}-e^{2\pi inx_0}\right)\right| \geq \left(\frac{\delta^2}{16}-400d\gamma\right)|\Lambda_2||\Lambda_3|.$$

However

$$\begin{aligned}\sum_{m\in\Lambda_2}\left|\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\left(e^{2\pi imx_0}-e^{2\pi inx_0}\right)\right| &\leq 2|\Lambda_2||\Lambda_3|\sup_{n\in\Lambda_3+m}\left|e^{2\pi imx_0}-e^{2\pi inx_0}\right| \\ &\leq 4\pi\gamma c_2\epsilon|\Lambda_2||\Lambda_3| \\ &\leq 4\pi\gamma|\Lambda_2||\Lambda_3|\end{aligned}$$

and so

$$\sum_{m\in\Lambda_2}\left|\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\right| \;\geq\; \left(\frac{\delta^2}{16}-420d\gamma\right)|\Lambda_2||\Lambda_3|.$$

However by (1) we have that

$$\left|\sum_{m\in\Lambda_2}\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\right| \;\leq\; 400d\gamma|\Lambda_2||\Lambda_3|,$$

and so

$$\sum_{m\in\Lambda_2}\left(\left|\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\right|+\sum_{n\in\Lambda_3+m}(A(n)-\delta_2)\right) \;\geq\; \left(\frac{\delta^2}{16}-820d\gamma\right)|\Lambda_2||\Lambda_3|. \qquad (19)$$

Therefore, for some m, we have that

$$\sum_{n\in\Lambda_3+m}(A(n)-\delta_2) \;\geq\; \left(\frac{\delta^2}{32}-410d\gamma\right)|\Lambda_3|.$$

If we take γ small enough then this gives the required density increment. Indeed with $\gamma \leq 2^{-16}\delta^2 d^{-1}$ one has

$$\begin{aligned}|A\cap(\Lambda_3+m)| \;&\geq\; \left(\delta_2+\frac{\delta^2}{64}\right)|\Lambda_3| \\ &\geq\; \left(\delta+\frac{\delta^2}{128}\right)|\Lambda_3|\end{aligned}$$

by (7). The reader may care to observe how very similar all this is to the proof of Roth's Theorem given in Gowers [3], even down to the use of the trick in (19).

Recall that $\Lambda_3 = \Lambda_{\theta',\gamma c_2\epsilon,\gamma c_2 M}$, where we required that $\gamma \leq 2^{-16}\delta^2 d^{-1}$. Let $c_3 = \gamma c_2$. In order to make the main loop of our argument work, we also require that Λ_3 be regular. This can be achieved, by Lemma 12, with some $\gamma \geq 2^{-17}\delta^2 d^{-1}$. Then we will have, by (9), that

$$2^{-59}\delta^4 d^{-3} \;\leq\; c_3 \;\leq\; 2^{-56}\delta^4 d^{-3}. \qquad (20)$$

This completes our analysis of the First Case: we have found a reasonably large Bohr Neighbourhood Λ_3 on which the density of A increases noticably. Our attention must now turn to the much more difficult Second Case.

8 Obtaining a Density Increment in the Second Case, Part 1

Recall that in the Second Case one has (16) holding, namely

$$\int_0^1 \left|\hat{A}_1(x)^2-\delta_1^2\hat{\Lambda}_1(x)^2\right|\left|\hat{\Lambda}_2(-2x)\right|\,dx \;\geq\; \frac{1}{4}\delta_1^2|\Lambda_1||\Lambda_2|.$$

We start by playing about with this a little. One has

$$\begin{aligned}\int_0^1 \left|\hat{A}_1(x)^2 - \delta_1^2\hat{\Lambda}_1(x)^2\right| \left|\hat{\Lambda}_2(-2x)\right| dx &= \int_0^1 \left|\hat{A}_1(x) + \delta_1\hat{\Lambda}_1(x)\right| \left|\hat{\Lambda}_2(-2x)\right| \left|\hat{A}_1(x) - \delta_1\hat{\Lambda}_1(x)\right| dx \\ &\leq \left\|\hat{A}_1 + \delta_1\hat{\Lambda}_1\right\|_2 \left\|\left|\hat{\Lambda}_2(-2x)\right| \left|\hat{A}_1 - \delta_1\hat{\Lambda}_1\right|\right\|_2,\end{aligned}$$

this last step being an instance of the Cauchy-Schwarz inequality. But

$$\begin{aligned}\left\|\hat{A}_1 + \delta_1\hat{\Lambda}_1\right\|_2 &\leq \left\|\hat{A}_1\right\|_2 + \delta_1 \left\|\hat{\Lambda}_1\right\|_2 \\ &= |A_1|^{1/2} + \delta_1|\Lambda_1|^{1/2} \\ &\leq 2\delta_1^{1/2}|\Lambda_1|^{1/2}, \qquad (21)\end{aligned}$$

using Parseval's Identity. Hence in the Second Case we have that

$$\left\|\left|\hat{\Lambda}_2(-2x)\right| \left|\hat{A}_1 - \delta_1\hat{\Lambda}_1\right|\right\|_2 \geq \frac{1}{8}\delta_1^{3/2}|\Lambda_1|^{1/2}|\Lambda_2|. \qquad (22)$$

However we also have

$$\left\|\hat{A}_1 - \delta_1\hat{\Lambda}_1\right\|_2 \leq 2\delta_1^{1/2}|\Lambda_1|^{1/2}$$

exactly as in (21). Let $\mathcal{F}$ be the subset of $[0,1]$ consisting of those x for which $\left|\hat{\Lambda}_2(-2x)\right| \geq \frac{1}{32}\delta_1|\Lambda_2|$. Then we have, using Minkowski's Inequality,

$$\begin{aligned}\left\|\left|\hat{\Lambda}_2(-2x)\right| \left|\hat{A}_1 - \delta_1\hat{\Lambda}_1\right|\right\|_2 &\leq \left\|\left(\hat{A}_1 - \delta_1\hat{\Lambda}_1\right) \hat{\Lambda}_2(-2x)\chi_{\mathcal{F}}\right\|_2 + \left\|\left(\hat{A}_1 - \delta_1\hat{\Lambda}_1\right) \hat{\Lambda}_2(-2x)\chi_{[0,1]\setminus\mathcal{F}}\right\|_2 \\ &\leq |\Lambda_2| \left\|\left(\hat{A}_1 - \delta_1\hat{\Lambda}_1\right) \chi_{\mathcal{F}}\right\|_2 + \frac{1}{32}\delta_1|\Lambda_2| \left\|\hat{A}_1 - \delta_1\hat{\Lambda}_1\right\|_2 \\ &\leq |\Lambda_2| \left\|\left(\hat{A}_1 - \delta_1\hat{\Lambda}_1\right) \chi_{\mathcal{F}}\right\|_2 + \frac{1}{16}\delta_1^{3/2}|\Lambda_1|^{1/2}|\Lambda_2|.\end{aligned}$$

Comparing this with (22) we see that

$$\left\|\left(\hat{A}_1 - \delta_1\hat{\Lambda}_1\right) \chi_{\mathcal{F}}\right\|_2 \geq \frac{1}{16}\delta_1^{3/2}|\Lambda_1|^{1/2}. \qquad (23)$$

We now deem ourselves to have finished playing around, and set about the task of deriving a density increment from (23).

9 The Structure of $\mathcal{F}$

Perhaps the key insight of the whole paper is that the set $\mathcal{F}$, the set of all x for which $\left|\hat{\Lambda}_2(-2x)\right| \geq \frac{1}{32}\delta_1|\Lambda_2|$, can be described in a reasonably exact way. In my opinion it is slightly outrageous to expect that this should be the case. As the reader may agree, the proof of this statement is also rather outrageous. Or at least very clever.

Theorem 17 *Let* $\Lambda = \Lambda_{\theta,\epsilon,M}$ *be a regular Bohr Neighbourhood, where* $\theta \in \mathbb{R}^d$*. Let* x *be a real number for which* $\left|\hat{\Lambda}(x)\right| \geq \kappa|\Lambda|$*. Then there is a vector* $\mathbf{k} \in \mathbb{Z}^d$ *for which*

$$\|\mathbf{k}\|_\infty \ \leq \ \frac{2^{21}d^5}{\kappa^3}\left(\log\frac{1}{\epsilon}\right)^2\frac{1}{\epsilon}$$

and

$$|x + \mathbf{k}.\theta| \ \leq \ \frac{2^{21}d^5}{\kappa^3}\left(\log\frac{1}{\epsilon}\right)^2\frac{1}{M}.$$

It is quite easy to see that *if* x is close to $\mathbf{k}.\theta$ for some reasonably small $\mathbf{k}$ then $\hat{\Lambda}(x)$ is likely to be large. This result asserts that the converse is also true to some extent.

We now embark on the quite lengthy process of proving Theorem 17. Let χ denote the characteristic function of the interval $[-\epsilon, \epsilon]$. Then the Fourier Coefficient $\hat{\Lambda}(x)$ can be written as

$$\hat{\Lambda}(x) \ = \ \sum_{|n|\leq M}\left(\prod_{j=1}^{d}\chi(n\theta_j)\right)e^{2\pi inx}. \tag{24}$$

After some thought it is not unnatural to consider writing $\chi(t)$ as a Fourier Expansion $\chi(t) = \sum_r \hat{\chi}(r)e^{2\pi irt}$. Suppressing our worries about the validity of such an expansion, we then have

$$\hat{\Lambda}(x) \ = \ \sum_{\mathbf{k}\in\mathbb{Z}^d}\left(\prod_{j=1}^{d}\hat{\chi}(k_j)\right)\sum_{|n|\leq M}e^{2\pi in(x+\mathbf{k}.\theta)}. \tag{25}$$

Now

$$\left|\sum_{|n|\leq M}e^{2\pi in(x+\mathbf{k}.\theta)}\right| \ < \ \frac{2}{\|x+\mathbf{k}.\theta\|}$$

and $\hat{\chi}(r) \to 0$ as $r \to \infty$. These are the sort of inequalities that might enable us to deduce a result like Theorem 17, namely that $\|x + \mathbf{k}.\theta\|$ is small for some small $\mathbf{k}$, from (25). However they are far to weak to get any result of that kind, let alone one strong enough for our purposes. The way in which we improve on them is an extremely ingeneous "smoothing" technique. In this technique one replaces the characteristic function χ by a function which is very similar to it, but which has a Fourier Transform which decays hugely more quickly. Rather luckily the regularity of Λ is exactly what allows us to make this replacement. The construction of the smooth approximation to χ is relegated to an appendix as to give it here would interrupt the flow of the argument. We simply state for now the results that will be proved.

Theorem 18 *Let* $0 < \delta \leq \frac{1}{16}$*, and let* L *be a positive real number. Then there is a function* $\tau = \tau_{L,\delta} : \mathbb{R} \to \mathbb{R}$ *with the following properties. Firstly,* σ *approximates the characteristic function of the interval* $[-L, L]$ *in that* $0 \leq \tau \leq 1$ *and*

$$\tau(x) \ = \ \begin{cases} 0 & (|x| > (1+\delta)L) \\ 1 & (|x| < (1-\delta)L). \end{cases}$$

Secondly, τ has a rapidly decaying Fourier Transform. Specifically,

$$|\hat{\tau}(t)| \;\leq\; 16Le^{-(\delta L|t|)^{1/2}}$$

for all real t.

This Theorem has a rather non-trivial Corollary. This is derived from it using the Poisson Summation Formula, and again the details are relegated to an appendix.

Theorem 19 *Let $0 < \delta \leq \frac{1}{16}$ and let $N \geq \frac{1}{\delta}$. Then there is a function $\sigma = \sigma_{N,\delta} : \mathbb{Z} \to \mathbb{Z}$ with the following properties. Firstly, τ approximates the characteristic function of the set $\{-N, \dots, N\}$ in that $0 \leq \sigma \leq 1$ and*

$$\sigma(x) \;=\; \begin{cases} 0 & (|x| > (1+\delta)N) \\ 1 & (|x| < (1-\delta)N). \end{cases}$$

Secondly, σ has a rapidly decaying Fourier Transform. Specifically,

$$|\hat{\sigma}(t)| \;\leq\; 2^9 Ne^{-(\delta N\|t\|)^{1/2}}$$

for all $t \in \mathbb{T}$.

The subtle differences between these two results should be carefully noted. In particular the Fourier Transform is a different object in the two different theorems. Indeed

$$\hat{\tau}(t) \;=\; \int_{-\infty}^{\infty} \tau(x)e^{itx}\,dx,$$

whilst

$$\hat{\sigma}(t) \;=\; \sum_{m\in\mathbb{Z}} \sigma(m)e^{2\pi imt}.$$

Let us now try and replace χ by $\tau = \tau_{\epsilon,\delta}$ in (24) where δ will be chosen later. For good measure (and because it turns out to be crucial) we also replace the characteristic function $[-M, M]$, hardly noticeable in (24), by $\sigma = \sigma_{M,\delta}$. It is trivial to check that the condition $M \geq \frac{1}{\delta}$ required by Theorem 19 is satisfied provided that (12) is satisfied, which we assume is always the case. One has

$$\begin{aligned} \left|\hat{\Lambda}(x) - \sum_n \left(\prod_{j=1}^d \tau(n\theta_j)\right)\sigma(n)e^{2\pi inx}\right| \;&\leq\; \sum_n \left|\Lambda(n) - \sum_n \left(\prod_{j=1}^d \tau(n\theta_j)\right)\sigma(n)\right| \\ &\leq\; \left|\Lambda_{\theta,(1+\delta)\epsilon,(1+\delta)M} \setminus \Lambda_{\theta,(1-\delta)\epsilon,(1-\delta)M}\right|. \end{aligned} \tag{26}$$

Recalling that $\Lambda_{\theta,\epsilon,M}$ is regular, we have that

$$\begin{aligned} \left|\Lambda_{\theta,(1+\delta)\epsilon,(1+\delta)M} \setminus \Lambda_{\theta,(1-\delta)\epsilon,(1-\delta)M}\right| \;&\leq\; ((1+100d\delta) - (1-100d\delta))\,|\Lambda| \\ &\leq\; 200d\delta|\Lambda| \\ &\leq\; \frac{\kappa}{2}|\Lambda| \end{aligned} \tag{27}$$

provided that $\delta \le \frac{\kappa}{400d}$. Note that the condition $|\delta| \le \frac{1}{100d}$ required by the definition of regularity is satisfied automatically here. Hence if $\delta \le \frac{\kappa}{400d}$ then we have, from (26), (27) and the hypothesis that $|\hat{\Lambda}(x)| \ge \kappa|\Lambda|$,

$$\begin{aligned} \left| \sum_n \left(\prod_{j=1}^d \tau(n\theta_j) \right) \sigma(n) e^{2\pi i x n} \right| &\ge \frac{\kappa}{2}|\Lambda|. \\ &\ge \frac{\kappa}{2} M\epsilon^d. \end{aligned}$$

For this last step we used Lemma 6. The left hand side of this expression can be rewritten as follows.

$$\begin{aligned} \left| \sum_n \left(\prod_{j=1}^d \tau(n\theta_j) \right) \sigma(n) e^{2\pi i x n} \right| &= \sum_{\mathbf{k}\in\mathbb{Z}^d} \left(\prod_{j=1}^d |\hat{\tau}(k_j)| \right) \left| \sum_n \sigma(n) e^{2\pi i n(x+\mathbf{k}.\theta)} \right| \\ &= \sum_{\mathbf{k}\in\mathbb{Z}^d} \left(\prod_{j=1}^d |\hat{\tau}(k_j)| \right) \left| \hat{\sigma}\left(\|x+\mathbf{k}.\theta\| \right) \right|. \end{aligned}$$

Therefore we have

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} \left(\prod_{j=1}^d |\hat{\tau}(k_j)| \right) \left| \hat{\sigma}\left(\|x+\mathbf{k}.\theta\| \right) \right| \ge \frac{\kappa}{2} M\epsilon^d. \tag{28}$$

where of course $\tau = \tau_{\epsilon,\delta}$, $\sigma = \sigma_{M,\delta}$ and $\delta \le \frac{\kappa}{400d}$. Now τ and σ were chosen, in Theorems 18 and 19, to have very rapidly decaying Fourier Transforms. Combining (28) with the estimates for these transforms one gets

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} \exp - \left(\delta^{1/2} M^{1/2} \|x+\mathbf{k}.\theta\|^{1/2} + \sum_{j=1}^d (\delta\epsilon)^{1/2} |k_j|^{1/2} \right) \ge \frac{\kappa}{2^{4d+10}}.$$

Recall that (28) was valid for any $\delta \le \frac{\kappa}{400d}$. Taking δ to be as large as possible gives

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} \exp - \left(\left(\frac{\kappa M}{400d} \right)^{1/2} \|x+\mathbf{k}.\theta\|^{1/2} + \left(\frac{\kappa\epsilon}{400d} \right)^{1/2} \sum_{j=1}^d |k_j|^{1/2} \right) \ge \frac{\kappa}{2^{4d+10}}. \tag{29}$$

We will now derive Theorem 17 from (29) in the "obvious" way. Namely we will show that for smallish k_0 the contribution to the sum in (29) from those $\mathbf{k}$ with $|\mathbf{k}| \ge k_0$ is small. This will imply that the sum over those $\mathbf{k}$ with $|\mathbf{k}| \le k_0$ is largeish. From this we will derive Theorem 17, which says that there is $\mathbf{k}$ with both $|\mathbf{k}|$ and $\|x+\mathbf{k}.\theta\|$ small, in quantitative form.

For brevity let us write $\gamma = \frac{\kappa\epsilon}{400d}$. Then, for any k_0, we have

$$\begin{aligned}
\sum_{|\mathbf{k}|\geq k_0} \exp - \left(\left(\frac{\kappa M}{400d}\right)^{1/2} \|x + \mathbf{k}.\theta\|^{1/2} + \gamma^{1/2} \sum_{j=1}^{d} |k_j|^{1/2} \right) \\
&\leq \sum_{|\mathbf{k}|\geq k_0} e^{-\gamma^{1/2}\sum_{j=1}^{d}|k_j|^{1/2}} \\
&= \sum_{|\mathbf{k}|\geq k_0} \left(\prod_{j=1}^{d} e^{-(\gamma|k_j|)^{1/2}} \right) \\
&\leq d \left(\sum_{m=1}^{\infty} e^{-(\gamma m)^{1/2}} \right)^{d-1} \left(\sum_{m=k_0+1}^{\infty} e^{-(\gamma m)^{1/2}} \right). \qquad (30)
\end{aligned}$$

At this point we stop for a quick estimate.

Lemma 20 *Let c be a non-negative real number. Then*

$$\sum_{m=c+1}^{\infty} e^{-(\gamma m)^{1/2}} \;\leq\; \frac{2\left(1 + (\gamma c)^{1/2}\right) e^{-(\gamma c)^{1/2}}}{\gamma}.$$

Proof We use the estimate

$$\sum_{m=c+1}^{\infty} e^{-(\gamma m)^{1/2}} \;\leq\; \int_{c}^{\infty} e^{-(\gamma x)^{1/2}}\, dx.$$

Surprisingly the integral here can be evaluated explicitly as the expression claimed in the statement of the Lemma. □

Returning to (30), we use Lemma 20 twice to get

$$\sum_{|\mathbf{k}|\geq k_0} \exp - \left(\left(\frac{\kappa M}{400d}\right)^{1/2} \|x + \mathbf{k}.\theta\|^{1/2} + \gamma^{1/2} \sum_{j=1}^{d} |k_j|^{1/2} \right) \;\leq\; 2^{d+1} d \left(\frac{400d}{\kappa\epsilon}\right)^{d} e^{-\frac{1}{2}\left(\frac{\kappa\epsilon k_0}{400d}\right)^{1/2}}.$$

In view of (29) our aim now must be to pick k_0 as small as possible so that the right hand side of this last equation is at most $\frac{\kappa}{2^{4d+11}}$. Recalling that $\epsilon \leq 1$, $\kappa \leq 1$ and $d \geq 1$ it can be checked (with a little work) that this is true provided that

$$k_0 \;\geq\; \frac{2^{13} d^3}{\kappa\epsilon} \left(\log \frac{400d}{\kappa\epsilon}\right)^2. \qquad (31)$$

Suppose from now on that k_0 is chosen to equal the value on the right hand side in (31). Then, as we have remarked, it follows from (29) that

$$\sum_{|\mathbf{k}|\leq k_0} \exp - \left(\left(\frac{\kappa M}{400d}\right)^{1/2} \|x + \mathbf{k}.\theta\|^{1/2} + \left(\frac{\kappa\epsilon}{400d}\right)^{1/2} \sum_{j=1}^{d} |k_j|^{1/2} \right) \;\geq\; \frac{\kappa}{2^{4d+11}}. \qquad (32)$$

But

$$\sum_{|\mathbf{k}|\le k_0} \exp - \left(\left(\frac{\kappa M}{400d}\right)^{1/2} \|x+\mathbf{k}.\theta\|^{1/2} + \left(\frac{\kappa\epsilon}{400d}\right)^{1/2} \sum_{j=1}^{d} |k_j|^{1/2} \right)$$

$$\le \exp - \left(\left(\frac{\kappa M}{400d}\right)^{1/2} \min_{|\mathbf{k}|\le k_0} \|x+\mathbf{k}.\theta\|^{1/2} \right) \left(\sum_{m\in\mathbb{Z}} e^{-\left(\frac{\kappa\epsilon m}{400d}\right)} \right)^{d}$$

$$\le \exp - \left(\left(\frac{\kappa M}{400d}\right)^{1/2} \min_{|\mathbf{k}|\le k_0} \|x+\mathbf{k}.\theta\|^{1/2} \right) \cdot 2^d \left(\frac{400d}{\kappa\epsilon}\right)^{d},$$

the last step following by another application of Lemma 20. Hence from (32) we have that, for some $\mathbf{k}$ with $|\mathbf{k}| \le k_0$,

$$\exp - \left(\frac{\kappa M}{400d}\|x+\mathbf{k}.\theta\|\right)^{1/2} \ge \frac{\kappa}{2^{5d+11}}\left(\frac{\kappa\epsilon}{400d}\right)^{d} \ge \left(\frac{\kappa\epsilon}{400d}\right)^{4d}.$$

Hence, for this $\mathbf{k}$,

$$\|x+\mathbf{k}.\theta\| \le \frac{2^{12}d^3}{\kappa M}\left(\log\left(\frac{400d}{\kappa\epsilon}\right)\right)^2.$$

Since $\epsilon < 1/2$ a short computation gives that

$$\log\left(\frac{400d}{\kappa\epsilon}\right) \le \frac{12d}{\kappa}\log\frac{1}{\epsilon}.$$

A little tidying up gives, at last, Theorem 17.

It is perhaps worth remarking at this point that the key feature of Theorem 17 seems to be that $\left(\log\frac{1}{\epsilon}\right)^2$ grows more slowly than any power of $\frac{1}{\epsilon}$ as $\epsilon \to 0$. If this were not the case then we would get a substantially weaker bound for the final result. The key feature of the proof that enables us to obtain such sub-powerlike dependence seems to be that the Fourier Transforms of the functions in Theorems 18 and 19 decay exponentially.

10 Obtaining a Density Increment in the Second Case, Part 2

Recall that we are working under the assumption (23) that

$$\left\|\left(\hat{A}_1 - \delta_1\hat{\Lambda}_1\right)\chi_{\mathcal{F}}\right\|_2 \ge \frac{1}{16}\delta_1^{3/2}|\Lambda_1|^{1/2}, \tag{33}$$

where $\mathcal{F}$ is the set of all $x \in [0,1]$ for which

$$\left|\hat{\Lambda}_2(-2x)\right| \ge \frac{\delta_1}{32}|\Lambda_2|. \tag{34}$$

With a little technical manipulation we can apply Theorem 17, recalling that $\Lambda_2 = \Lambda_{\theta,c_2\epsilon,c_2M}$ where c_2 satisfies (9), to assert the following. There is $\mathbf{k} \in \mathbb{Z}^d$ with

$$\|\mathbf{k}\|_\infty \;\leq\; \frac{2^{165}d^{11}}{\delta^9}\left(\log\frac{1}{\epsilon}\right)^2\frac{1}{\epsilon} \tag{35}$$

such that

$$|-2x+\mathbf{k}.\theta| \;\leq\; \frac{2^{165}d^{11}}{\delta^9}\left(\log\frac{1}{\epsilon}\right)^2\frac{1}{M}. \tag{36}$$

In deriving these estimates we have used the facts that $\epsilon \leq 1/2$ and $\delta_1 > \frac{1}{2}\delta$.

Let η, κ be positive real numbers to be chosen later. If $\Lambda_4 = \Lambda_{\frac{\theta}{2},c_4\epsilon,c_4M}$ for c_4 sufficiently small then Λ_1 looks η-locally like Λ_4. More importantly for $n \in \Lambda_4$ and $x \in \mathcal{F}$ we will have $\|nx\| \leq \kappa$ and so $\left|e^{2\pi inx}-1\right| \leq 2\pi\kappa$. We will explain these statements in the following fully quantitative Proposition.

Proposition 21 (a) *If* $c_4 \leq d^{-2}2^{-39}\delta^2\eta$ *then* Λ_1 *looks* η*-locally like* Λ_4*.*

(b) *If*

$$c_4 \;\leq\; d^{-12}2^{-166}\delta^9\left(\log\frac{1}{\epsilon}\right)^{-2}\kappa$$

then $\|nx\| \leq \kappa$ *for all* $n \in \Lambda_4$ *and* $x \in \mathcal{F}$*.*

Proof (a) Recall from (5) that $\Lambda_1 = \Lambda_{\theta,c_1\epsilon,c_1M}$ where $c_1 \geq d^{-1}2^{-28}\delta^2$. We have

$$\Lambda_4 \;=\; \Lambda_{\frac{\theta}{2},c_4\epsilon,c_4M} \;\subseteq\; \Lambda_{\theta,2c_4\epsilon,2c_4M}.$$

Therefore, by Proposition 9, Λ_1 will look η-locally like Λ_4 if $2c_4 \leq \frac{c_1\eta}{400d}$. This is certainly true if the stated condition holds.

(b) Take $\mathbf{k}$ so that (35) and (36) are satisfied. Then for any n

$$|-2nx+n\mathbf{k}.\theta| \;\leq\; \frac{2^{165}d^{11}|n|}{\delta^9}\left(\log\frac{1}{\epsilon}\right)^2\frac{1}{M}.$$

Hence if $n \in \Lambda_4$ then

$$\begin{aligned}\left|nx-\sum_{j=1}^{d}k_j\frac{n\theta_j}{2}\right| \;&\leq\; \frac{2^{164}d^{11}|n|}{\delta^9}\left(\log\frac{1}{\epsilon}\right)^2\frac{1}{M}\\ &\leq\; \frac{2^{164}d^{11}c_4}{\delta^9}\left(\log\frac{1}{\epsilon}\right)^2.\end{aligned}$$

But

$$\left\|\sum_{j=1}^{d} k_j \frac{n\theta_j}{2}\right\| \;\leq\; c_4\epsilon \sum_{j=1}^{d} |k_j|$$
$$\leq\; c_4 \frac{2^{165} d^{12}}{\delta^9}\left(\log\frac{1}{\epsilon}\right)^2$$

and therefore

$$\|nx\| \;\leq\; \frac{2^{166} d^{12} c_4}{\delta^9}\left(\log\frac{1}{\epsilon}\right)^2,$$

from which the result follows. □

Now we saw in our analysis of the First Case that we can turn a statement about a Fourier Coefficient being large into a density increment by averaging over sets on which suitable exponentials $e^{2\pi inx}$ are nearly constant. In that case we had a single value $x = x_0$ to worry about. Here however we have, in (33), a statement about a whole family of Fourier Coefficients being *on average* large. The beauty of Theorem 17, and of the deductions (35) and (36) that we have made from it, is that we can pick a Bohr Neighbourhood Λ_4 on which *all* the exponentials $e^{2\pi inx}$ of interest are roughly constant. We can use this to derive a density increment from (33) in a manner which is, philosophically at least, the same as our deduction in the First Case.

To carry out this deduction we simply have to "follow our nose". We know that we want to take the expression $\hat{A}_1(x) - \delta_1\hat{\Lambda}_1(x)$, take some kind of average over translates of Λ_4 and then use the fact that certain exponentials are roughly constant on Λ_4. We begin by illustrating this qualitatively (i.e. without any κ's or η's). For $x \in \mathcal{F}$ we have

$$\sum_{n\in\Lambda_1} A(n)e^{2\pi inx} \;\approx\; \frac{1}{|\Lambda_4|}\sum_{m\in\Lambda_1}\sum_{n\in\Lambda_4+m} A(n)e^{2\pi inx}$$
$$\approx\; \frac{1}{|\Lambda_4|}\sum_{m\in\Lambda_1} e^{2\pi imx}\sum_{n\in\Lambda_4+m} A(n)$$
$$=\; \frac{1}{|\Lambda_4|}\sum_{m\in\Lambda_1} |(A-m)\cap\Lambda_4|\, e^{2\pi imx}.$$

Thus

$$\hat{A}_1(x) - \delta_1\hat{\Lambda}_1(x) \;\approx\; \sum_{m\in\Lambda_1} (\delta(m)-\delta_1)\, e^{2\pi imx}, \qquad (37)$$

where here and below we use $\delta(m)$ to denote the density

$$\frac{|(A-m)\cap\Lambda_4|}{|\Lambda_4|}.$$

The plan of the rest of this section is to derive an exact form of (37), and then to derive a density increment from it using our assuption (33) and Parseval's Identity.

Following our qualitative analysis above it is natural to write down the expression

$$
\begin{aligned}
\hat{A}_1(x) - \delta_1 \hat{\Lambda}_1(x) \;=\; & \left(\sum_{n\in\Lambda_1} A(n) e^{2\pi i n x} - \frac{1}{|\Lambda_4|} \sum_{m\in\Lambda_1} \sum_{n\in\Lambda_4+m} A(n) e^{2\pi i n x} \right) \\
& + \left(\frac{1}{|\Lambda_4|} \sum_{m\in\Lambda_1} \sum_{n\in\Lambda_4+m} A(n) \left(e^{2\pi i n x} - e^{2\pi i m x} \right) \right) \\
& + \left(\frac{1}{|\Lambda_4|} \sum_{m\in\Lambda_1} e^{2\pi i m x} \sum_{n\in\Lambda_4+m} A(n) \;-\; \delta_1 \sum_{m\in\Lambda_1} e^{2\pi i m x} \right) \qquad (38)
\end{aligned}
$$

We will refer to the three bracketed expressions here as $E_1(x)$, $E_2(x)$ and $E_3(x)$ respectively. By (33), (38) and Minkowski's Inequality we have

$$\frac{1}{16}\delta_1^{3/2}|\Lambda_1|^{1/2} \;\leq\; \|E_1(x)\chi_{\mathcal{F}}\|_2 + \|E_2(x)\chi_{\mathcal{F}}\|_2 + \|E_3(x)\chi_{\mathcal{F}}\|_2. \qquad (39)$$

We now proceed to estimate the three quantities on the right hand side.

To estimate $\|E_1(x)\chi_{\mathcal{F}}\|_2$ we use the fact that Λ_1 looks η-locally like Λ_4. For this part of the analysis the presence of $\mathcal{F}$ is irrelevant, and we show that $\|E_1(x)\|_2$ is small. Now we have

$$E_1(x) \;=\; \sum_n \frac{1}{|\Lambda_4|} e^{2\pi i n x} \left(|\Lambda_4| - Q(n) \right) A(n),$$

where

$$Q(n) = \left| \{ m \in \Lambda_1 | n - m \in \Lambda_4 \} \right|.$$

Hence, by Parseval's Identity,

$$\|E_1\|_2 \;=\; \frac{1}{|\Lambda_4|} \left(\sum_{n\in A} \left(|\Lambda_4| - Q(n) \right)^2 \right)^{1/2}. \qquad (40)$$

However Λ_1 looks η-locally like Λ_4, and so (by definition)

$$\sum_n |Q(n) - \Lambda_1(n)|\Lambda_4|| \;\leq\; \eta|\Lambda_1||\Lambda_4|.$$

This implies that

$$
\begin{aligned}
\sum_{n\in A} |Q(n) - |\Lambda_4|| \;&\leq\; \sum_{n\in\Lambda_1} |Q(n) - \Lambda_1(n)|\Lambda_4|| \\
&\leq\; \eta|\Lambda_1||\Lambda_4|.
\end{aligned}
$$

Hence, by (40),

$$\begin{aligned}
\|E_1\|_2 &= \frac{1}{|\Lambda_4|}\left(\sum_{n\in A}(|\Lambda_4| - Q(n))^2\right)^{1/2} \\
&\leq \frac{1}{|\Lambda_4|}\left(\sum_{n\in A}|Q(n) - |\Lambda_4||\right)^{1/2}\left(\sup_n |Q(n) - |\Lambda_4||\right)^{1/2} \\
&\leq \frac{1}{|\Lambda_4|}\cdot \eta^{1/2}|\Lambda_1|^{1/2}|\Lambda_4|^{1/2}\cdot (2|\Lambda_4|)^{1/2} \\
&\leq 2\eta^{1/2}|\Lambda_1|^{1/2}. \qquad (41)
\end{aligned}$$

To estimate $\|E_2(x)\chi_{\mathcal{F}}\|_2$ we really do need $\mathcal{F}$. Recall that, for $n \in \Lambda_4$ and $x \in \mathcal{F}$, we have

$$\left|e^{2\pi i n x} - 1\right| \leq 2\pi\kappa.$$

Hence we have

$$\begin{aligned}
E_2(x)\chi_{\mathcal{F}}(x) &= \left(\frac{1}{|\Lambda_4|}\sum_{m\in\Lambda_1}\sum_{n\in\Lambda_4+m}A(n)e^{2\pi i n x} - \frac{1}{|\Lambda_4|}\sum_{m\in\Lambda_1}\sum_{n'\in\Lambda_4}(A-m)(n')e^{2\pi i m x}\right)\chi_{\mathcal{F}}(x) \\
&= \frac{1}{|\Lambda_4|}\sum_{n\in\Lambda_4}\left(e^{2\pi i n' x} - 1\right)\sum_{m\in\Lambda_1}e^{2\pi i m x}(A-m)(n')\chi_{\mathcal{F}}(x)
\end{aligned}$$

and so

$$\begin{aligned}
|E_2(x)\chi_{\mathcal{F}}(x)| &\leq \frac{1}{|\Lambda_4|}\cdot 2\pi\kappa\cdot\sum_{n\in\Lambda_4}\left|\sum_{m\in\Lambda_1}e^{2\pi i m x}(A-m)(n)\right|\chi_{\mathcal{F}}(x) \\
&\leq \frac{2\pi\kappa}{|\Lambda_4|}\sum_{n\in\Lambda_4}\left|\sum_{m\in\Lambda_1}e^{2\pi i m x}(A-m)(n)\right|.
\end{aligned}$$

Taking $\mathcal{L}^2$-norms gives

$$\begin{aligned}
\|E_2(x)\chi_{\mathcal{F}}(x)\|_2^2 &\leq (2\pi\kappa)^2\cdot\frac{1}{|\Lambda_4|}\cdot\sum_{n\in\Lambda_4}\left\|\sum_{m\in\Lambda_1}e^{2\pi i m x}(A-m)(n)\right\|_2^2 \\
&= \frac{(2\pi\kappa)^2}{|\Lambda_4|}\sum_{n\in\Lambda_4}\sum_{m\in\Lambda_1}(A-m)(n) \\
&\leq (2\pi\kappa)^2|\Lambda_1|,
\end{aligned}$$

and so

$$\|E_2(x)\chi_{\mathcal{F}}(x)\|_2 \leq 2\pi\kappa|\Lambda_1|^{1/2}. \qquad (42)$$

Finally we turn to the estimation of $\|E_3(x)\chi_{\mathcal{F}}\|_2$. This is another case in which the $\mathcal{F}$ serves no purpose, and we have in fact

$$\|E_3(x)\|_2 \;=\; \left(\sum_{m\in\Lambda_1}(\delta(m)-\delta_1)^2\right)^{1/2}. \tag{43}$$

Putting together (39), (41), (42) and (43) gives

$$\left(\sum_{m\in\Lambda_1}(\delta(m)-\delta_1)^2\right)^{1/2} \;\geq\; \left(\frac{1}{16}\delta_1^{3/2}-2\eta^{1/2}-2\pi\kappa\right)|\Lambda_1|^{1/2}.$$

Hence if we take c_4 so small that $\eta \leq 2^{-14}\delta_1^3$ and $\kappa \leq 2^{-9}\delta_1^{3/2}$ we will have

$$\sum_{m\in\Lambda_1}|\delta(m)-\delta_1|^2 \;\geq\; 2^{-10}\delta_1^3|\Lambda_1|. \tag{44}$$

Applying Proposition 21 and recalling that $\delta_1 > \frac{\delta}{2}$, it is very easy to see that any c_4 with

$$c_4 \;\leq\; \delta^{11}2^{-177}d^{-12}\left(\log\frac{1}{\epsilon}\right)^{-2} \tag{45}$$

will do. Observe that getting κ small enough is the overriding constraint. By Lemma 12 we can choose some c_4 with

$$c_4 \;\geq\; \delta^{11}2^{-178}d^{-12}\left(\log\frac{1}{\epsilon}\right)^{-2} \tag{46}$$

to satisfy (45) and so that the resulting Λ_4 is regular.

From (44) we are almost home – indeed we can immediately guarantee the existence of an m for which $|\delta(m)-\delta_1| \geq 2^{-5}\delta^{3/2}$. What we need, however, is a *positive* density increment. For this we recall Lemma 14 which, when applied in the present context, gives

$$\left|\sum_{m\in\Lambda_1}(\delta(m)-\delta_1)\right| \;\leq\; \eta|\Lambda_1| \;\leq\; 2^{-14}\delta_1^3|\Lambda_1|. \tag{47}$$

Now from (44) we have

$$\max_{m\in\Lambda_1}|\delta(m)-\delta_1|\sum_{m\in\Lambda_1}|\delta(m)-\delta_1| \;\geq\; 2^{-10}\delta_1^3|\Lambda_1|.$$

Hence either $\delta(m) \geq 2\delta_1$ for some m (representing a quite vast density increment) or else

$$\sum_{m\in\Lambda_1}|\delta(m)-\delta_1| \;\geq\; 2^{-10}\delta_1^2|\Lambda_1|.$$

In this eventuality we use the classic trick that we have alreay seen – by (47) one has

$$\begin{aligned}\sum_{m\in\Lambda_1}\left(|\delta(m)-\delta_1|+\delta(m)-\delta_1\right) &\geq \left(2^{-10}\delta_1^2-2^{-14}\delta_1^3\right)|\Lambda_1| \\ &\geq 2^{-11}\delta_1^2|\Lambda_1|.\end{aligned}$$

It follows that, for some m, we have a density increment

$$\delta(m) \geq \delta_1+2^{-12}\delta_1^2.$$

Recalling from (6) that $\delta_1 \geq \delta - 2^{-13}\delta^2$, a short calculation gives at last a density increment

$$\delta(m) \geq \delta+2^{-14}\delta^2 \tag{48}$$

in the Second Case.

11 Conclusion

Let us now summarise what we have proved in the entire paper so far in the form of a Theorem.

Theorem 22 *Let $\Lambda = \Lambda_{\theta,\epsilon,M}$, where $\theta \in \mathbb{R}^d$, be a regular Bohr Neighbourhood. Suppose $A \subseteq \mathbb{Z}$ has density δ in Λ. then one of the following alternatives holds:*

(i) *A contains a nontrivial 3-term AP;*

(ii) *(Proposition 16 cases (i) and (ii)) There is a regular Bohr Neighbourhood $\Lambda' = \Lambda_{\theta,c\epsilon,cM}$ with $c \geq d^{-1}2^{-28}\delta^2$ on which some translate of A has density at least $\delta+2^{-20}\delta^2$;*

(iii) *(The First Case) There is a regular Bohr Neighbourhood $\Lambda' = \Lambda_{\theta',c\epsilon,cM}$ with $c \geq d^{-3}2^{-59}\delta^4$ and $\theta' \in \mathbb{R}^{d+1}$, on which some translate of A has density at least $\delta+2^{-7}\delta^2$;*

(iv) *(The Second Case) There is a regular Bohr Neighbourhood $\Lambda' = \Lambda_{\frac{\theta}{2},c\epsilon,cM}$ with*

$$c \geq \delta^{11}d^{-12}2^{-178}\left(\log\frac{1}{\epsilon}\right)^{-2}$$

on which some translate of A has density at least $\delta+2^{-14}\delta^2$;

(v) *(Failure of* (12)*)*

$$M\epsilon^d \leq \left(\frac{2^{61}d^2}{\delta^6}\right)^d.$$

This may be summarised more briefly (although less precisely) as follows.

Corollary 23 *Let $\Lambda = \Lambda_{\theta,\epsilon,M}$ be a regular Bohr Neighbourhood with $\theta \in \mathbb{R}^d$, and let $A \subseteq \mathbb{Z}$ have density δ on Λ. If A does not contain a nontrivial 3-term AP then one of the following must hold:*

(i) *There is a regular Bohr Neighbourhood* $\Lambda' = \Lambda_{\theta',\epsilon',M'}$ *on which some translate of* A *has density at least* $\delta + 2^{-20}\delta^2$*, where* $\theta' \in \mathbb{R}^{d'}$ *and*

$$\begin{aligned} d' &\leq d+1 \; ; \\ \epsilon' &\geq \delta^{11} d^{-12} 2^{-178} \left(\log \frac{1}{\epsilon}\right)^{-2} \epsilon \qquad \text{and} \\ M' &\geq \delta^{11} d^{-12} 2^{-178} \left(\log \frac{1}{\epsilon}\right)^{-2} M \; ; \end{aligned}$$

(ii)

$$M\epsilon^d \leq \left(\frac{2^{61} d^2}{\delta^6}\right)^d .$$

Finally we are ready for

Theorem 24 (Effective Version of Roth's Theorem) *Suppose that* $A \subseteq \{-N, \ldots, N\}$ *is a set of density* δ*. If*

$$\delta \;\geq\; 2^{28} \left(\frac{\log\log N}{\log N}\right)^{1/2}$$

then A *contains a nontrivial 3-term AP.*

Proof As we have pointed out the set

$$\Lambda_1 \;=\; [-N, N] \;=\; \Lambda_{1,\frac{1}{2},N}$$

is a regular Bohr Neighbourhood. Indeed

$$\left|\Lambda_{1,\frac{(1+\gamma)}{2},(1+\gamma)N}\right| \;=\; 2\lfloor(1+\gamma)N\rfloor + 1,$$

from which it follows easily that Λ_1 is regular. Assume that A does not contain a nontrivial 3-term AP. If N is large enough then, applying Corollary 23 repeatedly, we get a sequence $\left\{\Lambda_j = \Lambda_{\theta_j,\epsilon_j,M_j}\right\}$ of regular Bohr Neighbourhoods with the following properties:

(i)

$$\theta_j \in \mathbb{R}^{d(j)} \qquad \text{where} \qquad d(j) \leq j \; ;$$

(ii)

$$\epsilon_j \;\geq\; \delta^{11} d^{-12} 2^{-178} \left(\log \frac{1}{\epsilon_{j-1}}\right)^{-2} \epsilon_{j-1} \; ;$$

(iii)

$$M_j \;\geq\; \delta^{11} d^{-12} 2^{-178} \left(\log \frac{1}{\epsilon_{j-1}}\right)^{-2} M_{j-1} \; ;$$

(iv) Some translate of A has density at least δ_j on Λ_j, where the sequence $\{\delta_j\}$ satisfies

$$\delta_j \;\geq\; \delta_{j-1} + 2^{-20}\delta_{j-1}^2.$$

We want N to be large enough so that we can carry on the above process for long enough that $\delta_j > 1$, which will be a clear contradiction. Let us then examine the sequence δ_j. Clearly δ_j will reach 2δ after $\left\lceil \frac{2^{20}}{\delta} \right\rceil$ steps, then 4δ after a further $\left\lceil \frac{2^{20}}{2\delta} \right\rceil$ steps, and so on. Therefore δ_j will be greater than 1 for some j satisfying

$$\begin{aligned} j &\leq 2^{20}\left(\frac{1}{\delta} + \frac{1}{2\delta} + \dots\right) + \log\left(\frac{1}{\delta}\right) \\ &\leq \frac{2^{22}}{\delta}. \end{aligned}$$

Now if $j \leq 2^{22}\delta^{-1}$ then

$$\epsilon_j \;\geq\; 2^{-500}\delta^{23}\left(\frac{1}{\epsilon_{j-1}}\right)^{-2}\epsilon_{j-1} \qquad \text{and}$$

$$M_j \;\geq\; 2^{-500}\delta^{23}\left(\frac{1}{\epsilon_{j-1}}\right)^{-2}M_{j-1}.$$

Now $\epsilon_1 = \frac{1}{2}$ and $M_1 = N$. It is easy to check inductively that, for $j \leq 2^{22}\delta^{-1}$, we have

$$\epsilon_j \;\geq\; 2^{-600j}\delta^{27j} \qquad \text{and}$$

$$M_j \;\geq\; 2^{-600j}\delta^{27j}N.$$

Now we know that $\delta_j > 1$ for some $j \leq \frac{2^{22}}{\delta}$. For such a j we will have $\Lambda_j = \Lambda_{\theta,\epsilon,M}$ where

$$\epsilon \;\geq\; \left(\frac{\delta}{2}\right)^{2^{32}\delta^{-1}} \qquad \text{and}$$

$$M \;\geq\; \left(\frac{\delta}{2}\right)^{2^{32}\delta^{-1}} N.$$

This is a contradiction provided that N is sufficiently large that alternative (ii) of Corollary 23 does not hold, i.e. provided that

$$M\epsilon^j \;\geq\; \left(\frac{2^{61}j^2}{\delta^6}\right)^j.$$

A short calculation shows that this is the case if

$$N \;\geq\; \left(\frac{2}{\delta}\right)^{2^{56}\delta^{-2}}.$$

This concludes the proof. ∎

12 Appendix – Smoothly Approximating Intervals

In this appendix we prove Theorems 18 and 19. We repeat their statements now.

Theorem 25 *Let $0 < \delta \leq \frac{1}{16}$, and let L be a positive real number. Then there is a function $\tau = \tau_{L,\delta} : \mathbb{R} \to \mathbb{R}$ with the following properties. Firstly, τ approximates the characteristic function of the interval $[-L, L]$ in that $0 \leq \tau \leq 1$ and*

$$\tau(x) \ = \ \begin{cases} 0 & (|x| > (1+\delta)L) \\ 1 & (|x| < (1-\delta)L). \end{cases}$$

Secondly, τ has a rapidly decaying Fourier Transform. Specifically,

$$|\hat{\tau}(t)| \ \leq \ 16Le^{-(\delta L|t|)^{1/2}}$$

for all real t.

Theorem 26 *Let $0 < \delta \leq \frac{1}{16}$ and let $N \geq \frac{1}{\delta}$. Then there is a function $\sigma = \sigma_{N,\delta} : \mathbb{Z} \to \mathbb{Z}$ with the following properties. Firstly, τ approximates the characteristic function of the set $\{-N, \dots, N\}$ in that $0 \leq \sigma \leq 1$ and*

$$\sigma(x) \ = \ \begin{cases} 0 & (|x| > (1+\delta)N) \\ 1 & (|x| < (1-\delta)N). \end{cases}$$

Secondly, σ has a rapidly decaying Fourier Transform. Specifically,

$$|\hat{\sigma}(t)| \ \leq \ 2^9 Ne^{-(\delta N\|t\|)^{1/2}}$$

for all $t \in \mathbb{T}$.

We start by constructing a smooth bump function. This construction is presumably standard, but the author has not located it in the literature. It seems rather more natural than the usual construction of smooth bump functions via the pathological function $f(x) = e^{-\frac{1}{x^2}}$. Furthermore it is possible to quantify the smoothness of our function by giving a very strong estimate for its Fourier Transform.

Proposition 27 *There is a non-negative function $F : \mathbb{R} \to \mathbb{R}$ with Supp$(F) \subseteq [-1, 1]$ and $\|F\|_1 = 1$ whose Fourier Transform satisfies the decay estimate $|\hat{F}(\tau)| \leq 2^8|\tau|^{1/2}e^{-|\tau|^{1/2}}$ for $|\tau| \geq 4$.*

Proof Let $I_j = \left[-\frac{1}{4j^2}, \frac{1}{4j^2}\right]$ and let f_j be the characteristic function of I_j weighted so that its integral is 1, i.e. $f_j = 2j^2\chi_{I_j}$. Let $g_k = f_1 * \cdots * f_k$ be the convolution of the first k functions f_j. We have immediately that

$$\|g_k\|_1 \ = \ \prod_{j=1}^{k} \|f_j\|_1 \ = \ 1.$$

Furthermore we can compute

$$\hat{f}_j(\tau) \;=\; \frac{4j^2}{\tau}\sin\left(\frac{\tau}{4j^2}\right),$$

and hence

$$\hat{g}_k(\tau) \;=\; \prod_{j=1}^{k}\frac{4j^2}{\tau}\sin\left(\frac{\tau}{4j^2}\right). \tag{49}$$

Here, as is natural, we have interpreted the value of $\frac{\sin x}{x}$ at $x = 0$ to be 1.

Now if $|x| \le 1$ then

$$\begin{aligned}\left|\frac{\sin x}{x} - 1\right| &= \left|\frac{x^2}{3!} - \frac{x^4}{5!} + \ldots\right| \\ &\le |x|^2.\end{aligned}$$

Therefore

$$\prod_{j=1}^{k}\frac{4j^2}{\tau}\sin\left(\frac{\tau}{4j^2}\right) \;=\; \prod_{j=1}^{k}(1+\epsilon_j)$$

where

$$|\epsilon_j| \;\le\; \frac{\tau^2}{16j^4}.$$

Since $\sum_j |\epsilon_j| < \infty$, the product in (49) converges absolutely and $\hat{g}_k(\tau) \to h(\tau)$ for some function h.

Let us now split the product for $\hat{g}_k$ and write, for any $m \le k$,

$$\begin{aligned}|\hat{g}_k(\tau)| &= \left|\prod_{1\le j<m}\frac{4j^2}{\tau}\sin\left(\frac{\tau}{4j^2}\right)\right|\left|\prod_{m\le j\le k}\frac{4j^2}{\tau}\sin\left(\frac{\tau}{4j^2}\right)\right| \\ &\le \left|\prod_{1\le j<m}\frac{4j^2}{\tau}\sin\left(\frac{\tau}{4j^2}\right)\right| \\ &\le \left(\frac{4}{|\tau|}\right)^m (m!)^2. \end{aligned}\tag{50}$$

Also, rather more trivially, $|\hat{g}_k(\tau)| \le 1$ for all τ.

By putting $m = 2$ in (50) we see that the functions $|\hat{g}_k(\tau)|$ for $k \ge 2$ are simultaneously dominated by an integrable function. It follows from the Bounded Convergence Theorem that

$$\int_{-\infty}^{\infty}\hat{g}_k(\tau)e^{i\tau x}\,d\tau \longrightarrow \int_{-\infty}^{\infty}h(\tau)e^{i\tau x}\,d\tau \tag{51}$$

for all real x. However, by a suitable version of the Fourier Inversion Theorem (see, for example, Körner [4]), we have that

$$\int_{-\infty}^{\infty} \hat{g}_k(\tau) e^{i\tau x}\, d\tau \;=\; 2\pi g_k(x)$$

and so (51) tells us that

$$g_k(x) \longrightarrow \frac{1}{2\pi}\hat{h}(-x)$$

for all real x.

What we have shown so far is that our functions g_k, which we suspect by (50) to have rather rapidly decaying Fourier Transforms for large k, tend pointwise to some function $F(x) = \frac{1}{2\pi}\hat{h}(-x)$. This function, then, would seem to be a natural candidate for our smooth bump function. We prove that it is such a function in two Lemmas.

Lemma 28 *The function F just constructed is non-negative, supported in $[-1,1]$ and has $\|F\|_1 = 1$.*

Proof Non-negativity is clear. It is easy to see that g_k is supported in $[-\eta_k, \eta_k]$, where

$$\eta_k \;=\; \frac{1}{4}\left(1 + \frac{1}{2^2} + \cdots + \frac{1}{k^2}\right) \;\leq\; 1.$$

Hence F is supported in $[-1,1]$ as claimed. Finally we have

$$\begin{aligned} g_k(x) &= f_1 * (f_2 * \cdots * f_k) \\ &= \int_{-\infty}^{\infty} f_1(x-y)(f_2 * \cdots * f_k)(y)\, dy \\ &\leq \|f_1\|_\infty \|f_2 * \cdots * f_k\|_1 \\ &= 2 \end{aligned}$$

for all real x. Since each g_k is supported in $[-1,1]$ it follows from the Bounded Convergence Theorem that $\|g_k\|_1 \to \|F\|_1$. Therefore $\|F\|_1 = 1$. □

During the proof of the last Lemma we saw that the sum of the lengths of the intervals I_j was at most 2. It turns out that one gets the smoothest bump functions by making $\sum_j |I_j|$ converge as slowly as possible. We have tried to do reasonably well, whilst retaining simplicity, by taking $|I_j| = \frac{2}{j^2}$.

Lemma 29 *The Fourier Transform of F satisfies the decay estimate*

$$|\hat{F}(\tau)| \;\leq\; 2^8 |\tau|^{1/2} e^{-|\tau|^{1/2}}$$

for all $|\tau| \geq 4$.

Proof It follows immediately from the Bounded Convergence Theorem that $\hat{g}_k(\tau) \to \hat{F}(\tau)$ pointwise. Therefore we have, from (50), that

$$|\hat{F}(\tau)| \ \le \ \left(\frac{4}{|\tau|}\right)^m (m!)^2 \tag{52}$$

for any positive integer $m = m(\tau)$ we care to choose. A crude version of Sterling's Formula gives the inequality

$$m! \ \le \ 8m^{m+\frac{1}{2}}e^{-m}.$$

Substituting this into (52) yields

$$|\hat{F}(\tau)| \ \le \ 64m\left(\frac{4m^2}{e^2|\tau|}\right)^m \tag{53}$$

for any positive integer m. Supposing that $|\tau| \ge 4$, take $m = \left\lfloor \frac{1}{2}|\tau|^{1/2} \right\rfloor$. Writing $m = \frac{1}{2}|\tau|^{1/2} - \eta$, where $0 \le \eta \le 1$, we get

$$\begin{aligned} |\hat{F}(\tau)| \ &\le \ 64 \cdot \frac{1}{2}|\tau|^{1/2} \cdot \left(\frac{4}{e^2|\tau|}\right)^{\left(\frac{1}{2}|\tau|^{1/2}-\eta\right)} \left(\frac{1}{4}|\tau|\right)^{\left(\frac{1}{2}|\tau|^{1/2}-\eta\right)} \\ &= \ 32|\tau|^{1/2} \cdot e^{2\eta} \cdot e^{-|\tau|^{1/2}} \\ &\le \ 2^8|\tau|^{1/2}e^{-|\tau|^{1/2}}. \end{aligned}$$

Proposition 27 is proved. □

Most of the hard work has now been done and we can prove Theorem 25 relatively quickly. First of all define the function G_δ by $G = F_\delta * \chi_{[-1,1]}$, where $F_\delta(x) = \frac{1}{\delta}F\left(\frac{x}{\delta}\right)$. We have $\hat{F_\delta}(\tau) = \hat{F}(\delta\tau)$ and $\hat{\chi}_{[-1,1]}(\tau) = \frac{2\sin\tau}{\tau}$. Therefore if $\delta|\tau| \ge 4$ we have

$$\begin{aligned} |\hat{G}_\delta(\tau)| \ &= \ |\hat{F}(\delta\tau)|\left|\frac{2\sin\tau}{\tau}\right| \\ &\le \ 2^9\left(\frac{\delta}{|\tau|}\right)^{1/2} e^{-(\delta|\tau|)^{1/2}} \\ &\le \ 2^8\delta e^{-(\delta|\tau|)^{1/2}} \\ &\le \ 16e^{-(\delta|\tau|)^{1/2}}, \end{aligned} \tag{54}$$

the last step following from the assumption that $\delta \le \frac{1}{16}$. Since trivially $|\hat{G}_\delta(\tau)| \le 2$ for all τ, it is easy to see that the bound in (54) is actually valid for all real τ. Now set $\tau_{\delta,L}(x) = G_\delta\left(\frac{x}{L}\right)$ to finish off the proof of Theorem 25. □

Finally we turn to the proof of Theorem 26. It turns out that we have essentially constructed $\sigma = \sigma_{\delta,N}$ already. Define σ as a function on $\mathbb{Z}$ by $\sigma(n) = G_\delta\left(\frac{n}{N}\right)$, and extend it to a function ρ on

$\mathbb{R}$ in the obvious way by defining $\rho(x) = G_\delta\left(\frac{x}{N}\right)$. The Fourier Transform $\hat{\sigma}$ lives on $\mathbb{T}$ and is defined by

$$\hat{\sigma}(\tau) \;=\; \sum_n \sigma(n) e^{2\pi i n\tau}.$$

This certainly looks like some sort of discrete approximation to the Fourier Transform of ρ, namely

$$\hat{\rho}(\tau) \;=\; \int_{-\infty}^{\infty} \sigma(x) e^{2\pi i x\tau}\, dx.$$

We know from earlier investigations that this is rapidly decaying. However it still seems rather unreasonable to me that such a vague comparison can be made completely explicit, as we now show using the Poisson Summation Formula. Let us start by recalling the version of the Poisson Summation Formula given in [4].

Theorem 30 *[Poisson Summation Formula] Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a continuous function for which $\sum_{m=-\infty}^{\infty} |\hat{f}(m)|$ converges and $\sum_{n=-\infty}^{\infty} |f(2\pi n + x)|$ converges uniformly on $[-\pi, \pi]$. Then for any real x we have*

$$\sum_{m\in\mathbb{Z}} \hat{f}(m) e^{imx} \;=\; 2\pi \sum_{n\in\mathbb{Z}} f(2\pi n + x).$$

The proof of this result is not hard, and essentially consists of checking that both sides of the claimed equation have the same Fourier Coefficients when considered as functions on $\mathbb{T}$. The Poisson Summation Formula can be better understood in a fully general context in which arbitrary locally compact abelian groups are considered (as opposed to the rather simple group $\mathbb{R}$, with its discrete subgroup $\mathbb{Z}$, as appear here).

Let us apply the Poisson Summation Formula with $f = \hat{\rho}$. By the Fourier Inversion Formula we have

$$\hat{\hat{\rho}}(x) \;=\; 2\pi\rho(-x) \;=\; 2\pi\rho(x).$$

This and our earlier estimates make it clear that the conditions Theorem 30 are satisfied in this case. Applying it, and recalling that $\sigma = \rho$ on $\mathbb{Z}$, we have

$$\begin{aligned}
\hat{\sigma}(x) &= \sum_{m\in\mathbb{Z}} \sigma(m) e^{2\pi i m x} \\
&= \sum_{m\in\mathbb{Z}} \rho(m) e^{2\pi i m x} \\
&= \frac{1}{2\pi} \sum_{m\in\mathbb{Z}} \hat{\hat{\rho}}(m) e^{2\pi i m x} \\
&= \sum_{n\in\mathbb{Z}} \hat{\rho}\left(2\pi(n + x)\right) \\
&= N \sum_{n\in\mathbb{Z}} \hat{G}_\delta\left(2\pi N(n + x)\right) \\
&= N \sum_{n\in\mathbb{Z}} \hat{G}_\delta\left(2\pi N(n + \|x\|)\right). \qquad (55)
\end{aligned}$$

We know that $\hat{\rho}$ is rapidly decaying, and so the dominant terms are those with with $|n| \leq 1$. The sum of these terms outweighs the sum of all the other terms. We conclude our exposition by working out the details. From (55) and our estimates for $\hat{G}_\delta$ we have

$$\begin{aligned} |\hat{\sigma}(x)| &= N \sum_{n \in BbbZ} \hat{G}_\delta \left(2\pi N(n + \|x\|)\right) \\ &\leq 16N \sum_{n \in \mathbb{Z}} e^{-(2\pi\delta N|n+\|x\||)^{1/2}} \\ &\leq 16N \sum_{n \in \mathbb{Z}} e^{-(\delta N|n+\|x\||)^{1/2}} \\ &\leq 16N \left(5e^{-(\delta N\|x\|)^{1/2}} + 2\sum_{m \geq 2} e^{-(\delta N m)^{1/2}} \right). \end{aligned} \tag{56}$$

Now from Lemma 20 we have, if $\gamma \geq 1$, that

$$\begin{aligned} \sum_{m \geq 2} e^{-(\gamma m)^{1/2}} &\leq \frac{2(1+\gamma^{1/2})e^{-\gamma^{1/2}}}{\gamma} \\ &\leq 8e^{-\frac{3}{4}\gamma^{1/2}} \\ &\leq 8e^{(\gamma\|x\|)^{1/2}} \end{aligned}$$

for any real number x. Now we assumed that $N \geq \frac{1}{\delta}$ and so we may combine this with (56) to get that

$$|\hat{\sigma}(x)| \leq 2^9 N e^{-(\delta N\|x\|)^{1/2}}$$

for all x. This concludes the proof of Theorem 26. □

References

[1] Bourgain, J. *On Triples in Arithmetic Progression*, To appear in Geometric and Functional Analysis.

[2] Croft, H.T. *Some Problems*, Eureka 1968

[3] Gowers, W.T. *Additive and Combinatorial Number Theory*, Part III Course 1999.

[4] Körner, T.W. *Fourier Analysis*, Part III Course 1999.