GAFA, Geom. funct. anal.
Vol. 12 (2002) 584 – 597
1016-443X/02/030584-14 $ 1.50+0.20/0

© Birkhäuser Verlag, Basel 2002

GAFA Geometric And Functional Analysis

ARITHMETIC PROGRESSIONS IN SUMSETS

B. Green

Abstract

We prove several results concerning arithmetic progressions in sets of integers. Suppose, for example, that α and β are positive reals, that N is a large prime and that $C, D \subseteq \mathbb{Z}/N\mathbb{Z}$ have sizes γN and δN respectively. Then the sumset $C + D$ contains an AP of length at least $e^{c\sqrt{\log N}}$, where $c > 0$ depends only on γ and δ. In deriving these results we introduce the concept of hereditary non-uniformity (HNU) for subsets of $\mathbb{Z}/N\mathbb{Z}$, and prove a structural result for sets with this property.

0 Notation

Let $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$ be the group of residues modulo N, where N is an odd prime. We write $\omega = e^{2\pi i/N}$; although this quantity depends on N, there is no danger of confusion as we will always work with a fixed N. If $f : \mathbb{Z}_N \to \mathbb{C}$ is any function then we define its rth Fourier coefficient, $\hat{f}(r)$, by

$$\hat{f}(r) = \sum_x f(x)\omega^{rx}.$$

If $A \subseteq \mathbb{Z}_N$ then we also use the letter A to denote the characteristic function of A, defined by $A(x) = 1$ if $x \in A$ and $A(x) = 0$ if $x \notin A$. If A, B are two sets then we define their *convolution* $A * B$ by

$$(A * B)(x) = \sum_y A(y)B(x - y)\,.$$

This is simply the number of solutions to $a + b = x$. We define the sumset $A + B$ to be the set containing everything of the form $a + b$ with $a \in A$, $b \in B$. Observe that the functions $(A + B)(x)$ and $(A * B)(x)$ are very different in general, although they have the same support.

The author is supported by an EPSRC research grant, and has also enjoyed the hospitality of Princeton University for some of the period during which this work was carried out.

Our definitions of Fourier transform and convolution are (perhaps up to the occasional sign) very standard and so we leave it to the reader to work out Parseval's identity and the interaction of Fourier transforms and convolutions, both of which will feature in what follows.

1 Introduction

Around ten years ago several papers appeared in which the progressive enrichment of structure resulting from repeated set addition was studied. Bourgain, for example, proved the following in [Bou1].

Theorem 1 (Bourgain)**.** *Let $C, D \subseteq \mathbb{Z}$ have cardinalities γN and δN respectively. Then there is an absolute constant $c > 0$ such that $C + D$ contains an AP of length at least*

$$\exp\left(c((\gamma\delta \log N)^{1/3} - \log\log N)\right).$$

Freiman, Halberstam and Ruzsa [FHR], building on a technique of Bogolubov [Bo], showed that even longer progressions result when one adds three or more sets together. For example,

Theorem 2 (Freiman–Halberstam–Ruzsa)**.** *Let $A \subseteq \mathbb{Z}_N$ have $|A| = \alpha N$. Then $A + A + A$ contains an AP of length at least $c\alpha N^{c\alpha^3}$.*

In this paper we improve both of these results. First the technically detailed statements:

Theorem 3. *Let $C, D \subseteq \mathbb{Z}$ have cardinalities γN and δN respectively. Then there is an absolute constant $c > 0$ such that $C + D$ contains an AP of length at least*

$$\exp\left(c((\gamma\delta \log N)^{1/2} - \log\log N)\right).$$

Theorem 4. *Let $A \subseteq \mathbb{Z}_N$ have $|A| = \alpha N$. Then $A + A + A$ contains an AP of length at least*

$$2^{-24}\alpha^5\big(\log(1/\alpha)\big)^{-2}N^{\alpha^2/250\log(1/\alpha)}. \tag{1}$$

As we have stated them, Theorems 3 and 4 look rather technical. Stating the theorems in a weaker form removes this illusion. In Theorem 3, regard γ and δ as fixed. Then the theorem says that $C + D$ contains a progression of length $e^{c(\log N)^{1/2}}$, compared to Bourgain's $e^{c(\log N)^{1/3}}$. In Theorem 4, the quantity (1) looks rather less fearsome when written, crudely, as $N^{\alpha^{2+\epsilon}}$.

Theorems 3 and 4 may be contrasted with the best known lower bounds for these questions. Regarding Theorem 3, Ruzsa [Ru] gave an ingenious

construction of a set $A \subseteq \mathbb{Z}_N$ with $|A| = \frac{1}{2} - \epsilon$, but with $A + A$ not containing any APs of length $\exp(C_\epsilon(\log N)^{2/3+\epsilon})$. As for Theorem 4, a relatively straightforward construction (see [FHR]) shows that $A + A + A$ need not contain an AP of length $2N^{\log(1/\alpha)}$. It is conjectured in [FHR] that this is close to the truth.

It turns out that Theorem 3 gives information about van der Waerden numbers. If $h \geq 2$ and $l_1, \ldots, l_h \geq 3$ are integers then write $W(h; l_1, \ldots, l_h)$ for the smallest integer n such that, however we partition $\{1, \ldots, n\}$ into h colours $C_1, \ldots, C_h$, there is some j such that C_j contains an AP of length l_j. The existence of W is non-trivial, and upper bounds for these numbers are notoriously difficult to come by (see [G] for example). The most famous instances of this problem are probably those concerning $W(h; 3, 3, \ldots, 3)$ and $W(2; l, l)$. In this paper we consider the quantity $W(2; 3, k)$, which seems to have been given rather less attention in the literature. This is a touch surprising as the study of the corresponding Ramsey number, $R(3, k)$, has proved to be very fruitful. Perhaps the reason for this neglect is that the best bounds currently known for $W(2; 3, k)$ follow trivially from Roth's theorem, which we state now. (The reader may recognise this result as the special case $k = 3$ of Szemerédi's theorem.)

Theorem 5 (Roth)**.** *Let $\delta > 0$ be a real number. Then there is a minimal $N_3(\delta)$ with the following property. If $N \geq N_3(\delta)$ and if $A \subseteq \{1, \ldots, N\}$ has size at least δN then A contains an AP of length 3.*

The best bound currently known is $N_3(\delta) \leq e^{C\delta^{-2}\log(1/\delta)}$, due to Bourgain [Bou2]. It is rather easy to see that $W(2; 3, k) \leq N_3\left(k^{-1}\right)$. Indeed colour $\{1, \ldots, N\}$ red and blue. Then either the set of red numbers has density at least k^{-1}, guaranteeing a red AP of length 3, or else the set of blue numbers has density at least $1 - k^{-1}$, guaranteeing a blue AP of length k for trivial reasons. Thus we have

Theorem 6 (Bourgain)**.** $W(2; 3, k) \leq e^{Ck^2 \log k}$.

Our arguments constitute a much simpler proof of a statement which is weaker than this only in the power of $\log k$. Furthermore this seems to be the first occasion on which a strong bound for a series of van der Waerden numbers has been given without establishing a bound for the corresponding density problem.

We close this section with a brief outline of the rest of the paper. In §2 we introduce the concept of hereditary non-uniformity for subsets of $\mathbb{Z}/N\mathbb{Z}$. We state a result about such sets, Theorem 7, the proof of which forms the main substance of this paper. In §3 we show how this implies

Theorem 3. The next two sections are devoted to proofs. In §4 we assemble some important tools. In §5, which may be regarded as the heart of the paper, we prove Theorem 7. In §6 we introduce the concept of restricted hereditary non-uniformity. We state and prove a structure theorem for sets with this property, and show how it implies Theorem 4. In §7 we deduce the bound for $W(2;3,k)$.

2 HNU Sets and Their Structure

Let $A \subseteq \mathbb{Z}_N$ and let $\alpha \in (0,1)$ be a real number. We say that A is α-hereditarily non-uniform, which we shall abbreviate as α-HNU, if for every subset $S \subseteq A$ one has

$$\sup_{r \neq 0} |\hat{S}(r)| \geq \alpha |S| \,.$$

The main result of this paper is the following theorem about HNU sets.

Theorem 7. *Suppose that $\alpha \geq 4000 \log\log N/(\log N)^{1/2}$ and that A is α-HNU. Then A^c, the complement of A, contains an AP of length at least $e^{\alpha\sqrt{\log N}/32}$.*

We will prove this in a short while, but our first priority is to motivate the definition of HNU.

3 Sumsets

Theorem 3 follows immediately from Theorem 7 and the following.

PROPOSITION 8. *Let $C, D \subseteq \mathbb{Z}_N$ have $|C| = \gamma N$ and $|D| = \delta N$. Let A be the complement of $C + D$. Then A is $\sqrt{\gamma\delta}$-HNU.*

Proof. If $S \subseteq A$ then $\sum_x S(x)(C * D)(x) = 0$. Writing this in terms of Fourier coefficients gives

$$\sum_r \hat{S}(r)\overline{\hat{C}(r)\hat{D}(r)} = 0 \,,$$

from which we get

$$\sum_{r \neq 0} |\hat{S}(r)||\hat{C}(r)||\hat{D}(r)| \geq |S||C||D|$$

by the triangle inequality. From this we get

$$|S||C||D| \leq \sup_{r \neq 0} |\hat{S}(r)| \sum_r |\hat{C}(r)||\hat{D}(r)|$$

$$\leq \sup_{r\neq 0}|\hat{S}(r)|\Big(\sum_r |\hat{C}(r)|^2\Big)^{1/2}\Big(\sum_r |\hat{D}(r)|^2\Big)^{1/2}$$
$$= \sup_{r\neq 0}|\hat{S}(r)|\cdot(\gamma\delta)^{1/2}N^2\,.$$

The claim follows immediately. □

4 Tools

One of the most important ingredients of our arguments is the following large deviation inequality of Bernstein. The original paper [B] of Bernstein dates from 1924, and the result has been discussed in many other works since then. The reader may find a proof of the following variant in [Gr1].

LEMMA 9. *Let $X_1,\dots,X_n$ be independent complex-valued random variables with $\mathbb{E}X_i = 0$ and $\mathbb{E}\,|X_j|^2 = \sigma_j^2$. Write $\sigma^2 = \sigma_1^2+\cdots+\sigma_n^2$, and suppose that $|X_j|\leq 1$ uniformly in j. Suppose that $\sigma^2 \geq 6nt$. Then we have the inequality*

$$\mathbb{P}\big(|\overline{X}|\geq t\big)\leq 4e^{-n^2t^2/8\sigma^2}\,.$$

Our second prerequisite concerns so-called Bohr neighbourhoods. Define a function $|.|:\mathbb{Z}_N\to\mathbb{R}$ as follows. Select a residue x' in $\{-N/2,\dots,N/2\}$ which is congruent to x modulo N and define $|x|$ to be $|x'/N|$. Now let $\Gamma\subseteq\mathbb{Z}_N$ and let $\epsilon>0$ be a real number. We write $\mathcal{B}(\Gamma,\epsilon)$ for the set of all $x\in\mathbb{Z}_N$ for which $|\gamma x|\leq\epsilon$ for all $\gamma\in\Gamma$. The letter $\mathcal{B}$ is for H. Bohr; we call such an object a *Bohr neighbourhood*. The following lemma may be proved by a straightforward application of the pigeonhole principle:

LEMMA 10. *Suppose that $|\Gamma| = d$. Then $\mathcal{B}(\Gamma,\epsilon)$ contains an AP of length at least $\epsilon N^{1/d}$.*

The final result to be discussed in this section is the following theorem of Chang [C] concerning the structure of the points at which a set can have large Fourier coefficients.

LEMMA 11 (Chang). *Let $B\subseteq\mathbb{Z}_N$ have cardinality βN and let $\mathcal{R}$ be the set of all $r\in\mathbb{Z}_N$ for which $|\hat{B}(r)|\geq\rho|B|$. Then there is a set Λ, $|\Lambda|\leq 250\rho^{-2}\log(1/\beta)$, such that $\mathcal{R}\subseteq\mathcal{E}(\Lambda)$.*

The notation $\mathcal{E}(\Lambda)$ refers to the set of all $\{-1,0,1\}$-linear combinations of elements of Λ. That is, if $\Lambda=\{\lambda_1,\dots,\lambda_m\}$ then $\mathcal{E}(\Lambda)$ consists of everything of the form $\sum_j\varepsilon_j\lambda_j$ with $\varepsilon_j\in\{-1,0,1\}$ for all j.

In [C] this lemma is derived from an inequality of Rudin [R]. A full proof of Lemma 11 including all necessary background may be found in

[Gr2], and in [Gr3] examples are given which show that it is sharp.

5 The Structure of HNU Sets

Our aim is to prove Theorem 7, which states that if A is α-HNU then A^c contains a long AP.

The basic method of attack will be roughly as follows. Let β be a parameter to be chosen later. Let $B \subset A$ be such that $\sup_{r\neq 0}|\hat{B}(r)|$ is minimised subject to $|B| = \beta N$. We will attempt to produce a "better" set B' by removing t random elements of B and adding t random elements of $\mathbb{Z}_N$. In general this will not be a subset of A but we give a procedure for "deforming" B' so that $B' \subseteq A$. It turns out that such a deformation is possible unless A^c contains a long AP.

Let, then, $B \subseteq A$ have $\sup_{r\neq 0}|\hat{B}(r)|$ minimal subject to $|B| = \beta N$. Let the value of this minimum be $\eta|B|$ and observe that $\eta \geq \alpha$ because A is α-HNU.

Let

$$\mathcal{R} = \left\{r : |\hat{B}(r)| \geq \eta|B|/2\right\}.$$

LEMMA 12. *The Bohr neighbourhood $\mathcal{B}(\mathcal{R}, \eta/64)$ contains an AP, P, of length at least*

$$\frac{\eta^3}{2^{14}\log(1/\beta)} N^{\eta^2/250\log(1/\beta)}.$$

Proof. By Lemma 11 every element of $\mathcal{R}$ is in the ± 1-linear span of a set Λ of size at most $m = 250\eta^{-2}\log(1/\beta)$. Therefore

$$\mathcal{B}(\Lambda, \eta/64m) \subseteq \mathcal{B}(\mathcal{R}, \eta/64),$$

so the result follows from Lemma 10. □

LEMMA 13. *For at least $(1-\eta/16)N$ values of x we have*

$$\left|(x+P)\cap B\right| \leq \tfrac{16\beta}{\eta}|P|.$$

Proof. Suppose not. Then we would have

$$\begin{aligned} |P||B| &= \sum_x \left|(x+P)\cap B\right| \\ &> \tfrac{\eta N}{16}\cdot\tfrac{16\beta}{\eta}|P| \\ &= |P||B|, \end{aligned}$$

a contradiction. □

Call the set C of such x *good*; the above lemma tells us that $|C| \geq (1-\eta/16)N$. As C is very large, it cannot have any really huge Fourier coefficients. Indeed if $r \neq 0$ then

$$\begin{aligned} |\hat{C}(r)| &= |\hat{C^c}(r)| \\ &\leq |C^c| \\ &\leq \eta N/16 \\ &\leq \eta|C|/8\,. \end{aligned} \tag{2}$$

We are now going to choose a subset $D \subseteq C$ of size t (where t, as with so many other variables, will be chosen later). We will do this by picking elements of C at random with probability $t/|C|$. It turns out that, provided t is large enough, D inherits from C the property of not having any really large Fourier coefficients.

LEMMA 14. *Let $t \geq 2^{14}\eta^{-2}\log N$. Then there is a subset $D \subseteq C$ with size t such that $\sup_{r\neq 0}|\hat{D}(r)| \leq \eta t/4$.*

Proof. As promised, choose a set $E \subseteq C$ by letting each $x \in C$ be in E with probability $p = t/|C|$, these choices being independent. The Fourier coefficient $\hat{E}(r)$ is then a sum of $|C|$ independent random variables $X_j^{(r)} = E(x)\omega^{rx}$ with means $p\hat{C}(r)$ and variances at most p. It follows from Lemma 9 and (2) that

$$\begin{aligned} \mathbb{P}\big(|\hat{E}(r)| \geq \eta t/6\big) &\leq \mathbb{P}\big(|\hat{E}(r) - \mathbb{E}\hat{E}(r)| \geq \eta t/24\big) \\ &< 4e^{-\eta^2 t/5000}. \end{aligned}$$

By the same token

$$\mathbb{P}\big(||E| - t| \geq \eta t/24\big) < 4e^{-\eta^2 t/5000}\,.$$

If $t \geq 2^{14}\eta^{-2}\log N$, then, there is a positive probability of all the above events happening. By adding or deleting at most $\eta t/12$ elements from E we get a set D satisfying the conclusion of the lemma. □

An almost identical argument proves the following.

LEMMA 15. *Let $\beta N \geq t \geq 2^{14}\eta^{-2}\log N$. Then there is a subset $X \subseteq B$ with $|X| = t$ and*

$$\left|\hat{X}(r) - \tfrac{t\hat{B}(r)}{|B|}\right| \leq \eta t/12$$

for all $r \neq 0$.

LEMMA 16. *Let S be the (multi)set $(B \setminus X) \cup D$. Then*

$$\sup_{r\in\mathcal{R}}|\hat{S}(r)| \leq \eta|S| - \eta t/6\,,$$

whilst

$$|\hat{S}(r)| \leq \tfrac{\eta|S|}{2} + \tfrac{\eta t}{3}$$

for all other $r \neq 0$.

Proof. We have

$$\begin{aligned}\hat{S}(r) &= \hat{B}(r) - \hat{X}(r) + \hat{D}(r)\\ &= \left(1 - \tfrac{t}{|B|}\right)\hat{B}(r) + Q\,,\end{aligned}$$

where $|Q| \leq \eta t/3$ by the previous two lemmas. If $r \in \mathcal{R}$ then $|\hat{B}(r)|/|B| \geq \eta/2$ by definition, and the first part of the result follows easily. For the second part of the result observe that if $r \notin \mathcal{R}$ then

$$\begin{aligned}|\hat{S}(r)| &\leq \left|1 - \tfrac{t}{|B|}\right| |\hat{B}(r)| + |Q|\\ &\leq \tfrac{\eta|S|}{2} + \tfrac{\eta t}{3}\,.\end{aligned}$$

This proves the lemma. □

Now let $D = \{d_1, \ldots, d_t\}$. Let D' be any set obtained by replacing d_j $(j = 1, \ldots, t)$ with $d_j + x_j$, where $x_j \in P$ (now might be a good opportunity for the reader to recall the definition of P).

LEMMA 17. *Suppose that* $t \leq \eta\beta N/10$. *Let* S' *be the* (*multi*)*set* $(B \setminus X) \cup D'$. *Then* $\sup_{r\neq 0} |\hat{S}'(r)| < \eta|S'|$.

Proof. We deal first with the easy case $r \notin \mathcal{R}$. However we change the elements of D the contribution to $\hat{S}(r)$ cannot vary by more than $2t$. It follows from Lemma 16 that, for $r \notin \mathcal{R}$,

$$|\hat{S}'(r)| \leq \tfrac{\eta|S'|}{2} + 5t < \eta|S'|\,.$$

Now suppose that $r \in \mathcal{R}$, and recall that $P \subseteq \mathcal{B}(\mathcal{R}, \eta/64)$. We have that

$$\begin{aligned}\left|\hat{S}'(r) - \hat{S}(r)\right| &\leq \sum_{j=1}^{t} \left|\omega^{r(d_j+x_j)} - \omega^{rd_j}\right|\\ &\leq t \sup_j |\omega^{rx_j} - 1|\\ &\leq \eta t/8\,.\end{aligned}$$

The result follows immediately from Lemma 16. □

If we could choose $x_1, \ldots, x_t$ so that S' was actually a set (as opposed to a multiset) and also so that $S' \subseteq A$ then we would have a contradiction of our earlier assumption about the minimality of B. It follows that there is no such choice of $x_1, \ldots, x_t$.

LEMMA 18. *There is some j such that $d_j + P$ is contained in $B \cup A^c$, except for at most t elements, which could lie in $A \setminus B$.*

Proof. Suppose not. Choose $x_1 \in P$ so that $d_1 + x_1 \in A \setminus B$. Choose $x_2 \in P$ so that $d_2 + x_2 \in A \setminus (B \cup \{d_1 + x_1\})$. Continue in this way; at the last stage we will still be able to choose $x_t \in P$ so that

$$d_t + x_t \in A \Big\backslash \Big(B \cup \bigcup_{j=1}^{t-1} \{d_j + x_j\} \Big).$$

This gives us an S' of the type that we argued couldn't exist. The lemma follows. □

Let $j \in [t]$ be such that the conclusion of this lemma holds. We are now closing in on a structure theorem for A^c. There is one crucial fact that we have yet to use – the fact that d_j lies in C, so that

$$\big|(d_j + P) \cap B\big| \leq \tfrac{16\beta}{\eta}|P|\,.$$

Lemma 18 now tells us that A^c contains a proportion at least $\big(1 - \frac{16\beta}{\eta}\big)$ of $d_j + P$, minus a possible t points. If we choose parameters so that

$$t \leq \tfrac{16\beta}{\eta}|P| \tag{3}$$

then A^c will in fact contain a proportion $\big(1 - \frac{32\beta}{\eta}\big)$ of $d_j + P$. Recall that at earlier stages we also required

$$t \geq 2^{14}\eta^{-2}\log N \tag{4}$$

and

$$t \leq \eta\beta N/10\,. \tag{5}$$

Let us suppose that parameters have also been chosen so that

$$|P| \geq \tfrac{\eta}{32\beta}\,. \tag{6}$$

Then it is very easy to see that any set, such as A^c, which contains more than $\big(1 - \frac{32\beta}{\eta}\big)$ of P must in fact contain a progression of length at least $\frac{\eta}{64\beta}$.

Before we actually do our big choosing of parameters, there is one very important remark to be made. This is the remark that if $|A| < \beta N$ then it is impossible to even define B. However in this case A^c contains a progression of length at least $1/\beta > \eta/64\beta$ anyhow. Summarising then, we have

PROPOSITION 19. *Suppose that* (3), (4), (5) *and* (6) *are all satisfied. Then A^c contains an AP of length at least $\eta/64\beta$.*

A tedious calculation using Lemma 12 shows that one can take $t = 2^{14}\eta^{-2}\log N$ and

$$\beta = e^{-\eta\sqrt{\log N}/16},$$

and a further tedious calculation (recalling that $\eta \geq \alpha$) shows that this implies Theorem 7. Although we would not wish such calculations on anyone, we should like to draw the reader's attention the fact that it is not difficult to verify a bound of the correct *form*. Assigning explicit values to the constants is tedious.

6 RHNU Sets and Their Structure. Progressions in $A + A + A$

In this section we introduce the concept of restricted hereditary non-uniformity (RHNU) and prove a structure theorem for sets with this property. Despite appearing rather technical, the RHNU condition turns out to be substantially easier to work with than HNU.

Let $A \subseteq \mathbb{Z}_N$, let $\alpha \in (0,1)$ be a real number and let $F, G \subseteq \mathbb{Z}_N \setminus \{0\}$. We say that A is (α, F, G)-restricted HNU, which we shall abbreviate as (α, F, G)-RHNU, if $G \subseteq \mathcal{E}(F)$ and for every subset $S \subseteq A$ one has

$$\sup_{r\in G} |\hat{S}(r)| \geq \alpha|S|\,.$$

Theorem 20. *Suppose that A is (α, F, G)-RHNU. Then*

(i) *A^c contains a translate of $\mathcal{B}(F, \alpha/20|F|)$, minus at most $576\alpha^2\log|G|$ points;*

(ii) *A^c contains an AP of length at least $2^{-14}\alpha^3 N^{1/|F|}/|F|\log|G|$.*

Before proving this theorem we show how Theorem 4 may be deduced from it.

Proposition 21. *Let $A \subseteq \mathbb{Z}_N$ have $|A| = \alpha N$. Let C be the complement of $A + A + A$. Then there is a set F, $|F| \leq 250\alpha^{-2}\log(1/\alpha)$, and a set G, $|G| \leq \alpha^{-3}$, such that C is (α, F, G)-RHNU.*

Proof. If $S \subseteq C$ then $\sum_x S(x)(A * A * A)(x) = 0$. Writing this in terms of Fourier coefficients and using the triangle inequality and Parseval's identity just as in §3, we get

$$\sup_{r\neq 0} |\hat{A}(r)||\hat{S}(r)| \geq \alpha|A||S|\,.$$

Thus $|\hat{S}(r)| \geq \alpha|S|$ for some $r \neq 0$ such that $|\hat{A}(r)| \geq \alpha|A|$. Let G be the set of all $r \neq 0$ such that $|\hat{A}(r)| \geq \alpha|A|$. Paresval's identity shows that

$|G| \leq \alpha^{-3}$. Furthermore Lemma 11 tells us that G is contained in $\mathcal{E}(F)$ for some set F of size at most $250\alpha^{-2}\log(1/\alpha)$. The proposition follows immediately. □

Theorem 20, part (ii), applies and we see that $A+A+A$ contains a progression of length at least

$$2^{-24}\alpha^5\big(\log(1/\alpha)\big)^{-2}N^{\alpha^2/250\log(1/\alpha)}\,, \tag{7}$$

confirming Theorem 4.

We turn now to the proof of Theorem 20.

LEMMA 22. *There is a set $X \subseteq \mathbb{Z}_N$ with $|X| = 576\alpha^{-2}\log|G|$ and*

$$\sup_{r\in G}|\hat{X}(r)| \leq \alpha|X|/3\,.$$

Choose a set Y at random by picking each element of $\mathbb{Z}_N$ independently at random with probability $p = t/N$. The Fourier coefficient $\hat{Y}(r)$, $r \neq 0$, is a sum of N independent random variables $S_j^{(r)} = Y(j)\omega^{rj}$ with means 0 and variances at most p. It follows from Lemma 9 that

$$\mathbb{P}\big(|\hat{Y}(r)| \geq \alpha t/6\big) \leq 4e^{-\alpha^2 t/288}\,. \tag{8}$$

Similarly

$$\mathbb{P}\big(||Y| - t| \geq \alpha t/6\big) \leq 4e^{-\alpha^2 t/288}\,. \tag{9}$$

Thus if $t \geq 576\alpha^{-2}\log|G|$ there is a positive probability that Y satisfies (8) for all $r \in G$ and also (9). Take a specific Y satisfying these conditions. By adding or deleting at most $\alpha t/6$ elements from Y we can produce an X satisfying the conclusion of the lemma. □

Proof of Theorem 20. Let $B{=}\mathcal{B}(F,\alpha/20|F|)$. Observe that $B{\subseteq}\mathcal{B}(G,\alpha/20)$ because $G \subseteq \mathcal{E}(F)$. Let $X = \{x_1,\dots,x_{|X|}\}$ be as in the previous lemma and let $b_1,\dots,b_{|X|}$ be elements of B. Let S be the multiset $\{b_j + x_j \mid j = 1,\dots,|X|\}$. If $r \in G$ then we have

$$\begin{aligned}\big|\hat{S}(r) - \hat{X}(r)\big| &\leq \sum_{j=1}^{|X|}|\omega^{r(b_j+x_j)} - \omega^{x_j}|\\ &\leq |X|\sup_j|\omega^{rb_j} - 1|\\ &\leq \alpha|X|/3\,.\end{aligned}$$

It follows that

$$\sup_{r\in G}|\hat{S}(r)| \leq 2\alpha|S|/3\,.$$

Since A is (α, F, G)-RHNU, this means that there is no choice of the b_j for which S is a subset of A.

Consider, however, the possibility of using the following algorithm. Choose $b_1 \in B$ so that $b_1 + x_1 \in A$. Choose $b_2 \in B$ so that $b_2 + x_2 \in A \setminus \{b_1 + x_1\}$. Continue in this way; at the last stage choose $b_{|X|} \in B$ so that

$$b_{|X|} + x_{|X|} \in A \setminus \bigcup_{j=1}^{|X|-1} \{b_j + x_j\}.$$

If it worked, this algorithm would generate a choice of elements b_j of the type we argued couldn't exist. However if the algorithm does not work then there must be some choice of j for which $B + x_j$ is contained entirely within A^c, except possibly for $j-1$ elements. Part (i) of Theorem 20 follows immediately. Part (ii) of the theorem is an easy corollary of part (i), using Lemma 10. □

7 Van der Waerden Numbers

In this section we show how to deduce a bound for the off-diagonal van der Waerden number $W(2; 3, k)$ from Theorem 3. The deduction is basically straightforward, but there are some technical difficulties. Suppose then that we have coloured $\{1, \dots, N\}$ red and blue. Write A for the set of red numbers, and suppose that $|A| = \alpha N$. Let $P = \{a, a + d, \dots a + (m-1)d\}$ be any arithmetic progression in $\{1, \dots, N\}$. Write

$$P_1 = \{a + 2\lambda d \mid 1 \le \lambda \le m/4\},$$
$$P_2 = \{a + (2\lambda + 1)d \mid 0 \le \lambda \le m/4\},$$
$$P_3 = \{a + 2\lambda d \mid m/4 < \lambda \le m/2\}$$

and

$$P_4 = \{a + (2\lambda + 1)d \mid m/4 < \lambda \le m/2\}.$$

Thus P is the disjoint union of P_1, P_2, P_3, P_4.

LEMMA 23. *Suppose that $N \ge 2^{12}$ and that $\alpha \ge N^{-1/56}$. Then there is a progression $P \subseteq \{1, \dots, N\}$ of length at least $\sqrt{N}$ such that each of $A \cap P_j$ has size at least $\alpha|P|/8$.*

Suppose not. Let $P^{(0)} = \{1, \dots, N\}$. Then some $|A \cap P_i^{(0)}|$ is at most $\alpha N/8$, and so some $|A \cap P_j^{(0)}|$ must be at least $7\alpha N/24$. Let $P^{(1)} = P_j^{(0)}$. Then $P^{(1)}$ has size at least $N/6$ and the density of A on $P^{(1)}$ is at least

$14\alpha/13$, provided that $N \geq 48$ (these numbers arise from the slight difficulty caused by the $P_k^{(0)}$ not all having size *exactly* $N/4$). Now proceed inductively: at the tth stage we will have a progression $P^{(t)}$ of length at least $N/6^t$ on which A has density at least $(14/13)^t\alpha$, provided that $N \geq 48 \cdot 6^t$. If $t \geq 14\log(1/\alpha)$ then this would be impossible, and so we have a contradiction provided that $N/6^t \geq \sqrt{N}$ at this stage. It is easy to check that the conditions on N in the statement of the lemma ensure this. □

Let us, then, pass to a progression P with this property. Write $B = A \cap P$ and write $B_j = B \cap P_j$ for $j = 1, 2, 3, 4$. Suppose that A does not contain a 3-term AP. Then neither does B, and so B must be disjoint from $\frac{1}{2}(B_1 + B_3)$, the subset of P containing all elements of the form $(b_1 + b_3)/2$. We claim that this set contains a long arithmetic progression. To see this, rescale P to $\{1, \ldots, M\}$, where $M \geq \sqrt{N}$. Regard the rescaled B_1 and B_3 as subsets of $\mathbb{Z}_p$ for some prime $p \in (2M, 4M]$. We know that $|B_j| \geq \alpha p/32$ for all j, and so by Theorem 7 $B_1 + B_3$ contains an AP of length at least $L(N) = \exp(c(\alpha(\log N)^{1/2} - \log\log N))$. Because of our choice of p this will be a genuine AP, not just a mod p progression. Thus $P \setminus B$ contains a long progression, and hence so does the set of blue numbers in our original colouring.

To finish the argument we simply work out a few numbers. One can check that if $\alpha \geq C\log\log N/(\log N)^{1/2}$ then $L(N) \gg \log N$. Thus either there is a red 3-AP or a blue AP of length at least $c\log N$. If, however, α is smaller than this then there is a blue AP of length at least $(\log N)^{1/2}/C\log\log N$ for trivial reasons. Hence we have

Theorem 24. $W(2; 3, k) \leq e^{Ck^2(\log k)^2}$.

As we have remarked, this is weaker than the best known result by a logarithm in the exponent. We should also remark that using the argument of this section one could deduce the bound $W(2; 3, k) \leq e^{Ck^2(\log k)^3}$ from Bourgain's 1990 paper [Bou1].

Acknowledgements I would like to thank W.T. Gowers for his encouragement. In addition to the sources of funding mentioned earlier, I would like to thank Trinity College, Cambridge, for providing me with excellent working conditions.

References

[B] S. Bernstein, Sur une modification de l'inéqualité de Tchebichef, Annal. Sci. Inst. Sav. Ukr. Sect. Math. I (1924).

[Bo] N.N. Bogolubov, Some algebraic properties of almost periods, Zap. Kafedry Mat. Fiz. Kiev 4 (1939), 185 – 194 (in Russian).

[Bou1] J. Bourgain, On arithmetic progressions in sums of sets of integers, in "A Tribute to Paul Erdős", CUP (1990), 105–109.

[Bou2] J. Bourgain, On triples in arithmetic progression, GAFA, Geom. funct. anal. 9:5 (1999), 968–984.

[C] M.C. Chang, Polynomial bounds for Freimans's theorem, Duke Math. Jour., to appear.

[FHR] G.A. Freiman, H. Halberstam, I.Z. Ruzsa, Integer sum sets containing long arithmetic progressions, J. London Math. Soc. (2) 46 (1992), 193–201.

[G] W.T. Gowers, A new proof of Szemerédi's theorem, GAFA, Geom. funct. anal. 11:3 (2001), 465–588.

[Gr1] B.J. Green, Bernstein's inequality and Hoeffding's inequality, available at `http://www.dpmms.cam.ac.uk/~bjg23/expos.html`.

[Gr2] B.J. Green, Edinburgh lecture notes on Freiman's theorem, available at `http://www.dpmms.cam.ac.uk/~bjg23/preprints.html`.

[Gr3] B.J. Green, Some constructions in the inverse spectral theory of cyclic groups, Combinatorics, Probability and Computing, to appear; available at `http://www.dpmms.cam.ac.uk/~bjg23/preprints.html`.

[R] W. Rudin, Fourier Analysis on Groups, Wiley 1990 (reprint of the 1962 original).

[Ru] I.Z. Ruzsa, Arithmetic progressions in sumsets, Acta Arith. 60:2 (1991), 191–202.

Ben Green, Trinity College, Cambridge University, Cambridge CB2 1TQ, UK

Submitted: August 2001