Arithmetic Progressions in Sumsets and Van der Waerden Numbers

Ben Green

Abstract

We introduce the concepts of hereditary non-uniformity and restricted hereditary non-uniformity for subsets of $\mathbb{Z}/N\mathbb{Z}$, and prove theorems about such sets which imply various results about arithmetic progressions. In particular we show that if A is a subset of $\mathbb{Z}/N\mathbb{Z}$ with size αN, where $\alpha \in (0,1)$ is fixed, then $A+A$ contains an arithmetic progression of length at least $e^{c\sqrt{\log N}}$. This improves a result of Bourgain.

1. Locating Arithmetic Progressions. 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 [2].

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*

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

Freiman, Halberstam and Ruzsa [5], building on a technique of Bogolubov [1], 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. We will show

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*

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

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

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

These may be contrasted with the best known lower bounds for these questions. Regarding Theorem 3, Ruzsa [9] 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 $\exp\left(C_\epsilon(\log N)^{2/3+\epsilon}\right)$. As for Theorem 4, a relatively straightforward construction (see [5]) shows that $A + A + A$ need not contain an AP of length $2N^{\log(1/\alpha)}$. It is conjectured in [5] 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 [6] 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.

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 [3]. It is rather easy to see that $W(2; 3, k) \leq N_3(k^{-1})$. 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 §3 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 §4 we show how this implies Theorem 3. The next two sections are devoted to proofs. In §5 we assemble some important tools. In §6, which may be regarded as the heart of the paper, we prove Theorem 7. In §7 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 §8 we deduce the bound for $W(2; 3, k)$. Finally in an appendix we give proofs of some results which are essential for our arguments, but which are difficult to locate in the literature.

2. Notation. We will be making substantial use of Fourier Analysis on the group $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$. For convenience assume that N is prime, and write $\omega = e^{2\pi i/N}$. If $f : \mathbb{Z}_N \to \mathbb{C}$ then for each $r \in \mathbb{Z}_N$ we write

$$\hat{f}(r) = \sum_{x \in \mathbb{Z}_N} f(x)\omega^{rx}.$$

We will often use the same letter to denote a set and its characteristic function.

If $X = \{x_1, \dots, x_m\}$ is a subset of an abelian group then write $\overline{X}$ for the set of all elements of the form $\epsilon_1 x_1 + \dots + \epsilon_m x_m$, where $\epsilon_j \in \{-1, 0, 1\}$.

3. 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.

4. 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

$$\begin{aligned} |S||C||D| \ &\leq \ \sup_{r\neq 0} |\hat{S}(r)| \sum_r |\hat{C}(r)||\hat{D}(r)| \\ &\leq \ \sup_{r\neq 0} |\hat{S}(r)| \left(\sum_r |\hat{C}(r)|^2\right)^{1/2} \left(\sum_r |\hat{D}(r)|^2\right)^{1/2} \\ &= \ \sup_{r\neq 0} |\hat{S}(r)| \cdot (\gamma\delta)^{1/2} N^2. \end{aligned}$$

The claim follows immediately. □

5. Tools. In this section we assemble some tools which will be required in the proofs of Theorems 7 and 20. First of all we state a concentration inequality for sums of independent random variables.

Lemma 9 (Bernstein) *Let $X_1, \dots, X_n$ be independent complex-valued random variables with $\mathbb{E}X_i = 0$ and $\mathbb{E}|X_i|^2 = \sigma_i^2$. Write $\sigma^2 = \sigma_1^2 + \dots + \sigma_n^2$, and suppose that $|X_i| \leq 1$ uniformly in i. Suppose that $\sigma^2 \geq 4nt$. Then we have the inequality*

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

The second tool we require is the following result from [4] concerning the structure of the points at which a set can have large Fourier coefficients. We have already used this result in the proof of Proposition 21 above.

Lemma 10 (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 $\mathcal{R} \subseteq \overline{\Lambda}$.*

In [4] this is derived from an inequality of Rudin [8]. There does not seem to be a particularly convenient source for these results in the literature, so we will prove Lemma 10 in an appendix.

The next lemma is a rather standard one concerning Bohr Sets. If $\Gamma \subseteq \mathbb{Z}_N$ and $\epsilon > 0$ then we write $B(\Gamma, \epsilon)$ for the set of all $x \in \mathbb{Z}_N$ with $\|\gamma x\| \leq \epsilon$ for all $\gamma \in \Gamma$.

Lemma 11 *Suppose that $|\Gamma| = d$. Then $B(\Gamma, \epsilon)$ contains an AP of length at least $\epsilon N^{1/d}$.*

6. 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 Set $B(\mathcal{R}, \eta/64)$ 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 10 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

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

so the result follows from Lemma 11. $\square$

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

$$|(x+P)\cap B| \;\leq\; \frac{16\beta}{\eta}|P|.$$

Proof. Suppose not. Then we would have

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

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

$$\begin{aligned} |\hat{C}(r)| &= |\hat{C^c}(r)| \\ &\leq |C^c| \\ &\leq \eta N/16 \\ &\leq \eta|C|/8. \end{aligned} \tag{1}$$

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* $t \geq 5000\eta^{-2}\log N$. *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 Bernstein's inequality and (1) that

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

By the same token

$$\mathbb{P}\left(||E| - t| \;\geq\; \eta t/24\right) \;<\; e^{-\eta^2 t/5000}.$$

If $t \geq 5000\eta^{-2}\log N$, 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. $\Box$

An almost identical argument proves the following.

Lemma 15 *Let $\beta N \geq t \geq 5000\eta^{-2}\log N$. Then there is a subset $X \subseteq B$ with $|X| = t$ and*

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

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

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

for all other $r \neq 0$.

Proof. We have

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

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

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

This proves the lemma. $\Box$

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 $r \notin \mathcal{R}$,

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

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

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

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

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

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

$$t \leq \frac{16\beta}{\eta}|P| \tag{2}$$

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

$$t \ \geq \ 5000\eta^{-2}\log N \tag{3}$$

and

$$t \ \leq \ \eta\beta N/10. \tag{4}$$

Let us suppose that parameters have also been chosen so that

$$|P| \ \geq \ \frac{\eta}{32\beta}. \tag{5}$$

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

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 $1/\beta>\eta/32\beta$ anyhow. Summarising then, we have

Proposition 19 *Suppose that* (2), (3), (4) *and* (5) *are all satisfied. Then* A^c *contains an AP of length at least* $\eta/32\beta$.

A tedious calculation using Lemma 12 shows that one can take $t=5000\eta^{-2}\log N$ 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.

7. 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 $G\subseteq\overline{F}$ 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 the Bohr Set $B(F, \alpha/20|F|)$, 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 before, 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 Idenity shows that $|G| \leq \alpha^{-3}$. Furthermore a theorem of Chang, which we state formally as Lemma 10 below, tells us that G is contained in $\overline{F}$ 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

$$2^{-24}\alpha^5 \left(\log(1/\alpha)\right)^{-2} N^{\alpha^2/250\log(1/\alpha)}, \tag{6}$$

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 Bernstein's Inequality that

$$\mathbb{P}\left(|\hat{Y}(r)| \ \geq \ \alpha t/6\right) \ \leq \ e^{-\alpha^2 t/288}. \tag{7}$$

Similarly

$$\mathbb{P}\left(||Y| - t| \ \geq \ \alpha t/6\right) \ \leq \ e^{-\alpha^2 t/288}. \tag{8}$$

Thus if $t \geq 576\alpha^{-2}\log|G|$ there is a positive probability that Y satisfies (7) for all $r \in G$ and also (8). 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. $\square$

Proof of Theorem 20. Let $B = B(F, \alpha/20|F|)$. Observe that $B \subseteq B(G, \alpha/20)$ because $G \subseteq \overline{F}$. 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 $\{b_j + x_j | j = 1, \dots, |X|\}$. If $r \in G$ then we have

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

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 11. $\square$

8. 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

$$\begin{aligned}P_1 &= \{a + 2\lambda d \mid 1 \leq \lambda \leq N/4\}, \\ P_2 &= \{a + (2\lambda + 1)d \mid 0 \leq \lambda \leq N/4\}, \\ P_3 &= \{a + 2\lambda d \mid N/4 < \lambda \leq N/2\}\end{aligned}$$

and

$$P_4 \;=\; \{a + (2\lambda + 1)d \mid N/4 < \lambda \leq N/2\}.$$

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

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

Suppose not. Let $P^{(0)} = \{1, \ldots, N\}$. 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. $\Box$

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 $L(N) = \exp\big(c(\alpha(\log N)^{1/2} - \log\log N)\big)$. 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 logairthm 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 [2].

9. Miscellaneous Remarks. The author would like to make two remarks concerning the originality of Theorem 4. Firstly it is mentioned in [5] that a similar result would appear in a forthcoming paper of Ruzsa. So far as I am aware, this paper has not yet appeared. Secondly we believe that a result very similar to Theorem 4 would follow by combining the methods of [5] with Lemma 10. To the best of our knowledge this observation does not appear in the

literature either.

Appendix: Proof of Lemma 10. Lemma 10 was crucial to our whole argument but, as we remarked earlier, it takes some effort to locate and read a proof in the literature. Consequently we offer the reader the following.

Recall that we have a set $B \subseteq \mathbb{Z}_N$ with cardinality βN, and we write $\mathcal{R}$ for the set of all $r \in \mathbb{Z}_N$ for which $|\hat{B}(r)| \geq \rho|B|$. We wish to show that $\mathcal{R}$ has lots of structure, and we do this by proving that it does not contain a very large unstructured subset. Let $E = \{e_1, \ldots, e_m\}$ be a subset of $\mathbb{Z}_N$. We say that E is *dissociated* if the only solutions to

$$\epsilon_1 e_1 + \cdots + \epsilon_m e_m \ = \ \epsilon_1' e_1 + \cdots + \epsilon_m' e_m$$

with $\epsilon_i, \epsilon_i' \in \{-1, 0, 1\}$ are those in which $\epsilon_i = \epsilon_i'$ for all i. Let E be a maximal dissociated subset of $\mathcal{R}$. Let

$$f(x) \ = \ N^{-1} \sum_{n \in E} \hat{B}(n) \omega^{nx}.$$

Then the inner product $\langle f, B \rangle$ is not too small. In fact by Parseval's Identity we have

$$\langle f, B \rangle \ = \ N^{-2} \sum_{n \in E} |\hat{B}(n)|^2, \tag{9}$$

which is quite large because $E \subseteq \mathcal{R}$. Writing down the obvious lower bound now would, however, be fatal as the reader will see.

The lack of structure of E is such that one might expect the exponentials ω^{nx}, $n \in E$, to behave almost independently as x varies. This would make f rather small in general, and one might hope to contrast this information with (9).

To show that f behaves something like a sum of independent random variables we compare it with similar sums which *may* be chosen at random. This involves an application of Riesz products and is the key insight of the argument.

Let $\theta : E \to \mathbb{Z}_N$ be any function at all, and write f_θ for the twisted sum

$$f_\theta(x) \ = \ N^{-1} \sum_{n \in E} \hat{B}(n) \omega^{nx + \theta(n)}.$$

Write $p_\theta(x)$ for the Riesz product

$$p_\theta(x) \ = \ 2 \prod_{n \in E} \left(1 + \tfrac{1}{2} \left(\omega^{nx + \theta(n)} + \omega^{-(nx + \theta(n))}\right)\right).$$

Claim 1 *We have* $f_\theta = f * p_\theta$.

Proof of Claim. This can be established by a fairly straightforward computation. We have

$$f * p_\theta(x) \;=\; 2N^{-2} \sum_y \sum_{m \in E} \hat{B}(m) \omega^{m(x-y)} \prod_{n \in E} \left(1 + \tfrac{1}{2}\left(\omega^{ny+\theta(n)} + \omega^{-(ny+\theta(n))}\right)\right).$$

Multiplying out the product and changing the order of summation, one is confronted with sums of the form

$$\sum_y \omega^{(n_1+\cdots+n_r-n_1'-\cdots-n_s'-m)y},$$

where the n_i, n_i' are distinct elements of E and $m \in E$. The dissociativity of E implies that such a sum is zero unless m equals some n_i, in which case it equals N. The claim follows quickly. $\Box$.

We will actually use the relation $f = f_\theta * p_{-\theta}$, which can be proved in exactly the same way. Now the Riesz product p_θ is non-negative, and so $\|p_\theta\|_1$ is simply $N^{-1}\sum_x p_\theta(x)$. This sum may easily be calculated by expanding out another product and using dissociativity, and it turns that $\|p_\theta\|_1 = 2$. Thus by Young's Inequality and the Claim we have

$$\begin{aligned} \|f\|_p &= \|f_\theta * p_{-\theta}\|_p \\ &\leq \|f_\theta\|_p \|p_{-\theta}\|_1 \\ &= 2\|f_\theta\|_p \end{aligned} \tag{10}$$

for any even integer $p \geq 2$.

Why the restriction to even integers? Well if $p = 2k$ then we may write

$$\|f_\theta\|_{2k}^{2k} \;=\; N^{-1-2k} \sum_x \sum_{n_i \in E} \hat{B}(n_1) \ldots \hat{B}(n_k)\overline{\hat{B}(n_{k+1})} \ldots \overline{\hat{B}(n_{2k})} \times$$
$$\omega^{(n_1+\cdots+n_k-n_{k+1}-\cdots-n_{2k})x} \omega^{(\theta(n_1)+\cdots+\theta(n_k)-\theta(n_{k+1})-\cdots-\theta(n_{2k}))}.$$

Considering θ as a random function from E to $\mathbb{Z}_N$, we may take the expectation of this quantity over all θ. The expected value of the rightmost exponential is zero unless $(n_1, \ldots, n_k)$ is a permutation of $(n_{k+1}, \ldots, n_{2k})$, in which case it equals 1. Thus

$$\begin{aligned} \mathbb{E}\|f_\theta\|_{2k}^{2k} &\leq k! N^{-2k} \sum_{n_1,\ldots,n_k \in E} |\hat{B}(n_1)|^2 \ldots |\hat{B}(n_k)|^2 \\ &= k! \|f\|_2^{2k}. \end{aligned}$$

Thus for some particular choice of θ one has $\|f_\theta\|_{2k} \leq (k!)^{1/2k}\|f\|_2$, which implies by (10) that $\|f\|_{2k} \leq 2\sqrt{k}\|f\|_2$ for any integer $k \geq 1$. This implies that for any real number $p \geq 2$ one has

$$\|f\|_p \;\leq\; 3\sqrt{p}\|f\|_2. \tag{11}$$

Equation (11) shows that f is rather smooth, to the extent that one would expect if it was a sum of independent variables. Recall now (9). Let $p \geq 2$ be a real number to be chosen later, and let q be the conjugate index to p satisfying $\frac{1}{p} + \frac{1}{q} = 1$. We have

$$\begin{aligned} N^{-2} \sum_{n \in E} |\hat{B}(n)|^2 &\leq \langle f, B \rangle \\ &\leq \|f\|_p \|B\|_q \\ &\leq 3\sqrt{p}\beta^{1/q} \|f\|_2 \\ &= 3\sqrt{p}\beta^{1/q} \left(N^{-2} \sum_{n \in E} |\hat{B}(n)|^2 \right)^{1/2}. \end{aligned}$$

Cancelling the common factor from both sides and observing that $N^{-2} \sum_{n \in E} |\hat{B}(n)|^2 \geq \beta^2 \rho^2 |E|$ (since $E \subseteq \mathcal{R}$) we have

$$\beta\rho\sqrt{|E|} \leq 3\sqrt{p}\beta^{1/q}.$$

Choosing $p = \log(1/\beta)$ gives that

$$|E| \leq 70\rho^{-2} \log(1/\beta). \tag{12}$$

We have shown that $\mathcal{R}$ does not have a particularly large dissociated subset, and it is a short step from here to a proof of Lemma 10. Let $E = \{e_1, \ldots, e_m\}$ be a maximal dissociated subset of $\mathcal{R}$. The maximality implies that the addition of any new $r \in \mathcal{R}$ will spoil the dissociativity property. It is easy to see that this implies that each r is expressible as $\sum_j \eta_j e_j$, where $\eta \in \{-2, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, 2\}$. Lemma 10 follows on taking $\Lambda = \frac{1}{2}E \cup E \cup 2E$. □

We close by remarking that Lemma 10 is best possible for any values of β and ρ that might conceivably be useful. This is the main result of the author's preprint [7].

References

[1] Bogolubov, N.N. *Some Algebraic Properties of Almost Periods*, Zap. Kafedry Mat. Fiz. Kiev **4** (1939) 185 – 194 (Russian).

[2] Bourgain, J. *On Arithmetic Progressions in Sums of Sets of Integers*, in *A Tribute to Paul Erdös*, CUP 1990.

[3] Bourgain, J. *On Triples in Arithmetic Progression*, GAFA **9** (1999), no. 5, 968 – 984.

[4] Chang, M-C. *A Polynomial Bound in Freiman's Theorem*, preprint.

[5] Freiman G.A., Halberstam, H. and Ruzsa, I.Z. *Integer Sum Sets Containing Long Arithmetic Progressions*, J. London Math. Soc. (2) **46** (1992) 193 – 201.

[6] Gowers, W.T. *A New Proof of Szemerédi's Theorem*, to appear in GAFA.

[7] Green, B.J. *Fourier Transforms of Sets of Integers: Chang's Structure Theorem is Sharp*, preprint.

[8] Rudin, W. *Fourier Analysis on Groups*, Wiley 1990 (Reprint of the original 1962 edition).

[9] Ruzsa, I.Z. *Arithmetic Progressions in Sumsets*, Acta. Arith. **60** (1991) 191 – 202.