Some Applications of Harmonic Analysis to Arithmetic Combinatorics

Ben Green

Submission for the Smith-Raleigh-Knight Prize of the University of Cambridge

January 2001

Contents

1 **2**

1.1 Acknowledgements and Statement of Originality . . . 2
1.2 Introduction . . . 2
1.3 A Brief Introduction to Harmonic Analysis on $\mathbb{Z}_N$ . . . 3

2 The Number of Squares and $B_h[g]$-Sets **5**

2.1 Introduction . . . 5
2.2 Bounds for B_4 Sets – The First Part of the Argument . . . 9
2.3 A Lower Bound for the Number of Squares . . . 12
2.4 A Return to B_4 Sets . . . 19
2.5 Large Values of h . . . 20
2.6 New Bounds for $B_2[g]$-Sets Part I . . . 24
2.7 New Bounds for $B_2[g]$-Sets Part II . . . 25
2.8 Summary . . . 29

3 On Arithmetic Structures in Dense Sets of Integers **31**

3.1 Introduction . . . 31
3.2 A Short Proof of Sárközy's Theorem for Squares . . . 34
3.3 APs with Common Difference $x^2 + y^2$ . . . 37
3.4 Increasing the Density on a Special Subprogression . . . 47

4 Unfinished Business **54**

4.1 A Few Remarks on $B_h[g]$ Sets . . . 54
4.2 Further Remarks on Functions with Minimal $M(f)$ . . . 55
4.3 Arithmetic Non-Uniformity . . . 59
4.4 A Question of Verstraete . . . 63
4.4.1 Introduction . . . 63
4.4.2 The Number of Sumsets . . . 63
4.4.3 Large Sumsets . . . 65
4.5 Random Sets and a Result of Salem and Zygmund . . . 68

Chapter 1

1.1 Acknowledgements and Statement of Originality

Everything in this essay is original unless there is a clear statement to the contrary. I would like to thank my reseach supervisor Professor W. Tim Gowers for his assistance and encouragement. This work has been supported by an EPSRC research studentship, and during its preparation I have enjoyed the hospitality of Trinity College, Cambridge and Princeton University.

1.2 Introduction

This essay is in three parts. The first two are expanded versions of two papers,

- *The Number of Squares and $B_h[g]$-Sets* [17] and
- *On Arithmetic Structures in Dense Sets of Integers* [18].

The papers are both applications of harmonic analysis to arithmetic combinatorics, but beyond that they have rather little in common. For that reason I have kept the papers entirely separate so that each may be read independently of the other.

Some basic background material on harmonic analysis is common to both papers, so I have isolated this and presented it as an appetiser for the rest of the essay. I have tried to expand the introductions to the papers to the point where it should be possible for a non-expert to gain a reasonable overview of each area by reading them.

The third part of the essay is a miscellany of results and questions which indicate the direction in which my research is going, but which are not yet publication quality. By its very nature it is quite possible that this section contains some oversights, or some ideas which will turn out not to be interesting.

The reader may be forgiven for wondering just what *arithmetic combinatorics* is. This is not

an area listed in the Mathematics Subject Classification, and indeed so far as I am aware the only previous use of the term is by Terence Tao [44]. It would be fairly pointless to attempt a precise definition, so let me just say that I started describing my research as "arithmetic combinatorics" when I got bored of telling people that I worked "on the interface of additive number theory, combinatorial number theory and harmonic analysis".

1.3 A Brief Introduction to Harmonic Analysis on $\mathbb{Z}_N$

We shall make substantial use of Fourier analysis on finite cyclic groups, so we would like to take this opportunity to give the reader a swift introduction. If nothing else this will serve to clarify notation.

Let N be a fixed positive integer, and write $\mathbb{Z}_N$ for the cyclic group with N elements (this will no doubt annoy algebraic number theorists, who prefer to reserve this sort of notation for p-adic considerations). Let ω denote the complex number $e^{2\pi i/N}$. Although ω clearly depends on N, we shall not indicate this dependence in the rest of the paper, trusting that the value of N is clear from context. Let $f : \mathbb{Z}_N \to \mathbb{C}$ be any function. Then for $r \in \mathbb{Z}_N$ we define the Fourier transform

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

We shall repeatedly use two important properties of the Fourier transform. The first is Parseval's identity, which states that if $f : \mathbb{Z}_N \to \mathbb{C}$ and $g : \mathbb{Z}_N \to \mathbb{C}$ are two functions then

$$N \sum_{x \in \mathbb{Z}_N} f(x)\overline{g(x)} \;=\; \sum_{r \in \mathbb{Z}_N} \hat{f}(r)\overline{\hat{g}(r)}.$$

The second is the interaction of *convolutions* with the Fourier transform. If $f, g : G \to \mathbb{C}$ are two functions on an abelian group G we define the convolution

$$(f * g)(x) \;=\; \sum_{y \in G} f(y)\overline{g(y-x)}.$$

Observe that this notation is slightly non-standard. We shall use the plus symbol in place of what is often called the convolution, so that

$$(f + g)(x) \;=\; \sum_{y \in G} f(y)\overline{g(x-y)}.$$

Observe that in the particular case when f and g are the characteristic functions of two sets $A, B \subseteq G$ the convolution $f * g(x)$ can be interpreted as the number of solutions to $x = a - b$ with $a \in A$ and $b \in B$, and $(f + g)(x)$ is the number of solutions to $x = a + b$.

This seems like a good place to remark that we will often use the same letter to denote both

a set and its characteristic function. For example if $A \subseteq \mathbb{Z}$ then we set $A(x) = 1$ if $x \in A$ and $A(x) = 0$ otherwise.

One more key fact about Fourier transforms, which we shall use without further comment, is that

$$(f * g)\hat{}(r) \;=\; \hat{f}(r)\overline{\hat{g}(r)}.$$

Chapter 2

The Number of Squares and $B_h[g]$-Sets

2.1 Introduction

Let G be a torsion-free abelian group and let $h \geq 2$ and $g \geq 1$ be integers. A subset $A \subseteq G$ is called a $B_h[g]$ set if no $x \in G$ has more than g distinct representations in the form $a_1 + \cdots + a_h$ with $a_i \in G$, where two representations are regarded as distinct if they cannot be obtained from one another by a reordering of summands. We shall refer to a $B_h[1]$-set as simply a B_h-set.

We shall immediately specialise to the case $G = \mathbb{Z}$, where we shall look at $B_h[g]$-subsets of $\{1, \ldots, N\}$. B_2-subsets of $\{1, \ldots, N\}$ are often referred to in the literature as Sidon Sets. This is reasonable enough because they were first studied by Sidon in relation to a question about trigonometric series, but may lead to confusion because there is a quite different type of object in harmonic analysis which is also called a Sidon Set. Nevertheless, we shall occasionally use this term.

Sidon found himself wanting to know about the size of the largest B_2-subset of $\{1, \ldots, N\}$, a quantity we shall denote by $A(2, N)$. Doubtless he observed that trivially

$$A(2, N) \ \leq \ 2N^{1/2}. \tag{2.1}$$

To prove this let $A \subseteq \{1, \ldots, N\}$ be a B_2-set and double-count pairs (a, b) of elements of A. On the one hand there are exactly $|A|^2$ such pairs. However the B_2 property guarantees that the mapping

$$\psi : A \times A \longrightarrow [2N]$$

defined by $\psi((a, b)) = a + b$ is at most 2-1, and so indeed $|A|^2 \leq 4N$.

Apparently Sidon also showed that $A(2, N) \gg N^{1/4}$. I have not looked up his argument, because we have

Proposition 1 $A(2,N) \gg N^{1/3}$.

Proof Apply the greedy algorithm. Set $a_1 = 1$ and generate $a_2, a_3, \ldots$ inductively as follows. Given $a_1, \ldots, a_h$ with the B_2 property define a_{h+1} to be the least positive integer not of the form $a_i + a_j - a_k$. It is easy to see that $a_{h+1} \leq h^3 + 1$, and so after a certain time we will have a B_2-subset of $\{1, \ldots, N\}$ of size $\sim N^{1/3}$. □

It turns out that the trivial bound (2.1) gives the correct order of magnitude for $A(2,N)$, but to see this one has to be a lot cleverer. The necessary cleverness was provided by Singer in 1939 [41].

Theorem 2 (Singer) $A(2,N) \geq N^{1/2} + o(N^{1/2})$.

Proof Let p be a prime number and consider the finite field $K = \mathbb{F}_{p^2}$, together with the subfield $L \subseteq K$ isomorphic to $\mathbb{F}_p$. Let θ be a generator for the group $K^\times$, which is cyclic by standard results in elementary field theory. Let $\lambda_1, \ldots, \lambda_p$ be the elements of L and define integers $a_1, \ldots, a_p$ by

$$\theta^{a_j} \ = \ \theta + \lambda_j.$$

The a_i live in the group $G = \mathbb{Z}/(p^2-1)\mathbb{Z}$, and we claim that they form a Sidon subset of that group of cardinality p. Suppose indeed that $a_i + a_j = a_k + a_l$. This immediately implies that

$$(\theta + \lambda_i)(\theta + \lambda_j) \ = \ (\theta + \lambda_k)(\theta + \lambda_l),$$

and so

$$(\lambda_i + \lambda_j - \lambda_k - \lambda_l)\theta \ = \ \lambda_k\lambda_l - \lambda_i\lambda_j.$$

It now looks rather as if we have exhibited θ as an element of L, which would be extremely unfortunate. The only way in which we could have failed to perform such an exhibition would be if

$$\lambda_i + \lambda_j - \lambda_k - \lambda_l \ = \ \lambda_k\lambda_l - \lambda_i\lambda_j \ = \ 0,$$

which is easily seen to imply that $\{i,j\} = \{k,l\}$. Thus $\{a_1, \ldots, a_p\}$ is a Sidon subset of G, and may be regarded as a Sidon subset of $\{1, \ldots, p^2-1\}$ in a natural way. To conclude recall that the greatest prime less than $N^{1/2}$, p say, satisfies $p \geq N^{1/2} + o(N^{1/2})$. Performing the above construction with this prime gives the theorem. □

In 1941 Erdös and Turán [9] proved that

$$A(2,N) \ \leq \ N^{1/2} + N^{1/4} + 1, \tag{2.2}$$

which together with Singer's result shows that $A(2,N) \sim N^{1/2}$. Their argument is nicely described in [24]. The main results of [17] came by reformulating this argument in terms of harmonic analysis, and I will give this reformulation later on.

As something of an aside I should like to remark that I consider the problem of estimating

the error term $E_N = \left|A(2,N) - N^{1/2}\right|$ to be extremely interesting for two reasons. Firstly in the 60 years since Erdős and Turán published their paper no-one has managed so much as to prove

$$A(2,N) \ \leq \ N^{1/2} + cN^{1/4} + C$$

with $c < 1$. Secondly, despite the fact that there are several constructions which show that $A(2,N) \geq N^{1/2}(1+o(1))$ (see [35]), all of these involve prime numbers in some way. It seems then that the estimation of E_N has a very number-theoretical flavour.

Now let $A(h,g,N)$ denote the size of the largest $B_h[g]$-subset of $\{1,\dots,N\}$, and write $A(h,N) = A(h,1,N)$ for short. In stark contrast to the above discussion, the correct asymptotics for $A(h,g,N)$ have not been obtained for any pair $(h,g) \neq (2,1)$. In the remainder of this introduction we survey the known results before March 2000. Other accounts may be found in [8], [14], [20] and [39]. At the end of the paper we will summarise the current state of affairs.

Let us begin with lower bounds. It turns out that Theorem 2 can be generalised to give

$$A(h,g,N) \ \geq \ (1+o(1))N^{1/h} \tag{2.3}$$

for any $h \geq 2$. This was done by Bose and Chowla [3]. More recently various bounds have been given by Cilleruelo, Ruzsa and Trujillo [8] which yield a better result when $g > 1$. These lower bounds are also constructive and we will mention them a little more below.

In [17] we are concerned entirely with upper bounds. The same trivial counting argument that gave us (2.1) yields

$$A(h,g,N) \ \leq \ (gh \cdot h!)^{1/h} N^{1/h}. \tag{2.4}$$

Together with (2.3) this shows that $A(h,g,N)$ is comparable to $N^{1/h}$. In view of this it is quite natural to write

$$\alpha(h,g) \ = \ \limsup_{N\to\infty} A(h,g,N)N^{-1/h}$$

and $\alpha(h) = \alpha(h,1)$, so that (2.3) and (2.4) give

$$1 \ \leq \ \alpha(h,g) \ \leq \ (gh \cdot h!)^{1/h}.$$

In particular when $h = 2$ we have

$$\alpha(2,g) \ \leq \ 2\sqrt{g}. \tag{2.5}$$

Until recently this trivial bound had not been improved for any pair (h,g) with $g > 1$. However Cilleruelo, Ruzsa and Trujillo [8][1] show that

$$\frac{3\sqrt{2}}{4}\sqrt{g}(1+o_g(1)) \ \leq \ \alpha(2,g) \ \leq \ \frac{2\pi+4}{\sqrt{\pi^2+4\pi+8}}\sqrt{g}. \tag{2.6}$$

The constant appearing in the lower bound here is about 1.061, and that in the upper bound is roughly 1.864. For $g = 2$ Cillereulo [6] and Helm have independently given the bound

$$\alpha(2,2) \ \leq \ \sqrt{6}. \tag{2.7}$$

Cillereulo's proof is simple and combinatorial, but only generalises to give

$$\alpha(2,g) \ \leq \ 2\sqrt{g-\tfrac{1}{2}},$$

a slight improvement on (2.5).

In the case $g = 1$ the situation is slightly better. Here the technique of Erdös and Turán, which gave the correct asymptotics for $A(2,N)$, has a natural generalisation which gives a non-trivial upper bound for $A(h,N)$. The first result of this kind was obtained by Lindström [28] in 1969. He showed that

$$\alpha(4) \ \leq \ 8^{1/4}. \tag{2.8}$$

Generalising his technique, Jia [26] obtained the bound

$$\alpha(2k) \ \leq \ \left(k(k!)^2\right)^{1/2k}. \tag{2.9}$$

A modification of this approach gives a corresponding bound for odd values of h. Indeed the bound

$$\alpha(2k-1) \ \leq \ (k!)^{2/(2k-1)} \tag{2.10}$$

was obtained independently by Chen [5] and Graham [14]. We conclude this potted history by mentioning two further results. The first is the paper of Kolountzakis [27], in which the bound (2.9) is obtained by an interesting Fourier technique. We believe that this proof and that of Jia are morally the same, but the new perspective is interesting. Secondly it is worth remarking that Graham [14] obtained a slight improvement on (2.10) in the case $k = 2$. He proved that

$$\alpha(3) \ \leq \ \left(4-\tfrac{1}{228}\right)^{1/3}. \tag{2.11}$$

[1]The story of the interaction of this paper with [17] is quite interesting. I did not learn about [8] until after I had written a draft of [17]. After reading [8] it came to my attention that their methods complemented mine, and it then proved possible to obtain a number of stronger results by importing one or two of their ideas. I find it rather interesting that [8] appeared at roughly the same time as my paper, when there had not been much written on these questions for a very long time.

The argument is long and combinatorial.

Let us conlcude with a brief description of the main results of [17], which the rest of Chapter 2 is devoted to discussing. First of all we offer the first improvement on (2.8) since 1969, obtaining

$$\alpha(4) \ \leq \ 7^{1/4}. \tag{2.12}$$

We in fact obtain bounds for $\alpha(h)$ for all $h \geq 3$, and in particular we can improve (2.11) to

$$\alpha(3) \ \leq \ \left(\tfrac{7}{2}\right)^{1/3}. \tag{2.13}$$

In §2.6 we offer our own improvement to the upper bound (2.5). This improves on the results of [8] for $g \leq 68$. Finally in §2.7 we combine some of our ideas with some of the ideas in [8] to improve the bound (2.6) for all g.

2.2 Bounds for B_4 Sets – The First Part of the Argument

In this section we begin our treatment of B_4 sets which will lead to the bound (2.12). It is hoped that, after reading this section, the reader will have a good idea of the direction in which we are headed.

Like all previous approaches, our attack takes as motivation the original argument of Erdös and Turán from 1941 [9]. We now give this argument in the form that we use to get our generalisation. We leave it as a (slightly non-trivial) exercise for the reader to check that our argument and that of [9] are really the same.

Theorem 3 (Erdös – Turán) *For all N we have the bound*

$$A(2,N) \ \leq \ N^{1/2} + N^{1/4} + 1.$$

Hence, in view of Theorem 2, we have $\alpha(2) = 1$.

Proof Let $A \subseteq \{1, \ldots, N\}$ be a B_2-set. It is easy to check that we must have $A * A(x) \leq 1$ for all $x \neq 0$. Let u be a positive integer to be chosen later, and regard A as a subset of $\mathbb{Z}_{N+u}$. It is no longer the case that the *modular* version of $A * A$ satisfies $A * A(x) \leq 1$ for all x, but it *is* true that $A * A(x) \leq 1$ for $0 < |x| \leq u$. Let I be the characteristic function of $\{1, \ldots, u\}$, and write

$$E \ = \ \sum_{x \in \mathbb{Z}_{N+u}} A * A(x) I * I(x).$$

We count E in two ways. Firstly, by the discussion above, we have

$$\begin{aligned} E &\leq |A|u + \sum_{0<|x|\leq u} A*A(x)I*I(x) \\ &\leq |A|u + u^2. \end{aligned}$$

Secondly, by Parseval's identity, we have

$$\begin{aligned} E &= \frac{1}{N+u}\sum_r |\hat{A}(r)|^2|\hat{I}(r)|^2 \\ &\geq \frac{|A|^2u^2}{N+u}, \end{aligned}$$

where the hat symbol denotes Fourier transform in $\mathbb{Z}_{N+u}$. Comparing these upper and lower bounds for E, and putting $u = \lfloor N^{3/4} \rfloor$, gives the result. □

An interesting feature of the argument is that we only used the fact that $A*A(x) \leq 1$ for $0 < |x| \leq N^{3/4}$, which seems to be a much weaker statement than saying that A is B_2.

We now move on to B_4-sets. Let $A \subseteq \{1, \dots, N\}$ be a B_4-set.

Lemma 4 *For all $x \in \mathbb{Z}$ we have*

$$A*A*A*A(x) \;=\; (2A-2A)(x) \;\leq\; 4\left(1+|A|(A*A)(x)\right).$$

Proof Fix x. Both quantities $A*A*A*A(x)$ and $(2A-2A)(x)$ count the number of solutions to the equation

$$a_1 + a_2 - b_1 - b_2 \;=\; x \tag{2.14}$$

with $a_1, a_2, b_1, b_2 \in A$. If there are no solutions to this equation then the lemma is immediate.

We show first that if there is a solution to (2.14) in which $a_i \neq b_j$ for all i, j then (2.14) has at most 4 solutions. Indeed fix such a solution (a_1, a_2, b_1, b_2) and suppose that (a_1', a_2', b_1', b_2') is another solution. Then

$$a_1 + a_2 + b_1' + b_2' \;=\; a_1' + a_2' + b_1 + b_2.$$

But A is B_4, so the quadruples (a_1, a_2, b_1', b_2') and (a_1', a_2', b_1, b_2) are the same up to a reordering. It follows that $\{a_1', a_2'\} = \{a_1, a_2\}$ and $\{b_1', b_2'\} = \{b_1, b_2\}$, giving at most 4 possibilities.

If there is no solution of the kind discussed here then all solutions have $a_1 = b_1$, $a_1 = b_2$, $a_2 = b_1$ or $a_2 = b_2$. It is clear that, for each of these possibilities, (2.14) has $|A|(A*A)(x)$ solutions. Adding together everything above gives the result of the lemma. □

We remark that the important feature of the above bound is the number 4. On average,

$A * A(x)$ will be tiny.

Once again we embed the problem into a cyclic group where we can take finite Fourier transforms. Let $v \leq N$ be a parameter to be chosen later, and regard A as a subset of $\mathbb{Z}_{2N+v}$ in the obvious way. Since $2A \subseteq [1, \ldots, 2N]$, the *modular* difference set $2A - 2A$ will satisfy the bound of Lemma 4 for $|x| \leq v$. That is to say,

$$A * A * A * A(x) \ \leq \ 4\,(1 + |A|(A * A)(x)) \tag{2.15}$$

for $|x| \leq v$, where everything in sight lives in $\mathbb{Z}_{2N+v}$. Let $I = [1, \ldots, u]$ (as a subset of $\mathbb{Z}_{2N+v}$), where $u \leq v$ is another parameter to be chosen later. It will turn out that u needs to be significantly smaller than v. For the remainder of this section the hat symbol refers to Fourier transforms on $\mathbb{Z}_{2N+v}$. Using (2.15) we have the following key computation.

$$\begin{aligned}
\frac{1}{2N+v}\sum_r |\hat{A}(r)|^4|\hat{I}(r)|^2 &= \sum_x (A*A*A*A)(x)(I*I)(x) \\
&\leq 4u^2 + 4|A|\sum_x (A*A)(x)(I*I)(x) \\
&= 4u^2 + \frac{4|A|}{2N+v}\sum_r |\hat{A}(r)|^2|\hat{I}(r)|^2 \\
&\leq 4u^2 + \frac{4|A|}{2N+v}|A|^2\sum_r |\hat{I}(r)|^2 \\
&= 4u^2 + \frac{4|A|}{2N+v}|A|^2(2N+v)u \\
&\leq 4u^2 + \frac{2^{24}N^{7/4}u}{2N+v},
\end{aligned} \tag{2.16}$$

where we have used the fact that $v \leq N$ and the trivial bound $|A| \leq 100N^{1/4}$. The inequality (2.16) gives immediately that

$$8Nu^2 + 4u^2v + 2^{24}N^{7/4}u \ \geq \ \sum_r |\hat{A}(r)|^4|\hat{I}(r)|^2. \tag{2.17}$$

Using the trivial lower bound $|A|^4u^2$ for the right hand side of (2.17) together with $u = v = N^{7/8}$ gives the Lindström bound (2.8). However we are in a position to make a further improvement. The set A is extremely irregularly distributed in $\mathbb{Z}_{2N}$, being contained in $\{1, \ldots, N\}$. This information allows us to say something non-trivial about $\hat{A}(r)$ when r is small and *non-zero*. Much of our paper is concerned with exactly how much it is possible to deduce from this observation as regards (2.17). However the reader who is keen to see an improvement of (2.8) as quickly as possible may care to find a positive constant c such that either $|\hat{A}(\pm 1)| \geq c|A|$ or $|\hat{A}(\pm 2)| \geq c|A|$. Since I is a rather small interval the coefficients

$\hat{I}(\pm 1)$ and $\hat{I}(\pm 2)$ are as near to $|I|$ as makes no difference. Substituting into (2.17) and putting $u = v = N^{7/8}$ gives a bound of form

$$|A| \ \leq \ (8 - 2c^4)^{1/4} N^{1/4}(1 + o(1)).$$

2.3 A Lower Bound for the Number of Squares

Let $f : \{1, \dots, N\} \to \mathbb{R}$ be a function. Write $|f| = \sum_x f(x)$ and suppose that $|f| = N$. Define the quantity $M(f)$ by

$$M(f) \ = \ \sum_{\substack{a,b,c,d \\ a+b=c+d}} f(a)f(b)f(c)f(d) \ = \ \sum_x (f * f)(x)^2 \ = \ f * f * f * f(0).$$

Since it is quite standard to refer to a quadruple of integers (a, b, c, d) with $a + b = c + d$ as a *square*, we call $M(f)$ the *number of squares* for f. Let us now introduce Fourier analysis to the study of $M(f)$. We can regard f as a subset of $\mathbb{Z}_{2N}$ in a natural way, by identifying $\{1, \dots, N\}$ with the "first half" of $\mathbb{Z}_{2N}$. Furthermore when we do this the *modular* convolution $f * f(x)$ is precisely the same as the $\mathbb{Z}$-version, since the $\mathbb{Z}$-version of $f * f$ is supported in an interval of length $2N$. It follows that

$$M(f) \ = \ \sum_{x \in \mathbb{Z}_{2N}} f * f(x)^2 \ = \ \frac{1}{2N} \sum_{r \in \mathbb{Z}_{2N}} |\hat{f}(r)|^4. \tag{2.18}$$

We shall be concerned with the following problem.

Problem 5 *How small can $M(f)$ be, if $|f| = N$?*

The reader who has looked at §2.2 will realise, in view of (2.18), the pertinence of this problem to the issue of upper bounds for B_4-sets.

To get a feel for Problem 5 we prove a few easy results.

Lemma 6 *We have $M(f) \geq N^3/2$ for all f.*

Proof Observe that $\sum_x f * f(x) = N^2$ and that $f * f(x) = 0$ for $x \notin \{-N + 1, \dots, N - 1\}$. The result is now immediate from the Cauchy-Schwarz inequality. Alternatively, the result is trivial from (2.18). □

Lemma 7 *We can have $M(f) \leq 2N^3/3 + O(N^2)$.*

Proof Take f to be the characteristic function of $\{1, \dots, N\}$. □

Exactly as in §2.2 we realise, on closer examination of the second argument in Lemma 6,

that a lot of information has been thrown away. We have not considered any of the non-zero Fourier coefficents $\hat{f}(r)$ $(r \neq 0)$. Furthermore, since f is extremely irregularly distributed (being contained in $\{1, \ldots, N\}$, which only fills half of $\mathbb{Z}_{2N}$) we have every right to expect that the contribution from these coefficients will be significant. In fact, because of the way that f is distributed, we expect the non-zero Fourier coefficients $\hat{f}(r)$ with $|r|$ very small to make a significant contribution. Our objective now is to obtain a strong quantitative result from this observation.

There are various details in the following discussion that may seem a little unmotivated. Some justification for the approach is provided by the material in §4.2, but since this material is in no way essential to an understanding of the argument we have not included it here.

Let $f : \{1, \ldots, N\} \to \mathbb{R}$ be a function. Let v be a positive integer (which will be assigned various values later on) and regard f as a function on $\mathbb{Z}_{2N+v}$ in the natural way. In the following we will use the hat symbol ($\hat{}$) to denote Fourier transforms on $\mathbb{Z}_{2N+v}$. Exactly as in (2.18) we have

$$M(f) \;=\; \frac{1}{2N+v} \sum_r |\hat{f}(r)|^4. \tag{2.19}$$

The reason for introducing v will be clear to the reader who has studied §2.2.

We can now outline our approach. Let $\mathcal{H}_{2N+v}$ denote the set of all functions $H : \mathbb{Z}_{2N+v} \to \mathbb{R}$ such that

$$|H| \;=\; \sum_x H(x) \;=\; 0 \tag{2.20}$$

and

$$H(x) = 1 \qquad \text{for} \qquad x = 1, \ldots, N. \tag{2.21}$$

We will use functions $H \in \mathcal{H}_{2N+v}$ to obtain a lower bound on

$$E = \sum_{r \neq 0} |\hat{f}(r)|^4$$

in the following manner. Observe that

$$\sum_x f(x)H(x) \;=\; N.$$

Taking Fourier transforms gives, by Parseval's identity, that

$$\sum_r \hat{f}(r)\overline{\hat{H}(r)} \;=\; N(2N+v).$$

Using the triangle inequality and the fact that $\hat{H}(0) = 0$, this yields

$$\sum_{r \neq 0} |\hat{f}(r)||\hat{H}(r)| \;\geq\; N(2N+v).$$

By Hölder's Inequality we then have

$$\left(\sum_{r\neq 0}|\hat{f}(r)|^4\right)\left(\sum_{r\neq 0}|\hat{H}(r)|^{4/3}\right)^3 \;\geq\; N^4(2N+v)^4. \tag{2.22}$$

By choosing H suitably, we can use this to bound E below. The most natural way to look for H is to take a fixed continous function $U:[0,2]\to\mathbb{R}$, and then to define

$$H(x) \;=\; U\left(\frac{x}{N+\frac{v}{2}}\right).$$

By ensuring that $U(x)=1$ for $x\leq 1$ and that $\int_0^2 U(x)\,dx=0$, the function H thus obtained should (very nearly) satisfy (2.20) and (2.21). Furthermore it does not seem unreasonable that $\hat{H}(r)$, the discrete Fourier transform of H, might be related to the Fourier transform of U as a function on the real line.

One further device allows us to obtain a lower bound for

$$E(X) \;=\; \sum_{0<|r|\leq X}|\hat{f}(r)|^4$$

when X is much smaller than N. This is a smoothing technique, in which H is replaced by a similar function whose Fourier transform decays rather rapidly.

After that brief outline, we can state the main theorem of this section. In this theorem (and for the remainder of the section) the tilde symbol ($\tilde{}$) denotes the Fourier transform on $\mathbb{R}$. In other words if $F\in\mathcal{L}^1[0,1]$ then we write, for $\lambda\in\mathbb{R}$,

$$\tilde{F}(\lambda) \;=\; \int_{-\infty}^{\infty}F(x)e^{ix\lambda}\,dx.$$

Theorem 8 *Let $f:\{1,\ldots,N\}\to\mathbb{R}$ be a function with $|f|=N$, and let v,X be positive integers. Let f be regarded as a function on $\mathbb{Z}_{2N+v}$ in the natural way, and let the hat symbol denote Fourier transforms on that group. Let $p\in\mathcal{C}^1[0,1]$ be such that*

$$\int_0^1 p(x)\,dx \;=\; 2.$$

Then there is a constant C, depending only on p, such that

$$E(X) \;\geq\; \gamma(p)N^4\left(1-C\left(\frac{v}{N}+\frac{N^2}{v^2X}+\frac{X^2}{N}\right)\right),$$

where

$$\gamma(p) \;=\; 2\left(\sum_{r\geq 1}|\tilde{p}(\pi r)|^{4/3}\right)^{-3}.$$

Proof We remark that, for suitable choices of v and X, the dominant term in the above bound will be $\gamma(p)N^4$. Throughout this proof C will denote a constant which depends only on the fixed function p. We follow the traditional convention in analytic number theory of allowing the same letter C to denote different constants! In following the argument the reader may care to recall the brief outline we gave above.

For $x \in [0,2]$ define

$$U(x) \;=\; \begin{cases} 1 & (0 \le x < 1) \\ 1 - p(x-1) & (1 \le x < 2), \end{cases} \tag{2.23}$$

and set, for $x \in \mathbb{Z}_{2N+v}$,

$$G(x) \;=\; U\left(\frac{x}{N+\frac{v}{2}}\right). \tag{2.24}$$

Let I be the characteristic function of the interval $\{1, \ldots, \frac{v}{4}\}$ and write

$$H(x) \;=\; \left(\frac{4}{v}\right)^2 G * I * I(x). \tag{2.25}$$

The idea here is that G is a discretised version of U, and H is a smoothed version of G. Observe that G is equal to 1 for $x = 1, \ldots, N + \frac{v}{2}$, and that $I * I$ is supported in $\{-v/4, v/4\}$. Therefore H is equal to 1 for $x = \frac{v}{4}, \ldots, N + \frac{v}{4}$. It follows that

$$\sum_x f(x) H\left(x + \frac{v}{4}\right) \;=\; N.$$

Applying Parseval's identity and the triangle inequality gives

$$\sum_r |\hat{f}(r)||\hat{H}(r)| \;\ge\; 2N^2. \tag{2.26}$$

In order to use this we require a variety of estimates for $\hat{H}(r)$. These will be of two forms. The first estimate says that, when $|r|$ is small, $\hat{H}(r)$ can be estimated by approximating the sum

$$\hat{H}(r) \;=\; \sum_x H(x)\omega^{rx}$$

by an integral. Observe that here we are using ω to denote the quantity $e^{\frac{2\pi i}{2N+v}}$, because we are working with the group $\mathbb{Z}_{2N+v}$. The second estimate tells us that $\hat{H}(r)$ is small when $|r|$ is at all large. It is for the purpose of proving such a result that we are using the smoothed function H rather than G.

Lemma 9 *Let $r : [0,1] \to \mathbb{R}$ be piecewise continuously differentiable, and let M be an integer. Then*

$$\left| \int_0^1 r(x)\,dx \;-\; \frac{1}{M} \sum_{0 \le n < M} r\left(\frac{n}{M}\right) \right| \;\le\; \frac{\|r'\|_\infty}{M}.$$

Proof This is just an easy application of the Fundamental Theorem of Calculus. □

Observe that

$$|\hat{H}(r)| \;=\; \left(\frac{4}{v}\right)^2 |\hat{G}(r)||\hat{I}(r)|^2. \tag{2.27}$$

It follows that $|\hat{H}(r)| \leq |\hat{G}(r)|$ for all r. Furthermore when $r \neq 0$ we have

$$|\hat{G}(r)| \;=\; \left|\sum_{0\leq x<N+\frac{v}{2}} p\left(\frac{x}{N+\frac{v}{2}}\right)\omega^{rx}\right|.$$

Hence, using this and Lemma 9 with $r(x) = p(x)e^{\pi irx}$ and $M = N + \frac{v}{2}$ gives us what we called our first estimate.

Lemma 10 *Let $0 < |r| < N + \frac{v}{2}$. Then we have the inequality*

$$|\hat{H}(r)| \;\leq\; |\hat{G}(r)| \;\leq\; (N+v)|\tilde{p}(\pi r)| + C|r|.$$

Proof Indeed we can take $C = 4\left(\|p'\|_\infty + \|p\|_\infty\right)$. □

It is possible to prove by very similar means that

$$|\hat{H}(0)| \;\leq\; C. \tag{2.28}$$

We now turn our attentions to estimating $|\hat{H}(r)|$ for large $|r|$. To this end recall (2.27). It follows immediately from Lemma 10 that there is a constant C such that

$$|\hat{G}(r)| \;\leq\; CN \tag{2.29}$$

for all $|r| \leq N + \frac{v}{2}$. We shall also require an upper bound for $|\hat{I}(r)|$.

Lemma 11 *Let $0 < |r| \leq N + \frac{v}{2}$. Then*

$$|\hat{I}(r)| \;\leq\; \frac{3N}{|r|}.$$

Proof By summing a geometric progression it is easy to see that

$$\begin{aligned} |\hat{I}(r)| \;&\leq\; \frac{2}{|\omega^r - 1|} \\ &=\; \left(\sin\left(\frac{\pi r}{2N+v}\right)\right)^{-1}. \end{aligned}$$

It is also a simple matter to verify, for $\theta \in [-\pi/2, \pi/2]$, the inequality

$$|\sin\theta|^{-1} \;\leq\; |\theta|^{-1} + 1.$$

The lemma follows immediately. □

Lemma 12 *There is a constant C such that*

$$|\hat{H}(r)| \ \leq \ \frac{CN^3}{v^2|r|^2}.$$

Proof This follows quickly from (2.27), (2.29) and the previous lemma. □

After all that hard work, we can sit back and simply combine all our estimates with (2.26). A swift application of Hölder's inequality will then take us home.

Using Lemma 12 and the fact that $|f| = N$, we get that

$$\sum_{|r|>X} |\hat{f}(r)||\hat{H}(r)| \ \leq \ \frac{CN^4}{v^2X}.$$

Using this and (2.28) it follows from (2.26) that

$$\sum_{0<|r|\leq X} |\hat{f}(r)||\hat{H}(r)| \ \geq \ 2N^2\left(1 - C\left(\frac{1}{N} + \frac{N^2}{v^2X}\right)\right).$$

Bringing Lemma 10 to bear on this gives, after a little calculation,

$$\sum_{0<|r|\leq X} |\hat{f}(r)||\tilde{p}(\pi r)| \ \geq \ 2N\left(1 - C\left(\frac{v}{N} + \frac{N^2}{v^2X} + \frac{X^2}{N}\right)\right). \tag{2.30}$$

Since both f and p are real-valued we have that $|\hat{f}(r)| = |\hat{f}(-r)|$ and $|\tilde{p}(\pi r)| = |\tilde{p}(-\pi r)|$ for all r. Therefore (2.30) implies that

$$\sum_{1\leq r\leq X} |\hat{f}(r)||\tilde{p}(\pi r)| \ \geq \ N\left(1 - C\left(\frac{v}{N} + \frac{N^2}{v^2X} + \frac{X^2}{N}\right)\right). \tag{2.31}$$

The proof of Theorem 8 can now be concluded by a single application of Hölder's Inequality. □

Apart from the problem of choosing a suitable function p, we can use Theorem 8 to get a lower bound for $M(f)$.

Theorem 13 *Let $f : \{1, \ldots, N\} \to \mathbb{R}$ be a function with $|f| = N$. Let $p \in \mathcal{C}^1[0,1]$ be such that*

$$\int_0^1 p(x)\,dx \ = \ 2.$$

Then

$$M(f) \ \geq \ \left(\frac{1+\gamma(p)}{2}\right)N^4(1 + O(N^{-1/7})),$$

where

$$\gamma(p) \;=\; 2\left(\sum_{r\geq 1} |\tilde{p}(\pi r)|^{4/3}\right)^{-3}.$$

Proof Recall (2.19). Using this and Theorem 8 with $v = N^{6/7}$, $X = N^{3/7}$ gives the result. □

It only remains to choose a good function p, where "good" means that $\int_0^1 p(x)\,dx$ equals 2 and $\gamma(p)$ is as large as possible. Unfortunately we have not been able to give a best possible choice in closed form. A simple function that gives a good bound is

$$p(x) \;=\; \tfrac{5}{2} - 40\left(x - \tfrac{1}{2}\right)^4. \tag{2.32}$$

One can compute that

$$|\tilde{p}(\pi r)| \;=\; \begin{cases} \dfrac{40}{\pi^2|r|^2} - \dfrac{960}{\pi^4|r|^4} & r \text{ even, } r \neq 0 \\ \dfrac{240}{\pi^3|r|^3} - \dfrac{1920}{\pi^5|r|^5} & r \text{ odd,} \end{cases} \tag{2.33}$$

and then that

$$\gamma(p) \;=\; \frac{2\left(\frac{\pi^2}{40}\right)^4}{\left(S_1 + \left(\frac{6}{\pi}\right)^{4/3} S_2\right)^3}, \tag{2.34}$$

where

$$S_1 \;=\; \sum_{\substack{r\text{even} \\ r\geq 0}} \left|\frac{1}{r^2} - \frac{24}{\pi^2 r^4}\right|^{4/3} \tag{2.35}$$

and

$$S_2 \;=\; \sum_{\substack{r\text{odd} \\ r\geq 0}} \left|\frac{1}{r^3} - \frac{8}{\pi^2 r^5}\right|^{4/3}. \tag{2.36}$$

There seems to be little hope of an analytic expression for these series, but one can compute that $S_1 \approx 0.0839757$, $S_2 \approx 0.1219299$. It then follows from (2.34) that $\gamma(p) > 1/7$ (in fact it turns out that $\gamma(p) \approx 1/6.9994$).

To conclude this section we restate Theorems 8 and 13 with this particular choice of p.

Theorem 14 *Let $f : \{1, \ldots, N\} \to \mathbb{R}$ be a function with $|f| = N$, and let v, X be positive integers. Let f be regarded as a function on $\mathbb{Z}_{2N+v}$ in the natural way, and let the hat symbol denote Fourier transforms on that group. Write*

$$E(X) \;=\; \sum_{0<|r|<X} |\hat{f}(r)|^4.$$

Then there is an absolute constant C such that

$$E(X) \ \geq \ \tfrac{1}{7}N^4\left(1 - C\left(\frac{v}{N} + \frac{N^2}{v^2X} + \frac{X^2}{N}\right)\right).$$

Theorem 15 *Let $f : \{1, \dots, N\} \to \mathbb{R}$ be a function with $|f| = N$. Then we have*

$$M(f) \ \geq \ \tfrac{4}{7}N^4$$

for all sufficiently large N.

2.4 A Return to B_4 Sets

Recall that our knowledge of B_4 sets is currently encapsulated in (2.17), which we urge the reader to reconsider now. In the remarks following that equation we stated that our goal would be to say something about the Fourier coefficients $\hat{A}(r)$ with r non-zero and small. We now have this knowledge, in the form of Theorem 14.

In (2.17) we must contend with the presence of $|\hat{I}(r)|^2$. Notice, however, that $|\hat{I}(r)|$ will differ insignificantly from u if $|r| \ll N/u$. Indeed

Lemma 16

$$|\hat{I}(r)| \ \geq \ u - \frac{\pi|r|u^2}{N}.$$

Proof Let $\omega = e^{2\pi i/(2N+v)}$. Then we have

$$\left|u - \hat{I}(r)\right| \ \leq \ \sum_{x=1}^{u}|1 - \omega^{rx}| \ \leq \ \frac{2\pi|r|u^2}{2N+v} \ \leq \ \frac{\pi|r|u^2}{N},$$

which is all we need. □

Applying Theorem 14 with $f(x) = NA(x)/|A|$ gives

$$\begin{aligned}\sum_r |\hat{A}(r)|^4|\hat{I}(r)|^2 \ &\geq \ \sum_{|r|\leq X}|\hat{A}(r)|^4|\hat{I}(r)|^2 \\ &\geq \ \tfrac{8}{7}u^2|A|^4\left(1 - C\left(\frac{v}{N} + \frac{N^2}{v^2X} + \frac{X^2}{N}\right)\right)\left(1 - \frac{\pi Xu}{N}\right)^2,\end{aligned}$$

where C is an absolute constant. From (2.17) we now have

$$8N + 2^{24}\left(v + \frac{N^{7/4}}{u}\right) \ \geq \ \tfrac{8}{7}|A|^4\left(1 - C\left(\frac{v}{N} + \frac{N^2}{v^2X} + \frac{X^2}{N} + \frac{Xu}{N}\right)\right). \tag{2.37}$$

We must now choose suitable values for u, v and X, recalling that we require $u \leq v$. There are many such choices and one is $u = N^{13/17}$, $v = N^{16/17}$, $X = N^{3/17}$. With these values in (2.37) we get

Theorem 17

$$A(4,N) \ \leq \ 7^{1/4}N^{1/4}(1+o(1)).$$

We now turn our attention to B_3-sets. We shall be very brief here as there very little difference between this and the case of B_4-sets.

Lemma 18 *Let $A \subseteq \{1,\dots,N\}$ be a B_3-set. Then*

$$A*A*A*A(x) \ \leq \ 2|A|(1+(A*A)(x)).$$

Proof A counting argument very similar to that in Lemma 4 gives

$$A*A*A(x) \ \leq \ 2(1+|A|A(x)).$$

Using the fact that

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

the lemma follows immediately. □

The remainder of the derivation is almost exactly as before. One winds up with

Theorem 19

$$A(3,N) \ \leq \ \left(\tfrac{7}{2}\right)^{1/3} N^{1/3}(1+o(1)).$$

2.5 Large Values of h

In this section we again consider upper bounds for B_h-sets. Our aim is to convince the reader that the methods we have just been using for B_3 and B_4 sets generalise rather easily to the case $h \geq 5$. For any given value of $h \geq 5$, the difficulties we have experienced optimising our approach (i.e. choosing a good function p) are even more apparent. Therefore we shall not, in the sequel, discuss any such specific value of h. Rather we shall turn our attention to the behaviour of our methods as h becomes large. It turns out that rather simple ideas constitute essentially the best possible application of the methods of this paper.

Let $f : G \to \mathbb{R}$ be a function on an abelian group G. Then throughout this section we will write f^{*k} for the k-fold convolution of f with itself.

The following proposition, a sort of generalisation of Theorem 14, will be our main tool. This comes as little surprise.

Proposition 20 *Let k be a positive integer. Let $f : \{1, \dots, N\} \to \mathbb{R}^+$ be a function, and regard f as a function on $\mathbb{Z}_{kN+v}$ in the natural way. Here k is to be regarded as fixed (but large), and N, v are positive integers with $v \ll N$. Then we have*

$$\sum_{|r| \leq k/2} |\hat{f}(r)|^{2k} \;\geq\; \frac{1}{\sqrt{\pi}} k^{1/2} (1 - \varepsilon(k)) |f|^{2k}, \tag{2.38}$$

where $\varepsilon(k) \to 0$ as $k \to \infty$.

Proof The idea is rather simple, but (as the form of (2.38) might suggest) some fairly careful analysis is required. Throughout the following k will be taken sufficiently large. We have

$$\begin{aligned} |\hat{f}(r)| &= \left| \sum_x f(x) \omega^{r\left(x - \frac{N}{2}\right)} \right| \\ &\geq \sum_x f(x) \cos\left(\frac{2\pi r \left(x - \frac{N}{2}\right)}{kN + v} \right). \end{aligned}$$

Therefore if $|r| \leq k/2$ we have

$$|\hat{f}(r)| \;\geq\; |f| \cos\left(\frac{\pi r}{k}\right).$$

Now we simply compute

$$\begin{aligned} |f|^{-2k} \sum_{|r| \leq k/2} |\hat{f}(r)|^{2k} &\geq \sum_{|r| \leq k/2} \left| \cos \frac{\pi r}{k} \right|^{2k} \\ &\geq \sum_{|r| \leq k^{5/8}} \left| 1 - \frac{\pi^2 r^2}{2k^2} \right|^{2k} \\ &\geq \left| 1 - \frac{25}{k^{3/2}} \right|^{2k} \sum_{|r| \leq k^{5/8}} e^{-\pi^2 r^2 / k}. \end{aligned}$$

In this last step we have used the inequality

$$1 - x \;\geq\; e^{-x}(1 - x^2),$$

which holds for $x \leq 1$. Now one observes that

$$\left| 1 - \frac{25}{k^{3/2}} \right|^{2k} \longrightarrow 1$$

as $k \to \infty$, and that

$$\frac{1}{\sqrt{k}} \sum_{|r| \le k^{5/8}} e^{-\pi^2 r^2/k} \longrightarrow \int_{-\infty}^{\infty} e^{-\pi^2 x^2}\, dx \;=\; \frac{1}{\sqrt{\pi}}.$$

The proposition now follows. The conscientious reader may care to check that we can even take $\varepsilon(k) = 100k^{-1/8}$. □

Interestingly the above is essentially best possible. We take the opportunity to sketch a proof of this fact now.

Proposition 21 *The bound of Proposition 20 is best possible in that the constant* $1/\sqrt{\pi}$ *cannot be increased.*

Proof Let $\delta > 0$ and let χ be the characteristic function of the set A_δ, which consists of the integers $n \in \{-N/2, \dots, N/2\}$ with $|n| \ge (1-\delta)N/2$. Let $f(n) = \delta^{-1}\chi(n)$, so that $|f| = N(1 + o(1))$. For $x \in [-1/2, 1/2]$ define

$$g(x) \;=\; \begin{cases} \delta^{-1} & (|x| \ge (1-\delta)/2) \\ 0 & (\text{otherwise}) \end{cases}$$

Now g is the probability density function of a random variable X with mean 0 and variance

$$\sigma^2 \;=\; \frac{1-(1-\delta)^3}{12\delta} \;>\; \tfrac{1}{4}(1-\delta). \tag{2.39}$$

Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent identically distributed random variables with density g. Then, by a suitable version of the Central Limit Theorem (see [19]), the density function of

$$\frac{X_1 + \cdots + X_m}{\sigma\sqrt{m}}$$

tends to the standard normal $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ uniformly in x. In other words,

$$\sqrt{m} \cdot g^{*m}(x\sigma\sqrt{m}) \longrightarrow \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2}$$

uniformly in x. In particular for $m \ge m(\delta)$ we have, using (2.39), that

$$g^{*m}(0) \;\le\; \sqrt{\frac{2}{m\pi(1-\delta)}}. \tag{2.40}$$

Now let k be a fixed positive integer, and regard f as a function on $\mathbb{Z}_{kN}$ in the natural way (in Proposition 20 we worked with $\mathbb{Z}_{kN+v}$, but we choose to ignore this technicality here). Using the hat symbol to denote Fourier transforms on this group, we have

$$\begin{aligned} \sum_r |\hat{f}(r)|^{2k} &= kN \sum_x |f^{*k}(x)|^2 \\ &= kN|f^{*2k}(0)|. \end{aligned} \tag{2.41}$$

Note that the modular version of $|f^{*2k}(0)|$ is the same as the $\mathbb{Z}$-version because f is supported in $\{-N/2, N/2\}$. It is thus not hard to see that, as $N \to \infty$, we have

$$\frac{f^{*2k}(0)}{N^{2k-1}} \longrightarrow g^{*2k}(0).$$

It follows from (2.41) that

$$|f|^{-2k} \sum_r |\hat{f}(r)|^{2k} \longrightarrow k g^{*2k}(0),$$

again as $N \to \infty$. If $2k \geq m(\delta)$ (as defined earlier) then this implies, by (2.40), that

$$|f|^{-2k} \sum_r |\hat{f}(r)|^{2k} \ \leq \ \sqrt{\frac{k}{\pi(1-2\delta)}}$$

for N sufficiently large. Since δ can be chosen arbitrarily small, the proposition follows. $\Box$

We turn now to the business of actually using Proposition 20 to get information about B_h sets. We shall be extremely brief, as almost all of the relevant ideas have been covered in §2.2 (which it may help to recall at this point). First of all we require generalisations of Lemmas 4 and 18.

Lemma 22 *Let $h = 2k$ be a positive even integer, and let $A \subseteq \{1, \ldots, N\}$ be a B_h-set. Then, for any $x \in \mathbb{Z}$, we have that*

$$A^{*2k}(x) \ \leq \ (k!)^2 + k^2 |A| A^{*(2k-2)}(x).$$

Lemma 23 *Let $h = 2k-1$ be a positive odd integer, and let $A \subseteq \{1, \ldots, N\}$ be a B_h-set. Then, for any $x \in \mathbb{Z}$, we have that*

$$A^{*2k}(x) \ \leq \ |A| \left(k!(k-1)! + k(k-1) A^{*(2k-2)}(x) \right).$$

Regard A as a subset of $\mathbb{Z}_{kN+v}$, and let I be the characteristic function of $\{1, \ldots, u\}$ where $u \ll v \ll N$. In the case $h = 2k$, an appropriate generalisation of (2.17) (which may be proved in exactly the same way) is

$$k(k!)^2 N u^2 + (2k+4)! \left(N^{2-1/2k} u + u^2 v \right) \ \geq \ \sum_r |\hat{A}(r)|^{2k} |\hat{I}(r)|^2.$$

Applying Proposition 20 and Lemma 16 quickly gives

$$k(k!)^2 N + (2k+4)! \left(\frac{N^{2-1/2k}}{u} + v \right) \ \geq \ \frac{1}{\sqrt{\pi}} k^{1/2} (1 - \varepsilon(k)) |A|^{2k} \left(1 - \frac{\pi k u}{N} \right)^2,$$

where here (and in the following) $\varepsilon(k) \to 0$ as $k \to \infty$. Taking $u = N^{1-1/3k}$ and $v = N^{1-1/4k}$ gives the following improvement of (2.9).

Theorem 24 *Let $A \subseteq \{1, \dots, N\}$ be a B_{2k}-set. Then*

$$|A| \ \leq \ \left(\pi^{1/2} k^{1/2} (k!)^2 (1 + \varepsilon(k))\right)^{1/2k} N^{1/2k} (1 + o(1)).$$

The case $h = 2k - 1$ may be treated similarly and one winds up with

Theorem 25 *Let $A \subseteq \{1, \dots, N\}$ be a B_{2k-1}-set. Then*

$$|A| \ \leq \ \left(\pi^{1/2} k^{-1/2} (k!)^2 (1 + \varepsilon(k))\right)^{1/(2k-1)} N^{1/(2k-1)} (1 + o(1)).$$

2.6 New Bounds for $B_2[g]$-Sets Part I

In this section we apply the results of §2.3 to the problem of bounding $A(2, g, N)$ above. We show that our ideas lead to a non-trivial bound which is stronger than that of [8] for $g \leq 68$. Let $A \subseteq \{1, \dots, N\}$ be a $B_2[g]$-set. Let $0 < v \leq N$ be an integer to be chosen later, and regard A as a subset of $\mathbb{Z}_{2N+v}$ in the obvious way. Then we have

$$M(A) \ = \ \sum_x (A + A)(x)^2 \ \leq \ 2g \sum_x (A + A)(x) \ = \ 2g|A|^2. \tag{2.42}$$

On the other hand, using the hat symbol to denote Fourier transforms in $\mathbb{Z}_{2N+v}$, we have

$$(2N + v) M(A) \ = \ \sum_r |\hat{A}(r)|^4. \tag{2.43}$$

Now let X be another positive integer to be chosen later, and split the sum in (2.43) into the two parts

$$\Sigma_1 \ = \ \sum_{|r| \leq X} |\hat{A}(r)|^4$$

and

$$\Sigma_2 \ = \ \sum_{X < |r| \leq N + \frac{v}{2}} |\hat{A}(r)|^4.$$

Applying Proposition 14 with $f(x) = NA(x)/|A|$ gives

$$\Sigma_1 \ \geq \ \tfrac{8}{7} |A|^4 \left(1 - \frac{C}{N} \left(v + \frac{N^3}{v^2 X} + X^2\right)\right). \tag{2.44}$$

Furthermore by the Cauchy-Schwarz inequality, Parseval's identity and the trivial bound $|A| \leq (4g)^{1/2} N^{1/2}$ we have

$$\begin{aligned} \Sigma_2 \ &\geq \ \frac{1}{2N + v} \left(\sum_{X < |r| \leq N + \frac{v}{2}} |\hat{A}(r)|^2 \right)^2 \\ &\geq \ \frac{1}{2N + v} \left((2N + v)|A| - 2X|A|^2 \right)^2 \\ &\geq \ (2N + v) \left(|A| - 12gX \right)^2. \end{aligned} \tag{2.45}$$

Taking $X = N^{3/7}$, $v = N^{6/7}$ in the above and doing a little calculation with (2.42), (2.43), (2.44) and (2.45) gives the following result.

Theorem 26 *We have*

$$A(2,g,N) \ \leq \ \left(\tfrac{7}{2}g - \tfrac{7}{4}\right)^{1/2} N^{1/2}(1+o(1)).$$

This improves on the bound of [8] *for* $g \leq 68$*, and in particular*

$$A(2,2,N) \ \leq \ \left(\tfrac{21}{4}\right)^{1/2} N^{1/2}(1+o(1)).$$

2.7 New Bounds for $B_2[g]$-Sets Part II

In this section we discuss the paper [8]. We begin by translating the techniques of that paper into the language we have been using here, a relatively easy task. We then show that our methods complement those of [8], in the sense that we can get an improved bound on $A(2,g,N)$. We shall in fact prove the following result.

Theorem 27

$$A(2,g,N) \ \leq \ \left(\tfrac{17}{5}\right)^{1/2} g^{1/2} N^{1/2}(1+o(1)).$$

Observe that $(17/5)^{1/2} = 1.84391..$ (in fact our method gives the slightly better constant 1.84385). This is weaker than the bound of Theorem 26 for $g \leq 18$, but stronger than the bound of [8] for all g. We give this bound more as an illustration of the sort of techniques that might be useful in this problem rather than for the constant 17/5 itself.

We now translate the technique used in [8] to obtain the bound (2.6) into the language of our paper. Let $A \subseteq \{1,\dots,N\}$ be a $B_2[g]$ set, and regard A as a subset of $\mathbb{Z}_{2N}$. Define $f(x) = 2g - (A+A)(x)$, so that $0 \leq f(x) \leq 2g$ for all $x \in \mathbb{Z}_{2N}$. For any $r \neq 0$ the Fourier transform $\hat{f}(r)$ is simply $-\hat{A}(r)^2$. Hence we have, if $\omega = e^{2\pi i/2N}$,

$$\begin{aligned} |\hat{A}(r)|^2 &= \left|\sum_x (2g - (A+A)(x))\,\omega^{rx}\right| \\ &\leq \sum_x |2g - (A+A)(x)| \\ &= 4Ng - |A|^2. \end{aligned} \tag{2.46}$$

Noting that $A \subseteq \{1,\dots,N\}$, Cilleruelo, Ruzsa and Trujillo show that A must have a large non-zero Fourier coefficient $\hat{A}(r)$. The technique used to do this bears some resemblance to the techniques we used earlier to show that, under the same hypotheses, $\sum_{r\neq 0} |\hat{A}(r)|^4$

cannot be too small. One finds a non-negative function f, supported on $\{1,\dots,N\}$, for which $\sum_{r\neq 0}|\hat{f}(r)|$ is large compared to $|f|$. Observing that

$$\sum_x A(x)f(x+N) \;=\; 0,$$

one uses Parseval's identity to conclude that

$$\sum_{r\neq 0}|\hat{A}(r)||\hat{f}(r)| \;\geq\; |A||f|,$$

from which it follows that

$$\sup_{r\neq 0}|\hat{A}(r)|\cdot\sum_{r\neq 0}|\hat{f}(r)| \;\geq\; |A||f|.$$

We suppress a more detailed discussion, referring the interested reader to [8].

For the rest of the paper define

$$N_\infty(A) \;=\; \frac{1}{|A|}\sup_{r\neq 0}|\hat{A}(r)|$$

and

$$N_4(A) \;=\; \frac{1}{|A|^4}\sum_{r\neq 0}|\hat{A}(r)|^4.$$

The situation may be summarised by saying that [8] obtains information from $N_\infty(A)$, whereas we profited from consideration of $N_4(A)$. Since these are rather different objects, it is not altogether surprising that a stronger bound can be achieved by playing the two approaches off against one another. The remainder of the paper, which aims to show that this is indeed so, consists of three parts. In Step 1 we show that a lower bound on $N_\infty(A)$ can be used slightly more effectively than was done in (2.46). In Step 2 we show how a lower bound on $N_4(A)$ gives information in a rather simpler (but slightly weaker) way than in §2.6. This keeps the whole argument manageable. Finally, in Step 3, we show that considering $N_\infty(A)$ and $N_4(A)$ together gives stronger information, for large g, than separate consideration of either $N_\infty(A)$ (cf. [8]) or $N_4(A)$ (cf. §2.6).

Step 1. Let us reconsider the derivation (2.46). There was only one inequality, but it was rather crude. Our task here is to improve it.

Lemma 28 *Let L, M, R be integers with $L+1 \leq M/R$. Let $\mathcal{C}(M,R)$ be the set of all functions $f:\mathbb{Z}_L\to\{0,1,2,\dots\}$ with $|f|=M$ and $f(x)\leq R$ for all x. Then, for all $f\in\mathcal{C}(M,R)$,*

$$|\hat{f}(1)| \;\leq\; R\left|\frac{\sin\frac{\pi}{L}\left(\frac{M}{R}+1\right)}{\sin\frac{\pi}{L}}\right|.$$

Proof The result clearly generalises to $|\hat{f}(r)|$ with $r \neq 0$. It is saying nothing more than that $|\hat{f}(1)|$ is maximised, among all functions in $\mathcal{C}(M,R)$, when f is as concentrated as possible. We prove this using a sort of compression argument. Let f be a member of $\mathcal{C}(M,R)$ with $|\hat{f}(1)|$ maximal. Let $u, v \in \mathbb{Z}_L$ be such that $f(u) > 0$ and $f(v) < M$, and define a new function $g_{u,v} \in \mathcal{C}(M,R)$ by

$$\begin{array}{rcll} g(x) & = & f(x) & (x \neq u, v) \\ g(u) & = & f(u) - 1 & \\ g(v) & = & f(v) + 1. & \end{array}$$

Let $\omega = e^{2\pi i/L}$ and suppose that $a \in [0, L)$ is such that $\hat{f}(1) = |\hat{f}(1)|\omega^{ra}$ (a need not be an integer). Then one can check that

$$\begin{aligned} |\hat{g}(1)| &= \left| |\hat{f}(1)| + \omega^{v-a} - \omega^{u-a} \right| \\ &\geq |\hat{f}(1)| + \cos\left(\frac{2\pi(v-a)}{L}\right) - \cos\left(\frac{2\pi(u-a)}{L}\right). \end{aligned}$$

In words, $|\hat{g}(1)|$ is greater than $|\hat{f}(1)|$ if $|v-a| < |u-a|$, where distance is measured on the torus $[0, L)$ which contains $\mathbb{Z}_L$ as a subgroup. By the extremal property of f, this means that we cannot select u and v with $|v-a| < |u-a|$, $f(u) > 0$ and $f(v) < M$. In other words f is as concentrated about a as possible. □

We can now use this lemma in place of the inequality (2.46), with $f(x) = 2g - (A+A)(x)$, $L = 2N$ and $M = 2g$. One gets for any $r \neq 0$ that

$$|\hat{A}(r)|^2 \leq 2g \left| \frac{\sin \frac{\pi}{2N}\left(\frac{4Ng - |A|^2}{2g} + 1\right)}{\sin \pi/2N} \right|.$$

Writing (here and for the rest of the paper)

$$Q = \frac{|A|^2}{4Ng},$$

this simplifies to

$$|\hat{A}(r)|^2 \leq 2g \left| \frac{\sin\left(\pi Q - \frac{\pi}{2N}\right)}{\sin(\pi/2N)} \right|.$$

Recalling that $Q < 1$ uniformly in N (a consequence of Theorem 26) this reduces yet further to give

$$|\hat{A}(r)|^2 \leq \frac{4gN \sin \pi Q}{\pi}(1 + o(1)).$$

Finally this implies our strengthened version of (2.46), namely

$$\frac{\sin \pi Q}{\pi Q} \ \geq \ N_\infty(A)^2(1+o(1)). \tag{2.47}$$

Step 2. In this brief section we show that upper bounds for $B_2[g]$ sets are related to lower bounds for $N_4(A) = \sum_{r\neq 0} |\hat{A}(r)|^4/|A|^4$ in a very simple way (which is a little weaker than the approach taken in §2.6). Let $A \subseteq \{1,\dots,N\}$ be a $B_2[g]$ set and regard A as a subset of $\mathbb{Z}_{2N}$. Then

$$\begin{aligned} |A|^4(1+N_4(A)) &= \sum_r |\hat{A}(r)|^4 \\ &= 2N\sum_x (A+A)(x)^2 \\ &\leq 4Ng\sum_x (A+A)(x) \\ &= \frac{|A|^4}{Q}. \end{aligned}$$

Hence

$$Q \ \leq \ \frac{1}{1+N_4(A)} \ . \tag{2.48}$$

Step 3. Unfortunately this section is a touch computational. We trust the reader will accept our apologies for this. At this point we take the opportunity to recall equation (2.31), from which we deduced Theorem 8 by a simple application of Hölder's Inequality. Let $f : \{1,\dots,N\} \to \mathbb{R}$ be a function with $|f| = N$, and regard f as a function on $\mathbb{Z}_{2N+v}$. Let $p : [0,1] \to \mathbb{R}$ be a continuously differentiable function with $\int_0^1 p(x)\,dx = 2$. Then, using the hat and tilde symbols to denote Fourier transforms on $\mathbb{Z}_{2N+v}$ and $\mathbb{R}$ respectively, we had

$$\sum_{1\leq r\leq X} |\hat{f}(r)||\tilde{p}(\pi r)| \ \geq \ N\left(1 - C\left(\frac{v}{N} + \frac{N^2}{v^2X} + \frac{X^2}{N}\right)\right). \tag{2.49}$$

Take $f(x) = NA(x)/|A|$ and $p(x) = \frac{5}{2} - 40\left(x - \frac{1}{2}\right)^4$ (as we did in proving Theorem 8). With $X = N^{3/7}$ and $v = N^{6/7}$, (2.49) gives

$$\sum_{1\leq r<N} |\hat{A}(r)||\tilde{p}(\pi r)| \ \geq \ |A|(1+o(1)). \tag{2.50}$$

Suppose that $|\hat{A}(1)| = \alpha|A|$. Then, recalling (2.33), we get that

$$\sum_{2\leq r<N} |\hat{A}(r)||\tilde{p}(\pi r)| \ \geq \ |A|\left(1 - \alpha\left(\frac{240}{\pi^3} - \frac{1920}{\pi^5}\right)\right)(1+o(1)). \tag{2.51}$$

Suppose that $\alpha < \frac{7}{10}$, so that

$$1 - \alpha\left(\frac{240}{\pi^3} - \frac{1920}{\pi^5}\right) \ \geq \ 0.$$

Then we may apply Hölder's Inequality to (2.51) to obtain, after a few calculations, that

$$N_4(A) \ \geq \ \left(2\alpha^4 + \frac{2\left(\frac{\pi^3}{240}\right)^4\left(1-\alpha\left(\frac{240}{\pi^3}-\frac{1920}{\pi^5}\right)\right)^4}{\left(\left(\frac{\pi}{6}\right)^{4/3}S_1 + S_2 - \left(1-\frac{8}{\pi^2}\right)^{4/3}\right)^3}\right)(1+o(1)) \tag{2.52}$$

$$\approx \ \left(2\alpha^4 + 4.8607836\,(1-1.4662621\alpha)^4\right)(1+o(1)), \tag{2.53}$$

where S_1 and S_2 are the sums appearing in (2.35) and (2.36). Call the polynomial appearing in (2.53) $p(\alpha)$. Then one can check that $p'(\alpha) < 0$ for $\alpha \in [0, 0.47]$. Hence, putting $\alpha = 0.4124078$ in (2.53), we get that either

$$N_\infty(A) \ \geq \ \frac{|\hat{A}(1)|}{|A|} \ \geq \ 0.4124078 \tag{2.54}$$

or else

$$N_4(A) \ \geq \ 0.1765468\,(1+o(1)). \tag{2.55}$$

If (2.55) holds then, by (2.48), we have

$$Q \ \leq \ 0.8499448\,(1+o(1)).$$

If (2.54) holds then, by (2.47) and a little easy computation we have again that

$$Q \ \leq \ 0.8499448\,(1+o(1)).$$

Either way, it is easy to see (recalling that $Q = |A|^2/4Ng$) that Theorem 27 is true. □

2.8 Summary

In view of the proliferation of different upper bounds mentioned in Chapter 2 we think it necessary to provide a summary. It should also be noted that since I submitted [17] there have been several other papers on the subject of bounds for $A(h,g,N)$ by Cilleruelo [7], Plagne [32] and Plagne - Habsieger [21]. So far, however, none of the bounds in [17] seem to have been improved. Recall that we wrote $\alpha(h,g)$ for $\limsup_{N\to\infty} N^{-1/h}A(h,g,N)$ and $\alpha(h)$ for $\alpha(h,1)$.

(i) Bounds for $\alpha(h)$.

- $1 \leq \alpha(3) \leq (7/2)^{1/3}$ [3],[17]
- $1 \leq \alpha(4) \leq 7^{1/4}$ [3],[17]
- $1 \leq \alpha(h) \leq \frac{1}{2e}\left(h + \frac{3}{2}\log h + o_h(\log h)\right)$ [3],[17]

The last inequality is what one gets by applying some careful asymptotics to Theorems 24 and 25. Note, however, that the results stated in those theorems are rather more exact than the inequality appearing here.

(ii) Bounds for $\alpha(2, g)$.

- $4/\sqrt{7} \leq \alpha(2,2) \leq \sqrt{21}/2$ [21],[17]
- $\frac{3\sqrt{2}}{4}\sqrt{g}(1 + o_g(1)) \leq \alpha(2, g) \leq \left(\frac{7}{2}g - \frac{7}{4}\right)^{1/2}$ [8],[17]
- $\frac{3\sqrt{2}}{4}\sqrt{g}(1 + o_g(1)) \leq \alpha(2, g) \leq (17g/5)^{1/2}$ [8],[17]

Observe that the first of these bounds is stronger when $g \leq 19$, but that the second is superior for $g \geq 20$.

Chapter 3

On Arithmetic Structures in Dense Sets of Integers

3.1 Introduction

If one believes that mathematics is the study of patterns then it is of no surprise that the following result of Szemerédi [43] is often regarded as one of the highlights of all combinatorics.

Theorem 29 (Szemerédi) *Let $\alpha > 0$ be a real number and let k be a positive integer. Then there is $N_0 = N_0(k, \alpha)$ such that any subset $A \subseteq \{1, \ldots, N\}$ of size at least αN contains an arithmetic progression of length k, provided that $N \geq N_0(k, \alpha)$.*

Szemerédi's proof was long and combinatorial but just two years later Furstenburg provided a completely different proof of Theorem 29 using ergodic theory. Furstenburg's methods have proved extremely amenable to generalisation, and Furstenburg himself proved the following result.

Theorem 30 *Fix $\alpha > 0$ and let $A \subseteq \{1, \ldots, N\}$ have size αN. Then provided $N > N_1(\alpha)$ is sufficiently large one can find two distinct elements $x, x' \in A$ whose difference $x - x'$ is a perfect square.*

As with all such applications of ergodic theory Furstenburg's approach gave no bound on $N_1(\alpha)$. At about the same time Sárközy [37] proved the same result in a completely different manner. Sárközy's argument took inspiration from a much earlier paper of Roth [33] in which Szemerédi's Theorem was proved for progressions of length 3. The method used is analytic in spirit and does lead to an effective bound on $N_1(\alpha)$, albeit one which is a great distance from the conjectured truth.

The paper [37] was the first in a series of three, and in the final paper [38] of this series an analytic proof of a generalisation of Theorem 30 was outlined. This generalisation says that

the squares may, in the formulation of that theorem, be replaced by the set $\{p(d) : d \in \mathbb{N}\}$ where p is any polynomial which maps $\mathbb{N}$ to itself and has an integer root. To see that some restriction on the polynomial is necessary for such a result to hold, we invite the reader to construct a set with density $\frac{1}{3}$ containing no difference of the form $x^2 + 1$.

Since the late 1970s there have been several significant advances in our understanding of these and related questions. Firstly we mention the result of Bergelson and Leibman from 1996 [1], which is an example of how far Furstenburg's ergodic-theoretic methods have been able to take us.

Theorem 31 (Bergelson-Leibman) *Fix $\alpha > 0$ and let $A \subseteq \{1, \dots, N\}$ have size αN. Let $p_1, \dots, p_r$ be polynomials with $p(\mathbb{N}) \subseteq \mathbb{N}$ and $p(0) = 0$. Then provided $N > N_2(\alpha)$ is large enough (exactly how large will depend on the polynomials p_i as well as on α) we can find $a, d \in \mathbb{N}$ for which all r of the numbers $a + p_i(d)$ lie in A.*

This extends both Theorem 29 and Theorem 30 and implies, amongst other things, that dense subsets of the integers contain arbitrarily long arithmetic progressions whose common difference is a square. Secondly there is a result of Gowers [12], which gives the first bounds for Szemerédi's Theorem.

Theorem 32 (Gowers) *Let $k > 0$ be an integer. Then there is an effectively computable constant $c(k)$ such that any subset of $\{1, \dots, N\}$ of density at least $(\log \log N)^{-c(k)}$ contains a k-term arithmetic progression.*

Gowers' argument takes inspiration from Roth's paper [33]. As we have already remarked, Sárközy's methods also bear some resemblance to those of Roth. It is therefore natural to ask whether Gowers' techniques can be adapted to give bounds for questions related to the polynomial Szemerédi Theorem.

In §3.2 we give a new variant on Sárközy's proof of Theorem 30 which we believe to be substantially easier to understand than the original (though it gives a worse bound for $N_1(\alpha)$). Perhaps more importantly we demonstrate that this argument can be made to fit almost entirely into the general methodology of [12], which we shall outline below. It should be pointed out that in 1985 Srinivasan [42] gave an argument which seems to be rather simpler than that of Sárközy. The three arguments (those of Sárközy, Srinivasan and myself) differ slightly in substance and fairly substantially in notation, a factor which is very important in generalising the approach. The rest of Chapter 3 is devoted to a proof of the following result.

Theorem 33 *There is a constant c such that any subset of $\{1, \dots, N\}$ of density at least $(\log \log N)^{-c}$ contains a 3-term arithmetic progression whose common difference is of the form $x^2 + y^2$.*

The proof of this result involves finding a mutual generalisation of the methods of Gowers and Sárközy, together with a slightly surprising use of the Selberg Sieve.

The reader will find it hard to understand this section unless she has a working knowledge of the methods of Gowers, such as can be obtained by reading [11] or, even better, [12]. It would be a sizeable undertaking to summarise those papers in any detail here, and indeed there seems little point in doing so.

What we will do is offer a crude outline of the top-level structure of Gowers' proof of Szemerédi's Theorem. All our arguments in this chapter will have this broad structure at their heart. Suppose then that we have a set $A \subseteq \{1, \dots, N\}$ of size δN which contains no arithmetic progression of length k. Set $A_0 = A$, $\delta_0 = \delta$ and $N_0 = N$.

- At the ith stage of our argument we will have a set $A_i \subseteq \{1, \dots, N_i\}$ with density δ_i which contains no arithmetic progression of length k.
- The fact that A_i contains no progressions of length k implies that A_i is *non-random* in a certain rather precise sense involving certain Fourier coefficients being large. This means that A_i does not satisfy a property which is possessed by almost all sets of density δ_i. In Gowers' proof this property is known as $(k-2)$-uniformity.
- If a set is not $(k-2)$-uniform then, by a long and complicated argument, it is possible to show that A_i has density at least $\delta_i + \eta(\delta_i)$ on a fairly long progression P, where η is an increasing function of δ_i.
- Define A_{i+1} to be $A_i \cap P$ and rescale so that P has common difference 1. Set $\delta_{i+1} = \delta_i + \eta(\delta_i)$ and $N_{i+1} = |P|$.

Iterating this argument leads to an effective version of Szemerédi's Theorem since after a finite number of steps the density δ_i will exceed 1, a contradiction.

To conclude this introduction we introduce the concepts of uniformity and quadratic uniformity, the only types of uniformity that will feature in this chapter. These correspond to what we called 1-uniformity and 2-uniformity in the line above.

Let $A \subseteq \mathbb{Z}_N$ be a set of size αN and let $f = A - \alpha$ be its *balanced function*. We say that A is η-uniform if $\|\hat{f}\|_\infty \leq \eta N$. To define quadratic uniformity we need some extra notation. If $f : G \to \mathbb{C}$ is any function on an abelian group we write

$$\Delta(f;h)(x) \;=\; f(x)f(x-h).$$

Let f be the balanced function of A. Then A is said to be *quadratically η-uniform* if

$$\sum_h \|\Delta(f;h)\hat{}\,\|_\infty^2 \;\leq\; \eta N^3. \tag{3.1}$$

These definitions are very similar, though not completely identical, to those in [12]. For a detailed discussion of the basic properties of uniformity and quadratic uniformity the reader should consult [12]. It turns out that being quadratically uniform is a stronger requirement than being uniform (and so knowing a set fails to be quadraticlly uniform gives less information than knowing it fails to be uniform). We will need the following quantitative statement of this fact later on.

Proposition 34 *If A is quadratically η-uniform then it is $\eta^{1/4}$-uniform.*

Proof By (3.1) we have

$$\sum_h |\Delta(f;h)\hat{}(0)|^2 \ \leq \ \eta N^3.$$

Writing this statement out in full gives

$$\sum_{a+b=c+d} f(a)f(b)f(c)f(d) \ \leq \ \eta N^3$$

which implies, by Parseval's theorem, that

$$\sum_r |\hat{f}(r)|^4 \ \leq \ \eta N^4.$$

It follows immediately that $\|\hat{f}\|_\infty^4 \leq \eta N^4$, which implies the proposition. □

3.2 A Short Proof of Sárközy's Theorem for Squares

In this section we prove Theorem 30 using an iterative argument of the type outlined above. For obvious reasons we say that A has a *square difference* if there are distinct elements x and x' in A with $x - x'$ a square. Let us recall Theorem 30.

Theorem 30 *Fix $\alpha > 0$ and let $A \subseteq \{1, \dots, N\}$ have size αN. Then provided $N > N_1(\alpha)$ is sufficiently large A has a square difference.*

As the first part of our argument we shall show that a set $A \subseteq \{1, \dots, N\}$ with size αN which contains no square difference fails to be η-uniform for some reasonably large η. In proving this statement we shall use just one fact about the squares, an elementary proof of which may be found in [30].

Lemma 35 *Let $r_6(n)$ denote the number of ordered sextuples $(a_1, \dots, a_6)$ with $a_1^2 + \dots + a_6^2 = n$. Then*

$$n^2 \ \leq \ r_6(n) \ \leq \ 40n^2.$$

In fact we shall have no use for the lower bound in the lemma, but have included it to emphasise that r_6 is comparable to n^2. □

Now let S be the set of non-zero squares less than $N/2$, and let $B = A \cap [0, N/2]$. Regard S, A and B as subsets of $\mathbb{Z}_N$. Assume without loss of generality that $|B| \geq \alpha N/2$. If $A - A$ contains no square then certainly

$$\sum_{x,d} B(x)A(x+d)S(d) \;=\; 0,$$

where the sums are over $\mathbb{Z}_N$. It follows easily from Parseval's Theorem and the triangle inequality that

$$\sum_{r\neq 0} |\hat{S}(r)||\hat{B}(r)||\hat{A}(r)| \;\geq\; |\hat{S}(0)||\hat{B}(0)||\hat{A}(0)| \;\geq\; \tfrac{1}{4}\alpha^2 N^{5/2}. \tag{3.2}$$

The left-hand side here is at most

$$\sup_{r\neq 0} |\hat{A}(r)|^{1/6} \sum_r |\hat{S}(r)||\hat{B}(r)||\hat{A}(r)|^{5/6}$$

which, by Hölder's inequality, is at most

$$\sup_{r\neq 0} |\hat{A}(r)|^{1/6} \left(\sum_r |\hat{S}(r)|^{12}\right)^{1/12} \left(\sum_r |\hat{B}(r)|^2\right)^{1/2} \left(\sum_r |\hat{A}(r)|^2\right)^{5/12}. \tag{3.3}$$

Now (by Parseval again) $\sum_r |\hat{S}(r)|^{12}$ is equal to $N\sum_x R_6(x)^2$, where $R_6(x)$ is the number of solutions to $a_1^2 + \cdots + a_6^2 \equiv x (\mathrm{mod} N)$ with $a_i^2 \in (0, N/2]$. An easy exercise using Lemma 35 shows that $R_6(x) \leq 600N^2$ for all x, which gives the bound

$$\sum_r |\hat{S}(r)|^{12} \;\leq\; 2^{19}N^6. \tag{3.4}$$

Parseval also gives $\sum_r |\hat{A}(r)|^2 \leq N^2$ and $\sum_r |\hat{B}(r)|^2 \leq N^2$. Substituting these and (3.4) into (3.3) gives that there is $r \neq 0$ for which

$$|\hat{A}(r)| \;\geq\; 2^{-30}\alpha^{11/2}|A|. \tag{3.5}$$

In other words, A is non-uniform.

We remark that (3.4) is the fact about squares that we really need, but we feel that Lemma 35 is somehow rather easier to appreciate. In many parts of the literature where these issues are discussed, Lemma 35 is derived from (3.4), which is itself established by the traditional form of the Hardy-Littlewood method. However the proof of Lemma 35 in [30] is completely

elementary, and there is a fantastic (and well known) proof of a similar result using modular forms.

To complete our proof of Theorem 30 by the iteration method we are going to use (3.5) to show that A has increased density on a *square-difference* AP. If we pass to such a subprogression and then rescale this subprogression to have common difference 1, the resulting set A' will still not contain a square difference. It is this consequence of the set of squares being multiplicative that makes this particular instance of Sárközy's Theorem substantially easier than its extension to general polynomials. To find a subprogression of the desired type requires a further standard fact about the squares, which is a quantitative version of the fact that the squares are a Heilbronn Set (see [29]). As noted in [12] it is rather difficult to find a precise statement of this result in the literature. Fortunately however we can use Lemma 5.5 of [12] to simply write down the next lemma.

Lemma 36 *Let $a \in \mathbb{Z}_N$, let $t \leq N$, and suppose that $t \geq 2^{2^{128}}$. Then there is $p \leq t$ such that $|p^2 a| \leq t^{-1/16}N$.*

Now let r be such that $|\hat{A}(r)|$ is large. Put $t = N^{1/4}$ in Lemma 36 and let $p \leq N^{1/4}$ be such that $|p^2 r| \leq N^{127/128}$. Let B be the arithmetic progression $p^2, 2p^2, \ldots, Lp^2$ where $L = \frac{1}{20}N^{1/128}$. One calculates

$$\begin{aligned} |\hat{B}(r)| &\geq L \sup_{x=1,\ldots,L} \left(1 - \left|1 - \omega^{rp^2x}\right|\right) \\ &\geq L\left(1 - \frac{2\pi|rp^2|L}{N}\right) \\ &\geq L/2. \end{aligned}$$

From this and the fact that $|\hat{A}(r)| \geq 2^{-30}\alpha^{11/2}|A|$ we get that

$$\begin{aligned} N\sum_x |A \cap (B+x)|^2 &= \sum_r |\hat{A}(r)|^2|\hat{B}(r)|^2 \\ &\geq \left(1 + 2^{-62}\alpha^{11}\right)|A|^2L^2. \end{aligned} \tag{3.6}$$

Now certain of the translates $B+x$ have the rather unfortunate property of splitting into two smaller progressions when we "unravel" $\mathbb{Z}_N$ to recover $\{1, \ldots, N\}$. We shall call these values of x bad, and denote the set of good values by $\mathcal{G}$. Now $p \leq N^{1/4}$ and so B has diameter less than $N^{2/3}$. It follows that there are at most $N^{2/3}$ bad values of x, and so their total contribution to $N\sum_x |A \cap (B+x)|^2$ does not exceed $L^2N^{5/3}$. Assuming that $\alpha \geq 32N^{-1/39}$ (which it certainly will be) one sees from (3.6) that

$$N\sum_{x\in\mathcal{G}} |A \cap (B+x)|^2 \geq \left(1 + 2^{-63}\alpha^{11}\right)|A|^2L^2.$$

The left hand side here is at most

$$N \sup_{x \in \mathcal{G}} |A \cap (B+x)| \cdot |A| L,$$

from which we deduce that there is $x \in \mathcal{G}$ for which

$$|A \cap (B+x)| \ \geq \ \left(\alpha + 2^{-63}\alpha^{12}\right)|B|.$$

We have deduced, from the assumption that $A - A$ does not contain a square and that $N \geq 2^{2^{130}}$, that A has density at least $\alpha + 2^{-63}\alpha^{12}$ on a subprogression with length at least $\frac{1}{20}N^{1/128}$ and square common difference. Iterating this argument leads to the following result.

Proposition 37 *There is a constant C such that, if A is a subset of $\{1, \ldots, N\}$ with density at least $C(\log\log N)^{-1/11}$, then A contains two elements a, a' with $a - a'$ a non-zero square.*

In order to prove Sárközy's result in the simplest possible manner we have not worried too much about sacrificing the quality of the bound obtained. There is a variation of the above argument in which one shows that A is what I call *arithmetically non-uniform*, which means that some $|\hat{A}(r)|$ is large with r approximately equal to a rational with small denominator. This allows one to perform the iteration more efficiently, and doing this gives a bound of form $(\log N)^{-c}$ in Proposition 37. This argument will be discussed in §4.3. The current best known bound of $(\log N)^{-c \log\log\log\log N}$ for this problem is due to Pintz, Steiger and Szemerédi [31] and makes use of some rather involved Fourier arguments. There is still a massive gap in our knowledge, and I cannot resist closing this section with the following open problem.

Problem 38 *Let $\epsilon > 0$ and let $N > N_0(\epsilon)$ be sufficiently large. Does there exist a set $A \subseteq \{1, \ldots, N\}$ with $|A| \geq N^{1-\epsilon}$, such that A does not contain two elements that differ by a square?*

The best that is known is that one can take $\epsilon = 0.267$, a result due to Ruzsa [34].

3.3 APs with Common Difference $x^2 + y^2$

In this section we turn our attentions to the main business of this chapter, a proof of Theorem 33.

Theorem 33 *There is a constant c such that any subset of $\{1, \ldots, N\}$ of density at least $(\log\log N)^{-c}$ contains a 3-term arithmetic progression whose common difference is of the form $x^2 + y^2$.*

Our proof of Theorem 33 will be by the iteration method, and as such will fall into two parts. In the first part, containing most of our original work on the problem, we show that a quadratically uniform set contains roughly the expected number of progressions $(a, a +$

$d, a+2d)$ with $d = x^2 + y^2$. The second part of the proof follows [12] extremely closely. Our objective there is to show that a set which fails to be quadratically uniform has increased density on a long subprogression with square common difference. Unfortunately it seems that in order to get this all-important modification to Gowers' work we have to analyse his argument in slightly more detail than would perhaps have been wished.

We shall start our treatment in quite a general setting. Let D be a subset of $\mathbb{N}$, and for $N \in \mathbb{N}$ regard $D_N = D \cap \{1, \dots, N\}$ as a subset of $\mathbb{Z}_N$ in the natural way. Suppose that for some $k \in \mathbb{N}$ we have

$$\sum_r \left|\hat{D}_N(r)\right|^{2k} \ \leq \ C|D_N|^{2k}$$

and

$$|D_{2N}| \ \leq \ C|D_N|,$$

where C is independent of N. Then we shall say that D is *uniformly k-dense*. The obvious non-trivial example in view of our earlier discussions is the set S of squares, which is uniformly 6-dense by (3.4). Our nomenclature is non-standard, but quite convenient.

If $A \subseteq \{1, \dots, N\}$ is a set of density α then we write $f = A - \alpha$ for its balanced function. Write $B = A \cap \{1, \dots, N/3\}$, let β be the density of B and let $g = B - \beta$ be its balanced function. Finally if D is any set then we say that an arithmetic progression $(x, x+d, x+2d)$ with $d \in D$ is a *D-progression*.

Proposition 39 *Suppose that A does not contain a D-progression, where D is uniformly k-dense. Then either*

$$\sup_{r \neq 0} |\hat{A}(r)| \ \gg \ \alpha^{2k} N \tag{3.7}$$

or else we have

$$\sum_x \sum_{d \in U} g(x) f(x+d) f(x+2d) \ \gg \ \alpha^3 N |U|. \tag{3.8}$$

where the sum is over $\mathbb{Z}_N$ and $U = D \cap \{1, \dots, N/3\}$.

Proof To begin with we show that the conclusion holds with room to spare if β is much smaller than expected. Suppose that $\beta \leq \alpha/12$, and let I be the characteristic function of $\{1, \dots, N/6\} \subseteq \mathbb{Z}_N$. Then

$$\frac{6}{N} \sum_x A(x)(I * I)(x + N/6) \ \leq \ \beta N \ \leq \ \alpha N/12.$$

Taking Fourier transforms gives

$$\sum_r \omega^{-rN/6} \hat{A}(r) |\hat{I}(r)|^2 \ \leq \ \alpha N^3/72,$$

which implies by the triangle inequality that

$$\sum_{r\neq 0}|\hat{A}(r)||\hat{I}(r)|^2 \ \geq \ \alpha N^3/72.$$

However Parseval's identity gives $\sum|\hat{I}(r)|^2 = N^2/6$, from which it follows immediately that

$$\sup_{r\neq 0}|\hat{A}(r)| \ \geq \ \alpha N/12.$$

Assume now that $\beta \geq \alpha/12$. It is easy to see that there is no *modular* progression of form $(x, x+d, x+2d)$, $d \in U$, in $B \times A \times A$, and so we have

$$\sum_{x\in\mathbb{Z}_N}\sum_{d\in U}B(x)A(x+d)A(x+2d) \ = \ 0.$$

Writing $A = f + \alpha$ and $B = g + \beta$ we may expand this as a sum of eight terms. Of these we have a term $T = \sum_x\sum_{d\in U}g(x)f(x+d)f(x+2d)$, a term $\alpha^2\beta N|U|$ and three terms which are identically zero. If $T \geq \alpha^2\beta N|U|/4$ then we are done. Failing this the triangle inequality implies that one of the other three terms must be at least $\alpha^2\beta N|U|/4$, which in turn is not less than $\alpha^3 N|U|/48$. These other three terms are very similar to one another, and brushing a little work under the carpet we suppose without loss of generality that

$$\alpha\sum_x\sum_{d\in U}g(x)f(x+d) \ \geq \ \alpha^3 N|U|/48. \tag{3.9}$$

This may be written in terms of Fourier coefficients as $\alpha N^{-1}\sum_r\hat{g}(r)\hat{f}(-r)\hat{U}(r)$. This sum may be estimated using Hölder's Inequality exactly as in §3.2. It is at most

$$\alpha N^{-1}\sup_r\left|\hat{f}(r)\right|^{1/k}\left(\sum_r\left|\hat{U}(r)\right|^{2k}\right)^{1/2k}\left(\sum_r|\hat{g}(r)|^2\right)^{1/2}\left(\sum_r\left|\hat{f}(r)\right|^2\right)^{(k-1)/2k} \tag{3.10}$$

To deal with this observe that since D is uniformly k-dense we have

$$\begin{aligned}
\sum_r|\hat{U}(r)|^{2k} &= N\sum_{\substack{a_1+\cdots+a_k\\=b_1+\cdots+b_k}}U(a_1)\dots U(a_k)U(b_1)\dots U(b_k)\\
&\leq N\sum_{\substack{a_1+\cdots+a_k\\=b_1+\cdots+b_k}}D_N(a_1)\dots D_N(a_k)D_N(b_1)\dots D_N(b_k)\\
&= \sum_r|\hat{D}_N(r)|^{2k}\\
&\ll |D_N|^{2k}\\
&\ll |U|^{2k}.
\end{aligned}$$

Using Parseval's identity on the other two factors in (3.10) we can bound that expression above by a constant multiple of

$$\sup_r |\hat{f}(r)|^{1/k} \cdot N^{1-1/k}|U|.$$

It follows from (3.9) that $\sup_r |\hat{f}(r)| \gg \alpha^{2k}N$, and we are done. □

Suppose now that the hypotheses of Proposition 39 are satisfied, so that we have a uniformly k-dense set D and a set $A \subseteq \{1, \dots, N\}$ with size αN containing no D-progression. We know that either (3.7) or (3.8) must hold. If (3.7) holds then we have the possibility of passing to various types of subprogression and performing an iteration argument. But how are we to derive information from (3.8)? A very natural approach is to use the Cauchy-Schwarz inequality, together with the fact that $\|g\|_\infty \leq 1$, to get that

$$\sum_x \left| \sum_d f(x+d)f(x+2d)U(d) \right|^2 \geq 2^{-12}\alpha^6 N|U|^2.$$

Multiplying out, rearranging and changing the summation variables, this implies that

$$\sum_h \sum_x \sum_d \Delta(f;h)(x)\Delta(f;2h)(x+d)\Delta(U;h)(d) \geq 2^{-12}\alpha^6 N|U|^2. \tag{3.11}$$

Our hope might now be to use this to show that, for many h, $\Delta(f;h)$ has a large Fourier coefficient. If, for many h, $\Delta(U,h)$ looked like a uniformly k-dense set then we could perhaps apply Hölder's Inequality much as in previous sections. It is easy to see that there is no chance of this happening for any D with $|D_N| \ll N^{1/2}$, because the sets $\Delta(U;h)$ are on average very thin. It is this fact which precludes such an approach from giving information about any case of the polynomial Szemerédi Theorem. There is nothing to preclude such an approach from working if $|D_N|$ is at least $N^{1/2+c}$ however, in which case the average size of $\Delta(U;h)$ is at least about N^{2c}.

Turning our thoughts to Theorem 33 let D be the set of all sums of two squares, and let $A \subseteq \{1, \dots, N\}$ be a set of size αN containing no D-progressions. It is unfortunate, as regards what we have just said, that our knowledge of $\Delta(U;h)$ is rather restricted. In fact it is not even a trivial exercise to estimate $|D_N|$ (though it has been known since Landau that $|D_N| \sim cN(\log N)^{-1/2}$), or to prove that D is uniformly k-dense for some k. These last two difficulties can be eased by using, instead of D, the function $r(n)$ defined to be the number of representations of n as a sum of two squares. The problems with $\Delta(U;h)$ remain, however, and it seems to be necessary to use advanced machinery to estimate objects like $\sum_{n \leq N} r(n)r(n+h)$ (cf. [25], [4]).

There is however a much better way out. We begin with the trivial observation that if A does

not contain any D-progressions then it does not contain any D'-progressions for $D' \subseteq D$. Secondly we recall the very classical fact that E, the set of all primes of form $4k+1$, is contained in D. We shall prove below, as our first application of the Selberg Sieve, that E is uniformly 2-dense. Given this fact it follows from Proposition 39 and the analysis leading up to (3.11) that either

$$\sup_{r \neq 0} |\hat{A}(r)| \gg \alpha^4 N \tag{3.12}$$

or

$$\sum_h \sum_x \sum_d \Delta(f;h)(x+d)\Delta(f;2h)(x+2d)\Delta(V;h)(d) \gg \frac{\alpha^6 N^3}{(\log N)^2} \tag{3.13}$$

where $V = E \cap \{1, \dots, N/3\}$. Here, of course, we have observed that $|V| \gg N/\log N$. Now one might think that if $\Delta(U;h)$ was a difficult object to deal with then $\Delta(V;h)$ can only be worse. There can be no hope of proving that, for example, $\Delta(V;4)$ is a uniformly k-dense for any k because the question of whether this set is even infinite is famously unsolved.

It transpires, however, that one can control the *average* behaviour of $\Delta(V;h)$ sufficiently accurately for our purposes by using Sieve Theory. In order to do this it seems to be essential to pass to the set E, whose multiplicative structure is so well understood.

Let us give a brief resumé of the results from Sieve Theory that we need. The standard reference for this subject is [23] but in our unbiased opinion the best way to understand the necessary background is to read [15], which has been specially updated for this purpose. We will only be concerned with sieving polynomial sequences, which is the simplest situation covered by the Selberg Sieve. Let h be a polynomial with integer coefficients, and let $\mathcal{A}$ denote the sequence $\{h(1), \dots, h(N)\}$. Let $\mathcal{P}$ denote the set of primes. Then the Selberg Sieve gives upper bounds for $S(\mathcal{A}, \mathcal{P}, z)$, which is defined to be the number of $x \in \mathcal{A}$ which are not divisible by any prime $p \leq z$. This upper bound is given in terms of a function ω defined at primes p to be the number of elements in $\{h(1), \dots, h(p)\}$ which are divisible by p. To put it another way, the proportion of elements of $\mathcal{A}$ which are divisible by p is roughly $\omega(p)/p$. The key result we shall require is the following, which is Theorem 11 in [15] (we should also note that it can be read out of [23]).

Theorem 40 *Suppose that $\omega(p) \leq C$ for all primes p. Then*

$$S(\mathcal{A}, \mathcal{P}, N^{1/16C}) \ll N \prod_{p \leq N^{1/16C}} \left(1 - \frac{\omega(p)}{p}\right),$$

where the implied constant depends only on C.

The first corollary we shall require is the following.

Proposition 41 *Let $n \in \mathbb{N}$. Then the number of representations of n as the sum of two elements of E, $r_2(E,n)$ satisfies*

$$r_2(E,n) \ll \frac{n}{(\log n)^2} \prod_{p|n} \left(1 + \frac{1}{p}\right).$$

Proof Clearly $r_2(E,n) \leq r_2(\mathcal{P},n)$. Consider the polynomial $h(x) = x(n-x)$ and let $\mathcal{A} = \{h(1), \ldots, h(n)\}$. We wish to count the number of $x \leq n$ for which x and $n-x$ are both prime. For such x we either have $x \leq n^{1/2}$, $x \geq n - n^{1/2}$ or else $h(x)$ has no prime factor less than $n^{1/2}$. It follows that $r_2(\mathcal{P},n)$ is bounded above by $2n^{1/2} + S(\mathcal{A}, \mathcal{P}, n^{1/2})$.

It is easy to see that, in the notation of our potted introduction to sieve theory, one has $\omega(p) = 2$ for all p except when $p|n$, in which case $\omega(p) = 1$. Thus by Theorem 40 one has

$$r_2(E,n) \ll n^{1/2} + n \prod_{p \leq n^{1/32}} \left(1 - \frac{2}{p}\right) \prod_{\substack{p \leq n^{1/32} \\ p|n}} \left(1 - \frac{2}{p}\right)^{-1} \left(1 - \frac{1}{p}\right). \tag{3.14}$$

To deal with this we require two further facts. The first is standard, and the second is very easy to prove.

Fact 1 (Mertens) *There are absolute constants C_1, C_2 such that*

$$C_1 \log m \leq \prod_{p \leq m} \left(1 - \frac{1}{p}\right)^{-1} \leq C_2 \log m.$$

Fact 2 *The infinite product*

$$\prod_p \left(1 - \frac{\lambda}{p^2}\right)$$

converges for any λ.

Armed with these two facts one sees from (3.14) that

$$r_2(E,n) \ll \frac{n}{(\log n)^2} \prod_{p|n} \left(1 + \frac{1}{p}\right)$$

as claimed. □

This is of course a very standard deduction from the Selberg Sieve, but we have included it to ensure that the reader is happy with our notation. It turns out to be extremely convenient to write $\xi(n)$ for the quantity $\prod_{p|n}\left(1 + \frac{1}{p}\right)$ appearing here. In a short while we will use Proposition 41 to show that E is uniformly 2-dense. Before doing this however it is necessary to give a crude estimate for the moments $\sum_u \xi(u)^s$ of ξ.

Lemma 42 *Let $s \geq 1$ be real. Then*

$$\sum_u \xi(u)^s \ \leq \ 2^{2s2^{2s}} N.$$

Proof First observe that for any $x \leq 1$ one has

$$(1+x)^s \ \leq \ 1 + 2^s x.$$

Secondly, if $p_1, \ldots, p_r$ are distinct primes then for any C one has

$$\begin{aligned}
\prod_{i=1}^r \frac{C}{p_i} &= \prod_{p_i \leq C^2} \frac{C}{p_i} \prod_{p_i > C^2} \frac{C}{p_i} \\
&\leq C^{C^2} \prod_{p_i > C^2} \frac{1}{\sqrt{p_i}} \\
&\leq C^{C^2} \prod_{i=1}^r \frac{1}{\sqrt{p_i}}.
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_u \xi(u)^s &= \sum_u \prod_{p|u} \left(1 + \frac{1}{p}\right)^s \\
&\leq \sum_u \prod_{p|u} \left(1 + \frac{2^s}{p}\right) \\
&\leq 2^{s2^{2s}} \sum_u \sum_{d|u} \frac{1}{\sqrt{d}} \\
&\leq 2^{s2^{2s}} N \sum_{d \leq N} d^{-3/2}.
\end{aligned}$$

This implies the result. □

Proposition 43 *The set E of primes of form $4k+1$ is uniformly 2-dense.*

Proof Regard $E_N = E \cap \{1, \ldots, N\}$ as a subset of $\mathbb{Z}_N$, as we must do to even make sense of what it means to be uniformly 2-dense. It is easy to see that

$$\sum_r |\hat{E}_N(r)|^4 \ = \ N \sum_{a+b=c+d} E(a)E(b)E(c)E(d),$$

where the equation $a + b = c + d$ is taken in $\mathbb{Z}_N$. This is clearly at most

$$N \sum_{n \leq N} \left(r_2(E,n) + r_2(E, N+n)\right)^2,$$

which by Proposition 41 is at most a constant times

$$\frac{N^3}{(\log N)^4}\sum_{n\leq N}\left(\xi(n)+\xi(N+n)\right)^2.$$

The sum, by the Cauchy-Schwarz inequality, is at most $4\sum_{n\leq 2N}\xi(n)^2$, a quantity which we know to be $O(N)$ by Lemma 42. Thus

$$\sum_r |\hat{E}_N(r)|^4 \ll \left(\frac{N}{\log N}\right)^4.$$

Since $|E_N|\sim N/2\log N$ it follows that E is indeed uniformly 2-dense. □

As we remarked earlier it now follows from Proposition 39 that a set A containing no E-progression satisfies either (3.12) or (3.13). If (3.12) holds then progress is relatively easy, so for the time being we suppose that (3.13) holds. Our aim is to show that the sets $\Delta(V;h)$ (where $V = E\cap\{1,\dots,N/3\}$) behave like uniformly 2-dense sets in some average sense which there is no point in defining at present. This will then be enough to deduce (using an application of Hölder's Inequality similar to that in §3.2) that $\Delta(f;h)$ fails to be uniform for rather a lot of h, and in fact that f fails to be quadratically uniform. The key to all this will be the following slightly less standard deduction from the Selberg Sieve.

Proposition 44 *Let $r_h(n)$ denote the number of ways of expressing n as a difference of 2 elements of $\Delta(V;h)$. Then*

$$r_h(n) \ll \frac{N}{(\log N)^4}\frac{\xi(h)^2\xi(n)^2\xi(n+h)\xi(n-h)}{\xi((n,h))^3}.$$

Proof We bring the Selberg Sieve to bear on this by observing that $r_h(n)$ is at most the number of $x\leq N$ for which x, $x-n$, $x-h$ and $x-n-h$ are all prime. With the exception of at most $8N^{1/2}$ of these x the polynomial

$$h(x) = x(x-n)(x-h)(x-n-h)$$

has no prime factor less than $N^{1/2}$. Writing $\mathcal{A}=\{h(1),\dots,h(N)\}$ this implies that

$$r_h(n) \ll 8N^{1/2}+S(\mathcal{A},\mathcal{P},N^{1/2}).$$

We would clearly like to apply Theorem 40, but first we must think about $\omega(p)$. Suppose that $p\geq 3$. Then it is reasonably easy to see that $\omega(p)=4$ unless one of the following four possibilities occurs: (i) $p|n$; (ii) $p|h$; (iii) $p|(n+h)$ and (iv) $p|(n-h)$. Furthermore these possibilities are mutually exclusive unless p divides both n and h, in which case they all occur. If they do all occur then $\omega(p)=1$. If (i) or (ii) occurs then $\omega(p)=2$, and if (iii) or

(iv) occurs then $\omega(p) = 3$. If $p = 2$ the behaviour is more subtle, but we will not concern ourselves with this as each individual prime only contributes a bounded multiplicative factor to the sieve estimate of Theorem 40.

Theorem 40 (with $C = 4$) certainly applies to this situation, then, and with a modicum of effort one can verify that the key quantity

$$N \prod_{p \leq N^{1/64}} \left(1 - \frac{\omega(p)}{p}\right)$$

is equal, up to a product of terms of the form $\prod_p (1 - \lambda p^{-2})$, to the delightful expression

$$\prod_{p \leq N^{1/64}} \left(1 - \frac{4}{p}\right) \prod_{p|n} \left(1 - \frac{2}{p}\right)^{-1} \prod_{p|h} \left(1 - \frac{2}{p}\right)^{-1} \\ \times \prod_{p|(n+h)} \left(1 - \frac{1}{p}\right)^{-1} \prod_{p|(n-h)} \left(1 - \frac{1}{p}\right)^{-1} \prod_{p|(n,h)} \left(1 - \frac{1}{p}\right)^{3}$$

where the final five products are also constrained to be over $p \leq N^{1/64}$. Since an integer less than N cannot have more than 64 prime factors $p \geq N^{1/64}$, this restriction can be removed at the expense of introducing another bounded multiplicative constant. By Fact 1 and Fact 2 (see earlier in the paper) this expression is equal to the one appearing in the statement of the proposition, again up to a universal, bounded, multiplicative constant. □

Proposition 45 *For any h we have*

$$\sum_n r_h(n)^2 \ll \frac{N^3}{(\log N)^8} \xi(h)^4,$$

where the implied constant is independent of h.

Proof By Proposition 44 we have

$$\sum_n r_h(n)^2 \ll \frac{N^2}{(\log N)^8} \xi(h)^4 \sum_n \xi(n)^4 \xi(n+h)^2 \xi(n-h)^2.$$

By Hölder's Inequality the sum over n is at most $\sum_n \xi(n)^8$. We are therefore done by Lemma 42. □

The next result clarifies the sense in which $\Delta(V; h)$ is, on average, uniformly 2-dense. In the following proposition $\Delta(V; h)$ is regarded as a subset of $\mathbb{Z}_N$.

Proposition 46

$$\sum_r |\Delta(V;h)\hat{}(r)|^4 \ll \frac{N^4}{(\log N)^8}\xi(h)^4. \tag{3.15}$$

Proof Immediate from Proposition 45 and the fact that $\Delta(V,h)$ is supported in an interval of size $N/3$ (so that there is no problem with modular addition not being the same as "ordinary" addition). □

Now we come to use (3.13). We shall use it to treat each h separately, which we do by noting that from (3.13) follows

$$\sum_x \sum_d \Delta(f;h)(x)\Delta(f;2h)(x+d)\Delta(V;h)(d) \gg \gamma(h)\frac{\alpha^6 N^2}{(\log N)^2},$$

where $\sum_h \gamma(h) = N$. Taking Fourier Coefficients, this implies that

$$\sum_r \Delta(f;h)\hat{}(r)\Delta(f;2h)\hat{}(-r)\Delta(V;h)\hat{}(r) \gg \gamma(h)\frac{\alpha^6 N^3}{(\log N)^2}. \tag{3.16}$$

Applying Hölder gives

$$\sup_r |\Delta(f;h)\hat{}(r)|^{1/2}\left(\sum_r |\Delta(f;h)\hat{}(r)|^2\right)^{1/4}\left(\sum_r |\Delta(f;2h)\hat{}(r)|^2\right)^{1/2}\left(\sum_r |\Delta(E;h)\hat{}(r)|^4\right)^{1/4}$$
$$\gg \gamma(h)\frac{\alpha^6 N^3}{(\log N)^2}.$$

The first two bracketed expressions may be bounded above using Parseval, and the third is subject to the upper bound (3.15). One gets

$$\sup_r |\Delta(f;h)\hat{}(r)| \gg \frac{\gamma(h)^2\alpha^{12}N}{\xi(h)^2}. \tag{3.17}$$

Thus there is a function $\phi : \mathbb{Z}_N \to \mathbb{Z}_N$ such that

$$\sum_h |\Delta(f;h)\hat{}(\phi(h))|^2 \gg \alpha^{24}N^2\sum_h \frac{\gamma(h)^4}{\xi(h)^4}.$$

This would imply that A fails to be quadratically uniform if we could show that the sum here is $\gg N$. To do this, we recall that $\sum_h \gamma(h) = N$ and use Hölder, getting

$$\sum_h \frac{\gamma(h)^4}{\xi(h)^4} \gg N^4\left(\sum_h \xi(h)^{4/3}\right)^{-3}.$$

This is indeed $\gg N$ by Lemma 42.

We have now shown that if (3.13) holds then A is not quadratically $C\alpha^{24}$-uniform for some C. We also know that if A does not contain any E-progressions then either (3.12) or (3.13) must hold. However (3.12) is just the statement that A is not $C\alpha^4$-uniform for some C. Thus by Proposition 34 we may incorporate everything we have done so far into the following.

Proposition 47 *Suppose that $A \subseteq \{1, \dots, N\}$ has density α yet does not contain a progression $(x, x+d, x+2d)$ with d a sum of two squares. Then A is not quadratically $C\alpha^{24}$-uniform for some C.*

To spell it out, the conclusion of this proposition implies that

$$\sum_u |\Delta(f; u)(\phi(u))|^2 \gg \alpha^{24} N^3$$

for some function $\phi : \mathbb{Z}_N \to \mathbb{Z}_N$.

3.4 Increasing the Density on a Special Subprogression

Sets that fail to be quadratically uniform have an interesting structure, as was established by Gowers [12] in the course of proving Szemerédi's Theorem for progressions of length 4 (a slightly weaker result was established in [11]). In that paper a result along the following lines is proved.

Theorem 48 (Gowers' Inverse Theorem) *Suppose that $A \subseteq \{1, \dots, N\}$ has density α and that A fails to be quadratically $C_1\alpha^{m_1}$-uniform. Then there is an arithmetic progression L of length $|L| \gg N^{C_2\alpha^{m_2}}$ on which A has density at least $\alpha + C_3\alpha^{m_3}$, where C_2, C_3, m_2 and m_3 depend only on m_1 and C_1.*

Unfortunately this result does not quite suffice for the purposes of proving Theorem 33. The reason for this is that when one restricts to a subprogression and rescales, the property of not containing any triples $(x, x+d, x+2d)$ with d a sum of two squares is not preserved. A result that certainly would suffice in this context would be a version of Theorem 48 in which L has square common difference, and we now turn our attentions to proving such a result.

It would be nice if one could deduce such a result without going into the detailed workings of Gowers' proof, but alas we have found it necessary to go a fair way in. We think that the best way to explain things is to present our argument together with a brief overview of parts of [12]. We have tried to do this in a fairly minimal way, assuming that the reader has access to Chapters 6,7 and 8 of [12]. We shall require one additional ingredient, which is a version of the following simultaneous approximation result of Schmidt [40].

Theorem 49 (Schmidt) *Let $r_1, \dots, r_h \in \mathbb{Z}_N$ and let $t \leq N$. Suppose that $t \geq N_0(h)$ is sufficiently large. Then there is $p \leq t$ such that $|p^2 r_i| \leq t^{-1/3h^2} N$ for all i, $1 \leq i \leq h$.*

Unfortunately there does not seem to be any place in the literature where an explicit value of $N_0(h)$ is derived. It would be possible to work through the proof in [40] and derive such an explicit value (which would probably not be too large) but we will prove our own, substantially weaker, version of Theorem 49 with explicit constants.

Proposition 50 *Let $r_1, \dots, r_h \in \mathbb{Z}_N$ and let $t \leq N$. Suppose that $t \geq 2^{2^{7h+129}}$. Then there is $p \leq t$ such that $|p^2 r_i| \leq t^{-2^{-(7h+6)}} N$ for all i, $1 \leq i \leq h$.*

Proof We make extensive use of Lemma 36. Choose $p_1, p_2, \dots$ inductively so that

$$p_i \ \leq \ t^{2^{-(7i+1)}} \tag{3.18}$$

and

$$\left|p_i^2 p_{i-1}^2 \dots p_1^2 r_i\right| \ \leq \ t^{-2^{-(7i+5)}} N \tag{3.19}$$

for each $i = 1, \dots, h$. Let $p = p_1 p_2 \dots p_h$. It is easy to check that $p \leq t$. Furthermore

$$\begin{aligned} |p^2 r_i| \ &\leq \ |p_h|^2 |p_{h-1}^2| \dots |p_{i+1}^2| \left|p_i^2 \dots p_1^2 r_i\right| \\ &\leq \ t^{2^{-(7i+6)}} \cdot t^{-2^{-(7i+5)}} N \\ &\leq \ t^{-2^{-(7h+6)}} N \end{aligned}$$

as required. The repeated applications of Lemma 36 require certain things to be sufficiently large. The most difficult condition to satisfy is that arising from the last step in our inductive construction, where we require

$$t^{2^{-(7h+1)}} \ \geq \ 2^{2^{128}}.$$

This is where the restriction in the proposition comes from. □

Let us assume then that $f : \mathbb{Z}_N \to \mathbb{C}$ has $\|f\|_\infty \leq 1$ and that

$$\sum_k |\Delta(f; k)\hat{}\,(\phi(k))|^2 \ \geq \ 2\beta N^3 \tag{3.20}$$

for some function $\phi : \mathbb{Z}_N \to \mathbb{Z}_N$. We start with a trivial deduction from this, which is proved by a simple averaging argument: there is a set $B \subseteq \mathbb{Z}_N$ with $|B| \geq \beta N$ and

$$|\Delta(f; r)(\phi(r))| \ \geq \ \beta^{1/2} N \tag{3.21}$$

for all $k \in B$. We have then

$$\sum_{k \in B} |\Delta(f; k)\hat{}\,(\phi(k))|^2 \ \geq \ \beta^2 N^3. \tag{3.22}$$

The following is essentially Proposition 6.1 of [12].

Proposition 51 *Suppose that* (3.22) *holds. Then there are at least* $\beta^8 N^3$ *quadruples* (a, b, c, d) *in* B^4 *such that* $a + b = c + d$ *and* $\phi(a) + \phi(b) = \phi(c) + \phi(d)$.

The aim now is to use this result to show that ϕ has a certain weak multilinearity property. In the shorter paper of Gowers [11] this is done by an appeal to a theorem of Freiman, but in [12] it is shown that this is not quite necessary, and that a better bound results by directly adapting some ideas of Ruzsa that were used in a proof of Freiman's Theorem. Embedded in this circle of ideas is an important technique of Bogolubov. The following result is immediate from Corollary 7.6 in [12], which is deduced from Proposition 51 with some effort. Before stating it we need a definition. If $S \subseteq \mathbb{Z}_N$ and $\psi : S \to \mathbb{Z}_N$ is a function then we say that ψ is a *Freiman* k*-homomorphism* if

$$s_1 + \cdots + s_k \;=\; s_{k+1} + \cdots + s_{2k}$$

implies that

$$\psi(s_1) + \cdots + \psi(s_k) \;=\; \psi(s_{k+1}) + \cdots + \psi(s_{2k})$$

for all $s_1, \ldots, s_{2k} \in S$.

Proposition 52 *There is a set* $B' \subset B$ *with* $|B'| \geq 2^{-1849}\beta^{5821}N$ *such that the restriction of* ϕ *to* B' *is a Freiman homomorphism of order 8.*

Let us now quote Corollary 7.9 of [12]. If $K \subseteq \mathbb{Z}_N$ and $\eta \in (0, 1)$ write $B(K, \eta)$ for the set of all $n \in \mathbb{Z}_N$ such that $|nk| \leq \eta N$ for all $k \in K$. We call such a set a *Bohr Neighbourhood*.

Proposition 53 *Let* $D \subseteq \mathbb{Z}_N$ *have size* δN*, and suppose that* $\phi : D \to \mathbb{Z}_N$ *is a Freiman homomorphism of order 8. Then there is a set* K *with* $|K| \leq 16\delta^{-2}$ *such that the following is true. If* m *is any positive integer and* $d \in B(K, \delta/100m)$ *then there is* c *such that* $\phi(x) - \phi(y) = c(x - y)$ *whenever* $x - y$ *belongs to the set* $\{jd : -m \leq j \leq m\}$.

Write P_0 for the arithmetic progression $\{d, 2d, \ldots, md\}$. In the last proposition let $D = B'$, $\delta = 2^{-1849}\beta^{5821}$ and suppose that $d \in B(K, \delta/100m)$ for the appropriate set K (which will depend on B' and ϕ). Choose a translate $P = P_0 + z$ for which $|P \cap B'| \geq \delta m$, and let $H = P \cap B'$. If x and y lie in H then $x - y$ is in $\{jd : -m \leq j \leq m\}$ and so $\phi|_H$ is the restriction of a linear function from $\mathbb{Z}_N$ to itself, by Proposition 53.

The reader may, at this point, wonder whether we are attempting an outright plagiarism of [12]. It is high time, therefore, that we interjected with a result of our own. The reader will have to make do with Proposition 50.

We will soon find ourselves dealing with some reasonably large numbers, and it is convenient to have a shorthand notation for them. If $n \gg 1$ is a parameter then we write $C_0(n)$ for any polynomial in n. $C_1(n)$ will mean a function of type $2^{p(n)}$, where p is a polynomial, and finally $C_2(n)$ will be a function of type $2^{2^{p(n)}}$. We will feel at liberty to use these symbols

several times, sometimes in the same formula, to denote different functions.

Proposition 50 allows us, provided $|K|$ and $\delta/100m$ satisfy a certain inequality, to conclude that d may be taken to be a small square number. Take $t = N^{1/4}$ in Proposition 50 and recall that $|K| \leq 16\delta^{-2}$. Then there is $p \leq N^{1/4}$ such that $p^2 \in B(K, \delta/100m)$ provided that

$$\frac{\delta}{100m} \;\geq\; N^{-1/C_1(\delta^{-1})}$$

and that $N \geq C_2(\delta^{-1})$. For N greater than some $C_2(\delta^{-1})$ we can pick $m = N^{1/C_1(\delta^{-1})}$ so that this is satisfied by a colossal margin.

Putting everything together, and recalling that δ is $2^{-1849}\beta^{5821}$, we have the following.

Proposition 54 *Suppose that* (3.22) *holds and that* $N \geq C_2(\beta^{-1})$. *Then there is a progression* $P \subseteq \mathbb{Z}_N$ *with common difference* p^2, *where* $p \leq N^{1/4}$, *and length at least* $N^{1/C_1(\beta^{-1})}$, *with the following property. There is* $H \subseteq P \cap B$ *with* $|H| \geq 2^{-1849}\beta^{5821}|P|$ *and* $\lambda, \mu \in \mathbb{Z}_N$ *such that* $\phi(s) = \lambda s + \mu$ *for all* $s \in H$.

Time now to put the "most citations" record beyond all reasonable doubt with a deduction from Proposition 8.1 of [12]. This is easily derived using the previous proposition and (3.21).

Proposition 55 *Suppose that* (3.20) *holds and that* $N \geq C_2(\beta^{-1})$. *Then there is an arithmetic progression* P *with common difference* p^2, *where* $p \leq N^{1/4}$, *and length at least* $N^{1/C_1(\beta^{-1})}$ *and quadratic polynomials* $\psi_0, \psi_1, \dots, \psi_{N-1}$ *such that*

$$\sum_s \left| \sum_{z \in P+s} f(z)\omega^{-\psi_s(z)} \right| \;\geq\; 2^{-1850}\beta^{5822}N|P|.$$

Before stating and proving the next Lemma we need a version of Lemma 36 for fourth powers. Once again we can simply read it from [12].

Lemma 56 *Let* $a \in \mathbb{Z}_N$, *let* $t \leq N$ *and suppose that* $t \geq 2^{2^{512}}$. *Then there is* $p \leq t$ *such that* $|p^4 a| \leq t^{-1/128}N$.

Proposition 57 *Let* $\psi(x) = ax^2 + bx + c$ *be a quadratic polynomial with coefficients in* $\mathbb{Z}_N$. *Let* $L \leq N$, *and let* $P \subseteq \mathbb{Z}_N$ *be an arithmetic progression of length* L *with square common difference. Let* $W \leq L^{2^{-21}}$. *Then there is a partition of* P *into subprogressions* $R_1, \dots, R_m$, *with* R_i *having length between* W *and* $2W$ *and square common difference, such that* $\mathrm{Diam}(\psi(R_i)) \leq L^{-2^{-20}}N$ *for all* i.

Proof We may rescale and assume that $P = \{1, \dots, L\}$ with no loss of generality. For any x_0, λ, d we have

$$\psi(x_0 + \lambda d^2) - \psi(x_0) \;=\; a\lambda^2 d^4 + (2ax_0 + b)\lambda d^2.$$

By Proposition 56 we may choose $d \leq L^{1/8}$ such that $|ad^4| \leq L^{-2^{-10}}N$. Partition $\{1, \ldots, L\}$ into progressions $P_1, \ldots, P_l$ with common difference d^2 and lengths lying between $L^{2^{-12}}$ and $L^{2^{-11}}$. The diameter of ψ on P_i then satisfies

$$\mathrm{Diam}(\psi(P_i)) \;\leq\; L^{-2^{-11}}N + \mathrm{Diam}(\theta_i(P_i)), \tag{3.23}$$

where θ_i is a linear polynomial depending on i. Fix i, and for ease of notation rescale P_i to $\{1, \ldots, K\}$ (by the square scaling factor d^{-2}) where $K \geq L^{2^{-12}}$. Suppose that $\theta_i(x) = rx + s$ under this rescaling. Clearly

$$\theta_i(x_0 + \mu e^2) - \theta_i(x_0) \;=\; r\mu e^2.$$

By Lemma 36 we may choose $e \leq K^{1/4}$ so that $|re^2| \leq K^{-1/64}N$. Divide P_i into further subprogressions E_j with common difference e^2 and lengths lying between $K^{1/256}$ and $K^{1/128}$. On these subprogressions we will have $\mathrm{Diam}(\theta_i(E_j)) \leq K^{-1/128}N$. Do this for each i, rescale the resultant progressions by the factor d^2, and then perform a further subdivision to satisfy the technical condition on the lengths of the R_i in the statement. Recalling (3.23) we get the result. □

Call an arithmetic progression $Q \subseteq \mathbb{Z}_N$ *nice* if it does not wrap in $\mathbb{Z}_N$, by which we mean that the length $L(Q)$ and the common difference $d(Q)$ satisfy $L(Q)d(Q) \leq N$. Observe that the progression P found in Proposition 54 is nice, because we constructed it to have small common difference. We observe that any translate of a nice progression is nice, as is any subprogression.

Let us now combine Propositions 55 and 57 in the obvious way.

Proposition 58 *Suppose that* (3.20) *holds and that* $N \geq C_2(\beta^{-1})$. *Then there is* $W = N^{1/C_1(\beta^{-1})}$, *nice arithmetic progressions* $R_{i,s}$, $1 \leq i \leq m, 1 \leq s \leq N$, *each having length between* W *and* $2W$ *and square common difference, and quadratic polynomials* $\psi_0, \ldots, \psi_{N-1}$ *such that*

$$\sum_i \sum_s \left| \sum_{z \in R_{i,s}} f(z)\omega^{\psi_s(z)} \right| \;\geq\; 2^{-1850}\beta^{5822} \sum_{i,s} |R_{i,s}|. \tag{3.24}$$

Furthermore every point of $\mathbb{Z}_N$ *lies in the same number of* $R_{i,s}$ *and*

$$N^{-1}\mathrm{Diam}(\psi_s(R_{i,s})) \;\leq\; W^{-2}. \tag{3.25}$$

Let us use this proposition immediately. For each i, s choose $z_0 \in R_{i,s}$. Then using (3.25) and the fact that $\|f\|_\infty \leq 1$ we have

$$\begin{aligned} & \left| \sum_{z \in R_{i,s}} f(z) \left(\omega^{\psi_s(z)} - \omega^{\psi_s(z_0)} \right) \right| \\ \leq \; & 2\pi W^{-2} \sum_{z \in R_{i,s}} |f(z)| \\ \leq \; & 4\pi W^{-1} \\ \leq \; & 2^{-1851} \beta^{5822} |R_{i,s}| \end{aligned}$$

provided that $W \geq C_0(\beta^{-1})$. This will be the case if N is at least some suitably large $C_2(\beta^{-1})$ (though N might need to be larger than it had to be before). It now follows from (3.24) that

$$\sum_i \sum_s \left| \sum_{z \in R_{i,s}} f(z) \right| \; \geq \; 2^{-1851} \beta^{5822} \sum_{i,s} |R_{i,s}|.$$

We are now almost home. The fact that every point lies in the same number of $R_{i,s}$ implies that

$$\sum_i \sum_s \sum_{z \in R_{i,s}} f(z) \; = \; 0,$$

and so

$$\sum_i \sum_s \left(\left| \sum_{z \in R_{i,s}} f(z) \right| + \sum_{z \in R_{i,s}} f(z) \right) \; \geq \; 2^{-1851} \beta^{5822} \sum_{i,s} |R_{i,s}|.$$

From this it follows immediately that

$$\sum_{z \in R} f(z) \; \geq \; 2^{-1852} \beta^{5822} |R|$$

for some $R = R_{i,s}$.

The one remaining obstacle is the fact that whilst R is nice, it might still straddle 0 in $\mathbb{Z}_N$. Thus when $\mathbb{Z}_N$ is "unwrapped", R might become two arithmetic progressions R_1 and R_2. Both of these will still have square common difference, so we can be hopeful of extricating ourselves. The following lemma, proved by a simple averaging argument, covers the situation.

Lemma 59 *Suppose that $\|f\|_\infty \leq 1$ and that $\sum_{z \in R} f(z) \geq \eta |R|$, where R is a nice arithmetic progression in $\mathbb{Z}_N$. Suppose that $R = R_1 \cup R_2$, where neither R_1 nor R_2 straddles 0. Then, for some $i \in \{1, 2\}$, R_i has length at least $\eta |R|/3$ and $\sum_{z \in R_i} f(z) \geq \eta |R_i|/3$.*

Finally, we can deduce the following variant of Gowers' Inverse Theorem.

Theorem 60 *Let $f : \mathbb{Z}_N \to \mathbb{C}$ have $\|f\|_\infty \leq 1$, and suppose that there is a function $\phi : \mathbb{Z}_N \to \mathbb{Z}_N$ such that*

$$\sum_r |\Delta(f;r)(\phi(r))|^2 \geq \beta N^3.$$

Suppose that $N \geq C_2(\beta^{-1})$. Then there is a genuine arithmetic progression $S \subseteq \mathbb{Z}$ with square common difference and length at least $N^{1/C_1(\beta)}$ such that

$$\sum_{x \in S} f(x) \geq 2^{-7677}\beta^{5822}|S|.$$

The sudden change in the powers of two here comes from the fact that we replaced 2β in (3.20) with β.

By the main result of §3.3 we have shown that if A contains no triples $(a, a+d, a+2d)$ with $d = x^2 + y^2$ then A has density $\alpha + O(\alpha^{139728})$ on a progression of size $N^{1/C_1(\alpha^{-1})}$ with square common difference. The new set A' must have the same property - it cannot contain an arithmetic progression with sum-of-two-squares common difference. How often can this argument be iterated?

After $O(\alpha^{-139728})$ iterations we will have reached density 1, by which time the length of our subprogression will still be of the form $N' = N^{1/C_1(\alpha^{-1})}$. In order for the iteration step to work we require that $N' \geq C_2(\alpha^{-1})$ at all times. We therefore have a contradiction provided that $N^{1/C_1(\alpha^{-1})} \geq C_2(\alpha^{-1})$, and it is easy to see that this is implied by some bound $\alpha \gg (\log\log N)^{-c}$. This concludes the proof of Theorem 33. □

According to my calculations, $c = 10^{-6}$ is admissible.

Chapter 4

Unfinished Business

As mentioned in the introduction this section is to be considered in a different light to the earlier two. The results and discussions here make no pretence to completeness or quality, but I hope that some of them will appear in more polished form in my thesis.

Some of the following sections may be read independently of Chapters 2 and 3, but others depend on one or other of these chapters. We hope it will be obvious when this is the case.

4.1 A Few Remarks on $B_h[g]$ Sets

In [17] we were concerned only with the problem of finding upper bounds for $B_h[g]$-subsets of $\{1, \dots, N\}$. However as noted in §2.1 the notion of $B_h[g]$-set makes sense on any subset of an abelian group and perhaps the most natural questions concern $B_h[g]$-subsets of $\mathbb{Z}_N$, particularly when N is prime. Indeed one feels that in passing from $\mathbb{Z}_N$ to $\{1, \dots, N\}$ one has somehow only introduced "boundary effects" that should have little bearing on the apparently deeper underlying arithmetic questions. Our work in [17] was concerned with improving the estimates for these boundary effects, and unfortunately our methods are unable to give any new information about $B_h[g]$-subsets of $\mathbb{Z}_N$. Let $M(h, g, N)$ denote the size of the largest $B_h[g]$-subset of $\mathbb{Z}_N$. Clearly then $M(h, g, N) \leq A(h, g, N)$. So far as I am aware, the best known upper bounds on $M(h, N) = M(h, 1, N)$ come from simple applications of Lemmas 22 and 23 (which, as one can easily check, hold equally well in the modular case). That is to say, one has

$$M(2k, N) \ \leq \ (k!)^{1/k} N^{1/2k} (1 + o(1))$$

and

$$M(2k-1, N) \ \leq \ \left(k!(k-1)!\right)^{1/(2k-1)} N^{1/(2k-1)} (1 + o(1)).$$

In my opinion any significant lowering of these bounds would require a substantial advance in our understanding of these additive representation questions.

For higher values of g it seems that essentially no non-trivial upper bounds are known in the modular case. For $B_2[g]$ sets one could adapt the argument of §2.6 to get the bound

$$M(2,g,N) \ \leq \ (2g-1)^{1/2}N^{1/2}(1+o(1)),$$

a slight improvement on the trivial bound.

The corresponding question about lower bounds is even more interesting. So far as I am aware even the following problem is unsolved.

Problem 61 *Let p be a prime. Is there a B_2-subset of $\mathbb{Z}_p$ of cardinality $p^{1/2}(1+o(1))$?*

As noted in [35] there are constructions of asymptotically optimal B_2 sets in $\mathbb{Z}_N$ for $N = p^2-1$, $N = p^2+p+1$ and $N = p(p-1)$, where p is a prime number. However these constructions do not even shed light on Problem 61 for infinitely many p with our current state of knowledge (though it is quite possible that there are infinitely many primes p for which p^2+p+1 is also prime).

4.2 Further Remarks on Functions with Minimal $M(f)$

In this section, which is really an appendix to Chapter 2, we offer a miscellany of further results concerning Problem 5. These results throw some light on the methods used in §2.3, and we assume a fairly strong familiarity with that section.

We start with a very simple example showing that the upper bound of Lemma 7 is not tight, even if we restrict attention to functions taking only two values.

Lemma 62 $M_0 \leq \frac{99+\sqrt{5}}{158}N^3 + O(N^2)$.

Proof Let A_α be the set obtained by removing an interval of length αN from $\{1,\dots,N\}$. Let f_α be the characteristic function of A_α, weighted by a factor $(1-\alpha)^{-1}$ so that $|f_\alpha| = N$. A slightly tedious computation enables one to show that

$$M(f_\alpha) \ = \ \frac{\frac{2}{3} - 3\alpha + 6\alpha^2 - 5\alpha^3}{(1-\alpha)^4}N^3 + O(N^2).$$

The lemma is obtained by minimising over all $\alpha \in [0,1]$. The value $(99+\sqrt{5})/158 \approx 0.6407$ is not the exact minimum, but it is curiously close. □

Define the quantity $S(\alpha, N)$ to be the smallest possible value of $M(f)$ among all functions f of the form $f(x) = NA(x)/|A|$, where $A \subseteq \{1,\dots,N\}$ is a set of size αN. The following questions might be of some interest.

Problem 63 *What does $S(\alpha, N)$ look like for fixed α as $N \to \infty$? What about if $\alpha = cN^{-\gamma}$? What do the extremal sets look like?*

We observe that $S(N^{-1/2}, N) \sim N^3$. Indeed this is an obvious lower bound, and the example of a Sidon Set of size $N^{1/2}(1 - o(1))$ together with a few extra points shows that it is best possible.

We now turn to a study of the extremal functions in Problem 5. We shall show that for each N there is a unique extremal function F, and that this function is completely characterised by the property that $F * F * F$ is constant on $\{1, \dots, N\}$.

Lemma 64 *Let $\mathcal{S}$ denote the set of all $f : \{1, \dots, N\} \to \mathbb{R}$ with $|f| = N$. Then there is a function $F \in \mathcal{S}$ such that $M(f) \geq M(F)$ for all $f \in F$.*

Proof This is a relatively straightforward compactness argument. Let $\mathcal{T}$ denote the set of all $f : \{1, \dots, N\} \to [-N, N]$ with $|f| = N$. Then $\mathcal{T} \subseteq \mathcal{S}$, and it is easy to see that given any function $f \in \mathcal{S}$ we may find a function $g \in \mathcal{T}$ with $M(f) \geq M(g)$. Indeed we can always take g to be either f itself or the characteristic function of $\{1, \dots, N\}$. Therefore it suffices to prove the lemma for $\mathcal{T}$. However $\mathcal{T}$ is compact in the topology induced by the $\| \ \|_\infty$-norm, and the function $M : \mathcal{T} \to \mathbb{R}$ is continuous. The result follows immediately. □

For the function F found here we write $M(F) = M_0$.

Lemma 65 *Let F be a function with the property in Lemma 64. Then $F * F * F$ is equal to M_0/N on $\{1, \dots, N\}$.*

Proof Let $g : \{1, \dots, N\} \to \mathbb{R}$ be any function with $\sum_x g(x) = 0$. Then for any ϵ we must have $M(F + \epsilon g) \geq M_0$. A small computation gives that

$$\begin{aligned} \frac{d}{d\epsilon} \sum_x \left((F + \epsilon g) * (F + \epsilon g)(x)\right)^2 \Big|_{\epsilon=0} &= 4 \sum_x (F * F)(x)(F * g)(x) \\ &= 4 \sum_x (F * F * F)(x) g(x). \end{aligned}$$

This expression must equal 0, and simple linear algebra tells us that this is the case precisely when $F * F * F$ equal to some constant $c(F)$ on $\{1, \dots, N\}$. It is easy to compute the value of $c(F)$ in terms of M_0. Indeed

$$M_0 \ = \ \sum_x (F * F)(x)^2 \ = \ \sum_x (F * F * F)(x) F(x) \ = \ c(F) N. \tag{4.1}$$

Lemma 66 *There is a unique extremal function in Problem 5.*

Proof For the proof of this lemma we let the hat symbol ($\hat{}$) denote the Fourier transform on $\mathbb{Z}$. Thus for $\theta \in \mathbb{T} = [0,1]$ we set

$$\hat{f}(\theta) \;=\; \sum_x f(x)e(x\theta),$$

where $e(t) = e^{2\pi it}$. With this notation it follows, using standard facts about Fourier transforms, that

$$M(f) \;=\; \int_0^1 |\hat{f}(\theta)|^4 \, d\theta.$$

Let $f, g : \{1, \ldots, N\} \to \mathbb{R}$ be two functions with $|f| = |g| = N$. Then

$$\begin{aligned} M\left(\frac{f+g}{2}\right) &= \int_0^1 \left|\frac{\hat{f}(\theta)+\hat{g}(\theta)}{2}\right|^4 d\theta \\ &\leq \int_0^1 \left(\frac{|\hat{f}(\theta)|+|\hat{g}(\theta)|}{2}\right)^4 d\theta \qquad (4.2) \\ &= \frac{M(f)+M(g)}{2} - \frac{1}{16}\int_0^1 \left(|\hat{f}(\theta)|-|\hat{g}(\theta)|\right)^4 d\theta \\ &\qquad -\frac{3}{8}\int_0^1 \left(|\hat{f}(\theta)|^2-|\hat{g}(\theta)|^2\right)^2 d\theta \\ &\leq \frac{M(f)+M(g)}{2}. \qquad (4.3) \end{aligned}$$

Suppose now that $M(f) = M(g) = M_0$. Then, as $|(f+g)/2| = N$, we must have equality in both (4.2) and (4.3). It is not hard to see that this forces $\hat{f}(\theta) = \hat{g}(\theta)$ for all θ, and hence $f = g$ identically. This proves the lemma. □

For the remainder of this section we will work in $\mathbb{Z}_{2N}$. By now the reader is most probably bored with hearing that we shall regard functions

$$f : \{1, \ldots, N\} \to \mathbb{R}$$

as functions on $\mathbb{Z}_{2N}$, and that the hat symbol refers to Fourier transforms on $\mathbb{Z}_{2N}$.

Lemma 67 *Let $G : \{1, \ldots, N\} \to \mathbb{R}$ be a function with $|G| = N$. Let $\mathcal{H}_{2N}$ denote the set of all functions $H : \mathbb{Z}_{2N} \to \mathbb{R}$ with $|H| = 0$ and $H(x) = 1$ for $x = 1, \ldots, N$ (compare §2.3). Then we have*

$$\left(\sum_{r\neq 0} |\hat{G}(r)|^4\right)\left(\sum_{r\neq 0} |\hat{H}(r)|^{4/3}\right)^3 \;\geq\; 16N^8 \qquad (4.4)$$

for all $H \in \mathcal{H}_{2N}$. Suppose that in addition $G*G*G$ is constant on $\{1, \ldots, N\}$. Then equality can occur in (4.4).

Proof We begin by noting a few useful facts. Observe that, as in (4.1), $G * G * G$ is actually equal to $M(G)/N$ on $\{1, \dots, N\}$. Since G is supported in $\{1, \dots, N\}$, the "modular" convolution $G * G(x)$ is equal to the "$\mathbb{Z}$-" convolution $G * G(x)$ for $-N < x < N$. Also the modular $G * G * G(x)$ is equal to the $\mathbb{Z}$- version for $x \in \{1, \dots, N\}$.

The first part of the lemma (that is to say the inequality (4.4)) can be derived in exactly the same way that we obtained (2.22) in §2.3. For the second statement, we simply observe that one can take

$$H(x) \ = \ \frac{2NG * G * G(x) - N^3}{2M(G) - N^3}.$$

This concludes the proof of the lemma. □

This lemma has a number of consequences. First of all it vindicates the approach we used in §2.3 to obtain a lower bound for $M(f)$, at least up to and including the point where we applied Hölder's Inequality to derive (2.22). Secondly it allows us to prove

Corollary 68 *Suppose $G : \{1, \dots, N\} \to \mathbb{R}$ is a function with $|G| = N$ for which $G * G * G$ is constant on $\{1, \dots, N\}$. Then $G = F$.*

Proof It follows easily from Lemma 67 that G is extremal for Problem 5, the problem of minimising $M(G)$. Hence, by Lemma 66, $G = F$. □

Now let $\lambda > 0$ be a real number and N an integer. Denote by $\mathtt{P}(\lambda)$ the problem of minimising $\sum_r |\hat{f}(r)|^\lambda$ over all functions $f : \mathbb{Z}_{2N} \to \mathbb{R}$ with $\mathrm{Supp}(f) \subseteq \{1, \dots, N\}$ and $\sum_x f(x) = N$. A fair portion of [17], then, was been devoted to the problem $\mathtt{P}(4)$. Using ideas related to Lemma 67, it is possible to demonstrate the following result. We omit the proof.

Proposition 69 *Let F be the unique extremal function for $\mathtt{P}(4)$, and suppose that*

$$\sum_r |\hat{F}(r)|^4 \ = \ cN^4.$$

Then the function

$$W(x) \ = \ \frac{\frac{c}{2} - N^{-2}(F * F * F)(x)}{c - 1}$$

is extremal for $\mathtt{P}\left(\frac{4}{3}\right)$. Furthermore we have

$$\sum_r |\hat{W}(r)|^{4/3} \ = \ N^{4/3}\left(1 + \frac{1}{(c-1)^{1/3}}\right).$$

Finally we remark that 4 may be replaced throughout by an arbitrary even number $2k$ with little extra difficulty.

I find it extremely annoying that it has not been possible to solve these extremal problems

in a reasonably explicit way. In particular the use of the sub-optimal function p in (2.32) leaves open the rather tedious possibility of improving the results of [17] by simply finding a "better" function. I believe that it would be of some interest to solve $\mathtt{P}(4)$ reasonably explicitly, since the quantity $M(f)$ is rather natural from a combinatorial point of view. The question is also related, it turns out, to the problem of finding the best possible constants in what are known as Inequalities of Nikol'skĭi Type.

Let me conclude by relating a slightly amusing story. It is not hard to write a computer program to plot the extremal function F numerically using Lemma 65. I was helped in this respect by Alex Barnard, and the results led Tim Hunt to conjecture that the extremal function F is essentially a scaled and discretised version of the function $h(x) = \frac{1}{\pi}(1-x^2)^{-1/2}$ defined on $(-1,1)$. I was quite excited by this possibility because h is a rather interesting function. It is the invariant measure for the logistic map in dynamical systems, and it is also the pdf of the random variable defined as follows: drop a unit rod randomly on the plane $\mathbb{R}^2$, and take the length of its x-projection. If it were true that $h*h*h$ were constant on $(-1,1)$ then we would have the following curious result - if one drops a toy consisting of three unit rods freely linked at their endpoints, the probabililty density of the length of the x-projection is constant on $(-1,1)$.

Unfortunately I was eventually able to prove, using *Mathematica* and some harmonic analysis, that Hunt's Conjecture is false. The singularities of h at ± 1 made this a rather interesting problem, but I still consider it too computational and also a bit irrelevant to include here.

4.3 Arithmetic Non-Uniformity

The would-be reader of this section may care to familiarise herself with §3.2 before looking at this. As we remarked there one can quite easily remove a logarithm from the bound of Proposition 37. To do this we introduce the concept of arithmetic non-uniformity.

We start with a reappraisal of (3.2). Recall that, under the assumption that $A-A$ contained no squares, we showed that

$$\sum_{r\neq 0} |\hat{S}(r)||\hat{B}(r)||\hat{A}(r)| \ \geq \ \tfrac{1}{4}\delta^2 N^{5/2}. \tag{4.5}$$

Now the Fourier Transform $\hat{S}(r)$ has been well studied in connection with representing integers as sums of squares. In particular it has been known since Weyl in 1918 that $|\hat{S}(r)|$ is small unless r/N is close to a rational with small denominator. It should therefore be possible to simply ignore, in (4.5), the contribution from those r for which r/N is not close to a rational with small denominator. Repeating the analysis of §3.2 would then produce for us a large Fourier coefficient $|\hat{A}(r)|$ in which r/N *is* close to a rational with small denominator.

We shall say in this case that A is *arithmetically non-uniform*, although we shall not attempt to define this notion precisely. It turns out that an arithmetically non-uniform set can be restricted to a very long arithmetic progression (with square common difference) and that a substantially improved bound results.

Let us begin by recalling Weyl's Inequality for a quadratic polynomial. For the proof, a good reference is [29].

Proposition 70 (Weyl) *Let $p(x) = \alpha x^2 + \beta x + \gamma$ be a polynomial and suppose that $|\alpha - a/q| \leq q^{-2}$ for some rational number a/q in lowest terms. Then*

$$\left|\sum_{x=1}^{M} e(p(x))\right| \leq 200\left(Mq^{-1/2} + \sqrt{M\log q} + \sqrt{q\log q}\right).$$

Let us fix a set $A \subseteq \{1,\dots,N\}$ of density $\delta \geq N^{-1/8}$. By Dirichlet's Theorem on approximation by rationals there is, for every $r \in \{1,\dots,N\}$, an integer $q \leq 2^{-26}\delta^2 N/\log N$ and a rational a/q such that

$$\left|\frac{r}{N} - \frac{a}{q}\right| \leq \frac{2^{26}\log N}{qN\delta^2}. \tag{4.6}$$

The numbers we have used here are (of course) far from arbitrary: read on! We say that r is *arithmetic* if the number q in (4.6) is at most $2^{26}\delta^{-2}$. We leave it to the reader to verify the following consequence of Weyl's Inequality.

Proposition 71 *Suppose that $N \geq 2^{200}$, and suppose that r is not arithmetic. Then*

$$|\hat{S}(r)| \leq \tfrac{1}{8}\delta N^{1/2}.$$

Let the set of all arithmetic r be denoted by $\mathcal{R}$. Then, using Proposition 71 and Parseval's Identity, we have

$$\begin{aligned}\sum_{r\notin\mathcal{R}} |\hat{S}(r)||\hat{B}(r)||\hat{A}(r)| &\leq \sup_{r\notin\mathcal{R}}|\hat{S}(r)|\left(\sum_r |\hat{A}(r)|^2\right)^{1/2}\left(\sum_r |\hat{B}(r)|^2\right)^{1/2}\\ &\leq \tfrac{1}{8}\delta^2 N^{5/2}.\end{aligned}$$

Therefore, by (4.5), we have

$$\sum_{r\in\mathcal{R}\setminus\{0\}} |\hat{S}(r)||\hat{B}(r)||\hat{A}(r)| \geq \tfrac{1}{8}\delta^2 N^{5/2}.$$

Now one can simply follow through the opening argument of §3.2, changing a few constants, to show that there is $r \in \mathcal{R}\setminus 0$ for which $|\hat{A}(r)| \geq 2^{-36}\delta^{11/2}|A|$. In other words, A is arithmetically non-uniform.

Let us now use this information to find a very long arithmetic progression on which A has increased density. Once again we follow the arguments of §3.2. Just to recap, we have a number $q \leq 2^{26}\delta^{-2}$ and an integer a such that

$$\left|\frac{r}{N} - \frac{a}{q}\right| \leq \frac{2^{26}\log N}{qN\delta^2}.$$

Writing $L = 2^{-56}\delta^4 N/\log N$, we can easily check that

$$\left\|\frac{rq^2L}{N}\right\| \leq \frac{1}{16},$$

where the $\|.\|$ symbol refers to the distance to the nearest integer. The analysis following Lemma 36 can now be repeated to show that, if B is the progression $q^2, 2q^2, \ldots, Lq^2$, then $|\hat{B}(r)| \geq L/2$. We now continue almost exactly as before. First of all we have that

$$N\sum_x |A \cap (B+x)|^2 \geq \left(1 + 2^{-74}\delta^{11}\right)|A|^2L^2.$$

There are at most $N/4\log N$ bad values of x, and so their contribution to the left-hand term is at most $L^2N^2/4\log N$. If $\delta \geq 64(\log N)^{-1/13}$ then this is less than $2^{-75}\delta^{11}|A|^2L^2$, and so there is a good x for which

$$|A \cap (B+x)| \geq \left(\delta + 2^{-75}\delta^{12}\right)|B|.$$

Suppose that $\delta \geq C(\log N)^{-1/13}$ for some large C (we actually require C to be at least 64 in order for one of our earlier conclusions to hold). If we iterate the above argument (assuming that $A - A$ does not contain a square) then at each stage δ will increase to $\delta + O(\delta^{12})$ and N will decrease by a factor of at most $c(\log N)^2$. After $O(\delta^{-11})$ iterations we will have reached density 1, yet N will still be large enough to make all our arguments work, a contradiction. Modulo a few details we have shown

Proposition 72 *There is a constant C such that, if A is a subset of $\{1, \ldots, N\}$ with density at least $C(\log N)^{-1/13}$, then A contains two elements a, a' with $a - a'$ a non-zero square.*

Although it is of interest to remove a logarithm as we have done here, there is another reason for introducing arithmetic non-uniformity. This is that it is possible to use the idea to prove the following generalisation of Sárközy's Theorem (which we mentioned in §3.1), which is in turn a special case of Theorem 31. I have not read [38] in detail but would not be surprised if my argument bears the same relation to it as the argument of §3.2 does to [37].

Theorem 73 *Let p be a polynomial with $p(\mathbb{Z}) \subseteq \mathbb{Z}$ and $p(0) = 0$. Let $\delta > 0$. Then there is $N_0 = N_0(\delta, p)$ such that if $A \subseteq \{1, \ldots, N\}$ has size δN then A contains two distinct elements x, x' which differ by a number of the form $p(d)$, provided only that $N \geq N_0$.*

To write out a fully detailed proof of this statement would be a daunting task. We will sketch how the proof goes in the simple case $p(n) = n(n+1)$ (I intend to work this out in detail at some future point).

Proposition 74 *Suppose that a is a positive integer and that N is "much bigger" than a. Let S be the characteristic function of the set $\{n(an+1) : n = 1, \dots, \lfloor (N/2a)^{1/2} \rfloor\}$. Regard S as a function on $\mathbb{Z}_N$ in the natural way. Then*

$$\sum_r |\hat{S}(r)|^{12} \leq 2^{50}|S|^{12}. \tag{4.7}$$

Remarks This is an analogue of (3.3). One way of proving it would be a via a direct application of the techniques of Hardy and Littlewood, but this is rather involved and unfortunately I have not managed to find a place in the literature where a result of this form which is uniform in a is derived. It is also possible that the modular forms proof of Lemma 35 could be adapted to prove something to the effect that the number of representations of n as

$$(ax_1^2 + x_1) + (ax_2^2 + x_2) + (ax_3^2 + x_3) + (ax_4^2 + x_4) + (ax_5^2 + x_5) + (ax_6^2 + x_6)$$

is $\ll n^2/a^3$ uniformly in small a. This would imply the proposition by the arguments that we used in §3.2 to derive (3.3) from Lemma 35. One could use Proposition 74 much as we have done above (with a few extra details to account for the presence of a) to prove a result of the following kind.

Proposition 75 *Let a be small in terms of n and suppose that $A \subseteq \{1, \dots, N\}$ has density δ yet contains no pair of elements x, x' with $x - x' = n(an+1)$ for some n. Then A is arithmetically non-uniform.*

Using this proposition one could conclude that A had increased density on a longish arithmetic progression P with small common difference q. Let $A' = A \cap P$. It is easy to see that A' cannot contain two elements which differ by a number of form $n(qan+1)$. This gives us the basis for an iteration in the spirit of the previous arguments we have described. Such an iteration would proceed along the following lines. Set $A_0 = A$ and $a_0 = 1$.

- At the ith stage we have a set A_i of density δ_i which does not contain a pair of elements which differ by $n(a_i n + 1)$, where a_i is not too big;
- A_i is therefore arithmetically non-uniform and so we may pass to a subprogression P_i with small common difference q_i on which A has increased density δ_{i+1};
- Let $A_{i+1} = A_i \cap P_i$ and $a_{i+1} = q_i a_i$.

As with previous arguments the density increases to 1 in finite time, which will give a contradiction provided N was large enough in terms of δ to make all the iteration steps work.

As the reader will see the fact that the squares are multiplicative made the arguments of Chapter 3 possible, and in other cases like the above one has to work harder.

Arithmetic non-uniformity has some intruiging possibilities. One is the fact that it may be used, in a manner similar to that outlined in the above sketch (and using information about the distribution of primes in arithmetic progressions with small common difference), to prove the following result of Sárközy [38].

Theorem 76 *Let $\delta > 0$ and let $A \subseteq \{1, \ldots, N\}$ have size δN, where $N \geq N_0(\delta)$. Then A contains two distinct elements x, x' which differ by $p-1$, where p is a prime.*

As far as I am aware there is no proof of this result using ergodic theory, and I believe that a natural generalisation of this result, namely that the arithmetic progresison in Szemerédi's Theorem may be required to have common difference $p-1$, is open.

Particularly in the context of this last remark there is the interesting possibility that there is a useful notion of "arithmetic quadratic non-uniformity". This might somehow tie in with the methods of Gowers to allow one to pass to a subprogression with small common difference in certain situations. At the moment, however, this is pure speculation.

4.4 A Question of Verstraete

4.4.1 Introduction

In this section we change gears completely and discuss a question asked by Jacques Verstraete [45] in connection with some work of his on cycles in graphs. It has transpired that the questions formulated below are not directly relevant to his work, but I consider them extremely interesting anyway.

We are concerned with sumsets in $\mathbb{Z}_p$, where $p > 2$ is a prime number.

Question 77 *How many subsets of $\mathbb{Z}_p$ have the form $A + A$, where $A \subseteq \mathbb{Z}_p$?*

We begin by obtaining some modest results on this question. Once one has started to think about sets of the form $A+A$, various other questions present themselves. For example whilst it is trivial to write down a very small set which is not of the form $A + A$ – in fact any 2-element set will serve – it seems to be much harder to write down a very large set of this type. In the second part of this note we investigate the quantity $E(p)$, which we define to be the minimal e for which there exists $S \subseteq \mathbb{Z}_p$ of cardinality $p - e$ which is not of form $A + A$.

4.4.2 The Number of Sumsets

Denote by $R(p)$ the number of subsets of $\mathbb{Z}_p$ of form $A + A$. By way of a partial answer to Question 77, we obtain the following bounds on $R(p)$.

Proposition 78

$$2^{\lfloor p/3 \rfloor} \ \leq \ R(p) \ \leq \ 2^{p(1+o(1))/2}.$$

Proof We deal with the lower bound first, as this is rather easy. Let $m = \lfloor p/3 \rfloor$ and let B be any subset of $\{1, \ldots, m\}$, regarded as a subset of $\mathbb{Z}_p$ in the obvious way. Consider the set $A = B \cup \{2m\}$. It is easy to see that $(A + A) \cap \{2m + 1, p\}$ is simply the set B translated by $2m$. This means that B can be completely recovered from $A + A$, and it follows that by this construction we get 2^m different sets of form $A + A$.

For the upper bound we count sets of the form $A+A$ by dividing into two cases. The number of sets of form $A + A$ with $|A + A| \geq 19p/20$ is of course at most

$$\sum_{n \geq 19p/20} \binom{p}{n},$$

and this is easily seen to be much less than $2^{p/2}$ for large p using Stirling's formula. Suppose on the other hand that $|A + A| \leq 19p/20$. Then C, the complement of $A + A$ in $\mathbb{Z}_p$, contains an arithmetic progression of length $k(p) \to \infty$ by Szemerédi's Theorem (Theorem 29). Since p is an odd prime it follows that for some λ and μ we have

$$((\lambda A + \mu) + (\lambda A + \mu)) \ \cap \ [-\tfrac{1}{2}k(p), \tfrac{1}{2}k(p)] \ = \ 0. \tag{4.8}$$

We should point out, lest there be any confusion, that λA means $\{\lambda x \mid x \in A\}$, whilst $A + \mu$ refers to $\{x + \mu \mid x \in A\}$. Let us now proceed to count sets D for which $D + D$ does not meet $[-L, L]$. To this end it is rather helpful to use a probabilistic language in which each $x \in \mathbb{Z}_p$ lies in D with probability $1/2$. Divide $\{1, \ldots, \lfloor p/2 \rfloor\}$ into intervals I_t of lengths between $L/4$ and $L/2$, and effect the mirror image decomposition of $\{-1, \ldots, -\lfloor p/2 \rfloor\}$ into intervals $-I_t$. There will obviously be at most $2p/L$ different values of t. The condition that $D + D$ does not meet $[-L, L]$ forces either $D \cap I_t$ or $D \cap -I_t$ to be empty for each t. For fixed t the probability of this occurring is at most $2^{1-|I_t|}$. Furthermore these events, as t ranges over all its values, are independent. It follows that the probability $q(L)$ that $(D + D) \cap [-L, L] = 0$ satisfies

$$\begin{aligned} q(L) \ &\leq \ \prod_t 2^{1-|I_t|} \\ &\leq \ 2^{2p/L} \cdot 2^{-\sum_t |I_t|} \\ &= \ 2^{2p/L - p/2}. \end{aligned}$$

The number of such sets D is thus at most $2^{2p/L + p/2}$. It follows that the number of sets A satisfying (4.8) for some choice of λ, μ is no more than

$$p^2 2^{\frac{4p}{k(p)} + \frac{p}{2}}.$$

Since this is $2^{p(1+o(1))/2}$, Proposition 78 follows. □

4.4.3 Large Sumsets

We now turn our attentions to the quantity $E(p)$ mentioned in the introduction. Recall that $E(p)$ is defined to be the largest e for which there exists a subset $S \subseteq \mathbb{Z}_p$, $|S| = p - e$, which does not have the form $A + A$. Our first result is that, perhaps a touch surprisingly, *all* sufficiently large subsets of $\mathbb{Z}_p$ have the form $A + A$. To prepare for this we recall a standard argument of Dirichlet.

Lemma 79 *Let $\delta > 0$, let $p \geq p_0(\delta)$ and let $k < \frac{\log p}{\log(2/\delta)}$. Suppose that $r_1, \ldots, r_k \in \mathbb{Z}_p$. Then there is $\lambda \in \mathbb{Z}_p$ for which $\|\lambda r_i\| \leq \delta p$ for all i.*

Proof Let $m > 0$ and let $x_1, \ldots, x_k \in \mathbb{Z}_p$. Write $T(x_1, \ldots, x_k)$ for the box

$$[x_1, x_1 + m) \times [x_2, x_2 + m) \times \cdots \times [x_k, x_k + m)$$

in $\mathbb{Z}_p^k$. For any λ the point

$$P(\lambda) \ = \ (\lambda r_1, \ldots, \lambda r_k)$$

lies in exactly m^k of these boxes, and so on average each box contains $p(m/p)^k$ of the $P(\lambda)$. Take $m = \lfloor \delta p \rfloor$ and let p be large enough that $m \geq \delta p/2$. Then if $p(\delta/2)^k > 1$ we can conclude that some box $T(x_1, \ldots, x_k)$ contains two points $P(\lambda_1)$ and $P(\lambda_2)$, and so that $\lambda = \lambda_1 - \lambda_2$ satisfies the conclusion of the lemma. □

Proposition 80

$$E(p) \ \gg \ \log p.$$

Proof Let $r = \lceil p/9 \rceil$. Let B be any subset of $\{r + 1, \ldots, 2r - 1\}$ and consider the set

$$A \ = \ \{0, \ldots, r\} \cup B \cup \{2r, \ldots, 3r\} \cup \{6r\}.$$

What does $A + A$ look like? It is easily seen that it contains everything in $\{0, \ldots, 7r\} \cup \{8r, \ldots, 9r\}$, but no elements in $\{7r + 1, \ldots, 8r - 1\}$ except for things of form $b + 6r$ with $b \in B$. Thus we see that if S is contained in an interval of length $r - 1$ then $\mathbb{Z}_p \setminus S$ has the form $A + A$. However it is clearly enough that some dilate λS has this property, and this will be the case for any S with $|S| \leq \frac{1}{4} \log p$ by Lemma 79. □

The best upper bound for $E(p)$ that I have found so far is a very long way from the lower bound above. It is also slightly inaccurate to say "I have found" since parts of the argument were supplied by Gowers [13] not long after reading an earlier version in which I had proved $E(p) = o(p)$ by invoking some results of Bourgain which are far too powerful for this purpose. We start with a lemma of a fairly standard type which says that "random sets have small Fourier coefficients".

Lemma 81 *Let $\mu \geq 256 \log p/p$ and choose a subset $B \subseteq \mathbb{Z}_p$ by picking elements randomly and uniformly with probability μ. Then with probability at least $1 - 2p^{-1}$ we have*

$$\sup_{r \neq 0} |\hat{B}(r)| \ \leq \ 4\mu^{1/2} p^{1/2} (\log p)^{1/2}$$

and

$$||B| - \mu p| \ \leq \ 4\mu^{1/2} p^{1/2} (\log p)^{1/2}.$$

Proof We use the following large-deviation inequality due to Bernstein [2], a proof of which can be found in [16] (actually there we only derived a version of Berstein's Inequality for real-valued random variables, but it is a simple matter to derive a fully complex result from this).

Theorem 82 (Bernstein) *Let $X_1, \ldots, X_n$ be independent complex-valued random variables with $\mathbb{E}X_i = 0$ and $\mathbb{E}\,|X_i|^2 = \sigma_i^2$. Write $\sigma^2 = \sigma_1^2 + \cdots + \sigma_n^2$, and suppose that $|X_i| \leq 1$ uniformly in i. Then we have the inequality*

$$\mathbb{P}(|\overline{X}| \geq t) \ \leq \ \exp\left(-\frac{n^2 t^2}{4\sigma^2}\left(1 - \frac{2nt}{\sigma^2}\right)\right).$$

Let us first apply this to estimate the probability that $\hat{B}(r)$ is large when $r \neq 0$. Observing that $\hat{B}(r)$ is a sum of independent random variables $B(x)\omega^{rx}$ which are bounded by 1 and have variance μ, one calculates that

$$\mathbb{P}\left(|\hat{B}(r)| \ \geq \ 4\mu^{1/2} p^{1/2} (\log p)^{1/2}\right) \ \leq \ 2p^{-2} \tag{4.9}$$

provided that $\mu \geq 256 \log p/p$.

In almost exactly the same way we can show that

$$\mathbb{P}\left(||B| - \mu p| \ \geq \ 4\mu^{1/2} p^{1/2} (\log p)^{1/2}\right) \ \leq \ 2p^{-2}. \tag{4.10}$$

The lemma now follows by combining (4.9) for $r = 1, \ldots, p-1$ with (4.10). □

We use Lemma 81 to deduce

Proposition 83 *Let $k \leq p$ be an integer. Then there is a subset $D \subseteq \mathbb{Z}_p$ with size exactly k and*

$$\sup_{r \neq 0} |\hat{D}(r)| \ \leq \ 16k^{1/2} (\log p)^{1/2}. \tag{4.11}$$

Proof Suppose first that $k \geq 256 \log p$. Then we may apply Lemma 81 with $\mu = k/p$. Pick a set B satisfying the conclusions of the lemma. $|B|$ does not differ from k by more than $4k^{1/2}(\log p)^{1/2}$, and so we may create a new set D with size exactly k by adding or deleting

a few points. Since $\|\hat{B} - \hat{D}\|_\infty \leq \|B - D\|_1$ the Fourier transform of D satisfies (4.11). If on the other hand $k \leq 256 \log p$ then (4.11) is trivially satisfied for *any* $D \subseteq \mathbb{Z}_p$ with size k. $\Box$

The relevance of all this to the bounding of $E(p)$ begins to become apparent with the next proposition.

Proposition 84 *Let $k \leq p$ be an integer, and let $D = D(k)$ be a set satisfying the conclusion of Proposition 83. Then $\mathbb{Z}_p \setminus D$ does not contain any sumset $A + A$ for which $|A| \geq 16p(\log p)^{1/2}k^{-1/2}$.*

Proof Suppose indeed that $A + A \subseteq \mathbb{Z}_p \setminus D$. This implies that $\sum_x (A + A)(x)D(x) = 0$, which can be written in terms of Fourier coefficients as $\sum_r \hat{A}(r)^2\hat{D}(-r) = 0$. Applying the triangle inequality gives

$$\begin{aligned} |A|^2 k &= |\hat{A}(0)|^2|\hat{D}(0)| \\ &\leq \sum_{r \neq 0} |\hat{A}(r)|^2|\hat{D}(r)| \\ &\leq 16k^{1/2}(\log p)^{1/2} \sum_r |\hat{A}(r)|^2 \\ &= 16k^{1/2}p(\log p)^{1/2}|A|, \end{aligned}$$

this last step following from Parseval's identity. The proposition follows immediately. $\Box$.

We will now obtain an upper bound for $E(p)$ in the following manner. Applying Proposition 84 for suitable k will give us a large set $E = \mathbb{Z}_p \setminus D$ which does not contain $A + A$ for A large. It follows that for any subset $F \subseteq E$ which has the form $A + A$, A must be small. However (if all the parameters are chosen correctly) there are far fewer sets of this form than there are large subsets of E. In particular, there is some large subset of E which does not have the form $A + A$ at all.

Before stating our main proposition we require a slightly tedious estimate for certain binomial coefficients.

Lemma 85 *Suppose that $r + s \leq p/4$, $r, s > 0$. Then*

$$\binom{p}{r} < \binom{p-s}{r+s}.$$

Proof We have

$$\binom{p}{r} = \left\{ \prod_{i=0}^{s-1} \frac{p-i}{p-r-i} \cdot \frac{r+s-i}{p-r-s-i} \right\} \binom{p-s}{r+s}.$$

The proof now consists in checking that each term of the product is less than 1, the sort of exercise traditionally left to the reader. $\Box$

This has the following immediate corollary, in which we write $\mathcal{N}(a,b)$ for the number of subsets of an a-set of size at most b.

Corollary 86 *Under the conditions of Lemma 85,*

$$\mathcal{N}(p,r) \ < \ \mathcal{N}(p-s,r+s).$$

Proof Using the lemma we have that

$$\begin{aligned}\mathcal{N}(p,r) &= \sum_{0\le i\le r}\binom{p}{i}\\ &< \sum_{0\le i\le r}\binom{p-s}{i+s}\\ &\le \mathcal{N}(p-s,r+s).\end{aligned}$$

Proposition 87

$$E(p) \ \le \ 16p^{2/3}(\log p)^{1/3}.$$

Proof Let $k \le p$ be an integer to be chosen later. Choose $D = D(k)$ as in Proposition 83. Letting $E = \mathbb{Z}_p \setminus D$, Proposition 84 tells us that E does not contain $A + A$ for any set A with $|A| \ge m$, where

$$m \ = \ m(k) \ = \ 16p(\log p)^{1/2}k^{-1/2}.$$

Now Corollary 86 tells us that there are more subsets of E of size at least $p - m - 2k$ than there are subsets $A \subseteq \mathbb{Z}_p$ with $|A| \le m$, because $\mathcal{N}(p,m) < \mathcal{N}(p-k,m+k)$. It follows immediately that there is $F \subseteq E$ with $|F| \ge p - m - 2k$ which is not of the form $A + A$. Taking $k = 4p^{2/3}(\log p)^{1/3}$ gives the result. □

The above bound is the best I can prove, except for the constant 16. One way of improving it might be to find sets with even flatter spectrum than Proposition 83 guarantees. In considering this possibility one is led very naturally to the considerations of the following section.

4.5 Random Sets and a Result of Salem and Zygmund

In Lemma 81 we showed, roughly speaking, that the largest nontrivial Fourier coefficient of random subset $A \subseteq \mathbb{Z}_p$ of size k is at most $k^{1/2}(\log p)^{1/2}$. Is this close to best possible?

We begin with a trivial observation, which is that $\sup_{r\ne 0}|\hat{A}(r)|$ is at least about $k^{1/2}$ by Parseval's identity, with equality precisely when all of the $\hat{A}(r)$ have roughly the same magnitude. In a random set one might expect that the $\hat{A}(r)$ do all have about the same magnitude, but this turns out to be false. It turns out, in fact, that Lemma 81 is close to the truth.

Theorem 88 (Salem-Zygmund) *Let $f : \mathbb{Z}_N \to \{\pm 1\}$ be a random function. Define the Fourier transform $\hat{f}(r)$ of f by*

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

where $\omega = e^{2\pi i/N}$. Then there is an absolute constant c such that

$$\lim_{N\to\infty} \mathbb{P}\left(\max_r \left|\hat{f}(r)\right| \;\geq\; c(N\log N)^{1/2}\right) \;=\; 1.$$

This result was first obtained by Salem and Zygmund in [36] and is roughly equivalent to the statement that Lemma 81 is best possible, up to a constant, when $\mu = 1/2$. In this section we shall furnish a short, largely combinatorial proof of Theorem 88.

Before we start, let us offer a heuristic which may help the reader to see why Theorem 88 is no great surprise. The Fourier transform $F_r = \hat{f}(r)$ is a sum of random vectors in $\mathbb{C}^2$ and (at least for large N) this sum is not biased in any particular direction. Therefore one expects that F_r will be distributed rather like a 2-dimensional normal random variable with mean 0. It is not hard to calculate the expected value of $|F_r|^2$ using Parseval's identity, or directly: it is simply N. Therefore $N^{-1/2}F_r$ ought to behave like a (2-dimensional) $N(0,1)$ random variable X, having p.d.f.

$$\psi(x) \;=\; \frac{1}{\pi}e^{-|x|^2}$$

and distribution

$$\mathbb{P}(|X| \geq t) \;=\; e^{-t^2}.$$

Suppose that the $N^{-1/2}F_r$ were actually independent random variables, all having the same distribution as X. The probability $\mathbb{P}\left(\max|F_r| \leq \alpha(N\log N)^{1/2}\right)$ would then be $(1-N^{-\alpha^2})^N$, which tends to 0 if $\alpha < 1$ and to 1 if $\alpha > 1$. In this idealised situation, then,

$$\mathbb{P}\left(((1-\epsilon)N\log N)^{1/2} \;\leq\; \max_r |F_r| \;\leq\; ((1+\epsilon)N\log N)^{1/2}\right) \longrightarrow 1 \qquad (4.12)$$

for any $\epsilon > 0$.

Salem and Zygmund showed, then, that the lower bound part of this intuition is correct up to a constant. Halász [22] later showed that (4.12) does in fact hold. Both proofs involve establishing quantitive versions of results similar to the Central Limit Theorem.

To achieve the lower bound in (4.12) Halász proves a quantitative version of the Central Limit Theorem specific to the case at hand for the distribution of F_r, and then for the joint distribution of F_r and F_s when $r \neq \pm s$. It turns out that one gets a weak form of independence which is enough to substitute for the heuristic reasoning above.

In all that follows we shall assume for convenience that N is prime. Our mode of attack

hinges on consideration of the higher moments

$$M_{2k} \;=\; N^{-1}\sum_r |\hat{f}(r)|^{2k}.$$

Crucially M_{2k} can be written as

$$M_{2k} \;=\; \sum_{\substack{a_1+\cdots+a_k\\=b_1+\cdots+b_k}} f(a_1)\dots f(a_k)f(b_1)\dots f(b_k). \tag{4.13}$$

We will use this first of all to show that $\mathbb{E}M_{2k}$ is quite large, and then to demonstrate that (at least for moderately sized k) M_{2k} is concentrated about its mean. These two facts together will allow us to prove Theorem 88 with the not unreasonable constant $c = 9/25$.

Lemma 89 *Let $k \leq \log N$. Then $\mathbb{E}M_{2k} \geq \frac{1}{2}k!N^k$ for N sufficiently large.*

Proof Look at each individual term $T(a_i, b_i) = f(a_1)\dots f(b_k)$ of the sum in (4.13). For any x the random variable $X = f(x)$ has the property that $\mathbb{E}X^{2j+1} = 0$, whilst $X^{2j} = 1$ identically. It follows that the expected value of $T(a_i, b_i)$ is zero unless the $2k$ elements a_i, b_i are *paired*, by which we mean nothing more than that we may partition them into k pairs of equal elements. If (a_i, b_i) is paired then $T(a_i, b_i)$ is in fact identically equal to 1, and so $\mathbb{E}M_{2k}$ is at least the number of paired solutions to $\sum a_i = \sum b_i$. Now any permutation $\pi : [n] \to [n]$ gives such a solution by choosing the a_i arbitrarily and setting $b_i = a_{\pi(i)}$. If, furthermore, we choose the a_i to be distinct then we get a solution that cannot arise from any permutation except π. Thus, summing over π, we get

$$\begin{aligned}\mathbb{E}M_{2k} &\geq k!N(N-1)\dots(N-k+1)\\ &\geq \tfrac{1}{2}k!N^k\end{aligned}$$

for N sufficiently large. □

Now we can write $M_{2k} = Y + Z$ where Y is the sum over all $T(a_i, b_i)$ with (a_i, b_i) paired and Z is the sum over all other $T(a_i, b_i)$. As we observed in the proof of Lemma 89, Y is identically equal to $\mathbb{E}M_{2k}$ whilst $\mathbb{E}Z = 0$. The next step in our argument is an examination of $\mathrm{Var}(Z)$.

Lemma 90 $\mathrm{Var}(Z) \leq (4k)!N^{2k-1}/2^{2k}(2k)!$.

Proof We can write down

$$\mathrm{Var}(Z) \;=\; \sum_{a_i,b_i,c_i,d_i}\left(\prod_{i=1}^k f(a_i)f(b_i)f(c_i)f(d_i)\right), \tag{4.14}$$

where the sum is over all $4k$-tuples (a_i, b_i, c_i, d_i) with $\sum a_i = \sum b_i$, $\sum c_i = \sum d_i$ for which neither (a_i, b_i) nor (c_i, d_i) is paired. We shall call the set of all such $4k$-tuples P. Each term of the sum in (4.14) has expectation zero unless the $4k$-tuple (a_i, b_i, c_i, d_i) is itself paired, in which case the term has expectation 1. The number of ways in which we may try to pair elements of a $4k$-tuple is precisely $(4k)!/2^{2k}(2k)!$. For each of these "potential pairings" p, we ask how many $4k$-tuples $(a_i, b_i, c_i, d_i) \in P$ are actually paired as described by p (we shall say, in this case, that the $4k$-tuple is paired by p).

If p restricts to a pairing on (a_i, b_i) then no elements of P are paired by p, by the definition of P. In all other cases we have $2k$ linear relations on the vector (a_i, b_i, c_i, d_i) coming from the potential pairing p, plus a further relation $\sum a_i = \sum b_i$ which is not in the linear span of the first $2k$. Since N is prime, and therefore $\mathbb{Z}_N^{4k}$ is a vector space, this means that at most N^{2k-1} $4k$-tuples are paired by p. Summing over all possible p gives the bound stated. $\square$

Now a simple application of Tschebicheff's Inequality together with Lemmas 89 and 90 gives that

$$\begin{aligned}
\mathbb{P}\left(M_{2k} \leq \tfrac{1}{4}k!N^k\right) &\leq \mathbb{P}\left(Z \leq -\tfrac{1}{4}k!N^k\right) \\
&\leq \frac{\mathrm{Var}(Z)}{\left(\frac{1}{4}k!N^k\right)^2} \\
&\leq \frac{16(4k)!}{2^{2k}(2k)!(k!)^2N}.
\end{aligned}$$

An ever so slightly tedious calculation involving Stirling's formula shows that this is at most $N^{-1/30000}$ when $k \leq \frac{22}{61}\log N$, the key point here being that $22/61 < (4\log 2)^{-1}$. Now suppose that $M_{2k} \geq \frac{1}{4}k!N^k$. Parseval's identity then yields

$$\begin{aligned}
N^2 \max_r |\hat{f}(r)|^{2k-2} &= \max_r |\hat{f}(r)|^{2k-2} \sum_r |\hat{f}(r)|^2 \\
&\geq \sum_r |\hat{f}(r)|^{2k} \\
&\geq \tfrac{1}{4}k!N^{k+1},
\end{aligned}$$

from which we see that

$$\max_r |\hat{f}(r)| \geq N^{1/2}\sqrt{\frac{k}{e}}\left(1 + o_k(1)\right).$$

We know, from the considerations above, that this holds with probability $\to 1$ when $k = \frac{22}{61}\log N$, and it is easy to see that this implies Theorem 88 with $c > 9/25$.

There are a number of questions about random sets that I would like to be able to answer. Firstly it has come to my attention that the arguments of this section are quite similar to

some fairly old methods used to establish results about the distribution of eigenvalues of random matrices. It is vaguely possible that some of the newer technology of random matrix theory could allow me to give a much cleaner proof of Halász's result [22].

Of relevance to §4.4.3 is the existence of "better than random" subsets of $\mathbb{Z}_p$ which have very flat spectrum. In particular I do not believe that the answer to either of the following two questions is known.

Question 91 *Does there exist an absolute constant C such that for any real number $\alpha \in (0,1)$ there is a family of sets $A_p \subseteq \mathbb{Z}_p$ (one for each prime) with $|A_p| = \lfloor \alpha p \rfloor$ and*

$$\sup_{r \neq 0} \left| \hat{A}_p(r) \right| \leq C|A_p|^{1/2}$$

for all $p > p_0(\alpha)$?

Question 92 *Does there exist an absolute constant C such that for any real number $\alpha \in (0,1)$ there is a family of sets $A_p \subseteq \mathbb{Z}_p$ (one for each prime) with $|A_p| = \lfloor p^\alpha \rfloor$ and*

$$\sup_{r \neq 0} \left| \hat{A}_p(r) \right| \leq C|A_p|^{1/2}$$

for all $p > p_0(\alpha)$?

Very recently I have made some progress on Question 91. My best result so far is the following.

Proposition 93 *Let $\delta \in (0,1)$. For each sufficiently large prime p there is a set $A_p \subseteq \mathbb{Z}_p$ with size $\lfloor \delta p \rfloor$ whose spectrum is considerably flatter than that of a random set of the same size. Specifically,*

$$\sup_{r \neq 0} |\hat{A}_p(r)| \ll p^{1/2} \log\log p.$$

I shall not discuss this result any further here, as it represents work that has only just begun.

Bibliography

[1] Bergelson, V. and Leibman, A. *Polynomial Extensions of Van der Waerden's and Szemerédi's Theorems*, J. Amer. Math. Soc. **9** (1996), no 3, 725 – 753.

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

[3] Bose, R.C. and Chowla, S. *Theorems in the Additive Theory of Numbers*, Comment. Math. Helv. **37** (1962/1963), 141 – 147.

[4] Chamizo, F. *Correlated Sums of* $r(n)$, J. Math. Soc. Japan **51** (1999), no. 1, 237 – 252.

[5] Chen, S. *On the Size of Finite Sidon Sequences*, Proc. Amer. Math. Soc. **121** (1994) 353–356.

[6] Cilleruelo, J. *An Upper Bound for* $B_2[2]$ *Sequences*, Journal of Combinatorial Theory, Series A **89** (2000) 141–144.

[7] Cilleruelo, J, preprint.

[8] Cilleruelo, J., Ruzsa, I. Z. and Trujillo, C. *Upper and Lower Bounds for Finite* $B_h[g]$ *Sequences,* $g > 1$, to appear in Journal of Number Theory.

[9] Erdős, P. and Turán, P. *On a Problem of Sidon in Additive Number Theory and On Some Related Problems*, Journal of the London Mathematical Society **16** (1941) 212–215.

[10] Furstenburg, H. *Ergodic behaviour of diagonal measures and a theorem of Szemeredi on arithmetic progressions*, J. Analyse. Math. **31** (1977) 204 – 256.

[11] Gowers, W.T. *A New Proof of Szemerédi's Theorem for Progressions of Length 4*, Geom. Funct. Anal. **8** (1998), no. 3, 529 –551.

[12] Gowers, W.T. *A New Proof of Szemerédi's Theorem*, preprint.

[13] Gowers, W.T. *Personal Communication.*

[14] Graham, S.W. B_h *Sequences*, in Analyic Number Theory Vol 1 (Allerton Park, IL, 1995) 431–449, Progress in Mathematics **138**, Birkhäuser, Boston MA 1996.

[15] Green, B. J. *Notes on Sieve Theory: The Selberg Sieve*, available at `http://www.dpmms.cam.ac.uk/~bjg23/expos.html`.

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

[17] Green, B.J. *The Number of Squares and* $B_h[g]$ *Sets*, submitted to Acta Arithmetica.

[18] Green, B.J. *On Arithmetic Structures in Dense Sets of Integers*, preprint.

[19] Grimmett, G.R. and Stirzaker, D.R. *Probability and Random Processes*, Clarendon Press, Oxford 1992.

[20] Guy, R.K. *Unsolved Problems in Number Theory* Second Edition, Springer 1994.

[21] Habsieger, L. and Plagne, A. *Ensembles* $B_2[2]$*: L'Étau Se Resserre*, preprint.

[22] Halász, G. *On a result of Salem and Zygmund concerning random polynomials*, Studia Sci. Math. Hungar. **8** (1973) 369 – 377.

[23] Halberstam, H. and Richert, H. -E. *Sieve Methods*, Academic Press 1974.

[24] Halberstam, H. and Roth, K.F. *Sequences* Second Edition, Springer 1983.

[25] Iwaniec, H. *Introduction to the Spectral Theory of Automorphic Forms*, Biblioteca de la Revista Matemática Iberoamericana, 1995.

[26] Jia, X. *On* B_{2k} *Sequences*, Journal of Number Theory **48** (1994) 183–196.

[27] Kolountzakis, M.N. *The Density of* $B_h[g]$ *Sequences and the Minimum of Dense Cosine Sums*, Journal of Number Theory **56** (1996) 4–11.

[28] Lindström, B. *A Remark on* B_4*-Sequences*, Journal of Combinatorial Theory **7** (1969) 276–277.

[29] Montgomery, H.L. *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, CBMS Regional Conference Series in Mathematics 84, AMS 1994.

[30] Nathanson, M.B. *Elementary Methods in Number Theory*, Springer 2000.

[31] Pintz, J., Steiger, W.L. and Szemerédi, E. *On Sets of Natural Numbers whose Difference Set Contains No Squares*, J. London Math. Soc. (2) **37** (1988) 219 – 231.

[32] Plagne, A. *A New Upper Bound for the Cardinality of $B_2[2]$-Sets*, to appear in Jour. Comb. Th. Ser. A.

[33] Roth, K.F. *On Certain Sets of Integers*, J. London Math Soc **28** (1953) 104 – 109.

[34] Ruzsa, I.Z. *Difference Sets Without Squares*, Period. Math. Hungar. **15** (1984), no. 3, 205 – 209.

[35] Ruzsa, I.Z. *Solving a Linear Equation in a Set of Integers I*, Acta. Arith. **65** (1993), no. 3, 259 – 282.

[36] Salem, R. and Zygmund, A. *Some properties of trigonometric series whose terms have random signs*, Acta. Math. **91** (1954) 254 – 301.

[37] Sárközy, A. *On Difference Sets of Sequences of Integers I*, Acta. Math. Acad. Sci. Hungar **31** (1978), nos. 1 – 2, 125–149.

[38] Sárközy, A. *On Difference Sets of Sequences of Integers III*, Acta Math. Acad. Sci. Hungar **31** (1978), nos. 3 – 4 355 – 386.

[39] Sárközy, A. and Sós, V.T. *On Additive Representation Functions*, in The Mathematics of Paul Erdös, Vol. 1, Springer 1997.

[40] Schmidt, W.M. *Small Fractional Parts of Polynomials*, Regional Conference Series in Mathematics **32**, AMS 1977.

[41] Singer, J. *A Theorem in Finite Projective Geometry and Some Applications to Number Theory*, Trans. Amer. Math. Soc. **43** (1938), 377 – 385.

[42] Srinivasan, S. *On a Result of Sarkozy and Furstenburg*, Nieuw. Arch. Wisk (4) **3** (1985), no. 3, 275 – 280.

[43] Szemerédi, E. *On Sets of Integers Containing No k Elements In Arithmetic Progression*, Acta. Arith **27** (1975), 199 – 245.

[44] Tao, T. *From Rotating Needles to Stability of Waves: Emerging Connections Between Combinatorics, Analysis and PDE*, to appear, Notice. Amer. Math. Soc.

[45] Verstraete, J. *Personal Communication.*