ADDITIVE COMBINATORICS (WINTER 2005)

Andrew Granville

Introduction

For A, B subsets of an additive group Z, we define $A+B$ to be the sumset $\{a+b : a \in A, b \in B\}$, and kA to be the k-fold sum $A+A+\cdots+A$ of A. We also let $A-B=\{a-b : a \in A, b \in B\}$ and $b+A=\{b\}+A$ for a single element set $\{b\}$, a *translate* of A. Note that $A-A$ is *not* 0 unless $|A|=1$. We let $k \diamond A=\{ka : a \in A\}$, a *dilate* of A. There are many obvious properties of "+" that can be checked like commutativity, associativity and the distributive law $A+(B \cup C)=(A+B) \cup (A+C)$.

Prove that $k \diamond A \subseteq kA$ and classify when they are equal. Prove that $|b+A|=|A|$. Show that $|A| \leq |A+B| \leq |A||B|$. Describe the situations when we get equality. Improve this last upper bound for $|A+A|$ and for $|A-A|$.

We shall be most interested in understanding the size and structure of sumsets which are subsets of the integers $\mathbb{Z}$, often working with $\mathbb{Z}_{\geq m}$ (where $A_{\geq m}$ denotes the integers in A that are $\geq m$), or of $\mathbb{Z}/N\mathbb{Z}$ for some positive integer N. That every integer is the sum of four squares of integers can be written down as $4\mathbb{Z}^2=\mathbb{Z}$ (where here $\mathbb{Z}^2$ denotes the squares of the integers); if $h(A \cup \{0\}) \supseteq \mathbb{Z}_{\geq m}$ for some m then we say that A is a *basis of order h* for the integers, and thus $\mathbb{Z}^2$ is a basis of order 4. The Goldbach conjecture can be written as $2\mathbb{P}=2 \diamond \mathbb{Z}_{\geq 2}$ where $\mathbb{P}$ is the set of primes, or even $3(\mathbb{P} \cup \{0\})=\mathbb{Z}_{\geq 2} \cup \{0\}$ (verify this is indeed equivalent). The twin primes conjecture states that for every even integer k there are infinitely many pairs of primes $p, p+2k$: verify that this can be rewritten as $\mathbb{P}_{\geq m}-\mathbb{P}_{\geq m}=2 \diamond \mathbb{Z}$ for all m.

In this course we will be primarily studying what it means that $A+A$ is small; that is, if this is so then what does it imply about A. We shall find that this implies that A has a readily describable structure which can then be applied to various problems. There are many open problems in this field that invite investigation; for example to fully understand the structure of $A+B$ when this sumset is "small", in the case that B is significantly smaller than A.

It is easy to see that if $A \subset \mathbb{Z}$ then $|2A| \leq |A|(|A|+1)/2$, since the distinct elements of $2A$ are a subset of $\{a_i+a_j : \ 1 \leq i \leq j \leq |A|\}$; moreover $2A$ can be this large, for example with $A=\{1,2,2^2,2^3,\ldots,2^{n-1}\}$. Prove that this is so, and give an infinite class of examples described simply by the growth of the elements of A. Moreover show that if we select a set A of

These notes borrow from those of Tim Gowers and Jacques Verstraete, Terry Tao and Van Vu, Ben Green, and various published books

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

n integers "at random" from $\{1, \dots x\}$ with $x \geq n^{4+\epsilon}$ then $|2A|$ will equal $|A|(|A|+1)/2$ with probability $\to 1$ as $n \to \infty$.

So we have proved that "typically" $|2A|$ is large and that it is only very special circumstances that it is small. A key, but easy, result for getting a feel for our subject is the following:

Lemma 1. *If A and B are finite subsets of $\mathbb{Z}$ then $|A+B| \geq |A|+|B|-1$. Equality holds if and only if A and B are each complete finite segments of an arithmetic progression to the same modulus.*

Proof. Write the elements of A as $a_1 < a_2 < \dots a_r$, and those of B as $b_1 < b_2 < \dots b_s$. Then $A+B$ contains the $r+s-1$ distinct elements

$$a_1 + b_1 < a_1 + b_2 < a_1 + b_3 < \cdots < a_1 + b_s < a_2 + b_s < a_3 + b_s < \cdots < a_r + b_s.$$

If it contains exactly $r+s-1$ elements then these must be the same, in the same order, as $a_1+b_1 < a_2+b_1 < a_2+b_2 < a_2+b_3 < \cdots < a_2+b_s < a_3+b_s < \cdots < a_r+b_s$. Comparing terms, we have $a_1 + b_{i+1} = a_2 + b_i$ for $1 \leq i \leq s-1$; that is $b_j = b_1 + (j-1)d$ where $d = a_2 - a_1$. A similar argument with the roles of a and b swapped, reveals our result.

If $A+B$ is small, not as small as $|A|+|B|-1$ but not much bigger, then we might expect to be able to use a similar proof to prove a similar structure theorem. Try! After a little play one quickly finds that $A+B$ is small if A and B are both large subsets of complete finite segments of an arithmetic progression to the same modulus. A further interesting example is given by $A = B = \{1, 2, \dots, 10, 101, 102, \dots 110, 201, 202, \dots 210\}$, or its large subsets. One observes that this can be written as $1 + \{0, 1, 2, \dots, 9\} + \{0, 100, 200\}$, a translate of the sum of complete finite segments of two arithmetic progressions. More generally, define a *generalized arithmetic progression* $C = C(a_0, a_1, \dots a_k; N_1, N_2, \dots, N_k)$ as

$$C := \{a_0 + a_1 n_1 + a_2 n_2 + \cdots + a_k n_k : \ 0 \leq n_j \leq N_j - 1 \text{ for } 1 \leq j \leq k\}$$

where $a_0, a_1, \dots a_k$ are given integers, and $N_1, N_2, \dots N_k$ are given positive integers. Note that $C(a_0, a_1, \dots a_k; N_1, N_2, \dots, N_k) = a_0 + \sum_{i=1}^{k} a_i \diamond \{0, 1, \dots, N_i - 1\}$. This generalized arithmetic progression is said to have *dimension* k and *volume* $N_1 N_2 \dots, N_k$. Notice that

$$2C(a_0, a_1, \dots a_k; N_1, N_2, \dots, N_k) = C(2a_0, a_1, \dots a_k; 2N_1 - 1, 2N_2 - 1, \dots, 2N_k - 1).$$

so that $|2C| < 2^k |C|$.

If you try to find other sets A and B with $A+B$ small then it seems you will be out of luck. In the case that $A = B$ this is the extraordinary insight of Freiman [5]: he showed that $2A$ can be "small" if and only if it is a "large" subset of a "low" dimensional generalized arithmetic progression of "not too big" volume[1]. This is the central result of this course. Freiman's 1962 proof is both long and difficult to understand. I believe that it is fair to say that the subject did not progress as much as it might have done since people had difficulty appreciating what Freiman had done. It was not until Ruzsa's 1994 proof of Freiman's result, which is extraordinarily elegant and insightful, that the subject exploded with new ideas and results. As we will see in this course, much of our development of the subject stems from the wealth of ideas in Ruzsa'a treatment, and it is for this reason that we call this main result "the Freiman-Ruzsa theorem".

[1] The terms inside quotation marks all need quantifying and this is not easy.

Densities

For a given set of integers A which we will write as $\{a_1 < a_2 < \dots\}$, define $A(n) = \#\{a \in A : \ 1 \le a \le n\}$. The Schnirelmann density of A is defined by

$$\sigma(A) := \inf_{n \ge 1} A(n)/n.$$

Note that $A(n) \ge n\sigma(A)$ for all $n \ge 1$. We also define the more usual, upper and lower densities,

$$\underline{d}(A) = \lim \inf_{n \to \infty} A(n)/n \quad \text{and} \quad \overline{d}(A) = \lim \sup_{n \to \infty} A(n)/n,$$

with $d(A) = \underline{d}(A) = \overline{d}(A)$ if they are equal.

Show that (i) $\sigma(A) = 1$ if and only if $A \supseteq \mathbb{Z}_{\ge 1}$; (ii) If $1 \notin A$ then $\sigma(A) = 0$; (iii) If $\sigma(A) = 0$ then $1 \notin A$ or $\underline{d}(A) = 0$; (iv) $\underline{d}(A) \ge \sigma(A)$.

Define $A_n := \{a - n : \ a \ge n, \ a \in A\}$. (v) Show that $\underline{d}(A_n) = \underline{d}(A)$ for all $n \ge 1$.

For all $\epsilon > 0$ there exists n such that $\sigma(A_n) \ge \underline{d}(A) - \epsilon$: For if not then let $n_0 = 0$ and define $n_{i+1} = n_i + m$ where $A_{n_i}(m) \le (\sigma(A_{n_i}) + \epsilon/2)m \le (\underline{d}(A) - \epsilon/2)m$. Then $A(n_k) = \sum_{i=1}^{k-1} A(n_{i+1}) - A(n_i) = \sum_{i=1}^{k-1} A_{n_i}(n_{i+1} - n_i) \le \sum_{i=1}^{k-1} (\underline{d}(A) - \epsilon/2)(n_{i+1} - n_i) = (\underline{d}(A) - \epsilon/2)n_k$, which contradicts the definition of $\underline{d}(A)$, since the values of n_k get arbitrarily large .

We are interested in adding sets and obtaining a large sum. We will usually assume that $0 \in A \cap B$ for this implies that $A \cup B \subseteq A + B$. In the next result we prove a simple consequence of the pigeohole principle but rephrased here in our terminology.

Lemma 2.1. *If $0 \in A \cap B$ and $\sigma(A) + \sigma(B) \ge 1$ then $A + B \supseteq \mathbb{Z}_{\ge 0}$.*

Proof. Suppose not and let n be the smallest positive integer for which $n \notin A + B$: then $n \notin A \cup B$, and A and $n - B$ are disjoint. Let $C = A \cup (n - B)$ so that $n - 1 \ge C(n-1) = A(n-1) + (n - B)(n-1) = A(n-1) + B(n-1) = A(n) + B(n) \ge \sigma(A)n + \sigma(B)n \ge n$, a contradiction to the supposition.

Schnirelmann's theorem. *If $1 \in A$ and $0 \in B$ then $\sigma(A+B) \ge \sigma(A) + \sigma(B) - \sigma(A)\sigma(B)$.*

A greedy proof. Given $a \in A_{\ge 1}$ we count the number of elements in $a + B \subseteq A + B$ just a little larger than a. That is, if $x > a$ then $(A+B)(x) - (A+B)(a-1) \ge (a+B)(x) - (a+B)(a-1) = B(x-a) + 1 \ge \sigma(B)(x-a) + 1$. Therefore if $1 = a_1 < a_2 < \dots < a_k \le n$ are the elements of $A \cap [1, n]$ then

$$\begin{aligned}(A+B)(n) &= \sum_{i=1}^{k-1}((A+B)(a_{i+1}-1) - (A+B)(a_i - 1)) + ((A+B)(n) - (A+B)(a_k - 1)) \\ &\ge \sum_{i=1}^{k-1}(\sigma(B)(a_{i+1} - a_i - 1) + 1) + (\sigma(B)(n - a_k) + 1) \\ &= \sigma(B)(x - k) + n = \sigma(B)n + (1 - \sigma(B))A(n) \ge (\sigma(A) + \sigma(B) - \sigma(A)\sigma(B))n.\end{aligned}$$

We can deduce that any set of positive density is a basis for the integers:

Corollary 2.2. *If $0 \in A$ and $\sigma(A) > 0$ then there exists h such that $hA \supseteq \mathbb{Z}_{\geq 0}$. We may take $h \leq 2\lceil (\log 2)/(-\log(1-\sigma(A))\rceil$.*

Proof. As $\sigma(A) > 0$ we know that $1 \in A \subset kA$ for all $k \geq 1$; and we deduce, by induction, from Schnirelmann's theorem, that if $0, 1 \in A$ then $1 - \sigma(kA) \leq (1-\sigma(A))^k$ for all $k \geq 1$. Now let k be the smallest integer for which $(1-\sigma(A))^k \leq 1/2$ so that $\sigma(kA) \geq 1/2$, and then $kA + kA \supseteq \mathbb{Z}_{\geq 0}$ by Lemma 2.1.

The most famous consequence of this is Schnirelmann's result that the primes form an additive basis, a first step along the road to Goldbach's conjecture. To prove this will require a little sieve theory to prove that $\underline{d}(2\mathbb{P}_{\geq 3}) \geq 1/8$. Use this to show that the primes form an additive basis of order ≤ 13.

A similar method can be used to prove Hilbert's theorem on Waring's problem, that for every integer k, there exists h such that every the kth powers of integers form an additive basis.

First though, we will make an analogous first step along the road to the twin prime conjecture.

Given $S \subset Z$ we define *the cube* $\overline{S} := \{\sum_{s \in S} \epsilon_s s : \ \epsilon_s \in \{-1, 0, 1\}$ for each $s \in S\}$, to have *dimension* $|S|$. Notice that $|\overline{S}| \leq 3^{|S|}$. If we have equality here then we call this a *proper cube*. Notice that a cube is a special case of a generalized arithmetic progression; and that our definition works for any additive group Z.

Proposition 2.3. *If $A \subseteq \mathbb{Z}_{\geq 1}$ with $\overline{d}(A) > 0$, then there exists a finite cube $\overline{S}$ for which $A - A + \overline{S} = \mathbb{Z}$. In fact there exists such a cube of dimension $\lfloor \log(1/\overline{d}(A))/\log 2 \rfloor$.*

Proof. If $A - A \neq \mathbb{Z}$ then there exists an integer m such that $m \notin A - A$, and so A and $m+A$ are disjoint. Let $A_1 = A \cup (m+A)$ so that $\overline{d}(A_1) = 2\overline{d}(A)$ and $A_1 - A_1 = A - A + \overline{\{m\}}$. If this is not $\mathbb{Z}$, we then define $A_2, A_3, \ldots$ and so on. However this construction cannot continue if $\overline{d}(A_k) > 1/2$ (since we cannot have $\overline{d}(A_{k+1}) > 1$), and therefore $A_k - A_k = \mathbb{Z}$. This gives our result with k chosen so that $2^k\overline{d}(A) > 1/2$.

From this and the result that $\underline{d}(2\mathbb{P}_{\geq 3}) \geq 1/8$ we deduce that there exists a cube $\overline{S}$ of dimension ≤ 3 for which $2 \diamond \mathbb{Z} = 2\mathbb{P}_{\geq 3} - 2\mathbb{P}_{\geq 3} + \overline{S}$.

One can continue this line of thinking in the well-known example that $4\mathbb{Z}^2 = \mathbb{Z}_{\geq 0}$; in this case it is known that every positive integer is represented many times as the sum of four squares and that there are infinitely which are not the sum of three squares (in fact, precisely the integers $\{4^k(8m-1) : k \geq 0, m \geq 1\}$). One might ask whether one can find a "thin" subset A of $\mathbb{Z}^2$, perhaps finite, such that $3\mathbb{Z}^2 + A = \mathbb{Z}_{\geq 0}$. From the classification of integers that do not belong to $3\mathbb{Z}^2$ it is easy to show that $3\mathbb{Z}^2 + \{0, 1, 4\} = \mathbb{Z}_{\geq 0}$. It is a challenge to find a "thin" set A for which $2\mathbb{Z}^2 + A = \mathbb{Z}_{\geq 0}$ – show that such a set A cannot be finite.

The prime k-tuplets conjecture

The prime number theorem tells us that there are $\sim x/\log x$ primes $\leq x$; or, put another way, if we randomly chose an integer near x then it is prime with probability $1/\log x$.

If we were to ask how often n and $n+k$ are prime when $n \leq x$ then we might guess that one can assume that the events that they are each prime is "independent" one another, and so this happens for about $1/\log^2 x$ of the integers n around x. However, in the case that $k = 1$ this heuristic fails to account for the fact that one of n and $n+1$ is always even, and thus n and $n+1$ cannot be simultaneously prime when $n > 2$. To take the divisibility of prime numbers into account, we note that the probability that neither of two randomly chosen numbers are divisible by p is $(1-1/p)^2$, whereas the probability that neither of n and $n+k$ are divisible by p if n is chosen randomly is $1-\omega(p)/p$ where $\omega(p) = 2$ unless p divides n, in which case $\omega(p) = 1$. Thus we expect $\sim (\prod_p (1 - \omega(p)/p)/(1 - 1/p)^2 + o(1))x/\log^2 x$ prime pairs $n, n+k$ with $n \leq x$, and this claim is well supported by computational evidence.

Given a set of irreducible polynomials $f_1(x), f_2(x), \ldots, f_k(x)$ we define $\pi_f(x)$ to be the number of integers $n \leq x$ for which $|f_1(n)|, |f_2(n)|, \ldots$ and $|f_k(n)|$ are all prime. Similarly, by defining $\omega_f(p)$ to be the number of integers n for which $f_1(n)f_2(n)\ldots f_k(n) \equiv 0 \pmod{p}$, we predict that $\pi_f(x) \sim (\prod_p (1 - \omega(p)/p)/(1 - 1/p)^k + o(1))x/d\log^k x$, where d is the product of the degrees of the f_j.

Define $r(n) = \#\{p, q \text{ primes} : p + q = n,\ p \equiv i \pmod{m}, q \equiv j \pmod{m}\}$ where m is the product of the primes $\leq y$ where $y = [\log n/\log\log n]$ with $(ij, m) = 1$, and let $k = i+j$. Obviously $r(n) = 0$ unless $n \equiv k \pmod{m}$, in which case we predict (in a similar manner to above) that $r(n) \sim mn/(\phi(m)\log n)^2$. Now the fundamental lemma of the sieve gives upper bounds for $r(n)$ that are slightly weaker than this: modifying the proof of Theorem 3.11 in [HR] in the obvious way one obtains that $r(n) \leq \{4+o(1)\}mn/(\phi(m)\log n)^2$. Now fix $\epsilon > 0$. Let A be the set of integers $n \equiv k \pmod{m}$ with $(1-\epsilon)x < n \leq x$ for which $r(n) > 0$. Then

$$\{4+o(1)\}|A|\frac{mx}{(\phi(m)\log x)^2} \geq \sum_{\substack{(1-\epsilon)x<n\leq x\\ n\equiv k \pmod{m}}} r(n) = \sum_{\substack{p\equiv i \pmod{m}\\ q\equiv j \pmod{m}\\ (1-\epsilon)x<p+q\leq x}} 1 \sim \frac{(2\epsilon-\epsilon^2)x^2}{2(\phi(m)\log x)^2},$$

by the prime number theorem for arithmetic progressions, so that $|A| \geq (1/4 - \epsilon/8 + o(1))\epsilon x/m$. Summing over all such intervals, and then letting $\epsilon \to 0$, we discover that $r(n) > 0$ for $> (1/4 - o(1))x/m$ integers $n \leq x$ for which $n \equiv k \pmod{m}$. Summing over all even k (with an appropriate pair i, j) we discover that $\underline{d}(2\mathbb{P}_{\geq 3}) \geq 1/8$, as claimed in the previous subsection.

Modify the above argument to show that if Q is a subset of the primes of positive density (that is, there exists a constant $c > 0$ such that $Q(x) \geq c\pi(x)$ if x is sufficiently large) then $\underline{d}(2Q) > 0$. Deduce that Q is an additive basis for the integers.

The Dyson transformation and its consequences

Freeman Dyson defined a transformation $A, B \to \delta_e(A), \delta_e(B)$ on a pair sets, which is useful in the context of adding sets: For any $e \in A$ we let $B_e := \{b \in B : \ b + e \not\in A\}$ and then

$$\delta_e(A) := A \cup (e + B) = A \cup (e + B_e) \quad \text{and} \quad \delta_e(B) := B \setminus B_e.$$

Notice that $B_e \subseteq B$ and $B_e \cap A = \emptyset$; and $e + \delta_e(B) \subseteq A \subseteq \delta_e(A)$. It is useful in the context of adding sets for two reasons:

- We have $A \cap (e + B) = \delta_e(A) \cap (e + \delta_e(B))$ and $A \cup (e + B) = \delta_e(A) \cup (e + \delta_e(B))$;
- Also $\delta_e(A) + \delta_e(B) \subseteq A + B$.

To see this last assertion note that if $a \in \delta_e(A), b \in \delta_e(B) \subseteq B$ then either $a \in A$ (in which case $a + b \in A + B$ trivially) or $a \in e + B_e$, that is $a = e + b'$ with $b' \in B_e$: however we then have $e + b \in A$ (as $b \in \delta_e(B)$), say $e + b = a' \in A$, and so $a + b = (e + b') + b = (e + b) + b' = a' + b' \in A + B$.

With this tool we can prove Mann's improvement of Schnirelmann's theorem and other central results in this subject.

Mann's theorem. *If* $0 \in A \cap B$ *then* $\sigma(A + B) \geq \min\{1, \sigma(A) + \sigma(B)\}$.

This follows immediately from the stronger

Proposition 2.4. *If* $0 \in A \cap B$ *then* $\frac{(A+B)(n)}{n} \geq \min\left\{1, \min_{1 \leq m \leq n} \frac{A(m)+B(m)}{m}\right\}$.

Proof. Let $\eta := \min_{1 \leq m \leq n}(A(m)+B(m))/m$. The result follows from the proof of Lemma 2.1 if $\eta \geq 1$, so we may assume that $\eta < 1$. It will be convenient in this proof to suppose that $A, B \subseteq [0, n]$, wlog. There is nothing to prove if $n = 0$ or $B(n) = 0$ or $A(n) = 0$. So we will proceed by induction on n and then on $B(n)$, and we may now assume that $n, B(n) \geq 1$. We select e minimal so that $e + B$ is not a subset of A (evidently e exists, for we may simply take consider e and b to be the largest elements of A and B respectively, so that $e + b \not\in A$). Note that $B_e(n) \geq 1$, so that $\delta_e(B)(n) < B(n)$ and thus we may proceed by the induction hypothesis as $(A+B)(n) \geq (\delta_e(A)+\delta_e(B))(n)$ (since $\delta_e(A)+\delta_e(B) \subseteq A+B$), once we prove that $\delta_e(A)(m) + \delta_e(B)(m) \geq \eta m$ for all $1 \leq m \leq n$.

Now $\{b \in B_e : \ m - e < b \leq m\} \subseteq \{b \in B : \ m - e < b \leq m\}$ so we have

$$\begin{aligned} \delta_e(A)(m) + \delta_e(B)(m) &= (A(m) + B_e(m - e)) + (B(m) - B_e(m)) \\ &= A(m) + B(m) - (B_e(m) - B_e(m - e)) \\ &\geq A(m) + B(m) - (B(m) - B(m - e)) = A(m) + B(m - e). \end{aligned}$$

If $B(m) = B(m - e)$ then we are done, for example in the case $e = 0$. Otherwise let b_1 be the smallest element of B which is $> m - e$, so that $b_1 \leq m$, and let $0 \leq r := m - b_1 \leq e - 1$ $(\leq n - 1)$. Then $A(m) + B(m - e) = A(m) + B(b_1 - 1) = (A(b_1 + r) - A(b_1 - 1)) + (A(b_1 - 1) + B(b_1 - 1))$. Now $A(b_1 - 1) + B(b_1 - 1) \geq \eta(b_1 - 1)$ by hypothesis. Now if $a \in A$ with $a \leq r < e$ then $a + B \subset A$ so that $b_1 + a \in A$: therefore $A(b_1 + r) - A(b_1 - 1) \geq A(r) + 1$, and $A(r) = (A+B)(r)$ since every element of $A+B$ which is $< e$ must belong to A (by the definition of e). Putting this altogether, and using the induction hypothesis on n to note that $(A+B)(r) \geq \eta r$, we have $A(m) + B(m - e) \geq \eta r + 1 + \eta(b_1 - 1) = \eta m + (1 - \eta) \geq \eta m$.

Dyson generalization. *If $0 \in A_1 \cap A_2 \cap \cdots \cap A_k$ then*

$$\frac{(A_1 + A_2 + \cdots + A_k)(n)}{n} \geq \min \left\{ 1, \min_{1 \leq m \leq n} \frac{A_1(m) + A_2(m) + \cdots + A_k(m)}{m} \right\}.$$

A key question is how these results generalize to $\underline{d}$: one might guess that $\underline{d}(A+B) \geq \min\{1, \underline{d}(A) + \underline{d}(B)\}$, but this is wrong: The example $A = B = A+B = \{n \equiv 0 \pmod m\}$ shows that subgroups must be taken into consideration. The example $A = B = \{n \equiv 0$ or $1 \pmod m\}$ and $A+B = \{n \equiv 0, 1$ or $2 \pmod m\}$ shows that cosets of subgroups must also be taken into consideration. Evidently we first need to study addition of sets $\pmod m$, before going on to $\underline{d}$.

The Cauchy-Davenport theorem. *If A and B are non-empty subsets of $\mathbb{Z}/N\mathbb{Z}$ with $0 \in B$, and where $(b, N) = 1$ for all $b \in B \setminus \{0\}$ then $|A+B| \geq \min\{N, |A| + |B| - 1\}$.*

Proof. We need only prove this when $|A| + |B| - 1 \leq N$, for if $|A| + |B| - 1$ is larger then we simply take subsets $A' \subseteq A$ and $B' \subseteq B$ with $|A'| + |B'| - 1 = N$ and then $\mathbb{Z}/N\mathbb{Z} \supseteq A + B \supseteq A' + B' \supseteq \mathbb{Z}/N\mathbb{Z}$.

Proof by induction on $|B|$: if $|B| = 1$ then $B = \{0\}$ so $A+B = A$ and the result follows. For $|B| \geq 2$ select $b \in B \setminus \{0\}$. We claim that there exists $e \in A$ such that $e + B \not\subset A$ else $A + B = A$ and then summing the solutions of $a + b = a'$ over all $a \in A$ (and thus a' runs through the elements of A), we obtain $|A|b \equiv 0 \pmod N$. Now $(b, N) = 1$ by hypothesis and so $N||A|$ which is impossible since $1 \leq |A| \leq N - 1$.

The result holds for the pair $\delta_e(A), \delta_e(B)$ by the induction hypothesis (as $|\delta_e(B)| < |B|$), and then the result holds for the pair A, B by the properties of the Dyson-transform (verify this).

At first sight it would seem we could significantly weaken the hypothesis on B in the proof above to something like $(N, b_1, \ldots, b_r) = 1$ where $B = \{0, b_1, \ldots, b_r\}$. Explain why the proof fails in this situation.

Corollary 2.5. *If A and B are non-empty subsets of $\mathbb{Z}/p\mathbb{Z}$ where p is prime, then $|A + B| \geq \min\{p, |A| + |B| - 1\}$.*

Proof. Determine how to satisfy the hypotheses of the Cauchy-Davenport theorem in this case.

In fact one can prove that $|A + B| = |A| + |B| - 1 < p$ if and only if either
(i) A or B has just one element; or
(ii) $A = a_0 + d \diamond \{0, 1, \ldots, r-1\}, B = b_0 + d \diamond \{0, 1, \ldots, s-1\}$ for some $r + s \leq p - 1$; or
(iii) $A \cup (d - B)$ is a partition of $\mathbb{Z}/p\mathbb{Z}$ for some integer d.

In the proof above we see that if $|\delta_e(A) + \delta_e(B)| \geq |\delta_e(A)| + |\delta_e(B)|$ then $|A + B| \geq |A| + |B|$. Now if $|B| = 1$ then, trivially

Summing sets to get all of the integers

Proposition 2.6. *If A is an additive basis of order h then $A(n) \gg hn^{1/h}$. On the other hand there exists an additive basis A of order h with $A(n) \ll hn^{1/h}$.*

Proof. Suppose that $hA \supseteq \mathbb{Z}_{\geq m}$. Then $n + O(1) \leq hA(n) \leq A(n)^h/h! + O(A(n)^{h-1})$, and the first result follows from Stirling's formula. On the other hand one can show that

$$A = \bigcup_{j=0}^{h-1} (2^j \diamond B) \ \text{ where } \ B = \left\{ \sum_{i=1}^{\ell} 2^{e_i h} : \ 0 \leq e_1 < e_2 < \cdots < e_\ell, \ell \geq 0 \right\}$$

is a basis of order h (Fill in the details here as an exercise. Hint: Consider representing $2^h - 1$ as a sum of elements of A). Now if $2^{(k-1)h} \leq n \leq 2^{kh} - 1$ then $B(n) \leq B(2^{kh} - 1) = 2^k \leq 2n^{1/h}$ and so $A(n) \leq \sum_{j=0}^{h-1} 2(2^{-j}n)^{1/h} \leq 2(1 - 2^{-1/h})^{-1} n^{1/h} \ll hn^{1/h}$. Determine an asymptotic formula for the value of $B(n)$ and then of $A(n)$.

Evidently $(h-1)A$ in this example is nowhere close to being all of $\mathbb{Z}$. One might guess that bases are, in some sense, "complementary"; in that of you added enough together you would obtain $\mathbb{Z}$. Nothing could be further from the truth, as the following generalization of the above construction shows.

Cute example: For given integers $h \geq 2$ and k, we now construct additive bases $B_1, \ldots B_k$ of order h such that if $T = (h-1)(B_1 + B_2 + \cdots + B_k)$ then $T(n) \ll_{h,k} n^{1-1/(2h^k)}$:
Fix integer $g \geq (k(h-1))^2$and for any $S \subset \mathbb{Z}_{\geq 0}$ define $G(S) = \{\sum_{j \in S} e_j g^j : \ 0 \leq e_j \leq g-1\}$. Note that if $A \cap B = \emptyset$ then $G(A \cup B) = G(A) + G(B)$.

If $n = \sum_i n_i h^i$ in base n then let $N_{i,j}$ be the set of non-negative integers n with $n_i = j$, and note that $\cup_{j=0}^{h-1} N_{i,j}$ is a partition of $\mathbb{Z}_{\geq 0}$. Then define

$$B_i = \bigcup_{j=0}^{h-1} G(N_{i,j}), \ \text{ so that } (h-1)B_i = \bigcup_{j_i=0}^{h-1} (h-1)W(j_i) \ \text{ where } W(j_i) := \bigcup_{\substack{j=0 \\ j \neq j_i}}^{h-1} G(N_{i,j});$$

and therefore $T = \cup (h-1)(W(j_0) + W(j_1) + \cdots + W(j_{k-1}))$ where the union is taken over $0 \leq j_0, \ldots, j_{k-1} \leq h-1$. Fix $j_0, \ldots, j_{k-1}$ and write $U = (h-1)(W(j_0) + W(j_1) + \cdots + W(j_{k-1}))$, Now if $e \equiv j_0 + j_1 h + \cdots + j_{k-1} h^{k-1} \pmod{h^k}$ then $e \notin \cup_{j \neq j_i} N_{i,j}$ for any i, and so the least residue $\pmod{g^{e+1}}$ of an element of $W(j_1)$ is $\leq g^e - 1$. Therefore the least residue $\pmod{g^{e+1}}$ of an element of U is $\leq k(h-1)(g^e - 1)$, and so the coefficient of g^e in the base g expansion is $\leq k(h-1) - 1$. Therefore if $g^{(m-1)h^k} < n \leq g^{mh^k}$ for some integer $m \geq 1$ then

$$U(n) \leq U(g^{mh^k}) \leq (k(h-1)g^{h^k - 1})^m \leq g^{m(h^k - 1/2)} \leq g^{h^k - 1/2} n^{1-1/(2h^k)};$$

and then $T(n) \leq h^k g^{h^k - 1/2} n^{1-1/(2h^k)}$, as required.

It is often difficult, for given sets A and B, to determine whether $A+B=\mathbb{Z}$? (for example if $A=\{0\}\cup\{(p-1)/2 : \ p\in\mathbb{P}_{\geq 3})$. Or, for a given set A one might wish to find "thin" sets B for which $A+B=\mathbb{Z}$ (for example where $A=\{0\}\cup\mathbb{P}$).

An *essential component* is a set B such that if $1>\sigma(A)>0$ then $\sigma(A+B)>\sigma(A)$. Khintchin showed in 1933 that $\mathbb{Z}^2$ is an essential component, and Erdős and Landau showed that they do turn out to be more common than one might expect.

We shall consider how small a set B one can add to A to guarantee that $A+B\supseteq\mathbb{Z}_{\geq m}$.

Lemma 2.7. *Let $a_0\geq 0$ be the smallest element of A, and suppose that $n+1\geq m+a_0$. Then there exists a subset S of $[m,2n-1]$ such that $\{n+1,n+2,\dots,2n\}\subseteq A+S$ with $|S|\ll n\log(2A(n-m+1))/A(n-m+1)$.*

With this key lemma we deduce Lorentz's theorem:

Theorem 2.8. *If $0\in A$ then there exists $B\subset\mathbb{Z}_{\geq 0}$ for which $A+B=\mathbb{Z}_{\geq 0}$ and $B(n)\ll a_1+\sum_{n>m\geq 1}\log(A(m))/A(m)$, where a_1 is the smallest element of $A_{\geq 1}$.*

Deduce that there exists B with $B+\mathbb{P}=\mathbb{Z}_{\geq 2}$ and $B(n)\ll\log^3 n$. Also that there exists C with $C+\mathbb{Z}^2=\mathbb{Z}_{\geq 0}$ with $C(n)\ll\sqrt{n}\log n$.

We can also deduce a "mod n version:

Theorem 2.9. *For any subset A of $\mathbb{Z}/N\mathbb{Z}$ there exists $B\subseteq\mathbb{Z}/N\mathbb{Z}$ for which $A+B=\mathbb{Z}/N\mathbb{Z}$ and $|B|\ll N\log(2|A|)/|A|$.*

Note that $|A||B|\ll N\log N$ (obviously one needs $|A||B|\geq N$; I have no idea whether one can improve on this log factor).

Proof of Theorem 2.9. Represent A as a subset of $\{1,\dots,n\}$ and then apply Lemma 2.7 with $m=1$ and $B=S$.

Proof of Theorem 2.8. For any $j\geq 2$, we can take $n=2^j$ and $m=2^{j-1}+1$ in Lemma 2.7 to obtain $B_j\subseteq(2^{j-1},2^{j+1})$ such that $\{2^j+1,2^j+2,\dots,2^{j+1}\}\subseteq A+B_j$ and $|B_j|\ll 2^j\log(2A(2^{j-1}))/A(2^{j-1})\ll\sum_{i=2^{j-2}}^{2^{j-1}}\log(A(i))/A(i)$ as $\log(A(i))/A(i)$ is a decreasing function for $A(i)\geq 3$. We take $B=\{0,1,\dots,a_1\}$ together with the B_j for all j with $2^{j+1}\geq a_1$.

Proof of Lemma 2.7. (Greedy) Let $I_0=\{n+1,n+2,\dots,2n\}$. Given $I_j\subset I_0$ we select integer $s_j\in[m,2n-1]$ so that $J_j:=(A+s_j)\cap I_j$ is maximal, and then let $I_{j+1}=I_j\setminus J_j$. Note that if $i\in I_0$ and $s\in[m,2n-1]$ then $i-s\geq n+1-m$ and so

$$\begin{aligned}(2n-m)|J_j|\geq\sum_{s=m}^{2n-1}|(A+s)\cap I_j| &= \sum_{i\in I_j}|A\cap\{i-m,\dots,i-(2n-1)\}|\\ &=\sum_{i\in I_j}A(i-m)\geq|I_j|A(n+1-m).\end{aligned}$$

This implies that $|I_{j+1}|=|I_j|-|J_j|\leq|I_j|(1-A(n+1-m)/(2n-m))$. Select k to be the smallest integer $>(2n-m)\log(2A(n+1-m))/A(n+1-m)$, so that $|I_k|\ll n/A(n+1-m)$. We let $S=\{s_1,s_2,\dots,s_k\}\cup(I_k-a_0)$, and the result follows.

Constructions by probabilistic methods!?

Tilings: Discuss

Covering congruences. One can sometimes show that $A+B \neq \mathbb{Z}$ in spectacular fashion by showing that $A+B$ misses a complete arithmetic progression; that is $(A+B) \cap (c + N \diamond \mathbb{Z}_{\geq m}) = \emptyset$. In the case that $A = \mathbb{P}$ we may be able to write $B = B_1 \cup B_2 \cup \dots B_k$ such that $B_j \subset c_j + p_j \diamond \mathbb{Z}$, for certain distinct primes $p_1, \dots, p_k$. In this case $(A+B) \cap (c + p_1 \dots p_k \diamond \mathbb{Z}) \subset \{p_1, \dots, p_k\}$ where c is chosen so that $c \equiv c_j \pmod{p_j}$ for each j by the Chinese Remainder theorem, for if $a + b \in c + p_1 \dots p_k \diamond \mathbb{Z}$ with $b \in b_j$, say, then $p_j | a$ so that $a = p_j$ since $A = \mathbb{P}$.

Erdős invented this idea to show that a certain congruence class of odd integers cannot be written in the form $p + 2^k$ with $p \in \mathbb{P}$. To develop this proof yourself consider writing the set $B = \{2^k : k \geq 1\}$ as $B_2 \cup B_3 \cup B_4 \cup B_8 \cup B_{12} \cup B_{24}$ where $B_m = \{2^k : k \equiv c_m \pmod{m}\}$ with $c_2 = 0, c_3 = 0, c_4 = 1, c_8 = 3, c_{12} = 7, c_{24} = 23$, and go from there.

Kneser's Theorem. *If A and B are finite subsets of additive group Z with $|A|, |B| > 1$ for which $|A+B| < |A| + |B|$ then let H be the largest subgroup of Z for which $A+B$ is a union of cosets of H (note that such a subgroup always exists, namely $H = \{0\}$; and also that H can possibly be all of Z). Let $A_0 \subset A$ and $B_0 \subset B$ be minimal so that $A \subset A_0 + H$ and $B \subset B_0 + H$ (and therefore $A + B = A_0 + B_0 + H$). Then $|A_0 + B_0| = |A_0| + |B_0| - 1$ in Z/H, and if $A^* = (A_0 + H) \setminus A$ and $B^* = (B_0 + H) \setminus B$ then $|A^*| + |B^*| \leq |H| - 1$.*

It is worth noting that if we have $|A^*| + |B^*| \leq |H| - 1$ as above then we must have $A + B = A + B + H$ (that is, $A + B$ is a union of cosets of H), for if $a + b \in A + B$ then $|A \cap (a + H)| + |B \cap (b + H)| \geq 2|H| - |A^*| - |B^*| \geq |H| + 1$ and so for any $c \in a + b + H$ we obtain $c \in A + B$ by the pigeonhole principle (as in the proof of Lemma 2.1).

Let $r(n) = r_{A+B}(n)$ be the number of representations of n in the form $a + b, a \in A, b \in B$. By another application of the pigeonhole principle one can show that for each $n \in A + B$,

$$r_{A+B}(n) \geq |A| + |B| - |A + B|.$$

To prove Kneser's theorem we will need to develop the theory of Dyson transformations: We saw above that the transformation $A, B \to \delta_e(A), \delta_e(B)$ has the properties that $\delta_e(B) \subset B$, $A \subset \delta_e(A)$ with $|\delta_e(A)| + |\delta_e(B)| = |A| + |B|$ and $\delta_e(A) + \delta_e(B) \subseteq A + B$. We may assume, wlog, that $0 \in B$ (after a translation), and then $0 \in \delta_e(B)$. A non-trivial transformation exists unless $A + B \subset A$ (so that all $B_e = \emptyset$), in which case $A + H \subset A$ where $H =< B >$ is the semigroup generated by B (and if Z is finite then H is a subgroup of Z). If we start with any pair of sets A, B we can go through a "derived" sequence of Dyson transformations $A, B \to^{e_1} A_1, B_1 \to^{e_2} \cdots \to^{e_r} A_r, B_r$, and from the above properties we have that $0 \in B_r \subset B$ and $A_r \supset A$ with $|A_r| + |B_r| = |A| + |B|$ and $A_r + B_r \subseteq A + B$. If B is finite then this can continue for only finitely many steps (since $|B_r| \leq |B| - r$), and then we must have that $A_r + H \subset A_r$ where $H =< B_r >$. If A is also finite then $A_r + H = A_r$.

We also need a little technical lemma, for which we give a frustratingly long proof at the end of these notes. A keen student is invited to supply me with a short proof!

Lemma 2.10. *Let H_1, H_2 be finite subgroups of Z with $H_1 \cap H_2 = \{0\}$. Let C_j be a non-empty finite set of coset representatives for H_j, for $j = 1, 2$, and let $S_j = C_j + H_j$. Then either $S_1 \cup S_2 = S_1$ or S_2, or $|S_1 \cup S_2| \geq \min_{j=1,2} |S_j| + |H_j| - 1$.*

Proof of Kneser's Theorem. We will prove, by induction on $|T| \geq 1$, that for any non-empty subset $T \subset A + B$ there exists C, a finite set of coset representatives for some subgroup H, such that $T \subset C + H \subset A + B$, where $|A| + |B| \leq (|C| + 1)|H|$. Once this is proved then we take $T = A + B$ so that $A + B = C + H'$, and therefore $|A + B| \geq |A| + |B| - |H'|$ (since $|A + B| = |C||H'|$). Now select H to be the largest subgroup of Z for which $A + B$ takes the form $C' + H$, and note that $|A + B| \geq |A| + |B| - |H'| \geq |A| + |B| - |H|$. Now $(A + H) + (B + H) = A + B$, so we get the same largest subgroup, H, when we add $A + H$ and $B + H$, and then the inequality above is $|A + B| \geq |A + H| + |B + H| - |H|$. Writing $A + H = A_0 + H, B + H = B_0 + H, A + B = C_0 + H$ where A_0, B_0, C_0 are minimal such sets, then we see that $|C_0||H| = |A + B| \geq |A + H| + |B + H| - |H| = (|A_0| + |B_0| - 1)|H|$. Therefore $|C_0| \geq |A_0| + |B_0| - 1$: we may assume that $|C_0| = |A_0| + |B_0| - 1$, for if $|C_0| \geq |A_0| + |B_0|$ then $|A + B| = |C_0||H| \geq (|A_0| + |B_0|)|H| \geq |A + H| + |B + H| \geq |A| + |B|$. Thus $|A_0 + B_0| = |A_0| + |B_0| - 1$ in Z/H as claimed, and finally note that

$|A+B|=|H|(|A_0|+|B_0|-1)=|A|+|B|+|A^*|+|B^*|-|H|$ and so $|A^*|+|B^*|<|H|$, when $|A+B|<|A|+|B|$. Now to prove our claim:

For $T=\{a_0+b_0\}$ take $A,B\to A-a_0,B-b_0$ so, wlog, $T=\{0+0\}$ and $0\in A\cap B$. We now go through a derived sequence of Dyson transformations on the pair A,B to end up with a pair of sets A',B' for which $0\in B'\subset B$ and $|A'|+|B'|=|A|+|B|$ with $0\in A\subset A'=A'+B'=A'+H\subseteq A+B$ where $H=<B'>$. The result follows by taking C minimal so that $C+H=A'+H=A'$, giving that $|A|+|B|=|A'|+|B'|\le|C+H|+|H|=(|C|+1)|H|$, with equality if and only if $B'=H$.

For $|T|\ge 2$ we partition $T=T_1\cup T_2$ with $|T_1|,|T_2|\ge 1$ and apply the induction hypothesis to each part, obtaining $T_j\subset C_j+G_j\subset A+B$, where $|A|+|B|\le(|C_j|+1)|G_j|$ for $j=1,2$. If $C_1+G_1\subset C_2+G_2$ then we simply take $C=C_2$ and $H=G_2$ (and similarly if $C_2+G_2\subset C_1+G_1$). Otherwise let $H=G_1\cap G_2$ and $H_j=G_j/H$ for $j=1,2$. Put $C=(C_1+H_1)\cup(C_2+H_2)$ so that $C+H=(C_1+G_1)\cup(C_2+G_2)\supset T_1\cup T_2=T$. By Lemma 2.10 we have that $|C|\ge\min_{j=1,2}|C_j+H_j|+|H_j|-1$; and so for that value of j we have $|A|+|B|\le(|C_j|+1)|G_j|=(|C_j+H_j|+|H_j|)|H|\le(|C|+1)|H|$ as required.

Remarks: It would be good to have a result pertaining to the structure of A_0 and B_0. In particular if $\{0\}$ is the largest subgroup H of Z for which $A+B=A+B+H$ and $|A+B|=|A|+|B|-1$ then we should be able to classify the possible structures for A and B. Perhaps the classification is analagous to what one finds for $\mathbb{Z}/p\mathbb{Z}$, as in just after Corollary 2.5 above.

It would be good to have a direct proof of Kneser's theorem, using little more than induction and Dyson transformations, but I have not yet found out how to do this (but seem to be close).

Adding finite sets

Obviously, in any additive group G one has that $\max\{|A|,|B|\}\le|A+B|\le|A||B|$ and we begin by examining when these bounds can be attained. The reader should prove the following two results:

Proposition 6.1. *The following statements are equivalent: (i) $|A+B|=|A|$; (ii) $|A-B|=|A|$; (iii) $|A+mB-nB|=|A|$ for all $m,n\in\mathbb{Z}_{\ge 0}$; (iv) There exists a finite subgroup H of Z such that $B\subset b+H$ and $A=\bigcup_{i=1}^k(a_i+H)$, for certain $b,a_1,\dots,a_k\in Z$, with the a_i belonging to different cosets of H.*

It is useful to define $H(A):=\{h\in Z:\ h+A=A\}$, so that $A+H(A)=A$. Note that $H(A)$ must be a subgroup of Z.

Proposition 6.2. *The following statements are equivalent: (i) $|A+B|=|A||B|$; (ii) $|A-B|=|A||B|$; (iii) $(A-A)\cap(B-B)=\emptyset$.*

Show that if $|A|+|B|>|Z|$ then $Z=A+B=A-B$. Show that if $|A|+|B|=|Z|$ then it is possible that $A+B\neq Z$.

In general $A+B$ and $A-B$ do not have the same size:
Example: For $A=\{0,1,3\}$ we have $A+A=\{0,1,2,3,4,6\}$ and $A-A=\{-3,-2,-1,0,1,2,3\}$. Thus $|A+A|=6<|A-A|=7$. Note that we can take the *Cartesian product* of

any sets and consider the same questions. Thus for $B = A^{(k)} = \{0,1,3\}^k$ we have $|B+B| = 6^k < |B-B| = 7^k$.

Example: For $A = \{0,1,3,4,9,10,11,12,13,14,17,20,21,22\}$ we have all 45 integers in $[0,44]$ belong to $A+A$; however all 45 integers in $[-22,22]$ except ± 15 belong to $A-A$, and thus $|A-A| = 43 < |A+A| = 45$. Similarly if $B = A^{(k)}$ then $|B-B| = 43^k < |B+B| = 45^k$.

In the Freiman-Ruzsa theorem we are interested in the size of $A+A$ compared to A, so we define the *doubling constant* $D(A) := |2A|/|A|$. Note that $1 \le D(A) \le (|A|+1)/2$. The upper bound is attained exactly for Sidon sets.

Ruzsa defined a rather clever notion of distance:

$$d(A,B) = \log\left(\frac{|A-B|}{\sqrt{|A||B|}}\right).$$

Although the *Ruzsa distance* between two sets is not truly a distance function, since $d(A,A) \neq 0$ in general, it does have several desirable properties:

- Positivity: $d(A,B) \ge 0$ (since $|A-B| \ge \max\{|A|,|B|\}$)
- Symmetry: $d(A,B) = d(B,A)$
- Triangle inequality: $d(A,C) \le d(A,B) + d(B,C)$

We now prove that

$$|A-C||B| \le |A-B||B-C| \tag{6.0}$$

from which the triangle inequality follows. We do this by constructing an injection $(A-C)\times B \to (A-B)\times(B-C)$: For every element of $d \in A-C$ select a unique pair $a = a_d \in A, c = c_d \in C$ for which $d = a-c$. For each $b \in B$ we create the pair $(a-b, b-c) \in (A-B)\times(B-C)$. Now for any $(u,v) \in (A-B)\times(B-C)$ in the image of $(A-C)\times B$ we must have $d = a_d - c_d = (a_d - b) + (b - c_d) = u+v$ determined, and thus a_d and c_d, and therefore $b = a_d - u$.

Prove that $d(A,B) = 0$ puts us in the situation of Proposition 6.1.

We define the *Ruzsa diameter* $d(A,A) = \log(|A-A|/|A|)$; we also have $d(A,-A) = \log(D(A))$. By the triangle inequality we have $d(A,A) \le 2d(A,B)$ for any sets A,B or, in other words,

$$|A-A| \le |A-B|^2/|B|.$$

In particular $d(A,A) \le 2d(A,-A)$; that is $|A-A| \le |A+A|^2/|A|$.

Suppose that $r,s \ge 0$ are integers. By (6.0) we have, writing $(r+s)A = rA - (-sA)$, that $|(r+s)A||-A| \le |(r+1)A||sA-A|$; and that $|rA-sA||-A| \le |(r+1)A||(s+1)A|$. We deduce that

$$\frac{|(r+s)A|}{|A|} \le \frac{|(r+1)A|}{|A|}\,\frac{|(s+1)A|}{|A|}\,\frac{|2A|}{|A|};$$

and then, from a simple induction hypothesis, that

$$\frac{|nA|}{|A|} \le \left(\frac{|3A|}{|A|}\right)^{n-2}\left(\frac{|2A|}{|A|}\right)^{n-3} \quad \text{for all } n \ge 3.$$

Since $|2A|\leq |3A|$ we deduce that $|nA|/|A|\leq (|3A|/|A|)^{2n-5}$; that is, if $3A$ is small then so is nA for every $n\geq 2$. Deduce similar bounds for $|rA-sA|/|A|$ when $r,s\geq 2$, and then for all $r,s\geq 0$.

Our objective is to prove a similar result with '$2A$' in place of '$3A$'. To do this we will need to bound $|3A|/|A|$ in terms of $|2A|/|A|$, something we will postpone for a while. However since $|3A|/|A|\leq (|2A-A|/|A|)(|2A|/|A|)\leq (|2A-A|/|A|)^2$, and $|2A-A|/|A|\leq (|3A|/|A|)(|2A|/|A|)\leq (|3A|/|A|)^2$ we could just as well bound things in terms of $|2A-A|/|A|$. Above we saw that $|A-A|/|A|\leq (|2A|/|A|)^2$, so one objective is to bound $|2A|/|A|$ in terms of $|A-A|/|A|$.

Lemma 6.3. *(i) $d(A,B)=d(A+x,B)=d(-A,-B)$*
(ii) $d(A\times A',B\times B')\leq d(A,B)+d(A',B')$.
(iii) $d(A,B\cup B')\leq \max\{d(A,B),d(A,B')\}+\log 2$.
(iv) If $C\subset B$ then $d(A,C)\leq d(A,B)+\frac12\log(|B|/|C|)$.
(v) $d(A,B)-\frac12\log(|C||D|)\leq d(A+C,B+D)\leq d(A,B)-\frac12\log(|C+D|)$.

Proof. Homework

We define the *additive energy* between two sets A,B to be

$$E(A,B)=\#\{a+b=a'+b':\ a,a'\in A,\ b,b'\in B\}.$$

Note that $|A||B|\leq E(A,B)\leq |A||B|\min\{|A|,|B|\}$ with $E(A,B)=|A||B|$ if and only if we are in the situation of Proposition 6.2. We have $E(A,B)=E(B,A)=E(A+x,B+y)=E(A,-B)$. Simple counting arguments give $|A||B|=\sum_x r_{A+B}(x)=\sum_y r_{A-B}(y)$ and that

$$E(A,B)=\sum_x r_{A+B}(x)^2=\sum_y r_{A-B}(y)^2=\sum_z r_{A-A}(z)r_{B-B}(z). \tag{6.1}$$

We can therefore deduce, using the Cauchy-Schwarz inequality,

$$(|A||B|)^2=\left(\sum_x r_{A+B}(x)\right)^2\leq |A+B|E(A,B); \tag{6.2}$$

and that
(6.3)

$$E(A,B)^2=\left(\sum_z r_{A-A}(z)r_{B-B}(z)\right)^2\leq \sum_z r_{A-A}(z)^2\sum_z r_{B-B}(z)^2=E(A,A)E(B,B).$$

Lemma 6.3. $\max_x r_{A+B}(x)\geq \frac{|A||B|}{|A+B|}\max\left\{1,\left(\frac{|A+B|}{|A-B|}\right)^{1/2}\right\}$.

Proof. $|A+B|\max_x r_{A+B}(x)^2\geq \sum_x r_{A+B}(x)^2=E(A,B)$ by (6.1). Then $E(A,B)\geq (|A||B|)^2/\min\{|A+B|,|A-B|\}$ by (6.2), and the result follows.

Theorem 6.4. *$d(A,-B) \le 5d(A,B)$.*

Proof. Select x_1 so as to maximize $r_{A+B}(x)$ and then replace B by $B - x_1$. Notice that $r_{A+B}(0) = |-A \cap B|$, and $|A||B|/|-A\cap B|^2 \le \exp(d(A,B) + d(A,-B))$ by Lemma 6.3. Therefore, using the triangle inequality and Lemma 6.3(iv), we then obtain

$$\begin{aligned} d(A,-B) &\le d(A, -A\cap B) + d(-A\cap B, -B) \\ &\le d(A,B) + \frac{1}{2}\log\left(\frac{|B|}{|-A\cap B|}\right) + d(-A,-B) + \frac{1}{2}\log\left(\frac{|A|}{|-A\cap B|}\right) \\ &\le \frac{5}{2}d(A,B) + \frac{1}{2}d(A,-B), \end{aligned}$$

and the result follows.

Combining this with an earlier observation we deduce the key result that A has a small doubling constant if and only if A has a small Ruzsa diameter:

Corollary 6.5. *$\frac{1}{2}d(A,A) \le d(A,-A) \le 5d(A,A)$; that is, $|2A|/|A| \le (|A-A|/|A|)^5$.*

An essential notion in this area is to look not only at the set of all sums $A+B$ but also at a well-chosen subset:

Lemma 6.6. *There exists $S \subset A+B$ such that $\#\{a \in A, b \in B :\ a+b \in S\} \ge |A||B|/2$ with $|S| \ge \max\{|A|,|B|\}/2$ and for which $|A+B+nS| \le 2^n|A+B|^{2n+1}/|A|^n|B|^n$ for all $n \ge 0$.*

Proof. Let $S = \{s : r_{A+B}(s) \ge |A||B|/2|A+B|\}$ and prove the first two assertions. For any $c \in A+B+nS$ there exists $a_0 \in A, b_{n+1} \in B$ and $s_1,\dots,s_n \in S$ such that $c = a_0 + b_{n+1} + s_1 + \cdots + s_n$. Now for any given s_j there exists at least $|A||B|/2|A+B|$ solutions to $a_j + b_j = s_j$ and thus at least $(|A||B|/2|A+B|)^n$ sets $a_1,\dots,a_n \in A,\ b_1,\dots,b_n \in B$ with $a_1+b_1 = s_1,\dots,a_n+b_n = s_n$. Each such solution leads to an element $(t_1,\dots,t_{n+1}) \in (A+B)^{n+1}$ defined by $t_i = a_{i-1} + b_i$: we claim that these are distinct since we can recover $a_1,\dots,a_n,\ b_1,\dots,b_n$ given $a_0, b_{n+1}, s_1,\dots,s_n, t_1,\dots,t_{n+1}$. Therefore $(|A||B|/2|A+B|)^n|A+B+nS| \le |A+B|^{n+1}$ and the result follows.

Ruzsa's Covering Lemma. *$B \subset A - A + X$ for some $X \subset B$ with $|X| \le |A+B|/|A|$.*

Remark: Note that we are covering B by translates of $A-A$.

Proof. Choose $X \subset B$ maximal so that $\{A+x : x \in X\}$ are disjoint. Their union contains exactly $|A||X|$ elements, all inside $A+B$, and thus the bound on $|X|$. Now, if $b \in B$ then $A+b$ intersects $A+x$ for some $x \in X$, and so $b \in A-A+x$, and thus the result.

Corollary 6.7. *We have $|mA-nA|/|A| \le (|A-A|/|A|)(|2A-2A|/|A|)^{m+n-2}$, for any $m,n \ge 1$.*

Proof. Take $B = A - 2A$ in Ruzsa's covering lemma to get $2A - A \subset A - A + X$ for $X \subset 2A - A$ with $|X| \le |2A-2A|/|A|$. But then, adding A to both sides we obtain $3A - A \subset 2A - A + X \subset A - A + 2X$, and by induction $mA - nA \subset A - A + (m-1)X - (n-1)X$ for all $m,n \ge 1$. But this implies that $|mA-nA| \le |A-A||(m-1)X||(n-1)X| \le |A-A||X|^{m+n-2}$ and the result follows.

Corollary 6.8. *Let $< A >$ be the subgroup generated by A. There exists $X \subset 2A - A$ with $|X| \le |2A - 2A|/|A|$ such that $\langle A\rangle \subset A - A + \langle X\rangle$.*

Proof. As in the proof above we have $mA - nA \subset A - A + (m-1)X - (n-1)X$ for all $m, n \ge 1$. Now take the union of both sides over all $m, n \ge 1$ to get this result.

Green's variant. *$B \subset A - A + X$ for some $X \subset B$ with $|X| \le 2|A+B|/|A| - 1$, such that for all $b \in B$ there are $> |A|/2$ solutions to $b = a - a' + x$. In fact $B - B \subset A - A + X - X$.*

Proof. Let $X_0 = \emptyset$. We create $X_1, X_2, \dots$ by the following algorithm: Given X_j, if there exists $b \in B$ for which $|(b+A) \cap (X_j + A)| \le |A|/2$ then let $X_{j+1} = X_j \cup \{b\}$ and $b_{j+1} = b$, otherwise let $X = X_j$ and stop the algorithm. Thus for each $b \in B \setminus X$ there are $> |A|/2$ solutions to $b = a - a' + x$; and if $b \in X$ there are $\ge |A|$ solutions, namely those with $a' = a \in A$ and $x = b$.

Now $X_{j+1} + A = (X_j + A) \cup ((b_j + A) \setminus (X_j + A))$ so that $|X_{j+1} + A| = |X_j + A| + |b_j + A| - |(b_j + A) \cap (X_j + A)| \ge |X_j + A| + |A| - |A|/2 \ge (j+2)|A|/2$ by an appropriate induction hypothesis. Therefore $|B + A| \ge |X + A| \ge |A|(|X| + 1)/2$ as $X + A \subset B + A$ and we have now proved all parts of the first sentence.

Finally, for each of $b, b' \in B$ there are $> |A|/2$ solutions to $b = a - a' + x$, and also to $b' = a'' - a''' + x'$. Thus, by the pigeonhole principle, there are solutions with $a''' = a'$ so, subtracting, we obtain $b - b' = a - a'' + x - x' \in A - A + X - X$.

Theorem 6.9. $|2B - 2B| \le |A + B|^4 |A - A|/|A|^4$.

Proof. Let $z \in B - B$ so that $z = b_1 - b_2$ for some $b_1, b_2 \in B$. By Green's variant of Ruzsa's covering lemma there are $> |A|/2$ solutions (a_1, a_2, x) to $b_2 = x + a_1 - a_2$; that is $z = b_1 - b_2 = c - a_1 - x$ where $c = b_1 + a_2 \in A + B$. In other words $\#\{x \in X,\ c \in A + B,\ a_1 \in A:\ z = c - a_1 - x\} > |A|/2$ for any given $z \in B - B$. This is also true for $z' \in B - B$, so we have $z - z' = c - c' - d - x + x'$ where $d = a_1 - a_1'$. Note that is we are given z, z', c, c', d, x, x' we can recover a_1 and a_1' as $a_1 = c - x - z$ and $a_1' = c' - x' - z'$. Thus for any $z - z' \in 2B - 2B$ we have

$$\#\{x, x' \in X,\ c, c' \in A + B,\ d \in A - A:\ z - z' = c - c' - d - x + x'\} > |A|^2/4.$$

Therefore we deduce that $|2B - 2B||A|^2/4 < |X|^2|A+B|^2|A-A| < 4|A+B|^4|A-A|/|A|^2$; that is $|2B - 2B| < 16|A+B|^4|A-A|/|A|^4$.

To complete the proof we use the Cartesian product of sets. Thus, from this last inequality we deduce that, for any $k \ge 1$, we have $|2B - 2B|^k = |2B^{(k)} - 2B^{(k)}| < 16|A^{(k)} + B^{(k)}|^4 |A^{(k)} - A^{(k)}|/|A^{(k)}|^4 = 16(|A+B|^4|A-A|/|A|^4)^k$. Taking kth roots and letting $k \to \infty$ we get the result (since A and B are fixed and finite).

Take $B = A$ and $B = -A$ in Theorem 6.9, and then insert the first of these results in Corollary 6.7 using that $|A - A|/|A| \le (|2A|/|A|)^2$ to get:

Corollary 6.10. *$|2A - 2A|/|A| \le (|2A|/|A|)^4(|A - A|/|A|)$ and $|4A|/|A| \le (|A - A|/|A|)^5$. Also $|mA - nA|/|A| \le (|2A|/|A|)^{6m+6n-10}$, for any $m, n \ge 1$.*

Corollary 6.10 is good enough for our purposes. However a stronger version has been proved by Plünnecke:

Plünnecke-Ruzsa theorem. *We have $|mA - nA|/|A| \le (|2A|/|A|)^{m+n}$ for $m, n \ge 1$.*

The Hales-Jewett Theorem

In 1927 van der Waerden [20] answered a conjecture of Schur, by showing that if the natural numbers are partitioned into two sets then one set must contain arbitrarily long arithmetic progressions. One can ask to generalize this to r sets, and ask for explicit bounds: That is, for given positive integers k, r determine the least integer $W = W(k, r)$ such that no matter how the integers in $\{1, 2, \ldots, W\}$ are partitioned (or "coloured"), there is always a partition containing a k-term arithmetic progression (that is, there is always "a monochromatic k-term arithmetic progression").

The Hales-Jewett Theorem [8] provides a beautiful, highly combinatorial, way to prove this result; it can be thought of as a generalization of van der Waerden's problem. For given positive integers k, r we wish to find the least integer $d = d(k, r)$ such that if the elements of $\{1, 2, \ldots, k\}^d$ are r-colored then it contains a *monochromatic line*: For a given nonempty $S \subset \{1, 2, \ldots, k\}$ a *line* is a set of points of the form $L = \{\mathbf{x}_0 + t\mathbf{y}_S : \ t = 1, 2, \ldots k\}$ where $(\mathbf{y}_S)_i = 1$ if $i \in S$ and $(\mathbf{y}_S)_i = 0$ if $i \notin S$, and $(\mathbf{x}_0)_i = 0$ if $i \in S$. The Hales-Jewett Theorem asserts that $d(k, r)$ exists for all $k, r \geq 1$.

Show that $W(k, r) \leq k^{d(k,r)}$ by representing the integers up to $k^{d(k,r)}$ in base k, and then representing each such number by points in $\{1, 2, \ldots, k\}^{d(k,r)}$.

Shelah [14] recently gave a delightfully ingenious proof that $d(k, r)$ exists: We prove the result by induction on k. It is clear that $d(1, r) = 1$ for all r, so now suppose that $k \geq 2$ and we have a value for $d(k-1, s)$ for all $s \geq 1$. Take $M = d(k-1, r)$ and define $N_1 = r^{(k-1)^{M-1}}$ and $N_i = r^{(k-1)^{M-i}k^{N_1+\cdots+N_{i-1}}}$ for $i = 2, 3, \ldots, r$. We will prove that $d(k, r) \leq N := N_1 + N_2 + \cdots + N_M$.

An r-colouring κ of $\{1, 2, \ldots, k\}^N$ is a function $\kappa : \{1, 2, \ldots, k\}^N \to \{1, \ldots, r\}$. Our plan is to construct lines $L_1, \ldots, L_M$ with each L_j $(= \mathbf{x}_j + t_j\mathbf{y}_j) \subset \{1, 2, \ldots, k\}^{N_j}$, so that $\kappa(\mathbf{x}_1 + t_1\mathbf{y}_1, \mathbf{x}_2 + t_2\mathbf{y}_2, \ldots, \mathbf{x}_M + t_M\mathbf{y}_M)$ does not change value for any given set of values $t_1, t_2, \ldots, t_{i-1}, t_{i+1}, \ldots, t_M$ as t_i changes from $k-1$ to k. Now define a colouring of $\{1, \ldots, k-1\}^M$, so that the colour of $\kappa^*(t_1, \ldots, t_M) = \kappa(\mathbf{x}_1 + t_1\mathbf{y}_1, \ldots, \mathbf{x}_M + t_M\mathbf{y}_M)$; we know that there exists a monochromatic line $\{\mathbf{z}_0 + t\mathbf{w}_U \ : \ t = 1, 2, \ldots k-1\} \subset \{1, \ldots, k-1\}^M$ corresponding to some $U \subset \{1, \ldots, k-1\}$ by the definition of M. Define $S \subset \{1, \ldots, N\}$ to be the union of the subsets S_j, corresponding to the 1s in the vector $\mathbf{y}_j$, for each $j \in U$, yielding a monochromatic line $\{\mathbf{x}_0 + t\mathbf{y}_S \ : \ t = 1, 2, \ldots k-1\} \subset (L_1, \ldots, L_M) \subset \{1, \ldots, k-1\}^N$. By the construction of our lines $L_1, \ldots, L_M$, this implies the result.

We need to construct the lines $L_M, L_{M-1}, \ldots, L_1$ which we do in this order. In fact for, $J = M, M-1, \ldots, 1$ we assume that $L_i = \{\mathbf{x}_i + t_i\mathbf{y}_i : t_i = 1, 2, \ldots, k\}$ is given for each $J < i \leq M$ with the property that $\kappa(x, \mathbf{x}_{J+1} + t_{J+1}\mathbf{y}_{J+1}, \ldots, \mathbf{x}_M + t_M\mathbf{y}_M)$ is fixed as any t_j changes from $k-1$ to k, for any given $x \in \{1, 2, \ldots, k\}^{N_1+N_2+\cdots+N_J}$ and $j, J < j \leq M$.

For each $x_J \in \{1, 2, \ldots, k\}^{N_J}$ we define a colouring κ_{x_J} on $\{1, 2, \ldots, k\}^{N_1+N_2+\cdots+N_{J-1}} \times L_{J+1} \times \cdots \times L_M$ by $\kappa_{x_J}(y, l_{J+1}, \ldots, l_M) = \kappa(y, x_J, l_{J+1}, \ldots, l_M)$. Given that the value of κ does not change as the t_j change from $k-1$ to k, the number of vectors $(y, l_{J+1}, \ldots, l_M)$ with possibly independently chosen colourings is $\leq k^{N_1+N_2+\cdots+N_{J-1}}(k-1)^{M-J}$ and so the number of possibly different r-colourings is $\leq N_J$. There are $N_J + 1$ elements in $\{1, 2, \ldots, k\}^{N_J}$ consisting of $(k-1)$s followed by ks and so, by the pigeonhole principle, two must have the same colouring. Let us suppose that these have exactly g and h $(k-1)$s, respectively, where $g < h$. Then we define the line $L_J = \{\mathbf{x}_J + t_J\mathbf{y}_J : t_J = 1, 2, \ldots, k\}$, by taking $S_J = \{g+1, g+2, \dots, h\}$ with $(x_J)_i = k-1$ for $1 \le i \le g$ and $(x_J)_i = k$ for $h+1 \le i \le k$. From this construction it is clear that L_J has the required properties and the result follows.

This proof was a breakthrough in that it was the first to give primitive recursive bounds on the van der Waerden numbers, $W(k,r)$. The reader should deduce bounds, themselves, from the above proof; the necessary information is in the first paragraph of the proof.

Define $n_k(N)$ to be the smallest integer n such that every subset of $\{1, \dots, N\}$ of size n contains a k-term arithmetic progression. If one can get a good enough bound on $n_k(N)$ then this would obviously imply good bounds on $W(k,r)$.

Discrete Fourier transforms, I

Let $n \in \mathbb{N}$ and $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$. The *(discrete) Fourier transform* $\hat{f}$ of f is defined by

$$\hat{f}(r) = \sum_{s=0}^{N-1} f(s) e\left(\frac{rs}{N}\right),$$

where $e(t) = \exp(2i\pi t)$. This has inverse

$$f(s) = \frac{1}{N} \sum_{r=0}^{N-1} \hat{f}(r) e\left(\frac{-rs}{N}\right)$$

(Verify). The reader should verify that

$$\sum_r \hat{f}(r)\overline{\hat{g}}(r) = N \sum_r f(r)\overline{g}(r)$$

which in the special case $f = g$ gives Parseval's identity: $\sum_r |\hat{f}(r)|^2 = N \sum_r |f(r)|^2$.

The *convolution* $f * g$ *of* f *and* g is defined by

$$(f * g)(r) = \sum_{t-u=r} f(t)\overline{g(u)},$$

and the reader should verify that $\widehat{(f * g)} = \hat{f}\overline{\hat{g}}$ as well as

$$N \sum_r |(f * g)(r)|^2 = \sum_r |\hat{f}(r)|^2 |\hat{g}(r)|^2.$$

Taking $g = f$ we obtain $\sum_r |\hat{f}(r)|^4 = N \sum_{a+b=c+d} f(a)f(b)\overline{f(c)f(d)}$.

In an abuse of notation we let $A(.)$ be the characteristic function of the set A, that is $A(n) = 1$ if $n \in A$, and $A(n) = 1$ otherwise. We have used $A(x)$ to mean $\#\{a \in A : 1 \leq a \leq x\}$ above – we can only hope the reader does not get confused but reckons which definition we are using depending on context. Note that $\hat{A}(m) = \sum_{a \in A} e(am/N)$. When considering sumsets, like $A + B$, we define $r_{A+B}(n) = \#\{a \in A, b \in B : a + b = n\}$. Discrete Fourier transforms fit naturally in this context, for example $(A * B)(n) = r_{A-B}(n)$ so that, as one can verify,

$$\frac{1}{N} \sum_n |\hat{A}(n)|^2 |\hat{B}(n)|^2 = \sum_n |(A * B)(n)|^2 = E(A, B)$$

(see (6.1)). One also has that $\hat{A}(m)\hat{B}(m) = \sum_n r_{A+B}(n) e(mn/N)$ and this can be inverted to give $r_{A+B}(n) = (1/N) \sum_m \hat{A}(m)\hat{B}(m)e(-mn/N)$. A particular case of this is $r_{A-A}(n) = (1/N) \sum_m |\hat{A}(m)|^2 e(-mn/N)$, as well as $r_{kA-kA}(n) = (1/N) \sum_m |\hat{A}(m)|^{2k} e(-mn/N)$. In fact $r_{A-A}(n) = (1/N) \sum_m |\hat{A}(m)|^2 \cos(2\pi mn/N)$ since $\hat{A}(-m) = \overline{\hat{A}(m)}$.

Discrete Fourier transforms have traditionally appeared in understanding how well sets are distributed: Let $(t)_N$ denote the least non-negative residue of $t \pmod N$. We say that a set A is *uniformly distributed* $\pmod N$ if $\#\{a \in A : \alpha N < (ma)_N \leq \beta N\} \sim (\beta - \alpha)|A|$ for any $m \not\equiv 0 \pmod N$. This definition involves an asymptotic estimate "$\sim$", something we will see a lot more of. This "$\sim$" could be replaced by $\{1 + g(N)\}$ where $g(.)$ is some function for which $g(N) \to 0$ as $N \to \infty$.

This natural notion of uniform distribution fits in well with Fourier transforms:

The equidistribution theorem. *A is uniformly distributed* $\pmod N$ *if and only if $\hat{A}(m) = o(|A|)$ for all $m \not\equiv 0 \pmod N$.*

Our equidistribution theorem is the natural analogy to Weyl's famous equidistribution theorem for sequences of reals: Let $\{t\}$ be the fractional part of t (so, for example, $\{ma/N\} = (ma)_N/N)$, and suppose that $a_1, a_2, \dots$ is a given (and ordered) sequence of real numbers. Then Weyl showed that the a_j are *uniformly distributed mod one*, that is $\#\{n \leq N : \alpha < \{a_n\} \leq \beta\} \sim (\alpha - \beta)N$ as $N \to \infty$, if and only if for each integer m, we have $\sum_{n\leq N} e(ma_n) = o(N)$ as $N \to \infty$.

Show that if $a_1, a_2, \dots$ is uniformly distributed mod one, then so is $ka_1, ka_2, \dots$ for any integer $k \geq 1$.

The notion of uniform distribution surprisingly is related to the question: Do there exist solutions to $a + b = c$ with $a \in A, b \in B, c \in C$, three sets of residues $\pmod N$?

Proposition 3.1. *If A is uniformly distributed* $\pmod N$ *with $|A| \gg N$, and B, C are any other two sets mod N of size $\gg N$, then for any integers i, j, k coprime with N, and for any integer m we have that*

$$\#\{a \in A, b \in B, c \in C : ia + jb + kc \equiv m \pmod N\} \sim |A||B||C|/N.$$

Proof. We count the above set as

$$\sum_{\substack{a\in A, b\in B\\ c\in C}} \frac{1}{N}\sum_r e\left(\frac{r(ia+jb+kc-m)}{N}\right) = \frac{1}{N}\sum_r e\left(\frac{-rm}{N}\right)\hat{A}(ir)\hat{B}(jr)\hat{C}(kr).$$

The $r = 0$ term gives $|A||B||C|/N$. We regard the remaining terms as error terms, and bound them by their absolute values, giving a contribution

$$\leq \frac{1}{N}\max_{s\neq 0}|\hat{A}(s)|\sum_r |\hat{B}(jr)||\hat{C}(kr)| \leq \frac{1}{N}\max_{s\neq 0}|\hat{A}(s)|\left(\sum_t |\hat{B}(t)|^2\right)^{1/2}\left(\sum_u |\hat{C}(u)|^2\right)^{1/2}$$
$$= \frac{1}{N}\max_{s\neq 0}|\hat{A}(s)|(N|B|N|C|)^{1/2} = (|B||C|)^{1/2}\max_{s\neq 0}|\hat{A}(s)| \leq N\max_{s\neq 0}|\hat{A}(s)|$$

using the Cauchy-Schwarz inequality. The result follows from the fact that $\hat{A}(m) = o(|A|)$ for all $m \not\equiv 0 \pmod N$ since this implies that $N\max_{s\neq 0}|\hat{A}(s)| = o(|A||B||C|/N)$.

We are interested in whether there are necessarily three elements of A in arithmetic progression; in fact we will examine solutions to $a + b = 2b'$ with $a \in A,\ b, b' \in B \subset A$ other than $a = b = b'$. In other words we want $\#\{a \in A,\ b, b' \in B : a + b \equiv 2b' \pmod N\} > |B|$. The above proof gives that $\#\{a \in A,\ b, b' \in B : a+b \equiv 2b' \pmod N\} \geq |A||B|^2/N - |B|\max_{s\neq 0}|\hat{A}(s)|$ which is $> |B|$ provided $\max_{s\neq 0}|\hat{A}(s)| < |A||B|/N - 1$.

Proof of the equidistribution theorem. Suppose that $\#\{a \in A :\ \alpha N < (ma)_N \leq \beta N\} \sim (\beta - \alpha)|A|$ for any $m \not\equiv 0 \pmod N$, for any $0 \leq \alpha < \beta \leq 1$. In particular if $\alpha N <$

$(ma)_N \leq \beta N$ then $e(ma/N) = e(\alpha) + O(|\beta - \alpha|)$, so subdividing $(0, N]$ into intervals $I_j := (jN/k, (j+1)N/k]$ for fixed large k, we find that

$$\hat{A}(m) = \sum_{j=0}^{k-1} \sum_{\substack{a \in A \\ (ma)_N \in I_j}} e(ma/N) = \sum_{j=0}^{k-1} \{1 + o(1)\}(|A|/k)(e(jN/k) + O(1/k)) \ll |A|/k.$$

Letting $k \to \infty$ we have that $\hat{A}(m) = o(|A|)$.

On the other hand for $J = [\delta N]$

$$\sum_{\substack{a \in A \\ 1 \leq (ma)_N \leq J}} 1 = \sum_{j=1}^{J} \sum_{a \in A} \frac{1}{N} \sum_r e\left(r\left(\frac{ma - j}{N}\right)\right) = \frac{J}{N}|A| + \frac{1}{N} \sum_{r \neq 0} \hat{A}(rm) \sum_{j=1}^{J} e\left(\frac{-rj}{N}\right).$$

If r runs through the non-zero integers in $(-N/2, N/2]$ then $|\sum_{j=1}^{J} e\left(\frac{-rj}{N}\right)| \ll N/|r|$. Thus the second term here is

$$\begin{aligned} &\ll \sum_{r \neq 0} \frac{|\hat{A}(rm)|}{r} \leq \sum_{0 \leq |r| \leq R} \frac{|\hat{A}(rm)|}{r} + \sum_{R < |r| \leq N/2} \frac{|\hat{A}(rm)|}{r} \\ &\leq (\log R) \max_{s \neq 0} |\hat{A}(s)| + \left(\sum_r |\hat{A}(rm)|^2\right)^{1/2} \left(\sum_{R < |r|} 1/r^2\right)^{1/2} \\ &\leq (\log R) \max_{s \neq 0} |\hat{A}(s)| + (|A|N/R)^{1/2} = o(N) \end{aligned}$$

if we let $R \to \infty$ slowly enough. The result follows.

Try to develop an analogous proof of Weyl's theorem; or a proof of Weyl's theorem as a corollary to our equidistribution theorem.

In what follows we will be interested in determining how big $\#\{a \in A : \alpha N < (ma)_N \leq \beta N\}$ can get when $|\hat{A}(r)| > cN$.

We have seen direct connections between how well a set is distributed and the size of its Fourier transforms. In the context of set addition we are interested in when $A + B$ is small, and/or perhaps when $r_{A+B}(n)$ is large for some n. We now see that $r_{A-A}(n)$ is very large if and only if the weight of the Fourier transform is concentrated on the $\hat{A}(m)$ with $(mn)_N$ small:

Proposition 3.2a. *Let $\eta > 0$ be small. Suppose that $A \subset \mathbb{Z}/N\mathbb{Z}$. If $r_{A-A}(n) > (1-\eta)|A|$ then*

$$\sum_{m : |(mn)_N/N| \leq \eta^{1/3}} |\hat{A}(m)|^2 \geq (1 - O(\eta^{1/3})) \sum_m |\hat{A}(m)|^2.$$

On the other hand if

$$\sum_{m : |(mn)_N/N| \leq \epsilon} |\hat{A}(m)|^2 \geq (1 - \delta) \sum_m |\hat{A}(m)|^2$$

then $r_{A-A}(n) \geq (1-\delta-O(\epsilon^2))|A|$.

Proof. We remark that $\sum_m |\hat{A}(m)|^2 = N|A|$. Since $\cos t$ is decreasing for $0 \leq |t| \leq \pi$ we know that $\cos(2\pi mn/N) < \cos(2\pi\theta)$ if $|(mn)_N/N| > \theta$ and so, using the formula above, $(1-\eta)|A|N < Nr_{A-A}(n)$ which is

$$\begin{aligned}&= \sum_m |\hat{A}(m)|^2 \cos(2\pi mn/N) \leq \sum_{m:|(mn)_N/N|>\theta} |\hat{A}(m)|^2 \cos(2\pi\theta) + \sum_{m:|(mn)_N/N|\leq\theta} |\hat{A}(m)|^2 \\ &= |A|N - (1-\cos(2\pi\theta)) \sum_{m:|(mn)_N/N|>\theta} |\hat{A}(m)|^2.\end{aligned}$$

Now selecting $\theta = \eta^{1/3}$, we deduce that $\sum_{m:|(mn)_N/N|>\eta^{1/3}} |\hat{A}(m)|^2 \ll \eta^{1/3}N|A|$.

In the other direction

$$\begin{aligned}Nr_{A-A}(n) = \sum_m |\hat{A}(m)|^2 \cos(2\pi mn/N) &\geq \sum_{m:|(mn)_N/N|\leq\epsilon} |\hat{A}(m)|^2 \cos(2\pi\epsilon) \\ &\geq \cos(2\pi\epsilon)(1-\delta)\sum_m |\hat{A}(m)|^2 = \cos(2\pi\epsilon)(1-\delta)|A|N.\end{aligned}$$

As a counterpart to this theorem we have the uncertainty principle, which tells us that a function's support and the support of its Fourier transform cannot both be small. More precisely we now show that if A has no elements in a long segment then $\hat{A}$ is concentrated near to 0.

Proposition 3.2b. *Suppose that* $A \subset \mathbb{Z}/N\mathbb{Z}$ *has no elements in the interval* $(x-L, x+L)$. *Then there exists* $m, 0 < m < (N/L)^2$ *such that* $|\hat{A}(m)| \geq (L/2N)|A|$.

Proof. We can assume wlog that $x = 0$ since $A \hat{-} x(m) = e(mx/N)\hat{A}(m)$. Let I be the interval $[0, L)$ and note that $(I-I) \cap A = \emptyset$, so $\sum_r |\hat{I}(r)|^2 \hat{A}(r) = 0$. Therefore

$$\begin{aligned}L^2|A| = |\hat{I}(0)|^2\hat{A}(0) &\leq \sum_{r\neq 0} |\hat{I}(r)|^2 |\hat{A}(r)| \\ &\leq \max_{0\leq|r|\leq R} |\hat{A}(r)| \sum_r |\hat{I}(r)|^2 + |A| \sum_{R<|r|\leq N/2} |\hat{I}(r)|^2 \\ &\leq NL \max_{0\leq|r|\leq R} |\hat{A}(r)| + |A|N^2/2R\end{aligned}$$

since $\hat{I}(r) \leq 1/|\sin(\pi r/N)| \leq N/2|r|$ for $|r| \leq N/2$. Taking $R = (N/L)^2$ the result follows.

The equidistribution theorem gives that if $|\hat{A}(m)| > c|A|$ then A is not uniformly distributed mod N. We seek a more explicit result than this:

Proposition 3.3. *Suppose* $A \subset \{1, \dots, N\}$. *For any* $m \neq 0$ *there exists* $\ell > |\hat{A}(m)|/6\pi$ *and a value* x *such that* $\#\{a \in A : x < (am)_N \leq x + \ell\} \geq (1 + |\hat{A}(m)|/4|A|)\ (|A|/N)\ \ell$. *If* $(m, N) = 1$ *and* $|\hat{A}(m)| > CN/\log N$ *for a large constant* $C > 0$ *then there exist integers*

b and r and length $J \gg \sqrt{N}/\log N$ such that $\#\{a \in A: \ a = b + jr, \ 0 \le j < J\} \ge (1 + |\hat{A}(m)|/4|A|)\ (|A|/N)\ J$.

Proof. Define $\delta = |A|/N$. Let

$$\Delta(n) := A(n) - \delta = \begin{cases} 1-\delta & \text{if } n \in A \\ -\delta & \text{otherwise} \end{cases}$$

so that $\hat{\Delta}(r) - \hat{A}(r) = -\delta \sum_n e(-rn/N) = 0$ if $r \ne 0$. Fix J large and let $I_j := \{n : (j-1)N/J < (mn)_N < jN/J\}$ for $j = 1, 2, \dots, J$. We observe that

$$\left| \sum_{n \in I_j} \Delta(n) e(mn/N) - e(j/J) \sum_{n \in I_j} \Delta(n) \right| \le \sum_{n \in I_j} |\Delta(n)||1 - e(1/J)| \le 2|I_J| \sin(\pi/J),$$

so that

$$\left| \hat{\Delta}(m) - \sum_{j=1}^{J} e(jN/J) \sum_{n \in I_j} \Delta(n) \right| \le 2N \sin(\pi/J);$$

and, of course, $\sum_{j=1}^{J} \sum_{n \in I_j} \Delta(n) = 0$. An easy consequence is that

$$\begin{aligned} |\hat{A}(m)| = |\hat{\Delta}(m)| &\le \sum_{j=1}^{J} \left\{ \sum_{n \in I_j} \Delta(n) + \left| \sum_{n \in I_j} \Delta(n) \right| \right\} + 2N \sin(\pi/J) \\ &\le 2J \max_j \sum_{n \in I_j} \Delta(n) + 2N \sin(\pi/J). \end{aligned}$$

Taking $J = [5\pi N/|\hat{A}(m)|]$ we deduce that $\max_j \sum_{n \in I_j} \Delta(n) \ge |\hat{A}(m)|/4J$, and the first part of the result follows.

There exist integers $r \ne 0, s$ with $0 \le |r|, s < \sqrt{N}$ such that $mr \equiv s \pmod N$ (hint: consider the values $i + jm \pmod N$, $0 \le i, j \le \sqrt{N}$). Select $1 \le a_i \le N$ so that $(ma_i)_N = x + i$ for $i = 1, 2, \dots, s$, so that $m(a_i + jr) \equiv x + i + js \pmod N$. Therefore if $0 \le x < x + ks < N$ then

$$\#\{a \in A : x < (am)_N \le x{+}ks\} = \bigcup_{i=1}^{s} S_i \text{ where } S_i := \{a \in A: \ a = (a_i{+}jr)_N, \ 0 \le j \le k{-}1\};$$

and so, selecting $k = [N/Js]$, we deduce from the above that there exists i for which $|S_i| \ge \delta k(1 + \eta|\hat{A}(m)|/|A|)$ for some fixed $\eta > 1/4$. We may assume $r > 0$ for, if not, rewrite S_i as $\{a \in A: \ a = (a_i' + j|r|)_N, \ 0 \le j \le k-1\}$ where $a_i' \equiv a_i + (k-1)r \pmod N$ with $1 \le a_i' \le N$.

Now define $j_0 = 0 < j_1 < ... < j_w = k$ with j_ℓ chosen minimal so that $[(a_i{+}j_\ell r)/N] = \ell$ for $1 \le \ell \le w - 1 = [(a_i + (k-1)r)/N]$. Note that for $b_\ell = (a_i + j_\ell r)_N$ we have

$$\{a \in A: \ a = (a_i + jr)_N, \ j_\ell \le j < j_{\ell+1}\} = \{a \in A: \ a = b_\ell + jr, \ 0 \le j < j_{\ell+1} - j_\ell\},$$

and that $j_{\ell+1} - j_\ell = N/r + O(1)$ for $1 \leq \ell \leq w-2$. If $w > 1$ then let $u = 0$ unless $j_1 < \min\{N/r, k\}/\log N$ in which case we let $u = 1$, and let $v = w$ unless $j_w - j_{w-1} < \min\{N/r, k\}/\log N$ in which case we let $v = w-1$. Therefore

$$\begin{aligned}\sum_{\ell=u}^{v-1} \#\{a \in A : \ a = (a_i + jr)_N, \ j_\ell \leq j < j_{\ell+1}\} &\geq \delta k \left(1 + \eta \frac{|\hat{A}(m)|}{|A|}\right) - 2\frac{\min\{N/r, k\}}{\log N} \\ &\geq \delta (j_v - j_u) \left(1 + \frac{|\hat{A}(m)|}{4|A|}\right).\end{aligned}$$

Therefore there exists an integer b and $z \geq \min\{N/r, k\}/\log N \gg \sqrt{N}/\log N$ such that $\#\{a \in A : \ a = b + jr, \ 0 \leq j < z\} \geq \delta z(1 + |\hat{A}(m)|/4|A|)$

3b. Roth's Theorem.

Roth [11] showed that any set of integers of positive upper density contains arithmetic progressions of length three. This was generalized by Szemerédi [17] who showed that any set of integers of positive upper density contains arbitrarily long arithmetic progressions. His proof used ingenious combinatorial techniques, and a later proof given by Fürstenburg [6] surprisingly used methods from ergodic theory.

Szemerédi actually proved more than this: Let $n_k(N)$ denote the smallest integer such that any subset of $n_k(N)$ integers from $\{1,\dots,N\}$ contains an arithmetic progression of length k. Szemerédi established that $n_k(N)=o(N)$ for each k, proving a conjecture of Erdős and Turán [4]. His proof used van der Waerden's Theorem and Szemerédi's Regularity Lemma, and so the upper bound on the order of $n_k(N)$ obtained can be no better than the bounds given by these theorems. However Roth's remarkable analytic proof that $n_3(N)=o(N)$ can be used to get an explicit upper bound, namely $n_3(N)\leq cN/\log\log N$ for some constant $c>0$, and offers the possibility of generalization.

Roth's Theorem. *There exists a constant $c>0$ such that any set of $cN/\log\log N$ integers from $\{1,\dots,N\}$ contains an arithmetic progression of length three.*

Proof. Suppose that $A\subset\{1,\dots,N\}$ does not contain an arithmetic progression of length three. We will show that there exists a subset A' of A which is a subset of an arithmetic progression, and denser than A. Then we create a Freiman homomorphism of order 2 between $A_1\subset\{1,\dots,N_1\}$ and A' in each case, so that A_1 does not contain an arithmetic progression of length three. In fact we will show that $\sqrt{N}/\log N \ll N_1\leq N$ and if $|A|=\delta N$ then $|A_1|\geq(1+\delta/20)\ \delta\ N_1$. Iterating this enough times we create a set A^* of density $>2/3$ in the integers up to some point, so that it contains three consecutive integers. Do this explicitly and deduce the result.

To prove our claim, fix $1/5>\eta>0$ and let P be the smallest prime $>N$: by the prime number theorem $P=N+O(N/\log^2 N)$. If $\#\{a\in A:\ 0<a\leq P/3\}\geq(1+\eta)|A|/3$ then let $A_1=A'=\{a\in A:\ 0<a\leq P/3\}$, or if $\#\{a\in A:\ 2P/3<a\leq P\}\geq(1+\eta)|A|/3$ then let $A'=\{a\in A:\ 2P/3<a<P\}=P-A_1$. Otherwise we must have $|B|\geq(1-2\eta)|A|/3$ where $B=\{a\in A:\ P/3<a\leq 2P/3\}$. Notice that if $b,c\in B$ then $0<2c-b<N$. In this case let $i=j=1,k=-2$ and $C=B$ in the remarks following Proposition 3.1, so that there exists $m\not\equiv 0\pmod P$ such that $|\hat{A}(m)|\geq|A||B|/P-1$, Thus, by Proposition 3.3, if we take $A_1=\{j:0\leq j<J\}$ and $A'=\{a\in A:\ a=b+jr,\ j\in A_1\}$ then $|A_1|=|A'|\geq(1+|\hat{A}(m)|/4|A|)\ (|A|/N)\ J$ where $J\gg\sqrt{N}/\log N$.

Heath-Brown [9] and Szemerédi [18] have recently improved this to $n_3(N)\ll N/(\log N)^c$, for some constant $c>0$. It is unclear what is the correct order of $n_3(N)$. There is a beautiful lower bound due to Behrend:

Behrend's Theorem. *There exists a set of* $\geq N/\exp(c\sqrt{\log N})$ *integers from* $\{1,\dots,N\}$ *which does not contain an arithmetic progression of length three.*

Proof. The set $\{(x_0,\dots,x_{n-1})\in\mathbb{Z}^n:\ \ 0\leq x_i\leq d-1\}$ contains $(d-1)^n$ elements. Moreover if $\mathbf{x}$ belongs to this set then $|\mathbf{x}|^2$ is a positive integer $\leq n(d-1)^2$. Thus there exists an integer k such that the set $S=\{\mathbf{x}:\ |\mathbf{x}|^2=k,\ 0\leq x_i\leq d-1\}$ has $\geq d^{n-2}/n$ elements.

Let $A=\{x_0+x_1(2d-1)+x_2(2d-1)^2+\dots+x_{n-1}(2d-1)^{n-1}:\ \mathbf{x}\in S\}$. If $a_1+a_2=2a_3$ where each $a_i=a_i(\mathbf{x}_i)=\sum_j x_{i,j}(2d-1)^j$ then prove by induction that $x_{1,j}+x_{2,j}=2x_{3,j}$ for each $j\geq 0$ (hint: consider $a_1+a_2\equiv 2a_3 \pmod{2d-1}$, etc.). Therefore $\mathbf{x_1}+\mathbf{x_2}=\mathbf{2x_3}$, which implies that these three points of S lie on the same line: however this is impossible since a line intersects the surface of a sphere in at most two points. Thus A contains no three term arithmetic progressions.

Now every element of A is $\leq (d-1)\sum_{j=0}^{n-1}(2d-1)^j<N:=(2d-1)^n/2$ and $|A|=|S|\geq d^{n-2}/n\geq N/(2^nd^2)$. Choosing n to be an even integer and $d=2^{n/2-1}$ we obtain $N<2^{n^2-1}$ and $|A|\geq 4N/2^{2n}$ which implies the result for arbitrarily large N. Prove the result for all sufficiently large N using this construction.

3c. How often can $|\hat{A}(m)|$ be large?. For fixed $\rho>0$ let $R:=\{r \pmod N):\ |\hat{A}(r)|>\rho|A|\}$. Then

$$|A|N=\sum_m|\hat{A}(m)|^2\geq\sum_{m\in R}\rho^2|A|^2,$$

so that $|R|\leq\rho^{-2}\alpha^{-1}$, where $\alpha:=|A|/N$; that is, it is bounded as a function of ρ and α. One can determine some structure in the set R:

Theorem 3.4. *Suppose that* $A\subset\mathbb{Z}/N\mathbb{Z}$*. The set* $R:=\{r \pmod N):\ |\hat{A}(r)|>\rho|A|\}$ *is contained in a cube of dimension* $\leq 2\rho^{-2}\log(N/|A|)$*.*

Remember that a cube of dimension k is a set $\overline{\Lambda}$ of the form $\{\epsilon_1\lambda_1+\dots+\epsilon_k\lambda_k:$ each $\epsilon_i\in\{-1,0,1\}\}$, for given $\Lambda=\{\lambda_1,\lambda_2,\dots,\lambda_k\}$. We say that Λ is *dissociated* if $\epsilon_1\lambda_1+\dots+\epsilon_k\lambda_k=0$ with each $\epsilon_i\in\{-1,0,1\}$ only for $\epsilon_1=\dots=\epsilon_k=0$.

The proof of Theorem 3.4 involves some seemingly ad hoc analysis, which we state in the next lemma. For now write the "general" finite trigonometric polynomial as $f(x)=\sum_j c_j\cos(2\pi(\lambda_jx/N+\beta_j))$ where each $c_j\in\mathbb{R}$ and $0\leq\beta_j<1$. Show:

Lemma 3.5.
(i) We have $e^{ty}\leq\cosh(t)+y\sinh(t)$ *for all* $t\in\mathbb{R}$ *and* $|y|\leq 1$*.*
(ii) $\cosh(u)\leq e^{u^2/2}$ *for any* $u\in\mathbb{R}$*.*
(iii) With $f(x)$ *as above,* $\sum_{x \pmod N}f(x)^2=(N/2)\sum_j c_j^2$*.*

We deduce

Proposition 3.6. *If* Λ *is dissociated then*

$$\frac{1}{N}\sum_x\exp(tf(x))\leq\exp\left(\frac{1}{N}\sum_x t^2f(x)^2\right)$$

Proof. Using (i) (of Lemma 3.5) the left side above is

$$\leq \frac{1}{N} \sum_x \prod_j (\cosh(tc_j) + \cos(2\pi(\lambda_j x/N + \beta_j)) \sinh(tc_j)).$$

Writing each $\cos\theta$ as $(e^{i\theta} + e^{-i\theta})/2$ the jth term in the product takes the form $\sum_{\epsilon_j \in \{-1,0,1\}} c(j,\epsilon_j) \exp(\epsilon_j \lambda_j x/N)$ for certain constants $c(j,\epsilon_j)$. Multiplying this out over the j, then summing over $x \pmod N$, we see that the only terms that can have a non-zero sum are those for which $\sum_j \epsilon_j \lambda_j = 0$: and thus $\epsilon_1 = \cdots = \epsilon_k = 0$ as Λ is dissociated. This term is easily obtained from the above product as

$$\frac{1}{N} \sum_x \prod_j \cosh(tc_j) = \prod_j \cosh(tc_j) \leq \exp\left(\frac{t^2}{2} \sum_j c_j^2\right) = \exp\left(\frac{t^2}{N} \sum_x f(x)^2\right)$$

using Lemma 3.5 (ii) and (iii).

With this preparation we can prove

Proposition 3.7. *Suppose that $A \subset \mathbb{Z}/N\mathbb{Z}$. Any dissociated subset of $R := \{r \pmod N : |\hat{A}(r)| > \rho|A|\}$ has size $\leq 2\rho^{-2}\log(N/|A|)$.*

Deduction of Theorem 3.4. Let Λ be a maximal dissociated subset of R. Then any $r \in R$ can be written in the form $r = \sum_i \epsilon_i \lambda_i$ so that $R \subset \Lambda$.

Proof of Proposition 3.7. If $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ is a dissociated subset of R, let $f(x) = \mathrm{Re}\left(\sum_j \hat{A}(\lambda_j) e(\lambda_j x/N)\right)$, which we can rewrite as $f(x) = \sum_j c_j \cos(2\pi(\lambda_j x/N + \beta_j))$ where $c_j = |\hat{A}(\lambda_j)|$ and β_j is chosen appropriately. Note also that $\hat{f}(r) = N\hat{A}(r)/2$ if $r \in \Lambda \cup -\Lambda$, and $\hat{f}(r) = 0$ otherwise, so that

$$\sum_x f(x)A(x) = \frac{1}{N} \sum_r \hat{f}(r)\overline{\hat{A}(r)} = \frac{2}{N^2} \sum_r |\hat{f}(r)|^2 = \frac{2}{N} \sum_x |f(x)|^2.$$

Therefore, for any t we have, using the arithmetic-geometric mean inequality and then Proposition 3.6,

$$\exp\left(\frac{2t}{N|A|} \sum_x |f(x)|^2\right) = \exp\left(\frac{t}{|A|} \sum_{x \in A} f(x)\right)$$
$$\leq \frac{1}{|A|} \sum_{x \in A} \exp(tf(x)) \leq \frac{N}{|A|} \exp\left(\frac{t^2}{N} \sum_x f(x)^2\right).$$

Taking $t = 1/|A|$ we deduce that $\sum_x f(x)^2 \leq N|A|^2 \log(N/|A|)$. On the other hand, by Lemma 3.5(iii), $(2/N)\sum_x f(x)^2 = \sum_j c_j^2 = \sum_j |\hat{A}(\lambda_j)|^2 \geq |\Lambda|(\rho|A|)^2$, and the result follows from combining these last two inequalities.

Exponential sums, I

Suppose that $\alpha, \beta \in \mathbb{R}$. As each term has absolute value ≤ 1, and by summing the geometric series, we obtain

$$\left|\sum_{n=0}^{N-1} e(\alpha n+\beta)\right| \leq \min\{N, 1/2\|\alpha\|\}. \tag{4.1}$$

It is also useful to note that for any given $y_1, y_2, \dots, y_k$ we have

$$\sum_i \min\{N, 1/\|y_i\|\} \ll N + \frac{\log(2N)}{\min_{i\neq j}\|y_i-y_j\|}. \tag{4.2}$$

Now suppose that $|\alpha - a/q| \leq 1/q^2$ where $(a,q)=1$. We use this to show that

$$\sum_{n=0}^{N} \min\left\{Q, \frac{1}{\|\alpha n+\beta\|}\right\} \ll (Q+q\log 2Q)(1+N/q). \tag{4.3}$$

First we prove it for $N < q/2$, since if $0 \leq m < n \leq N$ then $\|\alpha n+\beta\| - \|\alpha m+\beta\| \geq \|\alpha(m-n)\| \geq \|(m-n)a/q\| - |m-n|/q^2 \geq 1/q - (q/2)/q^2 = 1/(2q)$ and we apply (4.2). The result follows for general but cutting the sum up into intervals of length $< q/2$.

All this preparation leads to a remarkable lemma:

Lemma 4.1. *For any $\alpha \in \mathbb{R}$ and any quadratic polynomial $f(x) \in \mathbb{R}[x]$ we have*

$$\left|\sum_{n=0}^{N-1} e(\alpha f(n))\right| \ll N/q^{1/2} + ((q+N)\log 2N)^{1/2}$$

where $|\alpha - a/q| \leq 1/q^2$ with $(a,q)=1$.

Proof. If $f(x) = dx^2+bx+c$ then $f(x+h)-f(x) = h(2dx+dh+bh)$ so that, writing $m=n+h$ we obtain

$$\begin{aligned}\left|\sum_{n=0}^{N-1} e(\alpha f(n))\right|^2 &= \sum_{m,n=0}^{N-1} e(\alpha(f(m)-f(n)))\\ &= \sum_{h=-(N-1)}^{N-1} e(\alpha(d+b))h^2) \sum_{n=\max\{0,-h\}}^{\min\{N-1,N-1-h\}} e(2d\alpha hn)\\ &\leq \sum_{h=-(N-1)}^{N-1} \min\{N, 1/2\|2d\alpha h\|\}\end{aligned}$$

using (4.1), and the result follows from (4.3).

Challenge problem: Generalize this result to polynomials of arbitrary degree.

Theorem 4.2. *For any $\alpha \in \mathbb{R}$ and any sufficiently large M, there exists $m \leq M$ such that $\|\alpha m^2\| \ll 1/M^{1/5}$.*

Proof. Select prime N arbitrarily large ($> M^3$) and b with $|\alpha - b/N| \leq 1/(2N)$. Let $A = \{bm^2 : 1 \leq m \leq M\} \subset \mathbb{Z}/N\mathbb{Z}$ and let $L = N/2M^{1/5}$. If A contains an element in $(-L, L)$ then $\|\alpha m^2\| \leq \|bm^2/N\| + M^2/2N \leq 1/M^{1/5}$. Otherwise there exists $r, 0 < r < 4M^{2/5}$ such that $|\hat{A}(r)| \geq M^{4/5}/4$ by Proposition 3.2b. Select $q \leq M$ for which $|r\alpha - a/q| < 1/qM$ for some $(a, q) = 1$ so that $|\hat{A}(m)| \ll M/q^{1/2} + (M \log 2M)^{1/2}$ by Lemma 4.1, which implies that $q \ll M^{2/5}$. But then $\|\alpha(qr)^2\| \leq q^2 r|r\alpha - a/q| < qr/M \ll 1/M^{1/5}$.

Extend this result to $\|\alpha m^k\|$ for arbitrary positive integers k.

The Geometry of Numbers

A *lattice* in $\mathbb{R}^n$ is a subgroup generated by n linearly independent vectors, with basis $x_1, x_2, \dots, x_n$ say. The *fundamental parallellopiped* of Λ with respect to $x_1, x_2, \dots, x_n$ is the set $P = \{a_1x_1 + a_2x_2 + \dots + a_nx_n : 0 \leq a_i < 1\}$. The sets $x + P$, $x \in \Lambda$ are disjoint and their union is $\mathbb{R}^n$. The *determinant* $det(\Lambda)$ of Λ is the volume of P; in fact $det(\Lambda) = |det(A)|$, where $A = (x_1, x_2, \dots, x_n)$ and the x_i are column vectors with respect to the canonical basis for $\mathbb{R}^n$. A *convex body* K is a bounded convex open subset of $\mathbb{R}^n$. Show that $vol(K) = \lim_{t\to\infty} |\Lambda \cap tK| det(\Lambda)/t^n$. A key result is:

Blichfeldt's Lemma. *Let $K \subset \mathbb{R}^n$ be a measurable set, Λ a lattice and suppose $vol(K) > det(\Lambda)$. Then $K - K$ contains a non-zero lattice point.*

The proof is a challenge problem. This immediately gives

Minkowski's First Theorem. *If K is a centrally symmetric convex body with $vol(K) > 2^n det(\Lambda)$ then K contains a non-zero point of Λ.*

Proof. As K is convex and centrally symmetric, $K = \frac{1}{2}K - \frac{1}{2}K$. However, $vol(\frac{1}{2}K) > det(\Lambda)$, so the result follows by Blichfeldt's Lemma.

For a centrally symmetric convex body K define λ_k to be the infimum of those λ for which λK contains k linearly independent vectors of Λ. We call $\lambda_1, \lambda_2, \dots, \lambda_n$ the successive minima of K with respect to Λ. Let $b_1, b_2, \dots, b_n \in \mathbb{R}^n$ be linearly independent vectors with $b_k \in \lambda_k\overline{K} \cap \Lambda$ for each k. The proof of the next result, and much more, can be found in [15].

Minkowski's Second Theorem. *If $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ are the successive minima of convex body K with respect to Λ then $\lambda_1\lambda_2 \dots \lambda_n vol(K) \leq 2^n det(\Lambda)$.*

Let $r_1, r_2, \dots, r_k \in \mathbb{Z}/N\mathbb{Z}$ and $\delta > 0$ be given. We define the *Bohr neighbourhood*

$$B(r_1, r_2, \dots, r_k; \delta) := \{s \in \mathbb{Z}/N\mathbb{Z} : \|r_is/N\| \leq \delta \text{ for } i = 1, 2, \dots, k\};$$

that is, the least residue, in absolute value, of each $r_is \pmod N$ is $< \delta N$ in absolute value.

A generalized arithmetic progression is called *proper* if its elements are distinct are all distinct (that is $|C(a_0, a_1, \dots, a_k; N_1, N_2, \dots, N_k)| = N_1N_2 \dots N_k$).

It will be seen, in §6, that Bohr neighbourhoods can be used as a step in finding arithmetic progressions, using Fourier transforms.

Theorem 5.7. *If $0 < \delta < 1/2$ then the Bohr neighbourhood $B(r_1, \dots, r_k, \delta)$ contains a proper k-dimensional arithmetic progression of cardinality at least $(2\delta/k)^kN$.*

Proof. We have $s \in B(r_1, r_2, \dots, r_k; \delta)$ if and only if $(r_1s, r_2s, \dots, r_ks) + N\mathbb{Z}^k$ contains a point x with $\|x\|_\infty \leq \delta N$. Let Λ be the lattice generated by $N\mathbb{Z}^k$ and $(r_1, r_2, \dots, r_k)$: It can be shown that $det(\Lambda) = N^{k-1}$.

Let $K = \{(a_1, a_2, \dots, a_k) : -1 < a_i < 1\}$ and, as described above, obtain a basis $b_1, b_2, \dots, b_k$ of $\mathbb{R}^k$ with each $b_i \in \Lambda$ satisfying $\|b_i\|_\infty = \lambda_i$. Define s_i so that $b_i \in (r_1s_i, r_2s_i, \dots, r_ks_i) + N\mathbb{Z}^k$. By Minkowski's Second Theorem, $\lambda_1\lambda_2 \dots \lambda_k vol(K) \leq det(\Lambda) \cdot 2^k$ so that $\lambda_1\lambda_2 \dots \lambda_k \leq N^{k-1}$.

Let P be the k-dimensional arithmetic progression $\{\sum_{i=1}^{k} a_i s_i : |a_i| \le \delta N/k\lambda_i\}$. If $s \in P$ then, for each j,

$$\left\| \frac{r_j s}{N} \right\| \le \sum_{i=1}^{k} |a_i| \left\| \frac{r_j s_i}{N} \right\| \le \sum_{i=1}^{k} \frac{\delta N}{k\lambda_i} \left\| \frac{(b_i)_j}{N} \right\| \le \sum_{i=1}^{k} \frac{\delta N}{k\lambda_i} \frac{\|b_i\|_\infty}{N} = \delta.$$

Moreover since $\delta < 1/2$ we have that P is proper. Finally note that $|P| \ge \prod 2\delta N/k\lambda_i = (2\delta N/k)^k \cdot (\lambda_1\lambda_2 \dots \lambda_k)^{-1} \ge (2\delta/k)^k N$.

Corollary 5.13. *If A is a non-empty subset of $\mathbb{Z}$ and $|A+A| \le C|A|$, then $|kA| \le C^k|A|$ for each $k \ge 3$.*

Proof. Take $i = 1$ and $B = A$ in the preceding Corollary. This implies that there exists a non-empty $A' \subset A$ such that $|A' + kA| \le C^k|A'| \le C^k|A|$, but $|A' + kA| \ge |kA|$, so the result is proved.

Lemma 5.14. *Let $U, V, W \subset \mathbb{Z}$. Then $|U||V - W| \le |U + V||U + W|$.*

Proof. Define, for $x \in V - W$, $\phi(u, x) = (u + v(x), u + w(x))$ where $v(x) \in V$, $w(x) \in W$ satisfy $v(x) - w(x) = x$. Then ϕ is an injection $U \times (V - W) \to (U + V) \times (U + W)$.

Theorem 5.15. *Let $A, B \subset \mathbb{Z}$ such that $|A+B| \le C|A|$ and let k and l be natural numbers with $l \ge k$. Then $|kB - lB| \le C^{k+l}|A|$.*

Proof. Suppose $l \ge k \ge 1$. By Corollary 5.12, there exists $A' \subset A$ with $|A'+kB| \le C^k|A'|$. Again there exists $A'' \subset A'$ with $|A'' + lB| \le C^l|A''|$. Using Lemma 5.14, $|A''||kB - lB| \le |A'' + kB||A'' + lB| \le C^{k+l}|A'||A''|$ and the result follows on dividing by $|A''|$.

In the next chapter, we will see the use of Theorem 5.15. In essence, the arithmetic properties of kA for large k are easier to deal with than when k is small. Theorem 5.15 also allows one to deal with distinct set sums $A + B$ by converting the problem to a single set difference problem $kB - lB$.

subchapter: Sum-Product formulas, I. Proofs from "The Book"

All true mathematicians are motivated by elegant proofs, none more so than the great Paul Erdős. Erdős used to say that "the supreme being" kept a book which contained all of the most beautiful proofs of each theorem and just occasionally we mortals are allowed to glimpse this book, as we discover an extraordinary proof. In this section we shall see three such proofs from "The Book", all by Hungarians.

One of Paul Erdős's proofs from the book, comes in his "multiplication table theorem": Let $A = \{1, 2, \dots, N\}$; how big is $A \cdot A = \{ab : a, b \in A\}$? That is, how many distinct integers appear in the N-by-N multiplication table? It is trivial that $|A \cdot A| \leq N(N+1)/2$ (using the symmetry that $ab = ba$) but is it the case that $|A \cdot A|/N^2 \to$a limit as $N \to \infty$, and if so, what is that limit? Erdős proved that the limit exists and is 0, that is $|A \cdot A| = o(N^2)$. His proof rests on the beautiful result of Hardy and Ramanujan that all but $o(N)$ of the integers $n \leq N$ have $\{1 + o(1)\} \log\log N$ prime factors (counting multiplicity[2]). But then all but $o(N^2)$ of the products ab with $a, b \leq N$ have $\{2+o(1)\} \log\log N$ prime factors, whereas almost all integers up to N^2 have $\{1 + o(1)\} \log\log(N^2) = \{1 + o(1)\} \log\log N$ prime factors, and the result follows!

In the other direction, consider integers of the form $n = pm \leq N^2$ where $p \in (N^{2/3}, N]$ is prime and $m \leq N$. There are $\geq \{1+o(1)\}N^2/\log N$ such product by the prime number theorem, and any n can be represented in at most two ways as such a product, so that $|A \cdot A| \geq \{1/2+o(1)\}N^2/\log N$. In the case that A is an arithmetic progression $\{a+ib : 1 \leq i \leq N\}$ then we may assume without loss of generality that $(a, b) = 1$ (else we divide through by the common factor). If $b > 2N$ then $|A{\cdot}A| = N(N+1)/2$; for if $(a+ib)(a+jb) = (a+Ib)(a+Jb)$ then $a(i+j)+bij = a(I+J)+bIJ$ so that $a(i+j) \equiv a(I+J) \pmod b$ implying $i+j \equiv I+J$ $(\text{mod } b)$ and thus $i+j = I+J$ since $2 \leq i+j, I+J \leq 2N < b$, and so $ij = IJ$ and therefore $\{i, j\} = \{I, J\}$. Similarly if $a > N^2$ then $|A \cdot A| = N(N + 1)/2$. Finally if $b \leq 2N$ and $a \leq N^2$ then all elements of A are $\leq N^2+N(2N) = 3N^2$. Let B be the subset of A consisting of all integers in A with a prime factor in $(N/2, N]$. Note that all primes in $(N/2, N]$ that do not divide b, divide either one or two elements of A so that $|B| \geq \{1/2+o(1)\}N/\log N$. Any element of $A \cdot B$ is $\leq 9N^4$ so contains no more than 4 prime factors $> N/2$ and so cannot be written in more than eight ways as $ab, a \in A, b \in B$. Therefore $|A \cdot A| \geq |A \cdot B| \geq \{1/8 + o(1)\}N^2/\log N$.

One might expect to generalize this so that if $A + A$ is small then A has a lot of additive structure, that is it is a subset of a generalized arithmetic progression, and so $A \cdot A$ is large. In fact there should be some "play off" between the two in that if one is much smaller than the expected size then the other should not be; that is one might guess, as did Erdős and Szemerédi, that

$$|A + A| + |A \cdot A| \gg_\epsilon |A|^{2-\epsilon}$$

for any $\epsilon > 0$; or, more daringly like Solymosi, that

$$|A + B| + |A \cdot C| \gg_\epsilon |A|^{2-\epsilon} \text{ whenever } |A| = |B| = |C|.$$

There are several results of this type, for various values of ϵ, in the literature, but none more elegantly proved than the result of Elekes. This rests on a (generalization of a) result in combinatorial geometry of Szemerédi and Trotter, which in turn has recently been given a gorgeous proof via geometric and random graph theory by Székely:

[2] So that, for instance, 12 has 3 prime factors.

The Szemerédi-Trotter theorem. *We are given a set Υ of m curves in the complex plane such that*

- *Each pair of (distinct) curves in Υ meet in at most B_1 points;*
- *No more than B_2 curves in Υ contain any given pair of (distinct) points.*

For a given set, Π, of m points, define $X = X(\Upsilon, \Pi)$ to be the number of pairs (P, C) with $P \in \Pi, C \in \Upsilon$ where P lies on C. Then $X(\Upsilon, \Pi) \leq m + 4B_2 n + 4B_2 B_1^{1/3}(mn)^{2/3}$.

Proof. (Székely) The key idea is to determine how far away our set of curves and points are from being embeddable on the plane in the sense of graph theory, that is that the curves of Υ should only cross at points in Π. To convert this directly into a graph theory problem we replace each point of Π by a vertex of our graph G, and we join two vertices of G if and only if the corresponding points lie on the same curve $C \in \Upsilon$ with no other point of Π in-between (we call such pairs of points "neighbours on C"). In this definition G is a simple graph, even if two points are neighbours on several curves of Υ. We will also define a hypergraph G^* with the same vertex set as G, but as many edges between two vertices as the number of curves of Υ on which they are neighbours. Note that $X = e(G^*) + m \leq B_2 e(G) + m$ (where $e(H)$ and $v(H)$ are the number of edges and vertices, respectively, in H; note that $v(G) = n$). Therefore we will assume that $e(G) \geq 4n$ for otherwise we already have $X < 4B_2 n + m$ as desired.

We now define $Y = Y(G)$ to be the minimum, over all drawings of G in the complex plane, of the number of crossings of edges of G that occur at some point not in the vertex set of G; so $Y(G) = 0$ if G is planar. Note that if we remove these $Y(G)$ edges from G and as well as any isolated vertices, then the resulting new graph H is planar, with $e(H) = e(G) - Y(G)$ and $v(H) \leq v(G)$. Note also that $Y(G)$ can be no bigger than the sum over all pairs of curves in Υ, of the number of ways that those two curves can cross; that is $Y(G) \leq B_1 \binom{m}{2}$.

Now for any simple planar graph H one has the Euler characteristic formula that $f(H) - e(H) + v(H) = 2$ where $f(H)$ is the number of faces of H, and that any edge separates at most two faces whereas any face is surrounded by at least three edges so that $3f(H) \leq 2e(H)$. Therefore $6 + 3e - 3v = 3f \leq 2e$ so that $e(H) \leq 3v(H) - 6$. Combining this with the previous paragraph we deduce that $Y(G) \geq e(G) - 3v(G) + 6$.

Székely's extraordinary trick is to apply the result of the previous paragraph to randomly chosen subgraphs of G, resulting in an improvement of the above inequality! The random process involves deciding at random whether to select each vertex, independently, where the probability of each vertex being chosen is $p = 4v(G)/e(G)$ (which is ≤ 1 by the above assumption); and then retaining the edges from G that go between chosen vertices. If we call the resulting subgraph K then we see that the expected numbers of vertices in K is $pv(G)$, written $\mathbb{E}(v(K)) = pv(G)$ and also that $\mathbb{E}(e(K)) = p^2 e(G)$ and $\mathbb{E}(Y(K)) = p^4 Y(G)$. Substituting this into the bound attained above we have $p^4 Y(G) = \mathbb{E}(Y(K)) \geq E(e(K)) - 3\mathbb{E}(v(K)) + 6 = p^2 e(G) - 3pv(G) + 6$. With our choice of p, and the above bound for Y, this implies that $B_1 m^2/2 > Y(G) > e(G)^3/64v(G)^2$, so that $X \leq B_2(32B_1)^{1/3}(mn)^{2/3} + m$.

Prove that each term in the upper bound here is necessary by giving appropriate examples. For example for the third term let Υ be the set of lines $y = ax + b$ with $0 \leq a \leq A$ and $0 \leq b \leq AC$, and let Π be the set of points in the rectangle $\{0,1,\dots,C\}\times\{0,1,\dots 2AC\}$. A big subset of the points counted by X are given by the points on the lines with $0\le x\le C$.

Corollary (Elekes)**.** *If* $|B||C|\ge|A|$ *then* $|A+B|+|A\cdot C|\ge(64|B||C|/(|A|-1))^{1/4}(|A|-1)$.

Proof. Consider the set of points $\Pi=(A+B)\times(A\cdot C)$, and the set Υ of lines $y=c(x-b)$ for each $b\in B, c\in C$. Note that $B_1=B_2=1$. Each such line contains $|A|$ points, namely $\{(a+b,ac):a\in A\}$ and so $X\ge|A|m$ where $m=|B||C|$ and $n=|A+B||A\cdot C|$. Substituting this into the proof of the Szemerédi-Trotter theorem we obtain $(|A|-1)m\le n/4+(mn)^{2/3}/32^{1/3}$, from which we deduce that $|A+B||A\cdot C|\ge 2(|B||C|(|A|-1)^3)^{1/2}$ if $|B||C|\ge|A|$, and the result follows.

In the particular case that $|A|=|B|=|C|$ this gives $|A+B|+|A\cdot C|\ge|A|^{5/4}$, a first step to the above conjectures.

The Freiman-Ruzsa Theorem

Freiman's Theorem [5] describes the structure of a set A under the condition that $A+A$ has size close to that of A. We define a *generalised arithmetic progression* to be a sum P of ordinary arithmetic progressions (see Theorem 5.7). If P is a subset of a small generalised arithmetic progression then $|P+P|$ is close to $|P|$. Freiman's Theorem states the converse: if $|P+P|$ is close to P then P must be contained in a small generalized arithmetic progression.

We now proceed to the proof of Freiman's Theorem, using a remarkable and ingenious approach due to Ruzsa [12].

Let $A \subset \mathbb{Z}/s\mathbb{Z}$ or $A \subset \mathbb{Z}$ and $B \subset \mathbb{Z}/t\mathbb{Z}$.

Then $\phi : A \to B$ is called a *(Freiman) k-homomorphism* if whenever $x_1+x_2+\cdots+x_k = y_1+y_2+\cdots+y_k$, with $x_i, y_i \in A$, we have $\sum \phi(x_i) = \sum \phi(y_i)$. In addition, ϕ is called a *k-isomorphism* if ϕ is invertible and ϕ and ϕ^{-1} are k-homomorphisms.

Note that ϕ is a k-homomorphism if the map $\psi : (x_1, \dots, x_k) \mapsto \sum \phi(x_i)$ induced by ϕ is a well defined map $kA \to kB$, and a k-isomorphism if ψ is a bijection. Our interest will be in 2-isomorphisms, as these preserve arithmetic progressions – a set 2-isomorphic to an arithmetic progression is clearly an arithmetic progression. We use the following notation:

If $\phi : A \to B$ and $A' \subset A$, then $\phi|_{A'}$ denotes the restriction of ϕ to A'.

Lemma 6.1. *Let $A \subset \mathbb{Z}$ and suppose $|kA - kA| \leq C|A|$. Then, for any prime $N > C|A|$, there exists $A' \subset A$ with $|A'| \geq |A|/k$ that is k-isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$.*

Proof. We may suppose $A \subset \mathbb{N}$ and select a prime $p > k \max A$. Then the quotient map $\phi_1 : \mathbb{Z} \to \mathbb{Z}_p$ is a homomorphism of all orders, and $\phi_1|_A$ is a k-isomorphism. Now let q be a random element of $[p-1]$ and define $\phi_2 : \mathbb{Z}_p \to \mathbb{Z}_p$ by $\phi_2(x) = qx$. Then ϕ_2 is an isomorphism of all orders, and hence a k-isomorphism. Let $\phi_3(x) = x$ where $\phi_3 : \mathbb{Z}_p \to \mathbb{Z}$. Then for any j, $\phi_3|_{I_j}$ is a k-isomorphism where

$$I_j = \{x \in \mathbb{Z}_p : \frac{j-1}{k} p \leq x < \frac{j}{k} p - 1\}.$$

For, if $\sum_{i=1}^k x_i = \sum_{i=1}^k y_i \pmod p$ with $x_i, y_i \in I_j$, then $\sum_{i=1}^k x_i = \sum_{i=1}^k y_i$ in $\mathbb{Z}$. By the pigeonhole principle, there exist $A' \subset A$ with $|A'| \geq |A|/k$ (depending on q) and $\phi_2\phi_1[A'] \subset I_j$ for some j. Restricted to A', $\phi_3\phi_2\phi_1$ is a k-homomorphism. Finally, let ϕ_4 be the quotient map (a k-homomorphism) $\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$. Then with $\phi = \phi_4\phi_3\phi_2\phi_1$, $\phi(x) = qx \pmod p \pmod N$ and $\phi|_{A'}$ is a k-homomorphism, as it is the composition of k-homomorphisms.

The only way $\phi|_{A'}$ is not a k-isomorphism is if there are $a_1, a_2, \dots, a_k$, $a_1', a_2', \dots, a_k' \in A'$ such that $\sum_{i=1}^k \phi(a_i) = \sum_{i=1}^k \phi(a_i')$ but $\sum_{i=1}^k \phi(a_i) \neq \sum_{i=1}^k \phi(a_i')$.

Now $\sum_i a_i \neq \sum_i a_i'$ implies $\sum_i a_i \neq \sum_i a_i' \pmod p$ so we have $q(\sum_i a_i - \sum_i a_i') \pmod p$ is a multiple of N. The probability of this event is at most $|kA - kA|/N < 1$ since $|kA - kA| \leq C|A|$ and $N > C|A|$. So for some q, $\phi|_{A'}$ is a k-isomorphism.

The next theorem, due to Bogolyubov [3], shows that we may find long arithmetic progressions with small dimension in $2A - 2A$. The proof is surprisingly simple.

Theorem 6.2. *Let $A \subset \mathbb{Z}/N\mathbb{Z}$ with $|A| \geq \alpha N$. Then $2A - 2A$ contains an arithmetic progression of length at least $(\alpha^2/4)^{\alpha^{-2}} N$ and dimension at most α^{-2}.*

Proof. Let $g(x)$ be the number of ways of writing $x = (a - b) - (c - d)$ with $a, b, c, d \in A$. That is, $g = (A * A) * (A * A)$ and $x \in 2A - 2A$ if and only if $g(x) \neq 0$. Now $g(x) = N^{-1} \sum_r |\hat{A}(r)|^4 \omega^{rx}$,

by Lemma 2.2 (3). Let $K = \{r \neq 0 : \hat{A}(r) \geq \alpha^{3/2} N\}$. Then

$$\sum_{\substack{r \neq 0 \\ r \notin K}} |\hat{A}(r)|^4 \leq \max_{\substack{r \neq 0 \\ r \notin K}} |\hat{A}(r)|^2 \sum_r |\hat{A}(r)|^2 < \alpha^3 N^2 \cdot \alpha N^2 = \alpha^4 N^4.$$

Therefore, if x is such that $Re(\omega^{rx}) \geq 0$ for all $r \in K$, then

$$Re\Big(\sum_r |\hat{A}(r)|^4 \omega^{rx}\Big) > |\hat{A}(0)|^4 - \alpha^4 N^4 = 0.$$

Therefore $g(x) \neq 0$ and $2A - 2A$ contains the Bohr neighbourhood $B(K; 1/4)$ – $Re(\omega^{rs}) \geq 0$ if and only if $-N/4 \leq rs \leq N/4$. Now $\sum_{r \in K} |\hat{A}(r)|^2 \geq k\alpha^3 N^2$ and $\sum_{r \in K} |\hat{A}(r)|^2 \leq \alpha N^2$. By Theorem 5.7, $2A - 2A$ contains the required arithmetic progression.

We now present Ruzsa's proof of Freiman's Theorem.

Freiman's Theorem. *Let $A \subset \mathbb{Z}/N\mathbb{Z}$ be a set such that $|A + A| \leq C|A|$. Then A is contained in a d-dimensional arithmetic progression P of cardinality at most $k|A|$ where d and k depend on C only.*

Proof. By Theorem 5.15, $|8A - 8A| \leq C^{16}|A|$. By Lemma 6.1, A contains a subset A' of cardinality at least $|A|/8$ which is 8-isomorphic to a a set $B \subset \mathbb{Z}/N\mathbb{Z}$ with $C^{16}|A| < N \leq 2C^{16}|A|$, where N is prime and $C|A| < N \leq 2C|A|$, using Bertrand's Postulate. So $|B| = \alpha N$ with $\alpha \geq (16C^{16})^{-1}$. By Theorem 6.2, $2B - 2B$ contains an arithmetic progression of dimension at most α^{-2} and cardinality at least $(\alpha^2/4)^{\alpha^{-2}} N \geq (\alpha^2/4)^{\alpha^{-2}} |A|$. Since B is 8-isomorphic to A', $2B - 2B$ is 2-isomorphic to $2A' - 2A'$. Any set 2-isomorphic to a d-dimensional arithmetic progression is a d-dimensional arithmetic progression. Therefore $2A' - 2A'$, and hence $2A - 2A$, contains an arithmetic progression Q of dimension at most α^{-2} and cardinality $\gamma|A|$, where $\gamma \geq (\alpha^2/4)^{\alpha^{-2}}$. Now let $X = \{x_1, x_2, \ldots, x_k\} \subset A$ be maximal such that $x, y \in X$, $x \neq y$ imply $x - y \in Q - Q$. Equivalently, all the sets $x + Q$ are disjoint, so $X + Q = |X||Q|$. Since X is maximal, $A \subset X + (Q - Q)$ and X is contained in the k-dimensional arithmetic progression $R = \Big\{\sum_{i=1}^k a_i x_i : 0 \leq a_i \leq 1\Big\}$. Clearly $|R| \leq 2^k$. Therefore A is contained in the arithmetic progression $R + (Q - Q)$, of dimension at most $\alpha^{-2} + k$. We know that $X + (Q - Q) \subset A + (4A - 4A) = A + 2A - 2A + 2A - 2A$, and that $X + Q \subset A + 2A - 2A = 3A - 2A$.

So $|X + Q| \leq |3A - 2A| \leq C^5|A|$, by Theorem 5.15. So $k \leq C^5|A|/|Q| \leq C^5\gamma^{-1}$. Finally, $|Q - Q| \leq 2^{\alpha^{-2}}|Q|$, by d-dimensionality. So A is contained in an arithmetic progression of dimension at most $\alpha^{-2}C^5\gamma^{-1}$, and cardinality at most $2^k 2^{\alpha^{-2}}|Q| \leq 2^k 2^{\alpha^{-2}}|2A - 2A|$ $\leq kC^4 2^{\alpha^{-2}}|A|$.

The constants from this theorem can be chosen to be $d = exp(C^\alpha)$ and $k = expexp(C^\beta)$, where $\alpha, \beta > 0$ are absolute constants.

Using a refinement of the same approach, a better result can be obtained for set *differences* of the same set (see [2]):

Theorem 6.4. *Let C be a positive real number. Suppose A is a set of integers satisfying $|A - A| \leq C|A|$ and $|A| \geq \frac{\lfloor C \rfloor \lfloor C+1 \rfloor}{2(\lfloor C+1 \rfloor - C)}$. Then A is a subset of an arithmetic progression of dimension at most $\lfloor C - 1 \rfloor$ and cardinality at most $expexp(C^\gamma)$ where $\gamma > 0$ is an absolute constant.*

It is likely that a result with very much the same constants is true for $A + A$. These theorems can be generalized to theorems about abelian groups [4], [13].

We now turn to results concerning difference sets, which will eventually aid in finding four-term arithmetic progressions in the next chapter.

The Balog-Szemerédi-Gowers theorem

Lemma 6.5. *Let $A_1, A_2, \dots, A_m$ be subsets of an N element set S, let $\gamma > 0$ and suppose that $\sum_{i=1}^m |A_i| \geq \gamma m N$. Then there exists $B \subset \{1, \dots, m\}$, of cardinality at least $\gamma^5 m/2$, such that for at least ninety-five percent of pairs $(i,j) \in B \times B$, $|A_i \cap A_j| \geq \gamma^2 N/2$.*

Proof. The idea will be to show that this is true of a randomly chosen subset B, and thus there exist such B. So, let $x_1, x_2, \dots, x_5$ be chosen randomly and independently from S and define $B = \{i : \{x_1, x_2, \dots, x_5\} \subset A_i\}$. Then $\text{Prob}[i \in B] = (|A_i|/N)^5$ and thus the expected size of B is $\sum_{i=1}^m (|A_i|/N)^5 \geq m(\sum |A_i|/mN)^5 \geq \gamma^5 m$ (justify this inequality). Then, by the Cauchy-Schwartz inequality, $\mathbb{E}[|B|^2] \geq \mathbb{E}[|B|]^2 \geq \gamma^{10} m^2$.

If $|A_i \cap A_j| \leq \gamma^2 N/2$, then $\text{Prob}[i,j \in B] = \text{Prob}\{\{x_1, x_2, \dots, x_5\} \subset A_i \cap A_j\} < (\gamma^2/2)^5 = \gamma^{10}/32$. Define $C = \{i,j \in B \times B : |A_i \cap A_j| < \gamma^2 N/2\}$ so that $\mathbb{E}[|C|] < \gamma^{10} m^2/32$. Therefore $\mathbb{E}[|B|^2 - 24|C|] > \gamma^{10} m^2/4$, which implies that there exist $x_1, \dots, x_5$ for which $|B| > \gamma^5 m/2$ and $|C|/|B|^2 \geq 1/24 < 1/20$.

The following result is due to Balog and Szemerédi [1]. However the constants they gave were at least exponential in α; and it was Gowers who gave polynomial growth:

Theorem 6.6. *Let A be a subset of an abelian group, and $\alpha > 0$. Suppose that there are $\geq \alpha |A|^3$ quadruples $a, b, c, d \in A$ for which $a - b = c - d$. Then A contains a subset A' such that $|A'| \geq c|A|$ and $|A' - A'| \leq C|A|$ where c and C depend on α only (we can take $c = 2^{-12}\alpha^{10}$ and $C = 2^{26}\alpha^{-22}$).*

Proof. Let $|A| = n$ and $r(x) = (A * A)(x) = \#\{a, b \in A : x = a - b\}$. Trivially we have $\sum_x r(x) = n^2$ and $r(x) \leq r(0) = n$ as well as $r(x) = r(-x)$, and by hypothesis that $\sum_x r(x)^2 \geq \alpha n^3$. Let $X = \{x \neq 0 : \ r(x) \geq \alpha n/2\}$. Using the last inequality one can show that $|X| \geq \alpha n/2 - 1$.

We now construct a graph G with vertices corresponding to the elements of A, and the edges defined so that $a \sim_G b$ if and only if $a - b \in X$. Define $\Gamma_G(a) := \{b \in G : a \sim_G b\}$ to be the set of neighbours of a in G. Then $\sum_{a \in A} |\Gamma_G(a)| = \#\{a \neq b \in A : a \sim_G b\} = \sum_{x \in X} r(x) \geq |X|\alpha n/2 \geq \alpha n/2(\alpha n/2 - 1)$. Applying Lemma 6.5 with $m = N = n$ and $\gamma = \alpha^2/4$, we deduce that there exists $B \subset A$ with $|B| \geq \alpha^{10} n/2^{11}$ such that $|\Gamma_G(a) \cap \Gamma_G(b)| \geq \alpha^4 n/32$ for at least ninety-five percent of pairs $(a,b) \in B \times B$.

Now define a new graph H with vertex set B and edges defined so that $a \sim_H b$ if and only if $|\Gamma_G(a) \cap \Gamma_G(b)| \geq \alpha^4 n/32$. Let A' be the set of vertices of H with degree $\geq 3|B|/4$. Since the average degree in H is at least $(19/20)|B|$ we deduce that $|A'| \geq 4|B|/5$, and so $|A'| \geq \alpha^{10} n/(5 \cdot 2^9)$. Now define $R(x)$ to be the number of representations of x as $x_1 + x_2 + x_3 + x_4 - y_1 - y_2 - y_3 - y_4$ with $x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4 \in A$. We will show that if $x \in A' - A'$ then $R(x) \geq cn^7$ for a certain constant $\kappa > 0$ depending only on α. But then $\kappa n^7 |A' - A'| \leq \sum_{x \in A' - A'} R(x) \leq \sum_x R(x) = n^8$, and so $|A' - A'| \leq n/\kappa$ as desired.

Now for each $x \in A' - A'$ select $a, a' \in A'$ for which $x = a - a'$. If $a \sim_H b$ then there are at least $\alpha^4 n/32$ values of c for which $a - c, c - b \in X$. If $a - c \in X$ then there exists $\geq \alpha n/2$ pairs $x_1, y_1 \in A$ for which $a - c = x_1 - y_1$, and similarly there are $\geq \alpha n/2$ pairs $x_2, y_2 \in A$ for which $c - b = x_2 - y_2$. Taking all these representations for all such c we have $\geq (\alpha^4 n/32)(\alpha n/2)^2 = \alpha^6 n^3/2^7$ quadruples $x_1, y_1, x_2, y_2 \in A$ for which $a - b = x_1 - y_1 + x_2 - y_2$: To verify that these representations are distinct note that if we are given a, x_1, y_1 we can recover c as $a - x_1 + y_1$. If $a' \sim_H b$ then we can similarly find such representations for $b - a'$, and adding these together we find representations for $a - a'$ as desired which are all distinct as we can recover b in each case. Therefore $R(x) \geq |\Gamma_H(a) \cap \Gamma_H(a')|(\alpha^6 n^3/2^7)^2$; and we have $|\Gamma_H(a) \cap \Gamma_H(a')| \geq |\Gamma_H(a)| + |\Gamma_H(a')| - |H| \geq 3|B|/4 + 3|B|/4 - |B| = |B|/2$, and the result follows with $\kappa = \alpha^{22}/2^{26}$.

An alternate but useful version of this result is as follows:

Theorem 6.6A. *Let A be a subset of an abelian group and $S \subset A \times A$ with $|S| \geq \beta|A|^2$ for given $\beta > 0$. If $\#\{a + b : (a, b) \in S\} \leq (1/\beta)|A|$ then A contains a subset A' such that $|A'| \geq c|A|$ and $|2A'| \leq C|A|$ where c and C depend on β only (we can take $c = 2^{-12}\beta^{30}$ and $C = 2^{130}\beta^{-330}$).*

Proof. Let $B = \{a + b : (a, b) \in S\}$ and $r(x) = \#\{(a, b) \in S : a + b = x\}$, so that there are $\geq \sum_{x \in B} r(x)^2$ quadruples $a, b, c, d \in A$ for which $a - b = c - d$. Now by the Cauchy-Schwarz inequality we have $\sum_{x \in B} r(x)^2 \geq (\sum_{x \in B} r(x))^2/|B| = |S|^2/|B| \geq \beta^3|A|^3$. The construction of A' then follows from Theorem 6.6 with $\alpha = \beta^3$; and then $|A' + A'| \leq C^5|A|$ by Corollary 6.5.

Corollary 6.7. *Let $A \subset \mathbb{Z}^k$ with $|A| = m$ for which there are $\geq cm^3$ solutions to $a - b = c - d$ with $a, b, c, d \in A$. Then there exists a generalized arithmetic progression Q of cardinality at most Cm and dimension at most d such that $|A \cap Q| \geq cm$, where C and d depend only on c.*

Proof. In the previous result we have A' for which $|A' - A'| \ll |A'|$, and so $|2A'| \ll |A'|$ by Corollary 6.5. But then by the Freiman-Ruzsa Theorem there exists such a generalized arithmetic progression which contains A'.

The following result will be very useful in studying four-term arithmetic progressions in the next chapter.

Corollary 6.8. *Let $B \subset \mathbb{Z}/N\mathbb{Z}$ be a set of cardinality βN, and let $\phi : B \to \mathbb{Z}/N\mathbb{Z}$ be such that $\#\{a, b, c, d \in B : a - b = c - d \text{ and } \phi(a) - \phi(b) = \phi(c) - \phi(d)\} \geq \alpha N^3$. Then there exist constants γ and η, depending only on β and c, a $\mathbb{Z}/N\mathbb{Z}$-arithmetic progression P of cardinality at least N^γ and a linear function $\psi : P \to \mathbb{Z}/N\mathbb{Z}$ such that $\psi(s) = \phi(s)$ for at least $\eta|P|$ values of $s \in P$.*

Proof. Let $A = \{(a, \phi(a)) \in (\mathbb{Z}/N\mathbb{Z})^2 : a \in B\}$. We proceed as in Corollary 6.7 except now we need the Freiman-Ruzsa theorem in a finite group (see ***); in this case the relevant constants are all polynomial in α.

Write Q as $C(a_0, \ldots a_d; N_1, \ldots, N_d)$ with $N_1 N_2 \ldots N_d \geq cN$, and therefore there exists j for which $N_j \geq (cN)^{1/d}$ (we will take $j = d$ by re-ordering the N_i if necessary). Then we can think of Q as a union of arithmetic progressions mod a_d of length N_d with starting points given by $C(a_0, \ldots a_{d-1}; N_1, \ldots, N_{d-1})$. At least one of these, call it P must contain $(|A|/|Q|)N_d$ elements of A. But an arithmetic progression simply consists of points on a line and therefore the elements of $(a, \phi(a)) \in A \cap P$ must satisfy $\phi(a) = \lambda a + \nu$ for some $\lambda, \nu \in \mathbb{Z}/N\mathbb{Z}$.

Following Ruzsa's proof of Freiman's Theorem, we may take $\gamma = \alpha^K$ and $\eta = exp(-\alpha^{-K})$, where $K > 0$ is an absolute constant.

Additional Notes

We now look briefly at a selection of results concerning the structure of set sums themselves – this is the direct class of problems as opposed to the inverse problems of the type we have been studying until now. It is natural to consider sums of very dense sets. The following important theorem is due to Bourgain [6]:

Theorem 6.9. *Let $A, B \subset [N]$ of densities α and β respectively. Then $A+B$ contains an arithmetic progression of length at least exp$\big[c\alpha\beta\log N\big]^{1/3}$, where $c>0$ is an absolute constant.*

This theorem was shown to be almost best possible by an ingenious construction due to Ruzsa [28]:

Theorem 6.10. *Let $\varepsilon>0$. Then there exists $p(\varepsilon)$ such that for every prime $p>p(\varepsilon)$, there exists a symmetric set A of residues modulo p such that $|A|>(1/2-\varepsilon)p$ and $A+A$ contains no arithmetic progression of length exp$[\log p]^{2/3+\varepsilon}$.*

The set of subset sums

For $A=\{a_1<a_2<\cdots<a_k\}$ define the set of subset sums $A^{(*)}=\{\sum_{b\in B} b:\ B\subset A\}$. Erdős conjectured (for \$ 300) that $a_k \gg 2^k$ whenever the subset sums of A are distinct with $a_1\ge 1$ (note that $a_i=2^{i-1}$ is an example of such a sequence); and showed [9] that $a_k\ge 2^k/k$ for $k\ge 2$. Loosely speaking, Fraenkel [12] showed that the sequence of powers of two is the most dense sequence with subset sums distinct, under any reasonable definition of density: that is, for any such sequence, and any decreasing convex function f we have $\sum_{i=1}^n f(a_i)\le\sum_{i=1}^n f(2^i)$.

Using only elementary techniques Lev [23] showed the following strong structure theorem: If m is an integer for which $|A^{(*)}|\le m|A|-4m^3$ and $|A|\ge 8m^3$ then $A^{(*)}$ is a union of at most $m-1$ arithmetic progressions with the same common difference. He conjectured that this should hold with $|A^{(*)}|<m|A|-(m-1)^2$ which would be best possible; for if A is the multiset $\{1\}^{n-m+1}\cup\{x\}^{m-1}$ then $A^{(*)}=C(0,1,x;n-m+1,m-1)$ which can be shown to not be a union of $m-1$ arithmetic progressions with the same difference, if $n\ge 2m-2$ and x is large enough.

Let $A_k:=\{\sum_{b\in B} b : B\subset A,\ |B|=k\}$. Freiman [13] and, independently, Erdős and Sárközy [10] proved that if $A=\{1,2,\dots,N\}$ and $k>100(N\log N)^{1/2}$ then $A_k\supset\{rj : J\le j\le J+ck^2\}$ for some $c>0$.

Sumfree sets

A is a sum free set if there are no solutions to $a+a'=a''$ with $a,a',a''\in A$. If N is prime then one can easily show that there is a sumfree subset B of $A\subset\mathbb{Z}/N\mathbb{Z}$, with $|B|\ge(|A|-1)/3$: Select $x \pmod n$ so that $xA\cap(N/3,2N/3]$ is maximal, and then let $B=\{a\in A:\ ax\in(N/3,2N/3]\}$. Note that $b_1+b_2\ne b_3$ with $b_1,b_2,b_3\in B$ for if $b_i=xa_i$ then $a_1+a_2=a_3$ which is evidently impossible. Also $(N-1)|B|\ge$

$$\sum_{x=1}^{N-1}|xA\cap(N/3,2N/3]|=\sum_{a\in A^*}\#\{x : ax\in(N/3,2N/3]\}=|A^*|([2N/3]-[N/3])$$

where $A^* = A \setminus \{0\}$, and so $|B| \geq (|A|-1)([2N/3]-[N/3])/(N-1) \geq (|A|-1)/3$.

Alon and Kleitmann [1] improved this to the optimal coefficient (which cannot be 2/7 as in the notes).

Szemerédi's Theorem

We prove Szemerédi's Theorem [16] for four term arithmetic progressions, following Gowers [7]; in a proof that is analogous to our earlier proof of Roth's theorem. The difference is that a stronger notion of uniformity, namely *quadratic uniformity* is used here. Let $\mathbb{D} := \{z \in \mathbb{C} : |z| \leq 1\}$ be the unit disc, and throughout we consider functions $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$. If $A \subset \mathbb{Z}/N\mathbb{Z}$ then we may take $f(x) = f_A(x) = A(x) - |A|/N$ for example.

For $\alpha > 0$ we say that f is *α-uniform* if

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

Also call set A *α-uniform* if f_A is *α-uniform*.

Lemma 7.1. *Let $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$. If $\max_r |\hat{f}(r)| \leq \alpha^{1/2}N$ then f is α-uniform. Moreover, if f is α-uniform then $\max_r |\hat{f}(r)| \leq \alpha^{1/4}N$. Finally, if f is α-uniform then*

$$\sum_k \Big| \sum_s f(s)\overline{g(s-k)} \Big|^2 \leq \sqrt{\alpha} N^2 \|g\|_2^2$$

for any function $g : \mathbb{Z} \to \mathbb{C}$.

Proof. If $\max_r |\hat{f}(r)| \leq \alpha^{1/2}N$ then by using Parseval's identity we have

$$\sum_r |\hat{f}(r)|^4 \leq \max_r |\hat{f}(r)|^2 \sum_r |\hat{f}(r)|^2 \leq \alpha N^3 \sum_x |f(x)|^2 \leq \alpha N^4.$$

For the next part note that $|\hat{f}(r)|^4 \leq \sum_r |\hat{f}(r)|^4 \leq \alpha N^4$ and the result follows. For the final part, we know that if f is α-uniform then, using Parseval's identity twice and the Cauchy-Schwartz inequality,

$$\begin{aligned}\sum_k \Big| \sum_s f(s)\overline{g(s-k)} \Big|^2 &= \sum_k |(f * g)(k)|^2 = \frac{1}{N} \sum_r |\widehat{(f * g)}(r)|^2 \\ &= \frac{1}{N} \sum_r |\hat{f}(r)|^2 |\hat{g}(r)|^2 \leq \frac{1}{N} \Big(\sum_r |\hat{f}(r)|^4 \Big)^{1/2} \Big(\sum_r |\hat{g}(r)|^4 \Big)^{1/2} \\ &\leq \alpha^{1/2} N \sum_r |\hat{g}(r)|^2 = \alpha^{1/2} N^2 \sum_x |g(x)|^2.\end{aligned}$$

Define $\Delta(f;k)(x) = f(x)\overline{f(x-k)}$. We say that f is *quadratically α-uniform* if

$$\sum_k \sum_r |\hat{\Delta}(f;k)(r)|^4 \leq \alpha N^5.$$

We will show that sufficiently large quadratically α-uniform sets contain four-term arithmetic progressions. More generally we define

$$\Delta(f; k_1, k_2, \ldots, k_r) = \Delta(\Delta(f; k_1, k_2, \ldots, k_{r-1}); k_r).$$

The value is independent of the order in which we take the k_i. Moreover we have the identity

$$\sum_r |\hat{f}(r)|^4 = N \sum_{x,k,l} \Delta(f;k,l)(x), \tag{7.1}$$

as well as

$$\sum_k \sum_r |\hat{\Delta}(f;k)(r)|^4 = N \sum_{k,l} \Big| \sum_x \Delta(f;k,l)(x) \Big|^2. \tag{7.2}$$

Lemma 7.2. *If f is quadratically α-uniform, then f is $\alpha^{1/2}$-uniform.*

Proof. Take (7.1), apply Cauchy-Schwarz and then (7.2) to obtain

$$\Big(\sum_r |\hat{f}(r)|^4\Big)^2 \leq N^4 \sum_{k,l} \Big| \sum_x \Delta(f;k,l)(x) \Big|^2 = N^3 \sum_{k,r} |\hat{\Delta}(f;k)(r)|^4 \leq \alpha N^8.$$

We now repeat the Proposition 3.1 in this context:

Lemma 7.3. *Let $f_1, f_2, f_3 : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$. Suppose that f_3 is α-uniform. Then we have $\big|\sum_{a,d} f_1(a) f_2(a+d) f_3(a+2d)\big| \leq \alpha^{1/4} N^2$.*

Proof. If $S = \sum_{a,d} f_1(a) f_2(a+d) f_3(a+2d)$ then by Lemma 7.1 we have

$$\begin{aligned} |S| &= \Big| \sum_{a+c=2b} f_1(a) f_2(b) f_3(c) \Big| = \Big| \frac{1}{N} \sum_r \hat{f}_1(r) \hat{f}_2(-2r) \hat{f}_3(r) \Big| \\ &\leq \frac{1}{N} \max_r \hat{f}_3(r) \cdot \Big(\sum_r |\hat{f}_1(r)|^2\Big)^{1/2} \cdot \Big(\sum_r |\hat{f}_2(r)|^2\Big)^{1/2} \leq \frac{1}{N} \cdot \alpha^{1/4} N \cdot N^2 = \alpha^{1/4} N^2. \end{aligned}$$

Lemma 7.4. *Let $f_1, f_2, f_3, f_4 : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$ and $\alpha > 0$, with $(3, N) = 1$. Suppose f_4 is quadratically α-uniform. Then $|\sum_{a,d} f_1(a) f_2(a+d) f_3(a+2d) f_4(a+3d)| \leq \alpha^{1/8} N^2$.*

Proof. Let S be the sum we are estimating. Then, by the Cauchy-Schwarz inequality,

$$\begin{aligned} |S|^2 &\leq N \sum_a \Big| \sum_d f_1(a) f_2(a+d) f_3(a+2d) f_4(a+3d) \Big|^2 \\ &\leq N \sum_a |f_1(a)|^2 \, \Big| \sum_d f_2(a+d) f_3(a+2d) f_4(a+3d) \Big|^2 \\ &\leq N \sum_a \sum_{d,e} f_2(a+d) \overline{f_2(a+e)} f_3(a+2d) \overline{f_3(a+2e)} f_4(a+3e) \overline{f_4(a+3e)} \\ &= N \sum_k \sum_{b,d} \Delta(f_2;k)(b) \Delta(f_3;2k)(b+d) \Delta(f_4;3k)(b+2d), \end{aligned}$$

taking $k = d - e$ and $b = a + d$. Now define $\alpha(k) := N^{-4} \sum_r |\hat{\Delta}(f_4;k)(r)|^4$ so that $\Delta(f_4;k)$ is $\alpha(k)$-uniform, and $\sum_k \alpha(k) \leq \alpha N$ as f_4 is quadratically α-uniform. Therefore, by Lemma 7.3 (with f_j replaced by $\Delta(f_{j+1}, jk)$ for $j = 1, 2, 3$), we have

$$|S|^2 \leq N \sum_k \alpha(3k)^{1/4} N^2 \leq \alpha^{1/4}$$

by a simple optimization argument.

Theorem 7.5. *Let $A_1, A_2, A_3, A_4 \subset \mathbb{Z}/N\mathbb{Z}$ with $|A_i| = \delta_i N$. Suppose that A_3 is $\alpha^{1/2}$-uniform and A_4 is quadratically α-uniform. Then*

$$\left|\sum_{a,d} A_1(a)A_2(a+d)A_3(a+2d)A_4(a+3d) - \delta_1\delta_2\delta_3\delta_4 N^2\right| \leq 5\alpha^{1/8}N^2.$$

Proof. Set $f_i(x) = A_i(x) - \delta_i$. Replace the $A_i(\cdot)$ with $f_i(\cdot) + \delta_i$ in the above sum obtaining sixteen sums! We group the terms by the number of f_i involved: First, when there are no f_i involved we have the term $\sum_{a,d} \delta_1\delta_2\delta_3\delta_4 = \delta_1\delta_2\delta_3\delta_4 N^2$, which is thus the main term, and the rest are error terms. If we have just one f_i involved (say f_4) then the term is $\sum_{a,d} \delta_1\delta_2\delta_3 f_4(a+3d) = \delta_1\delta_2\delta_3 \sum_{b,d} f_4(b) = 0$ taking $b = a+3d$ for fixed d; and the same is true for the other such terms. If we have two f_i involved (say f_3 and f_4) then the term is $\sum_{a,d} \delta_1\delta_2 f_3(a+2d)f_4(a+3d) = \delta_1\delta_2 \sum_{b,c} f_3(c)f_4(b) = 0$ where $b = a+3d$ and then $c = b-d$, and the same is true for the other such terms. By Lemma 7.2 we know that A_4 is $\alpha^{1/2}$-uniform, as well as A_3: now if we have three f_i involved then at least one of them must be f_3 or f_4, so we may apply lemma 7.3 suitably modified. Finally if we have all f_i involved then we may apply lemma 7.4.

Corollary 7.6. *Fix $\delta > 0$. Suppose that $A \subset \{1, \ldots, N\}]$ with $|A| = \delta N$, and that A is quadratically α-uniform. If $\alpha \leq \delta^{32}/2^{72}$ and $N \geq 430/\delta^4$, then A contains an arithmetic progression of length four or we can find a subprogression where A has density at least $\frac{9}{8}\delta$.*

Proof. Let $A_1 = A_2 = A \cap [2N/5, 3N/5]$ and $A_3 = A_4 = A$. If $|A_1| \leq \delta N/10$ then either $A \cap [0, 2N/5]$ or $A \cap [3N/5, N)$ has at least $9\delta N/20$ elements and thus the second outcome. Now note that if $(a+(i-1)d)_N \in A_i$ for $i = 1, 2, 3$ and 4 then $a+(i-1)d \in A_i$ for $i = 1, 2, 3$ and 4. Thus the number of such four term arithmetic progressions is, by Theorem 7.5, $\geq (\delta^4/100 - 5\alpha^{1/8})N^2 > N$, so there exists such an arithmetic progression with $d \neq 0$.

We now turn to the case where f is *not* quadratically uniform. If A is the corresponding set of density δ, then we plan to show that A intersects a $\mathbb{Z}$-arithmetic progression $P \subset \{1, 2, \ldots, N\}$ of size at least N^d and such that $|A \cap P| \geq (\delta + \varepsilon)|P|$ where ε and d depend only on α and δ.

Lemma 7.7. *Suppose that f is not quadratically α-uniform. Then there exists a set B, of cardinality at least $\alpha N/2$, and a function $\phi : B \to \mathbb{Z}/N\mathbb{Z}$ such that*

$$\sum_{k \in B} |\hat{\Delta}(f;k)(\phi(k))|^2 \geq (\alpha/2)^{3/2}N^3.$$

Proof. Select $\phi(k)$ so that $|\hat{\Delta}(f;k)(\phi(k))|$ is maximized. Since f is not quadratically α-uniform, $\sum_k \sum_r |\hat{\Delta}(f;k)(r)|^4 > \alpha N^5$, and therefore must be a set B of more than $\alpha N/2$ values of k for which $\sum_r |\hat{\Delta}(f;k)(r)|^4 \geq \alpha N^4/2$. But then, by the second part of Lemma 7.1, we have $|\hat{\Delta}(f;k)(\phi(k))| \geq (\alpha/2)^{1/4}N$ for each $k \in B$ and the result follows.

Lemma 7.8. *Suppose that $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$, $B \subset \mathbb{Z}/N\mathbb{Z}$ and $\phi : \mathbb{Z}/N\mathbb{Z}$ is a function such that, for some $\alpha > 0$,*

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

Then there exist at least $\alpha^4 N^3$ quadruples $(a,b,c,d) \in B \times B \times B \times B$ such that $a+b=c+d$ and $\phi(a)+\phi(b)=\phi(c)+\phi(d)$.

Proof. We have that

$$\begin{aligned}\sum_{k \in B} |\hat{\Delta}(f;k)(\phi(k))|^2 &= \sum_{k \in B} \sum_{s,t} f(s)\overline{f(s-k)f(t)}f(t-k)e(\phi(k)(s-t)/N) \\ &= \sum_{k \in B} \sum_{s,u} f(s)\overline{f(s-k)f(s-u)}f(s-k-u)e(\phi(k)u/N) \\ &\leq \sum_{u,s} \Big| \sum_{k \in B} \overline{f(s-k)}f(s-k-u)e(\phi(k)u/N) \Big|,\end{aligned}$$

so $\alpha^2 N \leq \sum_u \gamma(u)$ by Cauchy-Schwarz, where

$$\gamma(u) := N^{-3} \sum_s \Big| \sum_{k \in B} \overline{f(s-k)}f(s-k-u)e(\phi(k)u/N) \Big|^2.$$

By the first part of Lemma 7.1 (with the roles of k and s swapped), we have that $e(\phi(k)u/N)$ is not $\gamma(u)^2$-uniform, and so $\sum_r \big|\sum_{k \in B} e((\phi(k)u + rk)/N)\big|^4 \geq \gamma(u)^2 N^4$ by definition. Now $\sum_u \gamma(u)^2 \geq \alpha^4 N$ since $\sum_u \gamma(u) \geq \alpha^2 N$, and therefore

$$\begin{aligned}\alpha^4 N^5 &\leq \sum_u \gamma(u)^2 N^4 \leq \sum_{u,r} \big| \sum_{k \in B} e((\phi(k)u + rk)/N)\big|^4 \\ &= \sum_{a,b,c,d \in B} \sum_u e\left(\frac{u(\phi(a)+\phi(b)-\phi(c)-\phi(d))}{N}\right) \sum_r e\left(\frac{r(a+b-c-d)}{N}\right) \\ &= N^2 \#\{(a,b,c,d) \in B^4 : \ a+b=c+d \text{ and } \phi(a)+\phi(b)=\phi(c)+\phi(d)\}\end{aligned}$$

Lemma 7.9. *Suppose that $\phi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ has at least αN^3 additive quadruples. Then there exist η, γ, depending only on α, and an arithmetic progression P of length at least N^γ such that for some λ and μ,*

$$\sum_{k \in P} |\hat{\Delta}(f;k)(\lambda k + \mu)|^2 \geq \eta N^2 |P|.$$

Proof. This follows from Corollary 6.8 with $\gamma = \alpha^K$ and $\eta = exp(-\alpha^{-K})$.

Lemma 7.10.1. *Let $f, g : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$. Let $A, B \subset \mathbb{Z}/N\mathbb{Z}$ and $C = A + B$. Then there exists r_x for each $x \in \mathbb{Z}/N\mathbb{Z}$ such that*

$$\sum_x \Bigg| \sum_{w\in x-A} g(w) e\left(\frac{\lambda w^2 + r_x w}{N}\right)\Bigg| \geq \frac{1}{N}\left(\frac{|B|}{|C|}\right)^{1/2} \sum_{a\in A}\Big|\sum_x f(x)\overline{g(x-a)}\ e\left(\frac{(2\lambda a+\mu)x}{N}\right)\Big|^2.$$

Proof. We have

$$\begin{aligned}
&\sum_{a\in A}\Big|\sum_x f(x)\overline{g(x-a)}\ e\left(\frac{(2\lambda a+\mu)x}{N}\right)\Big|^2 \\
&\qquad = \sum_{a\in A}\sum_{x,y} f(x)\overline{g(x-a)f(y)}g(y-a)\ e\left(\frac{(2\lambda a+\mu)(x-y)}{N}\right) \\
&\qquad = \frac{1}{|B|}\sum_{a\in A}\sum_{b\in B}\sum_{x,z} f(x)\overline{g(x-a)f(x-b-z)}g(x-a-b-z)\ e\left(\frac{(2\lambda a+\mu)(b+z)}{N}\right)
\end{aligned}$$

replacing y by $x-b-z$ for each $b \in B$. Now set $h_1(x) = f(x)e((\lambda x^2+\mu x)/N)$, $h_2(x) = g(x)e(\lambda x^2/N)$, $h_3(x) = f(x-z)e((\lambda x^2 + (\mu - 2\lambda z)x)/N)$, $h_4(x) = f(x-z)e((\lambda x^2 - 2\lambda z x)/N)$ for a given value of z, so that the above equals

$$\frac{1}{|B|}\sum_z e\left(\frac{\mu z}{N}\right)\sum_x h_1(x)\sum_{a\in A}\sum_{b\in B}\overline{h_2(x-a)h_3(x-b)}h_4(x-a-b).$$

The last two sums here can be rewritten as, with $C = A+B$,

$$\frac{1}{N}\sum_{a\in A}\sum_{b\in B}\sum_{c\in C}\overline{h_2(x-a)h_3(x-b)}h_4(x-c)\sum_r e\left(\frac{r(a+b-c)}{N}\right)$$

which is, in absolute value

$$\begin{aligned}
&\leq \frac{1}{N}\max_r\Bigg|\sum_{a\in A}\overline{h_2(x-a)}e\left(\frac{ra}{N}\right)\Bigg|\sum_r\Bigg|\sum_{b\in B}\overline{h_3(x-b)}e\left(\frac{rb}{N}\right)\Bigg|\Bigg|\sum_{c\in C}\overline{h_4(x-c)}e\left(\frac{-rc}{N}\right)\Bigg| \\
&\leq (|B||C|)^{1/2}\max_r\Bigg|\sum_{a\in A}h_2(x-a)e\left(\frac{-ra}{N}\right)\Bigg| = (|B||C|)^{1/2}\max_r\Bigg|\sum_{w\in x-A}g(w)e\left(\frac{\lambda w^2+rw}{N}\right)\Bigg|
\end{aligned}$$

by Cauchy-Schwarz, since $\sum_r |\sum_{d\in D} h(x-d)e(rd/N)|^2 = N\sum_{d\in D}|h(x-d|^2 \leq N|D|$. The result follows collecting the above estimates.

Lemma 7.10.2. *Let $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$. Let $\eta > 0$ and $P \subset \mathbb{Z}/N\mathbb{Z}$ be an $\mathbb{Z}/N\mathbb{Z}$-arithmetic progression with $|P| \leq \min\{3,\eta\}N/10$ such that*

$$\sum_{k\in P}|\hat{\Delta}(f;k)(2\lambda k+\mu)|^2 \geq \eta|P|N^2$$

for $\lambda, \mu \in \mathbb{Z}/N\mathbb{Z}$. Then, there exists a partition of $\mathbb{Z}/N\mathbb{Z}$ into translates $P_1, P_2, \dots, P_{q-1}$ of P or P with an endpoint removed, such that for each i we can find $r_i \in \mathbb{Z}/N\mathbb{Z}$ so that

$$\sum_i \left| \sum_{x \in P_i} f(x) e\left(\frac{\lambda x^2 + r_i x}{N} \right) \right| \geq \eta N/2.$$

Proof. Taking $g = f$ and $B = A = P$ in Lemma 7.10.1 so that $|B|/|C| > 1/2$, we obtain by our hypothesis,

$$\sum_x \left| \sum_{w \in x-P} f(w) e\left(\frac{\lambda w^2 + r_x w}{N} \right) \right| \geq \frac{\eta}{\sqrt{2}} |P| N.$$

If $P = \{v + is : 0 \leq i \leq m-1\}$ and suppose $N = qm + r$ with $0 \leq r \leq m-1$. With $v_j = (m+1)j$ for $0 \leq j \leq r$, and $v_j = mj + r$ for $r \leq j \leq q$ define $P_j = \{is : \ v_j \leq i < v_{j+1}\}$, so that $P_0, P_1, \dots P_{q-1}$ is a partition of $\mathbb{Z}/N\mathbb{Z}$. Therefore

$$\sum_y \sum_i \left| \sum_{w \in y+P_i} f(w) e\left(\frac{\lambda w^2 + r_{y,i} w}{N} \right) \right| \geq \frac{\eta}{\sqrt{2}} qmN - rN \geq \frac{\eta}{2} \ N^2$$

as $m < \min\{3, \eta\} N/10$, and the result follows by picking y so that the remaining sum is maximal, and changing the definition of $y + P_i$ to P_i.

Given $t \in \mathbb{Z}/N\mathbb{Z}$ define $|t| = N\|t/N\|$ to be the least residue of $t \pmod{N}$ in absolute value. For any set $S \subset \mathbb{Z}/N\mathbb{Z}$ define $\operatorname{diam}(S) = \max\{|x-y| : x, y \in S\}$.

Lemma 7.11. *Suppose that positive integers $l, m, r \leq N$ are given for which $l \leq (m/r)^{1/3}$. If P is a $\mathbb{Z}/N\mathbb{Z}$-arithmetic progression of length m and $\phi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ is a linear function then P can be partitioned into subprogressions $P_i, i \geq 1$ of lengths l or $l-1$, such that* $\operatorname{diam}(\phi(P_i)) \leq N/r$ *for each i.*

Proof. P maps linearly to $\{0, 1, 2, \dots m-1\}$ so we may take this as P, without loss of generality. By the pigeonhole principle, there exists $0 \leq i < j \leq rl$ such that $|\phi(j) - \phi(i)| \leq N/rl$; but as ϕ is linear we may take $d = j - i \leq rl$ such that $|\phi(d) - \phi(0)| \leq N/rl$. Set $Q = \{x, x+d, \dots, x+(l-1)d\}$. Then $|\phi(x+ld) - \phi(x)| \leq l|\phi(d) - \phi(0)| \leq N/r$ so $\operatorname{diam}(\phi(Q)) \leq N/r$. As each congruence class modulo d has size at least $m/d \geq m/rl \geq l^2$, we can split P into copies P_i of Q, differing in length by at most one.

In the next lemma, we apply Weyl's Theorem (Theorem 4.2):

Lemma 7.12. *Let $\phi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ be a quadratic polynomial and let P be a $\mathbb{Z}/N\mathbb{Z}$-arithmetic progression of length m. Then for any $l \leq m^{1/33}$, P can be partitioned into subprogressions P_i, $i \geq 1$, of lengths l or $l-1$, with* $\operatorname{diam}(\phi(P_i)) \ll N(l^{33}/m)^{1/16}$.

Proof. As in the previous result, P maps linearly to $\{0, 1, 2, \dots m-1\}$ so we may take this as P, without loss of generality, as long as we appropriately adjust ϕ. Then, for $\phi(x) = ax^2 + bx + c$ we can choose $d \leq D$ so that $|ad^2| \ll D^{-1/5} N$, by taking $\alpha = a/N$

in Theorem 4.2. Define $Q_i = \{i, i+d, \dots, i+(L-1)d\}$ with $l^3 \ll L \ll m/d$ and $0 \le i \le d-1$, and write $\phi(i+kd) - \phi(i+jd) = ad^2(k^2-j^2) + \psi_i(i+kd) - \psi_i(i+jd)$ where $\psi_i(x) := (2ai+b)(x-i)$ is a linear function. We now can apply Lemma 7.11, partitioning each Q_i into subprogressions R of length l or $l-1$ with $\mathrm{diam}(\psi_i(R)) \le N/r$ for each such R, provided $l \le (L/r)^{1/3}$. Then we have $\mathrm{diam}(\phi(R)) \ll (L^2/D^{1/5} + l^3/L)N$ taking $r \asymp L/l^3$. For given l, D this is minimized with $L = lD^{1/15}$, so we have $\mathrm{diam}(\phi(R)) \ll (l^2/D^{1/15})N$ provided $l^{30} \ll D \ll (m/l)^{15/16}$. Thus we may take $D \asymp (m/l)^{15/16}$ provided $l \ll m^{1/33}$.

Lemma 7.13. *Let $\phi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ be a quadratic polynomial and $m \le N$. For given $\eta > 0$ there exists integer $k \ll \eta^{-16/33} m^{32/33}$ such that $[0, m-1]$ can be partitioned into arithmetic progressions $P_1, P_2, \dots, P_k$, of lengths differing by at most 1, and such that, if $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{D}$ is any function with*

$$\left|\sum_{x=0}^{m-1} f(x) e\left(\frac{\phi(x)}{N}\right)\right| \ge \eta m, \quad \text{then} \quad \sum_{j=1}^{k} \Big|\sum_{x \in P_j} f(x)\Big| \ge \eta m/2.$$

Proof. By Lemma 7.12, for any $l \le m^{1/33}$ (taking $k = [m/l] + O(1)$), we find such $P_1, P_2, \dots, P_k$ for which $\mathrm{diam}(\phi(P_i)) \ll N(l^{33}/m)^{1/16}$. Select any $x_j \in P_j$ and note that if $x \in P_j$ then $|e(\phi(x)/N) - e(\phi(x_j)/N)| \ll \mathrm{diam}(\phi(P_i))/N \ll (l^{33}/m)^{1/16}$. Therefore $|\sum_{x \in P_j} f(x)| = |\sum_{x \in P_j} f(x) e(\phi(x)/N)| + O(l(l^{33}/m)^{1/16})$. Therefore, by the triangle inequality,

$$\eta m \le \sum_{j=1}^{k} \left|\sum_{x \in P_j} f(x) e\left(\frac{\phi(x)}{N}\right)\right| \le \sum_{j=1}^{k} \left|\sum_{x \in P_j} f(x)\right| + O(m(l^{33}/m)^{1/16}),$$

and the result follows provided $l \ll \eta^{16/33} m^{1/33}$

Szemerédi's Theorem. *There exists an absolute constant $c > 0$ such that if $A \subset \{1, 2, \dots, N\}$ with $|A| \gg N/(\log\log\log N)^c$ then A contains an arithmetic progression of length four.*

Proof. Regard A as a subset of $\mathbb{Z}/N\mathbb{Z}$. If A is quadratically $\alpha = \delta^{32}/2^{72}$-uniform, then the theorem is proved, by Corollary 7.6. Let $f(x) = A(x) - \delta$ and suppose f is not quadratically α-uniform. By Lemma 7.7, there exists a set B of cardinality at least $\alpha N/2$ and a function $\phi : B \to \mathbb{Z}/N\mathbb{Z}$ such that

$$\sum_{k \in B} |\hat{\Delta}(f; k)(\phi(k))|^2 \ge (\alpha/2)^{3/2} N^3.$$

By Lemma 7.8, ϕ has at least $(\alpha/2)^6 N^3$ additive quadruples and so, by Lemma 7.9, there exists an arithmetic progression P with $|P| \ge N^\gamma$ and

$$\sum_{k \in P} |\hat{\Delta}(f; k)(2\lambda k + \mu)|^2 \ge \eta |P| N^2$$

for some $\eta, \gamma > 0$ depending only on α.

By Ruzsa's proof of Freiman's Theorem, we may choose $\gamma = \alpha^K$ and $\eta \geq exp(-\alpha^{-K})$ where $K > 0$ is an absolute constant. By Lemma 7.10.2, we then have

$$\sum_i \left| \sum_{x \in P_i} f(x) e\left(\frac{\lambda x^2 + r_i x}{N}\right) \right| \geq \eta N/2.$$

where the P_i are as in Lemma 7.10.

Apply Lemma 7.13 in each P_i to obtain further progressions P_{ij}, of cardinalities differing by at most 1, and with average lengths $C\eta^{16/33}|P|^{1/33}$ (for some constant $C > 0$) and such that

$$\sum_i \sum_{j=1}^m \Big| \sum_{x \in P_{ij}} f(x) \Big| \geq \eta N/4.$$

A consequence of Lemma 7.12 is that we can insist that the P_{ij} are $\mathbb{Z}$-arithmetic progressions, except that (by Lemma 2.3) the average length of P_{ij} is $C|P|^{1/2.18.128}$, where $C > 0$ is a constant and no P_{ij} has more than twice this length.

Relabel the $P_{i,j}$s as $Q_1, Q_2, \ldots, Q_M$, where $M = N^{-\gamma/2.18.128}$ and the Q_i have average length $N^{\gamma/2.18.128}$. As $\sum f(x) = 0$, we have

$$\sum_i \Big(\Big| \sum_{x \in Q_i} f(x) \Big| + \sum_{x \in Q_i} f(x) \Big) \geq \eta N/4.$$

The contribution of Q_i with $|Q_i| \leq \sqrt{N/M}$ is at most $2N/\sqrt{M} \leq \eta N/8$, therefore there exists Q_i such that $|Q_i| \geq \sqrt{N/M}$ such that $\big| \sum_{x \in Q_i} f(x) \big| + \sum_{x \in Q_i} f(x) \geq \eta |Q_i|/8$. This implies that $\sum_{x \in Q_i} f(x) \geq \eta |Q_i|/16$.

So we have shown that there exists an arithmetic progression Q, of length at least $\sqrt{N/M} \geq N^{\gamma/4.18.128} = N^{\delta^c}$ such that $|A \cap Q| \geq (\delta + \exp[-\delta^c])|Q|$, where $c > 0$ is a constant. Rewriting this in terms of δ, a four-term arithmetic progression must be found when $\delta \geq (\log\log\log N)^{-c}$ for some $c > 0$.

Boring, technical Proofs

Proof of Lemma 2.10. Assume that the result in false and consider the counterexample with $|C_1|$ minimal.

We begin by supposing that there exists $c_1 \in C_1$ such that $c_1 + H_1 \subset S_2$. Note that the elements of $c_1 + H_1$ must belong to different cosets of H_2 else $c_1 + h_1 = c_2 + h_2$ and $c_1 + h_1' = c_2 + h_2'$ and so $h_1' - h_1 = h_2' - h_2 \in H_1 \cap H_2 = \{0\}$. Therefore we may write $S_2 = ((c_1 + H_1) \cup C_2') + H_2$. Now if for each $h_1 \in H_1$ there exists $h_2 \in H_2$ such that $(c_1 + h_1 + h_2) \notin S_1$ then $|S_2 \setminus S_1| \geq |H_1|$, and thus $|S_1 \cup S_2| \geq |S_1| + |H_1|$. Hence we may assume that there exists $h_1 \in H_1$ such that $(c_1 + h_1 + H_2) \subset S_1$. But S_1 is closed under addition by elements of H_1 and so $(c_1 + H_1 + H_2) \subset S_1$, and therefore we may write $S_1 = ((c_1 + H_2) \cup C_1') + H_1$. But now $S_j = (c_1 + H_1 + H_2) \cup S_j'$ where $S_j' = C_j' + H_j$, and the result follows from the induction hypothesis.

Now, since the elements of $c_2 + H_2$ must belong to different cosets of H_1 we deduce that $|(c_2 + H_2) \cap S_1| \leq |C_1|$ and so $|S_1 \cap S_2| \leq |C_1||C_2|$. Therefore $|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2| \geq |S_1| + |S_2| - |C_1||C_2|$.

We may now assume that for every $c_1 \in C_1$ there exists $h_1 \in H_1$ such that $c_1 + h_1 \notin S_2$; and therefore $|S_1 \cup S_2| \geq |S_2| + |C_1|$. Since all of the above arguments may be made with the roles of S_1 and S_2 exchanged, we also have $|S_1 \cup S_2| \geq |S_1| + |C_2|$.

Let us suppose that $|S_j| + |H_j| - 1 \geq |S_1 \cup S_2|$ for $j = 1, 2$. The equations of the last two paragraphs imply that $|H_1| - 1 \geq |C_2|(|H_2| - |C_1|)$ and $|H_2| - 1 \geq |C_1|(|H_1| - |C_2|)$, and then $|C_1| \leq |H_2| - 1$ and $|C_2| \leq |H_1| - 1$, respectively. Since all these terms are positive these can be combined: $|H_1| - 1 \geq |C_2|(|H_2| - 1) - |C_2|(|C_1| - 1) \geq |C_2||C_1|(|H_1| - |C_2|) - |C_2|(|C_1| - 1)$, so that $(|C_1||C_2| - 1)(|C_2| + 1 - |H_1|) \geq 0$. Therefore either $|C_1| = |C_2| = 1$, or $|C_2| = |H_1| - 1$ (and $|C_1| = |H_2| - 1$ by symmetry). In the first case $|S_1 \cup S_2| = |S_j| + |H_j| - |S_1 \cap S_2|$ for each j, so we get the result, with equality if and only if $S_1 \cap S_2 \neq \emptyset$. In the second case $|S_1 \cup S_2| \geq |S_1| + |S_2| - |C_1||C_2| = |S_j| + |H_j| - 1$ for each j, so we get the result, with equality if and only if $|S_1 \cap S_2| = |C_1||C_2|$. In this situation one has that $C_1 - C_2 \subset H_2 - H_1$; and that for every $c_1 \in C_1$ there exists a unique $h_1 \in H_1$ such that $c_1 + h_1 \notin S_2$. (Go on to fully classify when one gets equality in the lemma).

References

[BC] P.T. Bateman and S. Chowla, *Averages of character sums*, Proc. Amer. Math. Soc **1** (1950), 781-787.

[D] H. Davenport, *Multiplicative number theory*, Springer Verlag, New York, 1980.

[AK] N. Alon and R. Kleitman, *Sum free Sets of Integers*, A Tribute to Paul Erdős, (eds. Baker, Bollobás, Hajnal), Cam. Univ. Press, 1990, pp. 13–26.

[1] A. Balog and E. Szemerédi, *A statistical theorem of set addition*, Combinatorica **14** (1994), 263–268.

[BE] S. J. Benkoski and P. Erdos, Math. Comp **28** (1974), 617–623.

[2] Y. Bilu, *Structure of sets with small sumset, Structure Theory of Set Addition*, Astérisque **258** (1999), 77–108.

[3] N. N. Bogolyubov, Zap. Kafedry Mat. Fizi **4** (1939), 185.

[**] J. P. Bourgain, A Tribute to Paul Erdős, Cambridge Univ. Press, Cambridge, 1990, pp. 105–109.

[**] P. Erdős and A. Sárközy, *Arithmetic progressions in subset sums*, Disc. Math. **102** (1992), 249–264.

[4] P. Erdős and P. Turán, *On some sequences of integers*, J. London Math. Soc **11** (1936), 261–264.

[**] P. E. Fraenkel, *Integer sets with distinct subset sums*, Proc. Amer. Math. Soc **126** (1998), 3199–3200.

[**] G. A. Freiman, *New analytical results in subset sum problems*, Disc. Math. **114** (1993), 205–217.

[5] G. R. Freiman, *Foundations of a Structural Theory of Set Addition, Translations of Mathematical Monographs*, vol. 37, Amer. Math. Soc, Providence, R. I., USA.

[6] H. Fürstenburg, *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.

[7] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), 529–551.

[**] R. L. Graham, B. L. Rothschild, J. H. Spencer, *Ramsey Theory*, Wiley,, 1980.

[HR] H. Halberstam, H.-E. Richert, *Sieve methods*, Academic Press.

[8] A. W. Hales, R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222–229.

[**] Y. O. Hamidoune and J. A. D. da Silva, *Cyclic spaces for Grassman derivatives and additive theory*, Bull. London Math. Soc **26** (1994), 140–146.

[**] F. Hanson, J. M. Steele, F. Stenger, Proc. Amer. Math. Soc 66 (1977), 179–180.

[9] D. R. Heath-Brown, *Integer sets containing no arithmetic progressions*, J. London Math. Soc **35** (1987), 385–394.

[**] J. H. B. Kemperman, *On small sumsets in an abelian group*, Acta Math **103** (1960), 63–88.

[10] N. M. Korobov, *Exponential Sums and their Applications, Mathematics and its Applications* **80** (1992), Kluwer.

[**] V. Lev, *The structure of multisets with a small number of subset sums, Structure Theory of Set Addition*, Astérisque **258** (1999), 317–321..

[**] V. Lev, *On small sumsets in abelian groups, Structure Theory of Set Addition*, Astérisque **258** (1999), 317–321.

[**] E. Lipkin, *Subset sums of sets of residues, Structure Theory of Set Addition*, Astérisque **258** (1999), 187–193.

[11] K. F. Roth, *On certain sets of integers*, J. London Math. Soc **28** (1953), 245–252.

[**] I. Ruzsa, *Arithmetic progressions in sumsets*, Acta Arith. **60** (1991), 191–202.

[12] I. Ruzsa, *Generalized arithmetic progressions and sumsets*, Acta Math. Hungar. **65** (1994), , 379–388.

[13] I. Ruzsa, *An analog of Freiman's Theorem for abelian groups, Structure Theory of Set Addition*, Astérisque **258** (1999), 323–326.

[14] S. Shelah, *Primitive recursive bounds for van der Waerden Numbers*, J. Amer. Math. Soc **1** (1988), 683–697.

[**] C. F. Siegel, *Lectures on Geometric Number Theory.*

[**] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar **20** (1969), , 89–104.

[**] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. Hungar **27** (1975), 299–345.

[**] E. Szemerédi, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar **56** (1990), 155–158.

[**] G. Tenenbaum, *Introduction to analytic and probabilistic number theory, Cam. Studies in Advanced Math.*, vol. 46, Cam. Univ. Press, 1995.

[**] B. L. van der Waerden, *Beweis einer Baudetschen vermutung*, Nieuw Arch. Wisk **15** (1927), 212–216..

[**] R. C. Vaughan, *The Hardy-Littlewood Method, 2nd Ed. Cam. Tracts in Math.*, vol. 125, Cam. Univ. Press, 1997.

[**] I. M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, Trav. Inst. Math. Steklof **23** (1947).

[**] H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Annalen **77** (1913), 313–352..