AN INTRODUCTION TO ADDITIVE COMBINATORICS

Andrew Granville

Abstract. This is a slightly expanded write-up of my three lectures at the Additive Combinatorics school. In the first lecture we introduce some of the basic material in Additive Combinatorics, and in the next two lectures we prove two of the key background results, the Freiman-Ruzsa theorem and Roth's theorem for 3-term arithmetic progressions.

Lecture I: Introductory material

1. Basic Definitions.
2. If $A+B$ is small then A and B are...?
3. Densities.
4. The Dyson transformation.
5. Simple inequalities for sizes of sumsets.
6. The Freiman-Ruzsa theorem in groups, where the elements have bounded order.
7. The Balog-Szemeredi-(Gowers) theorem.
8. Discrete Fourier transforms.
9. Sum-Product formulas.

Lecture II: The Freiman-Ruzsa theorem

1. If $A+B$ is small what do A and B look like ?
2. The Geometry of numbers.
3. Structure of a Bohr set.
4. Freiman homomorphisms.
5. The Freiman-Ruzsa theorem.

Lecture III: Uniform distribution, Roth's theorem and beyond

1. Uniform distribution mod one.
2. Uniform distribution mod N.
3. Roth's theorem.
4. Large Fourier coefficients.
5. Four term arithmetic progressions.

Lecture I: Introductory material

I.1. Basic Definitions

Let A and B be subsets of G, an additive group. Typically we work with the integers $\mathbb{Z}$, or the integers mod N, that is $(\mathbb{Z}/N\mathbb{Z})$, though sometimes with other groups like $\mathbb{R}$ or $\mathbb{Z}^k$. The *sumset* of A and B is defined by

$$A+B := \{g \in G : \text{ There exist } a \in A, b \in B \text{ such that } g = a+b\}.$$

Typically we write $A+B = \{a+b : \ a \in A, b \in B\}$ with the understanding that elements are not repeated in $A+B$. For example, $\{1,2,3\}+\{1,3\} = \{2,3,4,5,6\}$.
The addition of sets, "+", is commutative if $(G,+)$ is commutative. It is also associative, and it is distributive over unions, that is, $A+(B\cup C) = (A+B)\cup(A+C)$.

Other important definitions include

$$\begin{aligned} kA &:= A+A+\cdots+A; \\ b+A &= \{b\}+A, \text{ a } \textit{translate} \text{ of } A; \\ A-B &= \{a-b : a \in A, b \in B\}; \\ k\diamond A &= \{ka : a \in A\}, \text{ a } \textit{dilate} \text{ of } A; \\ \text{and } \ A\diamond B &= \{ab : a \in A, b \in B\}. \end{aligned}$$

Having given all this notation we note that we will abuse it by writing $N\mathbb{Z}$ instead of $N\diamond\mathbb{Z}$, for the integers divisible by N.

Warm Up Exercises

1.1. Show that if $A-A = \{0\}$ then $|A| = 1$.

1.2. Show that $k\diamond A \subseteq kA$; when are they equal?

1.3. Show that $|A| \le |A+B| \le |A||B|$.

1.4. When do we have $|A+B| = |A||B|$?

If $a \in A$ then $A+\{a\} \subset A+A$ and $A-\{a\} \subset A-A$, so that $|A+A|, |A-A| \ge |A|$. Also $|A+A|, |A-A| \le |A\times A| \le |A|^2$.

1.5. Improve these upper bounds for $|A+A|$ and for $|A-A|$.

One of our main objectives is to study the size and structure of sumsets in $\mathbb{Z}$. Above we have considered finite sets, but there is an interesting history of results on summing infinite sets: Define $A_{\ge m} := \{n \in A : \ n \ge m\}$. A set of integers A is a *basis of order* h if $h(A\cup\{0\}) \supseteq \mathbb{Z}_{\ge m}$. We now give several well-known examples. Let $\mathbb{P}$ be the set of primes.

$$\begin{aligned} \text{Lagrange's theorem: } & \ 4\{n^2 : n \in \mathbb{Z}\} = \mathbb{Z}_{\ge 0} \\ \text{Goldbach's conjecture : } & \ 2\mathbb{P}_{\ge 3} = 2\diamond\mathbb{Z}_{\ge 3} \ \textit{ or } \ 3(\mathbb{P}\cup\{0\}) = \mathbb{Z}_{\ge 2}\cup\{0\} \\ \text{Generalized twin prime conjecture : } & \ \mathbb{P}_{\ge m} - \mathbb{P}_{\ge m} = 2\diamond\mathbb{Z} \ \text{ for all } m. \end{aligned}$$

Therefore Lagrange's theorem states that the squares form a basis of order 4, and Goldbach's conjecture postulates that the primes form a basis of order 3.

I.2. If $A+B$ is small then A and B are...?

Suppose that A and B are finite sets of integers, say A is $a_1 < a_2 < \cdots < a_r$, and B is $b_1 < b_2 < \cdots < b_s$. Then $A+B$ contains the $r+s-1$ distinct elements

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

so that

$$|A+B| \geq |A|+|B|-1. \tag{1.1}$$

Can we have equality in (1.1)? That is, what if $|A+B| = |A|+|B|-1$? We will write down another list $r+s-1$ distinct elements of $A+B$, namely

$$a_1+b_1 < a_2+b_1 < a_2+b_2 < \cdots < a_2+b_{s-1} < a_2+b_s < \cdots < a_r+b_s.$$

If $|A+B| = r+s-1$, then the terms in each list must be the same and so we have $a_1+b_2 = a_2+b_1$, and $a_1+b_3 = a_2+b_2$, etc., implying that $a_2-a_1 = b_2-b_1 = b_3-b_2 = \ldots$. In fact we can deduce that A and B are both arithmetic progressions with the same common difference; that is there exists a non-zero integer d such that

$$A = \{a+id : 0 \leq i \leq I-1\} \text{ and } B = \{b+jd : 0 \leq j \leq J-1\}.$$

Thus A and B are highly structured. However if A is a large subset of $\{a+id : 0 \leq i \leq I-1\}$ and B is a large subset of $\{b+jd : 0 \leq j \leq J-1\}$, then we expect that $|A+B| = |A|+|B|+\Delta$ for some small Δ, yet A and B may not have much internal structure. The key thing is that they are both large subsets of arithmetic progressions with the same common difference.

Another interesting case is given by

$$\begin{aligned} A &= \{1,2,\ldots,10,101,102,\ldots,110,201,202,\ldots,210\} \\ &= 1+\{0,1,\ldots,9\}+100\diamond\{0,1,2\}, \\ B &= 3+\{0,1,\ldots,7\}+100\diamond\{0,1,\ldots,4\}, \\ \text{and } A+B &= 4+\{0,1,2,\ldots,16\}+100\diamond\{0,1,\ldots,6\}, \end{aligned}$$

so that $|A| = 30, |B| = 40$ and $|A+B| = 119$. These are examples of a *generalized arithmetic progression* (GAP):

$$C := \{a_0+a_1n_1+a_2n_2+\cdots+a_kn_k : \ 0 \leq n_j \leq N_j-1 \text{ for } 1 \leq j \leq k\},$$

where $N_1, N_2, \dots, N_k$ are integers ≥ 2. This GAP is said to have *dimension* k and *volume* $N_1N_2 \dots N_k$; and is *proper* if its elements are distinct.

Most questions about the structure of A and B, when $A+B$ is small, are open! We study the structure of A when $A+A=2A$ is small (i.e., the case $B=A$). For a GAP C we have $|2C| < 2^k|C|$; and, indeed, if $A \subset C$ with $|A| \geq \delta|C|$ then

$$|2A| \leq |2C| < 2^k|C| \leq (2^k/\delta)|A|.$$

What about the converse? If $|2A|$ is a small multiple of $|A|$ then what possible A are there? A rather daring guess is that the only possible such A are large subsets of GAPs; and indeed this is the Freiman-Ruzsa theorem which we will prove in our next lecture.

The Freiman-Ruzsa theorem. *If $|2A|$ is "small" then A is a "large" subset of a GAP.*

Precise quantifiers in an explicit version of this result are complicated and best left till we study it in more detail.

I.3. Densities

The Schnirelmann density of a set A of integers is given by

$$\sigma(A) := \inf_{n \geq 1} \frac{\#\{a \in A : \ 1 \leq a \leq n\}}{n},$$

so that $A(n) \geq n\sigma(A)$ for all $n \geq 1$. It is easy to see, by the pigeonhole principle that if $0 \in A \cap B$ and $\sigma(A)+\sigma(B) \geq 1$ then $A+B \supseteq \mathbb{Z}_{\geq 0}$. By counting the elements in $A+B$ of the form $a_i + b_j$ with $a_i \leq a_i + b_j < a_{i+1}$, Schnirelmann proved that if $1 \in A$ and $0 \in B$ then

$$\sigma(A+B) \geq \sigma(A) + \sigma(B) - \sigma(A)\sigma(B).$$

This is more usefully rewritten as $(1-\sigma(A+B)) \leq (1-\sigma(A))(1-\sigma(B))$, since then we see that $(1-\sigma(hA)) \leq (1-\sigma(A))^h$. The last two results thus imply that if $1 \in A$ and $\sigma(A) > 0$ then A is a basis of order $2h$ where the integer h is chosen so that $(1-\sigma(A))^h \leq 1/2$. (Note that $\sigma(A) > 0$ implies that $1 \in A$; and that some condition like $1 \in A$ is necessary to avoid A, and hence hA, being a subset of the even integers.)

The lower density $\underline{d}(A)$ is defined by

$$\underline{d}(A) := \liminf_{n \to \infty} \frac{\#\{a \in A : \ 1 \leq a \leq n\}}{n},$$

We will prove that if $\underline{d}(A) > 0$ then for all $\epsilon \in (0, \underline{d}(A))$ there exists $r = r_\epsilon$ such that $\sigma(A_{(r)}) \geq \underline{d}(A) - \epsilon$, where $A_{(r)} = \{a - r : \ a \in A, \ a > r\}$. There exists an integer n_ϵ such that if $n \geq n_\epsilon$ then $\#\{a \in A : \ 1 \leq a \leq n\} \geq (\underline{d}(A) - \epsilon)n$. If there exists any $n \geq n_\epsilon$ with $\#\{a \in A : \ 1 \leq a \leq n\} < \underline{d}(A)n$ then there must be an $n \geq n_\epsilon$,

say $n = m_\epsilon$, with $\rho_\epsilon := \#\{a \in A : \ 1 \le a \le n\}/n$ minimal. Hence if $n > m_\epsilon$ then $\#\{a \in A : \ m_\epsilon < a \le n\} \ge \rho_\epsilon(n - m_\epsilon) \ge (\underline{d}(A) - \epsilon)(n - m_\epsilon)$. On the other hand if $\#\{a \in A : \ 1 \le a \le n\} \ge \underline{d}(A)n$ for all $n \ge n_\epsilon$ then either $\sigma(A) \ge \underline{d}(A)$, or there exists a maximal r_ϵ (which is necessarily $< n_\epsilon$) with $\#\{a \in A : \ 1 \le a \le r_\epsilon\} < \underline{d}(A)r_\epsilon$, and the result follows.

A straightforward sieve argument implies that at least 1/4 of the even integers can be written as the sum of two primes; that is $\underline{d}(2\mathbb{P}_{\ge 3}) \ge 1/8$. Using the argument of the previous paragraph, and Schnirelmann's theorem, one can prove that the primes are a basis of order 11 (or less). It can also be shown that the k-th powers of integers form an additive basis.

For a finite set of integers S define *the cube* $\overline{S}$ by

$$\overline{S} := \left\{ \sum_{s \in S} \epsilon_s s : \ \epsilon_s \in \{-1, 0, 1\} \ \text{ for all } s \in S \right\},$$

which is a GAP of dimension $|S|$ and volume $3^{|S|}$.

Theorem. *If A is a set of integers with $\underline{d}(A) > 0$, then there exists a finite set of integers S such that $A - A + \overline{S} = \mathbb{Z}$.*

Proof. If $A - A \ne \mathbb{Z}$ then there exists $m \notin A - A$, and so A and $m + A$ are disjoint. Let $A_1 = A \cup (m + A)$, so that $\underline{d}(A_1) = 2\underline{d}(A)$ and $A_1 - A_1 = A - A + \overline{\{m\}}$. If this is not $\mathbb{Z}$, define $A_2, A_3, \dots$ Therefore $|S| \le k$ where k is the largest integer for which $2^k \underline{d}(A) \le 1$.

Since $\underline{d}(2\mathbb{P}_{\ge 3}) \ge 1/8$, we can deduce that there exists a set S_1 of no more than three integers for which

$$\mathbb{Z} = 2\mathbb{P}_{\ge 3} - 2\mathbb{P}_{\ge 3} + \overline{S_1}.$$

It is interesting to determine how small a set one needs to "complete" a given set in this manner. Thus above we added $\overline{S_1}$ to $2\mathbb{P}_{\ge 3} - 2\mathbb{P}_{\ge 3}$ to obtain $\mathbb{Z}$, though we believe that $\mathbb{P}_{\ge 3} - \mathbb{P}_{\ge 3} + \{0, 1\} = \mathbb{Z}$. For sums of squares we have $4\{n^2 : n \in \mathbb{Z}\} = \mathbb{Z}_{\ge 0}$; and one can show that $3\{n^2 : n \in \mathbb{Z}\} + \{0, 2\} = \mathbb{Z}_{\ge 0}$. A challenge is to find "thin" sets B and C for which $2\{n^2 : n \in \mathbb{Z}\} + B = \mathbb{Z}_{\ge 0}$, and for which $\mathbb{P} + C = \mathbb{Z}_{\ge 0}$.

I.4. The Dyson transformation

Many of the early papers in "additive number theory" were characterized by complicated, seemingly ad hoc, arguments. However, once Freeman Dyson introduced a simple map between pairs of sets, researchers found new, cleaner arguments in many of the essential questions: For $e \in A$ let $B_e := \{b \in B : \ b + e \notin A\}$, and define the *Dyson transformation* of A, B with respect to e to be

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

Notice that $B_e \subseteq B$ and $(e+B_e)\cap A = \emptyset$. There are several other observations to be made besides:

$$e+\delta_e(B) \subseteq A \subseteq \delta_e(A), \text{ and } |\delta_e(A)|+|\delta_e(B)| = |A|+|B|;$$
$$A\cap(e+B) = e+\delta_e(B) = \delta_e(A)\cap(e+\delta_e(B)),$$
$$\text{and } A\cup(e+B) = \delta_e(A) = \delta_e(A)\cup(e+\delta_e(B)),$$
$$\text{as well as the non-trivial } \delta_e(A)+\delta_e(B)\subseteq A+B.$$

Using a sequence of Dyson transformations one can easily prove Mann's "best possible" improvement of Schnirelmann's theorem:

Mann's theorem. *If $0\in A\cap B$ then*

$$\sigma(A+B)\ge \min\{1,\sigma(A)+\sigma(B)\}.$$

Note that this result does not extend directly to questions about lower density; that is, $\underline{d}(A+B)\ge\min\{1,\underline{d}(A)+\underline{d}(B)\}$ is not true in general: For example, if $A=B=\{n\equiv 0 \text{ or } 1 \pmod m\}$ then $A+B=\{n\equiv 0,1 \text{ or } 2 \pmod m\}$. So, to understand set addition with respect to lower density, we certainly need to understand set addition mod N. Here the key result is

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

$$|A+B|\ge\min\{N,|A|+|B|-1\}.$$

Proof. By induction on $|B|$: If $|B|=1$ then $B=\{0\}$ so $A+B=A$ which is okay. We may assume that $1\le|A|\le N-1$. Now $A+B\ne A$ else for each $b\in B$, for all $a\in A$ there exists $a'\in A$ such that $a+b\equiv a' \pmod N$. Running through all $a\in A$ we obtain all $a'\in A$, and so taking the sum over all $a\in A$ we get $|A|b\equiv 0 \pmod N$. By selecting non-zero $b\in B$ we have $(b,N)=1$, and so N divides $|A|$, which is impossible.

So take $e\in A$ for which $e+b\notin A$. By the induction hypothesis the result holds for the pair $\delta_e(A),\delta_e(B)$ (which are non-empty since $A\subseteq\delta_e(A)$ and $0\in\delta_e(B)$), so that

$$|A+B| \ge |\delta_e(A)+\delta_e(B)|\ge\min\{N,|\delta_e(A)|+|\delta_e(B)|-1\}=\min\{N,|A|+|B|-1\}.$$

Corollary. *If $A,B\subseteq\mathbb{Z}/p\mathbb{Z}$ with p prime then $|A+B|\ge\min\{p,|A|+|B|-1\}$.*

There are just three cases in which we get equality (that is, $|A+B|=|A|+|B|-1$) when $A+B$ is a proper subset of $\mathbb{Z}/p\mathbb{Z}$:

- Either A or B has just one element (that is, $|A|=1$ or $|B|=1$); or
- A and B are segments of arithmetic progressions with the same common difference (that is, $A=a+d\diamond\{0,1,\dots,r-1\}$, and $B=b+d\diamond\{0,1,\dots,s-1\}$ for some $r+s\le p$); or
- A and B are selected maximally so that $d\notin A+B$ (that is, $A\cup(d-B)$ is a partition of $\mathbb{Z}/p\mathbb{Z}$ for some integer d).

I.5. Simple Inequalities for sizes of sumsets

The Freiman-Ruzsa theorem tells us that if $|A+A| < C|A|$ then A is a large subset of a d-dimensional GAP, G, for some d that can be bounded as a function of C. This implies that $A-A$ is a large subset of $G-G$, a GAP that is at most twice as large (in each direction) as G, and so $|A-A| \le 2^d|G| \le 2^d C'|A|$ for some constant C' which depends only on C. Similarly $kA-\ell A$ is a large subset of $kG-\ell G$, also a d-dimensional GAP, and so $|kA-\ell A| \le (k+\ell)^d C'|A|$.

In this section we derive consequences of this type directly, without using the relatively deep Freiman-Ruzsa theorem; that is, our objective is to prove that if $|A+A| < C|A|$ then $|kA-\ell A| \le C_{k,\ell}|A|$ for some constant $C_{k,\ell}$ which depends only on C, k, ℓ. We will see that there are several easy approaches to this problem. When we prove the Freiman-Ruzsa theorem during the next lecture, we will use such inequalities in our proof. We start with the most basic question of this type:

I.5.1. The relationship between $A+A$ and $A-A$. We will prove a little later that

$$\frac{1}{2} \le \log\left(\frac{|A+A|}{|A|}\right) \Big/ \log\left(\frac{|A-A|}{|A|}\right) \le 3; \tag{1.2}$$

we are interested in determining the strongest possible form of each of these inequalities. We give two examples

- For $A=\{0,1,3\}$ we have $A+A=\{0,1,2,3,4,6\}$ and $A-A=\{-3,-2,-1,0,1,2,3\}$, so that $|A+A|=6<|A-A|=7$.
- For $A=\{0,2,3,4,7,11,12,14\}$ we have $A+A=[0,28]\cap\mathbb{Z}\setminus\{1,20,27\}$ and $A-A=([-14,14]\cap\mathbb{Z})\setminus\{-13,-6,6,13\}$, so that $|A-A|=25<|A+A|=26$.

These isolated examples can be made into arbitrarily large examples by using the Cartesian product: The idea simply is to take $B=A^{(k)}=A\times\cdots\times A$, so in the first case $|B+B|=6^k<|B-B|=7^k$. One might object that B is not a subset of the integers but in fact the bijection $B\leftrightarrow C$ defined by $(a_0,\dots,a_{k-1})\longleftrightarrow a_0+a_1 7+\cdots+a_{k-1}7^{k-1}$ is also a bijection, when correctly interpreted, between the sets $B+B$ and $C+C$, and between $B-B$ and $C-C$. This map is called a Freiman 2-isomorphism (the "2" since it remains a bijection when we add two elements of our set); we will discuss this in detail in our next lecture. We thus conclude from our examples that the constant "$\frac12$" in (1.2) may not be increased beyond $\frac{\log(6/3)}{\log(7/3)}=.81806\dots$; and the constant "3" in (1.2) may not be decreased beyond $\frac{\log(26/8)}{\log(25/8)}=1.03442\dots$ A better example for the lower bound comes from taking $A=\{1,2,2^2,\dots,2^5\}$ so that $|A|=6$, $|A+A|=21<|A-A|=31$ as $\frac{\log(21/6)}{\log(31/6)}=.762843\dots$

I.5.2. Some first bounds. We begin by establishing that for any finite sets A,B,C inside an additive group G (whether commutative or not) we have

$$|A-C|\,|B| \le |A-B|\,|B-C|, \tag{1.3}$$

by showing that there is an injection $\phi : (A-C)\times B \to (A-B)\times(B-C)$: For each $\lambda \in A-C$ fix $a_\lambda \in A, c_\lambda \in C$ such that $a_\lambda - c_\lambda = \lambda$. Then define $\phi(\lambda, b) = (a_\lambda - b, b - c_\lambda)$. To see that this is an injection we show how to reconstruct λ and b given $a_\lambda - b$ and $b - c_\lambda$: First we have $\lambda = (a_\lambda - b) + (b - c_\lambda)$, so we obtain a_λ, c_λ, and thus b.

We now use (1.3) to obtain all sorts of useful inequalities:

— Taking $C = A$ gives $|A - A| \le |A - B|^2/|B|$.

— Then taking $B = -A$ gives $\frac{|A-A|}{|A|} \le \left(\frac{|A+A|}{|A|}\right)^2$, which is the lower bound in (1.2).

— Next taking $A = rA, B = -A$ with $C = -sA$ and then $C = sA$ implies:

$$|(r+s)A|\ |-A| \le |(r+1)A|\ |sA - A|,$$
$$|rA - sA|\ |-A| \le |(r+1)A|\ |(s+1)A|.$$

With the choices $r = n-2, s = 2$, and $r = 2, s = 1$, respectively, we obtain

$$\frac{|nA|}{|A|} \le \frac{|(n-1)A|}{|A|}\ \frac{|2A-A|}{|A|} \le \frac{|(n-1)A|}{|A|}\ \frac{|3A|}{|A|}\ \frac{|2A|}{|A|}.$$

We deduce that, for $n \ge 3$,

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

$$\text{and then that} \qquad \frac{|rA-sA|}{|A|} \le \left(\frac{|3A|}{|A|}\right)^{r+s-2}\left(\frac{|2A|}{|A|}\right)^{r+s-4} \quad \text{for all } r, s \ge 2.$$

This is almost what we asked for! We wanted bounds as a function of r, s and $|2A|/|A|$, and instead we have very easily obtained bounds in terms of these variables and $|3A|/|A|$. So the question becomes whether one can find an easy way to bound $|3A|/|A|$ in terms of $|2A|/|A|$? Certainly such bounds can be proved by straightforward combinatorial arguments, but we know of no proof that is quite so simple as that above. (Taking $r = 1, s = 2$ in the inequalities above, we see that we could replace $3A$ by $2A - A$ in these last few comments.) Relationships between these different quantities are explored in detail by Imre Ruzsa in his article in this volume [R3].

I.5.3. Representation numbers. Denote the number of representations of n as a sum $a + b, a \in A, b \in B$ by

$$r_{A+B}(n) := \#\{(a,b) : a \in A, b \in B, n = a + b\},$$

and similarly $r_{kA+\ell B}(n)$, etc. There are several straightforward but useful identities: First, by counting all ordered pairs $(a,b), a \in A, b \in B$ we obtain

$$|A||B| = \sum_x r_{A+B}(x) = \sum_y r_{A-B}(y).$$

The solutions to $a+b=a'+b'$ with $a,a'\in A, b,b'\in B$ are the same as the solutions to $a-b'=a'-b$, which are the same as the solutions to $a-a'=b'-b$, and so

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

Therefore we obtain, by the Cauchy-Schwarz inequality, that

$$(|A||B|)^2 = \left(\sum_x r_{A\pm B}(x)\right)^2 \le |A\pm B| E(A,B).$$

Also note that

$$E(A,B) \le \begin{cases} \max_x r_{A+B}(x) \sum_x r_{A+B}(x) = |A||B| \max_x r_{A+B}(x), \\ |A+B| \max_x r_{A+B}(x)^2. \end{cases}$$

Now we show that

$$r_{A+B}(x) \le \frac{|A-B|^2}{|A+B|}$$

by exhibiting, for a given value of $x\in A+B$, an injection from $R_{A+B}(x)\times(A+B)\to (A-B)\times(A-B)$, where $R_{A+B}(x)$ is the set of representations of x as $a+b, a\in A, b\in B$. So fix a representation $a+b=x$, and for any $\lambda\in A+B$ fix $a_\lambda\in A, b_\lambda\in B$ such that $a_\lambda+b_\lambda=\lambda$. The map $(a,b,\{a_\lambda,b_\lambda\})\to (a-b_\lambda, a_\lambda-b)$ is, indeed, an injection, because we can reconstruct our pre-image by noting that $\lambda = x+(a_\lambda-b)-(a-b_\lambda)$, from which we obtain a_λ and b_λ, then $a=(a-b_\lambda)+b_\lambda$ and $b=x-a$.

Combining the last three displayed equations we obtain

$$|A+B| \le \frac{|A-B|^2}{\max_x r_{A+B}(x)} \le \frac{|A-B|^2|A||B|}{E(A,B)} \le \frac{|A-B|^3}{|A||B|}.$$

Taking $B=A$ gives the upper bound in (1.2).

I.5.4. Disjoint unions. We start with an idea of Ruzsa that we shall see again.

Lemma 1. *There exists $X\subset B$ with $|X|\le |A+B|/|A|$ such that $B\subset A-A+X$.*

Proof. Choose $X\subset B$ to be as large as possible so that the sets $\{A+x:\ x\in X\}$ are disjoint. The union of these sets contains exactly $|A||X|$ elements, all in $A+B$, which implies that $|A|\cdot|X|\le |A+B|$.

Now if $b\in B$ then $(A+b)\cap(A+x)\ne\emptyset$ for some $x\in X$, else X would not have been maximal, so $b\in A-A+x$, and we are done.

Take $B = A - 2A$ in Lemma 1 to get $2A - A \subset A - A + X$ where $X \subset 2A - A$ with $|X| \le |2A - 2A|/|A|$ (replacing X by $-X$ for convenience). Add A to both sides to get

$$3A - A \subset 2A - A + X \subset A - A + 2X$$

and then, proceeding by induction, we obtain

$$mA - nA \subset A - A + (m-1)X - (n-1)X \quad \text{for all } m, n \ge 1. \tag{1.4}$$

Now, since each $|rX| \le |X|^r$, and as $|X| \le \frac{|2A-2A|}{|A|}$, we deduce that

$$\frac{|mA - nA|}{|A|} \le \frac{|A - A|}{|A|} \left(\frac{|2A - 2A|}{|A|} \right)^{m+n-2} \quad \text{for all } m, n \ge 1. \tag{1.5}$$

Another argument based on something similar to, but more complicated than, the above lemma (see Lemma 2.17, Proposition 2.18 and Corollary 2.19 of [TV]), leads to the inequality

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

Taking $B = A$ in this formula, and then the first inequality in (1.2), we deduce from (1.5) that

$$\frac{|mA - nA|}{|A|} \le \left(\frac{|2A|}{|A|} \right)^{6m+6n-10} \quad \text{for all } m, n \ge 1.$$

Finally, selecting $A = (n-1)A, C = -A, B = A - A$ in (1.3), and then substituting in (1.5) we obtain

$$\frac{|nA|}{|A|} \le \frac{|A - A|}{|A|} \left(\frac{|2A - 2A|}{|A|} \right)^{n} \le \left(\frac{|2A|}{|A|} \right)^{6n+2} \quad \text{for all } n \ge 1.$$

The strongest version of such an inequality that is known was first proved by Plünnecke [Pl], whose proof has been streamlined, over the years, by Ruzsa [R1] and others (though it is still too complicated to give here):

The Plünnecke-Ruzsa theorem. *For any $m, n \ge 0$ we have*

$$\frac{|mA - nA|}{|A|} \le \left(\frac{|2A|}{|A|} \right)^{m+n}.$$

We may rephrase this as: *If $|2A| \le C|A|$ then $|mA - nA| \le C^{m+n}|A|$.*

This result can be given in the slightly stronger form: If $|A+B| \le C|A|$ then $|mB - nB| < C^{m+n}|A|$ for all $m, n \ge 0$. Taking $B = A$ gives the above result. Taking $B = -A$ implies that the assumption $|A - A| \le C|A|$ yields the same conclusion, and therefore we may replace the "≤ 3" by "≤ 2" in (1.2).

I.6. The Freiman-Ruzsa theorem in groups, where the elements have bounded order.

Take the union of (1.4) over all $m, n \geq 1$ to obtain $\langle A\rangle \subset A - A + \langle X\rangle$. However $X \subset 2A - A \subset \langle A\rangle$ and so

$$\langle A\rangle = A - A + \langle X\rangle.$$

Suppose that $|2A| \leq C|A|$. Then $|X| \leq |2A - 2A|/|A| \leq C^4$ by the Plünnecke-Ruzsa theorem (we can get $\leq C^6$ if we only use the results that are proved above). That is, the GAP $\langle A\rangle$ belongs to a union of translates of the GAP $\langle X\rangle$, which has (bounded) dimension $\leq C^4$. If $A \subset G$, an abelian group in which the maximal order of any element is $\leq r$, then $|\langle X\rangle| \leq r^{|X|}$. Therefore

$$|\langle A\rangle| \leq |A - A|\ |\langle X\rangle| \leq C^2|A|r^{|X|} \leq (C^2 r^{C^4})|A|.$$

I.7. The Balog-Szemeredi-(Gowers) theorem .

In many applications one does not have that $A + B$ is small, but rather that there is a large subset $G \subset \{(a, b) : a \in A, b \in B\}$ which contains $\gg |A||B|$ elements, for which $S_G := \{a + b : (a, b) \in G\}$ is small. One then wishes to conclude something about the structure of large subsets of A and B. In the case that $|A| = |B|$ there is an important result of Balog and Szemeredi [BS], strengthened by Gowers [G1] (and subsequently by several others) with a much easier proof – Antal Balog's article in these proceedings [Ba] will discuss all this in detail. Here we simply state a version of this very flexible result, in order to get the flavour: Suppose that $|A| = |B| = n$ and that there exists $G \subset \{(a, b) : a \in A, b \in B\}$ containing $\geq \alpha n^2$ elements, for which $S_G := \{a + b : (a, b) \in G\} \leq n$. Then there exists $A' \subset A,\ B' \subset B$ with $|A'|, |B'| \geq (\alpha/16)n$ for which $|A' + B'| \leq (2^{23}/\alpha^5)\ n$, with

$$|G \cap \{(a', b') : a' \in A', b' \in B'\}| \geq (\alpha^2/128)n^2.$$

I.8. Discrete Fourier transforms

One of the most useful tools in additive combinatorics are Fourier transforms in $\mathbb{Z}/N\mathbb{Z}$: For a function $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ we define

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

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

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

One has
$$\sum_r \hat{f}(r)\overline{\hat{g}}(r) = N\sum_r f(r)\overline{g}(r).$$
Parseval's identity is the case $f = g$, namely $\sum_r |\hat{f}(r)|^2 = N\sum_r |f(r)|^2$.

We define the *convolution* of two functions to be
$$(f*g)(r) = \sum_{t-u=r} f(t)\overline{g(u)},$$
so that $\widehat{(f*g)} = \hat{f}\overline{\hat{g}}$, and
$$N\sum_r |(f*g)(r)|^2 = \sum_r |\hat{f}(r)|^2|\hat{g}(r)|^2.$$
Taking $g = f$ we obtain
$$\sum_r |\hat{f}(r)|^4 = N\sum_{a+b=c+d} f(a)f(b)\overline{f(c)f(d)}.$$

Let A be a subset of $\mathbb{Z}/N\mathbb{Z}$, and then define $A(n)$ to be the *characteristic* function of A; that is, $A(n) = 1$ if $n \in A$, and $A(n) = 0$ otherwise. Hence
$$\hat{A}(m) = \sum_{a\in A} e\left(\frac{am}{N}\right).$$
Noting that $(A*B)(n) = r_{A-B}(n)$ we deduce that
$$E(A,B) = \sum_n r_{A-B}(n)^2 = \sum_n |(A*B)(n)|^2 = \frac{1}{N}\sum_n |\hat{A}(n)|^2|\hat{B}(n)|^2.$$
We also have
$$\hat{A}(m)\hat{B}(m) = \sum_n r_{A+B}(n)e\left(\frac{mn}{N}\right),$$
which can be inverted to give
$$r_{A+B}(n) = \frac{1}{N}\sum_m \hat{A}(m)\hat{B}(m)\ e\left(\frac{-mn}{N}\right);$$
a special case of which is
$$r_{kA-kA}(n) = \frac{1}{N}\sum_m |\hat{A}(m)|^{2k}\ e\left(\frac{-mn}{N}\right).$$

I.9. Sum-Product formulas

I learnt to multiply by memorizing the multiplication tables; that is, we wrote down a table with the rows and columns indexed by the integers between 1 and N and the entries in the table were the row entry times the column entry.[1] Paul Erdős presumably learnt his multiplication tables rather more rapidly than the other students, and was left wondering: How many distinct integers are there in the N-by-N multiplication table? Note that if we take $A = \{1, 2, \dots, N\}$, then we are asking how big is $A \diamond A$? Or, more specifically, since the numbers in the N-by-N multiplication table are all $\le N^2$, what proportion of the integers up to N^2 actually appear in the table? That is,

$$\textit{Does } |A \diamond A|/N^2 \textit{ tend to a limit as } N \to \infty?$$

Erdős showed that the answer is, yes, and that the limit is 0. His proof comes straight from "The Book".[2] Erdős's proof is based on the celebrated result of Hardy and Ramanujan that "almost all" positive integers $n \le N$ have $\sim \log\log N$ (not necessarily distinct) prime factors (here "almost all" means for all but $o(N)$ values of $n \le N$): Hardy and Ramanujan's result implies that "almost all" products ab with $a, b \le N$ have $\sim 2\log\log N$ prime factors, whereas "almost all" integers $\le N^2$ have $\sim \log\log(N^2) \sim \log\log N$ prime factors! The result follows from comparing these two statements.

One can show that $|A \diamond A|$ is large whenever A is an arithmetic progression or, more generally, when A is a GAP of not-too-large dimension. This led Erdős and Szemerédi to the conjecture that for any $\epsilon > 0$, there exists $c_\epsilon > 0$ such that

$$|A + A| + |A \diamond A| \ge c_\epsilon |A|^{2-\epsilon}.$$

Even more, Solymosi conjectured that if $|A| = |B| = |C|$ then

$$|A + B| + |A \diamond C| \ge c_\epsilon |A|^{2-\epsilon}; \tag{1.6}$$

and proved this for $\epsilon = 8/11$ [S1]. We shall prove (1.6) for $\epsilon = 3/4$. We begin by stating the

Szemerédi-Trotter theorem. *We are given a set $\mathcal{C}$ of m curves in $\mathbb{R}^2$ such that*

- *Each pair of curves meet in $\le b_1$ points;*
- *Any pair of points lie on $\le b_2$ curves.*

For any given set $\mathcal{P}$ of n points, there are $\le m + 4b_2 n + 4b_1 b_2^{1/3}(mn)^{2/3}$ pairs (π, γ) with point $\pi \in \mathcal{P}$ lying on curve $\gamma \in \mathcal{C}$.

---

[1] In my primary school we took $n = 12$ which was the basic multiple needed for understanding U.K. currency at that time.

[2] Erdős claimed that the Supreme Being kept a book of all the best proofs, and only occasionally would allow any mortal to glimpse at "The Book".

Székely provided a gorgeous proof of this result, straight from *The Book*, via geometric and random graph theory. From this Elekes elegantly deduced that if $A, B, C \subset \mathbb{Z}$ then

$$|A+B| + |A \diamond C| \geq \frac{2}{3}(|B||C|)^{1/4}(|A|-1)^{3/4}. \tag{1.7}$$

Proof. Let $\mathcal{P}$ be the set of points $(A+B) \times (A \diamond C)$; and $\mathcal{C}$ the set of lines $y = c(x-b)$ where $b \in B$ and $c \in C$. In this case we have $b_1 = b_2 = 1$ with

$$m = |B||C| \quad \text{and} \quad n = |A+B|\ |A \diamond C|.$$

For fixed $b \in B$ and $c \in C$, all of the points $\{(a+b, ac) : a \in A\}$ in $\mathcal{P}$ lie on the line $y = c(x-b)$, so that

$$\#\{(\pi, \gamma) : \pi \in \mathcal{P} \text{ on } \gamma \in \mathcal{C}\} \geq |A|m.$$

Substituting this into the Szemerédi-Trotter theorem we obtain

$$(|A|-1)m \leq 4n + 4(mn)^{2/3}.$$

If $m > 64n^{1/2}$ then $(|A|-1)m \leq 4n + 4(mn)^{2/3} \leq (17/4)(mn)^{2/3}$ which yields $n^2 \geq (|A|-1)^3 m/77$; and if $m \leq 64n^{1/2}$ then $(|A|-1)m \leq 4n + 4(mn)^{2/3} \leq 68n$, which multiplied by the trivial $n \geq |A|^2$ yields the same. The result follows as $2/(77)^{1/4} > 2/3$.

Solymosi has proved (1.7) with several different counting arguments which do not involve the Szemerédi-Trotter theorem. Here we sketch one: Consider the set of distinct points $\{(a+b, ac) :\ a \in A, b \in B, c \in C\}$ in $\mathbb{R}^2$. We will suppose that we can partition $\mathbb{R}^2$ into a grid, with $|A|/3 + O(1)$ lines in each direction (that is lines of the form $x = r$ and of the form $y = s$), in which each box contains roughly equal numbers of points.[3] Now for each pair $b \in B, c \in C$ we will count the number of pairs of points $(b+a_i, ca_i), (b+a_{i+1}, ca_{i+1})$ which belong to the same box where A is the set $a_1 < a_2 < \cdots < a_n$. Since $b+a_i < b+a_{i+1}$ and $|c|a_i < |c|a_{i+1}$ we see that the set of points $\{(b+a, ca) :\ a \in A\}$ can lie in no more than $2|A|/3 + O(1)$ boxes. But then the number of pairs of points $(b+a_i, ca_i), (b+a_{i+1}, ca_{i+1})$ which belong to the same box is $\geq |A|/3 + O(1)$; so the total number of such pairs is $\gtrsim |A||B||C|/3$. Now for any two given points there is at most one triple b, c, i giving those two points else, taking the differences of x and y co-ordinates we have $a_{i+1} - a_i = a_{j+1} - a_j$ and $a_{i+1}/a_i = a_{j+1}/a_j$ which implies that $i = j$ and hence $b = b', c = c'$. Therefore the total number of such pairs is no more than the total number of pairs in our boxes. There are $\sim (|A|/3)^2$ boxes with $\sim \frac{|A+B|\ |A \diamond C|}{(|A|/3)^2}$ points in each box; and so with a total of $\sim \frac{|A+B|^2\ |A \diamond C|^2}{2(|A|/3)^2}$ pairs of points. Combining these remarks we deduce that

$$|A+B|^2 |A \diamond C|^2 \gtrsim \frac{2}{27}|A|^3\ |B|\ |C|;$$

[3] This is not quite as easy as it sounds!

which implies the slight improvement $|A+B|+|A\diamond C| \gtrsim (|B||C|)^{1/4}|A|^{3/4}$ over (1.7).

Sum-product inequalities have also been proved over finite fields (by Bourgain, Katz, Tao [BK], Konyagin, Chang, Glibichuk, ...): This was the basis for proving spectacularly strong bounds on exponential sums by Bourgain [B2], Bourgain, Glibichuk and Konyagin [BGK], Bourgain and Chang [BC] and others — see Kurlberg's article herein for a discussion of this proof [Ku]. These methods have been developed for non-abelian groups, in particular $\mathrm{SL}(2,\mathbb{Z}_p)$ by Helfgott [He], and then extended by Bourgain and Gamburd, Gowers, ... See Mei-Chu Chang's article herein for a discussion of these directions [C2]

Lecture II: The Freiman-Ruzsa theorem

II.1. If $A+B$ is small what do A and B look like ?

We have already seen $|A+B| \geq |A|+|B|-1$ and that if $|A+B| = |A|+|B|-1$ then there exists an integer $d \geq 1$ such that

$$A = \{a+id : 0 \leq i \leq I-1\} \text{ and } B = \{b+jd : 0 \leq j \leq J-1\}.$$

In other words A and B are segments of arithmetic progressions, both with the same common difference. Now if A' is a subset of A, and B' is a subset of B then $A'+B'$ is a large subset of $A+B$; so if A and B are segments of arithmetic progressions with the same common difference, then A' and B' can be chosen as large subsets with little particular structure and yet $A'+B'$ is relatively small. This construction generalizes to large subsets of generalized arithmetic progressions: A *generalized arithmetic progression* (GAP) is a set of integers of the form

$$C := \{a_0 + a_1n_1 + a_2n_2 + \cdots + a_kn_k : \ 0 \leq n_j \leq N_j - 1 \text{ for } 1 \leq j \leq k\}$$

where $N_1, N_2, \ldots, N_k$ are given integers ≥ 2. We say that this GAP has *dimension* k and *volume* $N_1N_2 \ldots N_k$. It is called *proper* if the elements are distinct; that is if there are $N_1N_2 \ldots N_k$ distinct elements in the GAP. Our key observation is that $|2C| < 2^k|C|$ for a GAP C of dimension k; so that if $A, B \subset C$ with $|A|, |B| \geq \delta|C|$ then

$$|A+B| \leq |2C| < 2^k|C| \leq (2^{k-1}\delta^{-1})(|A|+|B|).$$

What about the converse? If $|A+B|$ is a small multiple of $|A|+|B|$, what possibilities are there? Is it true that A and B are both large subsets of translates of the same low dimensional GAP? This question still remains open; we will restrict our attention to the case that $B = A$. In other words, if $|2A|$ is a small multiple of $|A|$ then is A necessarily a large subset of a low dimensional GAP? This question is answered by the wonderful

Freiman-Ruzsa theorem. *If $|2A|$ is "small" then A is a "large" subset of a GAP.*

This statement is a bit vague but, in essence, it is everything we asked for. The details are complicated and researchers have not yet found the best possible version so we leave all that until a little later.

This theorem was first announced by Freiman and gained broad distribution in his book [Fr]. Just to dare guess at such a classification result is an extraordinary achievement, and all proofs to date require much ingenuity. The proof in Freiman's book is deep, and is difficult to follow in places. Because of this, Freiman's result did not quickly gain the prominence it deserves in combinatorial number theory. However in 1994, Ruzsa [R2] came up with his own, much shorter and easier-to-follow proof, which caught many people's

imagination. It is Ruzsa's paper that heralded the outpouring of research into this exciting area. It is for these reasons that I feel it is fair to give both Freiman and Ruzsa credit for their extraordinary achievements by naming the theorem after them both[4]. The proof I give here is more-or-less that of Ruzsa, though incorporating some remarks from Ben Green's notes [G3].

II.2. The Geometry of Numbers

Given linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \mathbb{R}^k$, define a *lattice*

$$\Lambda := (\mathbf{x}_1 \diamond \mathbb{Z}) + (\mathbf{x}_2 \diamond \mathbb{Z}) + \cdots + (\mathbf{x}_k \diamond \mathbb{Z}).$$

We define $\det(\Lambda)$ to be the volume of

$$F := \{a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k : \ 0 \le a_i < 1 \text{ for all } i\},$$

the connection between Λ and F stemming from the fact that $F + \Lambda = \mathbb{R}^k$.

Blichtfeld's Lemma. *If $L \subset \mathbb{R}^k$ is measurable with $\mathrm{vol}(L) > \det(\Lambda)$ then $L - L$ contains a non-zero point of Λ.*

Suppose that K is a centrally symmetric and convex subset of $\mathbb{R}^k$, so that

$$K = \frac{1}{2} \diamond K - \frac{1}{2} \diamond K.$$

Since $\mathrm{vol}(\frac{1}{2} \diamond K) = \frac{1}{2^k}\mathrm{vol}(K)$, Blichfeldt's Lemma with $L = \frac{1}{2} \diamond K$ implies:

Minkowski I. *If $\mathrm{vol}(K) > 2^k \det(\Lambda)$ then K contains a non-zero point of Λ.*

Suppose we are given a lattice Λ in $\mathbb{R}^k$, as well as a closed, convex body $K \subset \mathbb{R}^k$. The *successive minima* $\lambda_1, \lambda_2, \dots, \lambda_k$ of K with respect to Λ are the smallest values λ_j such that $\lambda_j \diamond K$ contains j linearly independent elements of Λ.

Minkowski II. *Suppose that K is a centrally symmetric, closed, convex subset of $\mathbb{R}^k$ and Λ a lattice of rank k. With the definitions as above, there exist linearly independent vectors $\mathbf{b}_1, \dots, \mathbf{b}_k \in \Lambda$ where $\mathbf{b}_j$ lies on the boundary of $\lambda_j \diamond K$ for $j = 1, 2, \dots, k$. Moreover, $\lambda_1 \dots \lambda_k \mathrm{vol}(K) \le 2^k \det(\Lambda)$.*

Let $\|t\|$ be the distance from t to the nearest integer (that is $\|t\| := \min_{m \in \mathbb{Z}} |t - m|$), and then $\|(x_1, \dots, x_k)\| := \max_i \|x_i\|$. For given $r_1, \dots, r_k$ let $\mathbf{r} := (r_1, \dots, r_k)$. When $r_1, \dots, r_k \in \mathbb{Z}/N\mathbb{Z} \setminus \{0\}$ and $\delta > 0$, we define the *Bohr neighbourhood* to be

$$B(r_1, \dots, r_k; \delta) := \{s \in \mathbb{Z}/N\mathbb{Z} : \|\mathbf{r}s/N\| \le \delta\};$$

in other words $s \in B(r_1, \dots, r_k; \delta)$ if the least residue, in absolute value, of each $r_i s \pmod N$ belongs to the interval $[-\delta N, \delta N]$.

[4]Though some authors give credit only to Freiman.

II.3. Structure of a Bohr set

Theorem 1. *Suppose that N is prime with $r_1, \dots, r_k \in \mathbb{Z}/N\mathbb{Z}\setminus\{0\}$, and that $0 < \delta < 1/2$. Then $B(r_1, \dots, r_k; \delta)$ contains a k-dimensional GAP of volume $\geq (\delta/k)^k N$.*

Proof. Let Λ be the lattice generated by $\mathbf{r}$ and $N\mathbb{Z}^k$ so that $\det(\Lambda) = N^{k-1}$. Let $K = \{(t_1, t_2, \dots, t_k) : -1 \leq t_i \leq 1\}$, and then select $\lambda_1, \lambda_2, \dots, \lambda_k$ and $\mathbf{b}_1, \dots, \mathbf{b}_k \in \Lambda$ as in Minkowski II, so that we may write $\mathbf{b}_i = s_i\mathbf{r} + N\mathbb{Z}^k$ for some $s_i \pmod N$, for each i. Therefore

$$\|s_i r_j/N\| = \|(\mathbf{b}_i)_j/N\| \leq \|\mathbf{b}_i/N\| = \|\lambda_i/N\| \leq \lambda_i/N.$$

Let $P := \{\sum_{i=1}^k a_i s_i : |a_i| \leq \delta N/k\lambda_i\}$, which is a k-dimensional GAP. If $s = \sum_{i=1}^k a_i s_i \in P$ then, for each j,

$$\Big\| \frac{sr_j}{N} \Big\| \leq \sum_{i=1}^k |a_i| \Big\| \frac{s_i r_j}{N} \Big\| \leq \sum_{i=1}^k \frac{\delta N}{k\lambda_i} \ \cdot \ \frac{\lambda_i}{N} = \delta,$$

so $s \in B(r_1, \dots, r_k; \delta)$; that is $P \subset B(r_1, \dots, r_k; \delta)$.

Using Minkowski II, and since there are at least t integers in the interval $[-t, t]$ for all $t \geq 0$, we have

$$\mathrm{Vol}(P) > \prod_{i=1}^k \frac{\delta N}{k\lambda_i} \geq \left(\frac{\delta N}{k}\right)^k \frac{\mathrm{vol}(K)}{2^k \det(\Lambda)} = \left(\frac{\delta}{k}\right)^k N,$$

as $\mathrm{vol}(K) = 2^k$ and $\det(\Lambda) = N^{k-1}$.

Remark. Note that if $\delta < 1/2$ then P is proper.

Bogolyubov's Theorem. *Let $A \subset \mathbb{Z}/N\mathbb{Z}$ with $|A| = \alpha N$. Then $2A - 2A$ contains a GAP of dimension $\leq \alpha^{-2}$ and volume $\geq (\alpha^2/4)^{\alpha^{-2}} N$.*

Proof. Let R be the set of "large" Fourier coefficients, that is $R := \{r \pmod N : |\hat{A}(r)| \geq \alpha^{3/2} N\}$. By Parseval's identity we have

$$|A|N = \sum_r |\hat{A}(r)|^2 \geq \sum_{r \in R} |\hat{A}(r)|^2 \geq |R|\alpha^3 N^2$$

so that $|R| \leq \alpha^{-2}$. We also have

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

If $n \in B(R;1/4)$ then $\|rn/N\| \le 1/4$, and hence $\cos(2\pi rn/N) \ge 0$ for all $r \in R$. Using that $\cos(2\pi rn/N) \ge -1$ for all $r \notin R$, that $|\hat{A}(-r)| = |\hat{A}(r)|$, and that $0 \in R$, we obtain

$$r_{2A-2A}(n) := \frac{1}{N} \sum_{r \pmod N} |\hat{A}(r)|^4 \, e\left(\frac{rn}{N}\right) = \frac{1}{N} \sum_{r \pmod N} |\hat{A}(r)|^4 \, \cos\left(2\pi \, \frac{rn}{N}\right)$$
$$\ge \frac{1}{N}\left(|\hat{A}(0)|^4 - \sum_{r \notin R} |\hat{A}(r)|^4\right) > 0.$$

Therefore $2A - 2A$ contains $B(R;1/4)$, and hence contains the required arithmetic progression by Theorem 1.

II.4. Freiman homomorphisms

Suppose that A and B are both finite subsets of some ring like $\mathbb{Z}/s\mathbb{Z}$ or $\mathbb{Z}$ (perhaps different).

The map $\phi : A \to B$ is a *(*Freiman*) k-homomorphism* if

$$\phi(x_1) + \cdots + \phi(x_k) = \phi(y_1) + \cdots + \phi(y_k)$$

whenever $x_i, y_i \in A$ satisfy

$$x_1 + x_2 + \cdots + x_k = y_1 + y_2 + \cdots + y_k.$$

ϕ is a (Freiman) *k-isomorphism* if ϕ is invertible and ϕ and ϕ^{-1} are Freiman k-homomorphisms. (Henceforth we drop the adjective "Freiman".)

Examples: The reduction $\rho_p : \mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ is a k-homomorphism for all k. If $A = \{a_1 < \cdots < a_n < a_1 + p/k\}$ then $\rho_p\big|_A$ is a k-isomorphism.

If $(q,p) = 1$ then $\mu_{q,p} : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$, where $\mu_{q,p}(x) \equiv qx \pmod p$, is a k-isomorphism for all k.

The heart of our proof comes in the following remarkable lemma of Ruzsa which gives a Freiman isomorphism between a large subset of our given set A, and some subset of the integers mod N. This allows us to work inside the integers mod N, where there are more convenient tools.

Ruzsa's Lemma. *For any set of integers A and any prime $N > 2|kA - kA|$, there exists a subset A' of A, with $|A'| \ge |A|/k$, which is k-isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$.*

Proof. Select a prime $p > k(\max A - \min A)$, and any q with $(q,p) = 1$. Note that $\rho_p\big|_A$ is a k-isomorphism. By the pigeonhole principle, there exist $A' \subset A$ with $|A'| \ge |A|/k$ and

$$\mu_{q,p} \circ \rho_p\big|_{A'} \subset \left\{x \in \mathbb{Z}/p\mathbb{Z} : \text{There exists } n \in \left[\frac{j-1}{k}p, \frac{j}{k}p\right) \text{ with } n \equiv x \pmod p\right\}$$

for some j. Therefore, taking ρ_p^{-1} here to give a residue in $[0,p)$, we have that

$$\Psi := \rho_p^{-1} \circ \mu_{q,p} \circ \rho_p\big|_{A'}$$

is a k-isomorphism such that

$$r := \sum_{i=1}^k \Psi(a_i) - \sum_{i=1}^k \Psi(a_i') \ \in (-p,p),$$
$$\text{with } r \equiv q(\sum_{i=1}^k a_i - \sum_{i=1}^k a_i') \pmod p$$

for all $a_1, a_2, \dots, a_k, a_1', a_2', \dots, a_k' \in A'$. (The reader should verify that this is indeed a k-isomorphism.)

Now define

$$\Phi^{(q)} := \rho_N \circ \Psi = \rho_N \circ \rho_p^{-1} \circ \mu_{q,p} \circ \rho_p,$$

which is a k-homomorphism; so the question becomes: *Is $\Phi^{(q)}\big|_{A'}$ a k-isomorphism?* If not there exist integers $a_1, \dots, a_k, a_1', \dots, a_k' \in A'$ for which $r \neq 0$ but $r \equiv 0 \pmod N$. In this case define $b := \sum_{i=1}^k a_i - \sum_{i=1}^k a_i' \in kA - kA$, so that $qb \equiv r \pmod p$ and $b \not\equiv 0 \pmod p$. Hence q is of the form $r/b \pmod p$ where $r \in (-p,p)$ with $r \neq 0$ and $N|r$, and where $b \in kA - kA$ with $b \not\equiv 0 \pmod p$: The number of such q is therefore

$$\leq \#\{r \in (-p,p) : r \neq 0, N|r\} \ \times \ \#\{b \in kA - kA : b \not\equiv 0 \pmod p\}$$
$$\leq \frac{2(p-1)}{N} \ |kA - kA| < p - 1.$$

Therefore there must exist values of $q, 1 \leq q \leq p-1$ for which $\Phi^{(q)}\big|_{A'}$ is a k-isomorphism.

II.5. The Freiman-Ruzsa Theorem

We recall the following result from our previous lecture:

The Plünneke-Ruzsa theorem. *If A and B are finite sets of integers for which $|A+B| \leq C|A|$ then $|kB - \ell B| \leq C^{k+\ell}|A|$.*

We are now ready to state and prove our main result:

The Freiman-Ruzsa Theorem. *If $A \subset \mathbb{Z}/N\mathbb{Z}$ with $|A+A| \leq C|A|$ then A is contained in a GAP of dimension $\leq d(C)$ and volume $\leq \nu(C)|A|$.*

Here $d(C)$ and $\nu(C)$ are constants which depend only on the constant C. Following the works of Bilu [Bi] and Chang [C1] we know that we can take

$$d(C) = \lfloor C - 1 \rfloor, \text{ and } \ \nu(C) = e^{O(C^2(\log C)^3)}.$$

Proof. By the Plünneke-Ruzsa theorem with $k = \ell = 8$ we have $|8A - 8A| \leq C^{16}|A|$. Therefore, by Ruzsa's Lemma, there exists $A' \subset A$ with $|A'| \geq |A|/8$, which is 8-isomorphic

to some $B \subset \mathbb{Z}/N\mathbb{Z}$ where N is prime with $2C^{16}|A| < N \leq 4C^{16}|A|$ (such a prime may be selected by Bertrand's Postulate). So $|B| = \alpha N$ with $\alpha \geq 1/(32C^{16})$.

By Bogolyubov's Theorem, $2B - 2B$ contains a GAP of dimension $\leq \alpha^{-2}$ and volume $\gamma|A|$, with $1 \geq \gamma \geq (\alpha^2/8)^{\alpha^{-2}}$. Now B is 8-isomorphic to A' so $2B - 2B$ is 2-isomorphic to $2A' - 2A'$. Since any set which is 2-isomorphic to a d-dimensional GAP is itself a d-dimensional GAP, hence $2A' - 2A'$, and thus $2A - 2A$, contains a GAP Q of dimension at most α^{-2} and volume $\gamma|A|$.

Let S be a maximal subset of A for which the sets $s + Q, s \in S$ are disjoint, so that $|S + Q| = |S||Q|$. Since S is maximal, if $a \in A$ there exists $s \in S$ and $q_1, q_2 \in Q$ such that $a + q_1 = s + q_2$, and therefore

$$A \subset S + Q - Q \subset Q - Q + \sum_{s \in S} \{0, s\},$$

a GAP, of dimension $\leq |S| + \alpha^{-2}$ and volume $\leq 2^{|S|+\alpha^{-2}}\gamma|A|$. Now

$$S + Q \subset A + (2A - 2A) = 3A - 2A$$

so that

$$|S| = \frac{|S + Q|}{|Q|} \leq \frac{|3A - 2A|}{\gamma|A|} \leq \frac{C^5}{\gamma}$$

by the Plünneke-Ruzsa theorem. Tracing through the above proof, we find that the result follows with volume $\nu(C) = 2^{d(C)}$ where $d(C) = C^{C^{48}}$.

These bounds can be significantly improved by the following argument of Chang: The big bounds come as a consequence of the enormous size of S. We will improve this by replacing S and Q by S' and Q' where S' is significantly smaller than S, while Q' is a little bigger than Q: Let m be the smallest integer $\geq 2C$. Let $S_0 = S$ and $Q_0 = Q$. For any given $j \geq 0$, if $|S_j| \leq m$ then we stop the algorithm and let $r = j$. Otherwise we select any subset T_j of S_j of size m and let $Q_{j+1} = T_j + Q_j$. Now we select S_{j+1} to be a maximal subset of A for which the sets $s + Q_{j+1}, s \in S_{j+1}$ are disjoint, so that $|S_{j+1} + Q_{j+1}| = |S_{j+1}||Q_{j+1}|$. Note that this also implies that $|Q_{j+1}| = |T_j + Q_j| = |T_j||Q_j| = m|Q_j|$ for all j, so that $|Q_r| = m^r|Q|$. On the other hand $Q_{j+1} = T_j + Q_j \subset S_j + Q_j \subset A + Q_j$ for each j, so that $Q_r \subset rA + Q \subset (r + 2)A - 2A$, which implies that $|Q_r| \leq C^{r+4}|A|$ by the Plünneke-Ruzsa theorem. Therefore $2^r \leq (m/C)^r \leq C^4|A|/|Q| = C^4/\gamma$ by the last two equations.

Now $A \subset S_r + Q_r - Q_r \subset S_r + \sum_{j=0}^{r-1}(T_j - T_j) + (Q - Q)$, which is a GAP of dimension $\leq m(r + 1) + \alpha^{-2}$, and volume $\leq 3^{m(r+1)+\alpha^{-2}}\gamma|A|$. Tracing through, we find that $r \ll C^{32} \log C$ so that we can take $d(C) = C^{33} \log C$ and $\nu(C) = C^{O(C^{33})}$.

Lecture III: Uniform distribution, Roth's theorem and beyond

III.1. Uniform distribution mod one

We begin by discussing Hermann Weyl's famous criterion for recognizing uniform distribution mod one: Let $\{t\}$ be the fractional part of t, and $e(t) = e^{2i\pi t}$ so that $e(t) = e(\{t\})$. A sequence of real numbers $a_1, a_2, \dots$ is *uniformly distributed mod one* if, for all $0 \le \alpha < \beta \le 1$ we have

$$\#\{n \le N : \ \alpha < \{a_n\} \le \beta\} \sim (\beta - \alpha)N \ \text{ as } \ N \to \infty.$$

To determine whether a sequence of real numbers is uniformly distributed we have the following extraordinary, and widely applicable, criterion:

Weyl's criterion. *A sequence of real numbers $a_1, a_2, \dots$ is uniformly distributed mod one* if and only if *for every integer $b \neq 0$ we have*

$$\left| \sum_{n \le N} e(ba_n) \right| = o_b(N) \quad \textit{as} \quad N \to \infty. \tag{3.1}$$

In other words $\limsup_{N\to\infty} \ \frac{1}{N} |\sum_{n\le N} e(ba_n)| = 0$.

Note that if $a_1, a_2, \dots$ is uniformly distributed mod one then $ka_1, ka_2, \dots$ is uniformly distributed mod one for all non-zero integers k.

An interesting example is where $a_n = f(n)$ for some polynomial $f(t) \in \mathbb{R}[t]$. It can be shown, using Weyl's criterion, that the sequence $a_1, a_2, \dots$ is uniformly distributed mod one if and only if one or more of the coefficients of $f(t) - f(0)$ is irrational. Note that if all the coefficients of f are rational then there exists an integer $b > 0$ such that $b(f(t) - f(0)) \in \mathbb{Z}[t]$; but then each $e(bf(n)) = e(bf(0))$, and so (3.1) is not satisfied. If f is linear, that is $f = \gamma n + \delta$ with γ irrational then

$$\sum_{n\le N} e(ba_n) = e(b\delta) \sum_{n\le N} e(b\gamma n) = e(b(\gamma+\delta)) \cdot \frac{e(b\gamma N) - 1}{e(b\gamma) - 1},$$

the sum of a geometric progression, since $b\gamma$ is not an integer. Therefore

$$\left| \sum_{n\le N} e(ba_n) \right| \le \frac{2}{|e(b\gamma) - 1|} \asymp \frac{1}{\|b\gamma\|} \ll_b 1 = o_b(N)$$

as required, since $|e(t) - 1| \asymp \|t\|$, where $\|t\|$ denotes the distance from t to the nearest integer.

III.2. Uniform distribution mod N.

For a given set, A, of residues mod N, define

$$\hat{A}(b) := \sum_{n\in A} e\left(\frac{bn}{N}\right).$$

Let $(t)_N$ denote the least non-negative residue of $t \pmod N$ (so that $(t)_N/N = \{t/N\}$). The idea of uniform distribution mod N is surely something like: For all $0 \le \alpha < \beta \le 1$ and all $m \not\equiv 0 \pmod N$, we have

$$\#\{a \in A :\ \alpha N < (ma)_N \le \beta N\} \sim (\beta - \alpha)|A|. \tag{3.2}$$

One can only make sense of such a definition if $|A| \to \infty$ (since this is an asymptotic formula) but we are often interested in smaller sets A, indeed that are a subset of $\{1, 2, \dots, N\}$; so we will work with something motivated by, but different from, (3.2). Let us see how far we can go in proving an analogy to Weyl's criterion.

For given subset A of the residues mod N define

$$\text{Error}(A) := \max_{\substack{0\le x<x+y\le N\\ m\not\equiv 0 \pmod N}} \left|\frac{\#\{a \in A :\ x < (ma)_N \le x + y\}}{|A|} - \frac{y}{N}\right|.$$

Theorem 1. *Suppose that N is prime. Fix $\delta > 0$.*
(i) *If* $\text{Error}(A) \le \delta^2|A|$ *then* $|\hat{A}(m)| \ll \delta|A|$ *for any* $m \not\equiv 0 \pmod N$.
(ii) *If* $|\hat{A}(m)| \le \delta^2|A|$ *for all* $m \not\equiv 0 \pmod N$, *and* $|A| \ge N/e^{c/\delta}$ *then* $\text{Error}(A) \ll \delta|A|$, *for some absolute constant* $c > 0$.

Proof. For given integer $k \ge 1$, if $(ma)_N \in (x, x+N/k]$ then $e(ma/N) = e(x/N)+O(1/k)$. Therefore

$$\hat{A}(m) = \sum_{j=0}^{k-1} \sum_{\substack{a\in A\\ \frac{jN}{k}<(ma)_N\le\frac{(j+1)N}{k}}} e(ma/N) = \sum_{j=0}^{k-1} \sum_{\substack{a\in A\\ \frac{jN}{k}<(ma)_N\le\frac{(j+1)N}{k}}} (e(j/k) + O(1/k))$$
$$= \sum_{j=0}^{k-1} |A| \left(\frac{1}{k} + O(\text{Error}(A))\right) e(j/k) + O\left(\frac{|A|}{k}\right) \ll |A|(k\ \text{Error}(A) + 1/k).$$

The result follows by taking $k \asymp 1/\delta$.

In the other direction we have, for integers x, y with $0 \le x < x + y \le N$

$$\sum_{\substack{a\in A\\ x<(ma)_N\le x+y}} 1 = \sum_{j=1}^{y} \sum_{a\in A} \frac{1}{N} \sum_{r \pmod N} e\left(r\left(\frac{ma - x - j}{N}\right)\right)$$
$$= \frac{y}{N}|A| + \frac{1}{N} \sum_{\substack{r \pmod N\\ r\not\equiv 0}} \hat{A}(rm) e\left(\frac{-rx}{N}\right) \sum_{j=1}^{y} e\left(\frac{-rj}{N}\right).$$

If r runs through the non-zero integers in $(-N/2, N/2]$ then

$$\left| e\left(\frac{-rx}{N}\right) \sum_{j=1}^{y} e\left(\frac{-rj}{N}\right) \right| \ll \frac{N}{|r|},$$

and so the second term above is, as $|\hat{A}(-rm)| = |\hat{A}(rm)|$,

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

for $R \approx N/(\delta^2|A|)$.

To obtain an analogy to Weyl's criterion we think of an infinite sequence of pairs (A, N) with N prime and $N \to \infty$, where $|A| \gg N$. More precisely we have

Corollary. *For each prime N let A_N be a subset of the residues mod N with $|A_N| \gg N$. Then* $\mathrm{Error}(A_N) = o(1)$ *if and only if* $|\hat{A}_N(m)| = o(N)$ *for all* $m \not\equiv 0 \pmod N$.

One can therefore formulate an analogy to Weyl's criterion along the lines: The Fourier transforms of A are all small if and only if A and all of its dilates are uniformly distributed. (A *dilate* of A is the set $\{ma : a \in A\}$ for some $m \not\equiv 0 \pmod N$.) This idea is central to our recent understanding, in additive combinatorics, for proving that large sets contain 3-term arithmetic progressions; and finding appropriate analogies to this are essential to our understanding when considering k-term arithmetic progressions for $k \ge 3$. More on that later.

To give one example of how such a notion can be used, we ask whether a given set A of residues mod N contains a non-trivial 3-term arithmetic progression? In other words we wish to find solutions to $a + b = 2c$ with $a, b, c \in A$ where $a \neq b$.

Theorem 2. *If A is a subset of the residues $\pmod N$ where N is odd, for which $|\hat{A}(m)| < |A|^2/N - 1$ whenever $m \not\equiv 0 \pmod N$ then A contains non-trivial 3-term arithmetic progressions.*

Proof. Since $\frac{1}{N}\sum_r e(rt/N) = 0$ unless t is divisible by N, whence it equals 1, we have that the number of 3-term arithmetic progressions in A is

$$\sum_{a,b,c\in A} \frac{1}{N} \sum_r e\left(\frac{r(a+b-2c)}{N}\right) = \frac{1}{N} \sum_r \hat{A}(r)^2 \hat{A}(-2r).$$

The $r = 0$ term gives $|A|^3/N$. We regard the remaining terms as error terms, and bound them by their absolute values, giving a contribution (taking $m \equiv -2r \pmod N$)

$$\leq \frac{1}{N}\sum_r |\hat{A}(r)|^2 \cdot \max_{m\neq 0} |\hat{A}(m)| = |A| \max_{m\neq 0} |\hat{A}(m)|.$$

There are $|A|$ trivial 3-term arithmetic progressions (of the form a, a, a) so we have established that A has non-trivial 3-term arithmetic progressions when

$$|A|^3/N - |A| \max_{m\neq 0} |\hat{A}(m)| > |A|,$$

yielding the result.

Rather more generally we can ask for solutions to

$$ia + jb + kc \equiv \ell \pmod N \tag{3.3}$$

where $(ijk, N) = 1$ with $a \in A, b \in B, c \in C$ and $A, B, C \subset \mathbb{Z}/N\mathbb{Z}$. We count the above set as

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

The $r = 0$ term contributes $\frac{1}{N}\hat{A}(0)\hat{B}(0)\hat{C}(0) = \frac{|A||B||C|}{N}$. The total contribution of the other terms can be bounded above by

$$\begin{aligned}\frac{1}{N}\sum_{r\neq 0} |\hat{A}(ir)|\ |\hat{B}(jr)|\ |\hat{C}(kr)| &\leq \frac{1}{N}\max_{m\neq 0}|\hat{A}(m)| \sum_r |\hat{B}(jr)||\hat{C}(kr)|\\ &\leq \frac{1}{N}\max_{m\neq 0}|\hat{A}(m)| \left(\sum_t |\hat{B}(t)|^2\right)^{\frac{1}{2}}\left(\sum_u |\hat{C}(u)|^2\right)^{\frac{1}{2}}\\ &= \frac{1}{N}\max_{m\neq 0}|\hat{A}(m)|(N|B|N|C|)^{1/2} = (|B||C|)^{1/2}\max_{m\neq 0}|\hat{A}(m)|\end{aligned}$$

using the Cauchy-Schwarz inequality. Therefore there are $\geq |A||B||C|/2N$ solutions to (3.3) provided

$$|\hat{A}(m)| \leq \frac{(|B||C|)^{1/2}}{2N}\ |A| \ \text{ for every } m \not\equiv 0 \pmod N. \tag{3.4}$$

III.3. Roth's Theorem.

In 1953, Roth [Ro] proved that for any $\delta > 0$ if N is sufficiently large then any subset A of $\{1, \dots, N\}$ with more than δN elements contains a non-trivial 3-term arithmetic progression. We shall prove Roth's theorem in this section.

In 1975 Szemerédi [S2] generalized this to obtain non-trivial k-term arithmetic progressions. This was reproved by Furstenberg [2] in 1977, and there have been recent proofs by Gowers [G1, G2], Tao, and many others. See the article herein by Tao [Ta] for an inspiring discussion of these proofs; and Kra's article [Kr] for developments of Furstenberg's ideas.

In Roth's proof, as we will see below, one can take $\delta \approx 1/\log\log N$. This was improved (but remained unpublished until this volume) by Szemerédi [S4] to $\delta \approx 1/\exp(\sqrt{c\log\log N})$. In the late eighties, both Heath-Brown [HB] and Szemerédi [S3] showed one can take $\delta \approx 1/(\log N)^c$ for some small $c > 0$. The best result known, due to Bourgain [Bo1], is that one can take

$$\delta \approx \sqrt{\frac{\log\log N}{\log N}}.$$

To start our proof of Roth's theorem we note that the result is easy for $\delta > 2/3$ since then A contains a subset of the form $\{a, a+1, a+2\}$. For smaller δ we shall either prove directly that it has 3-term arithmetic progressions by the methods of the previous section, or that there is a large arithmetic progression of length N_1 which contains $\delta_1 N_1$ terms of A with $\delta_1 > \delta(1 + c\delta)$ for some $c > 0$. This can be used to construct a subset A_1 of the first N_1 integers, of size $\delta_1 N_1$, which must have a 3-term arithmetic progression by an appropriate induction hypothesis (as δ_1 is significantly larger than δ), so that A also does.

Replace N by the smallest prime $\geq N$ which can be done with negligible change in our hypothesis. Let us assume that A is a subset of the integers up to N, containing at least δN elements, but which has no three term arithmetic progression. We will suppose that we have proved Roth's theorem for any constant $\delta' > \delta(1 + c\delta)$.

- If $\#\{a \in A : \ 0 < a < \frac{N}{3}\} \geq (1 + c\delta)|A|/3$ then $A_1 := \{a \in A : \ 0 < a \leq \frac{N}{3}\}$.
- If $\#\{a \in A : \ \frac{2N}{3} < a < N\} \geq (1 + c\delta)|A|/3$ then $A_1 := \{N - a : a \in A, \ \frac{2N}{3} < a < N\}$.

In these cases $N_1 = [N/3]$, and the result follows from our hypothesis. Otherwise we let $B := \{a \in A : \ \frac{N}{3} < a < \frac{2N}{3}\}$, so that $|B| > (1 - 2c\delta)|A|/3$. There are no solutions to $a + b \equiv 2d \pmod N$ with $a \in A$ and $b, d \in B \subset A$, all distinct. For if $b, d \in B$ then $0 < 2d - b < N$ and so $a + b = 2d$, hence $a = b = d$ by our assumption that A has no non-trivial 3-term arithmetic progressions.

This implies that there must exist $m \not\equiv 0 \pmod N$ such that $|\hat{A}(m)| > \delta(1 - 2c\delta)|A|/6$ else we have many non-trivial solutions to (3.3) (with $i = j = 1, k = -2, \ell = 0$) by (3.4). But then A is not uniformly distributed mod N; in particular, $\text{Error}(A) \gg \delta^2|A|$ by Theorem 1(i). In other words there is some dilate of A and some long interval which does not contain the expected number of elements of the dilate A; in fact it is out by a constant factor. However we need slightly more than that: We need an interval that has

too many elements of A by a constant factor and so we make one more observation: Select an integer $\ell \gg 1/\delta$, and define

$$A_j := \left\{ a \in A : (ma)_N \in \left(\frac{jN}{\ell}, \frac{(j+1)N}{\ell} \right] \right\}$$

for $0 \le j \le \ell - 1$, so that if a is counted by A_j then $e(ma/N) = e(j/\ell) + O(1/\ell)$. Therefore

$$\hat{A}(m) = \sum_{j=0}^{\ell-1} \left(\#A_j - \frac{|A|}{\ell} \right) e\left(\frac{j}{\ell}\right) + O\left(\frac{|A|}{\ell}\right),$$

implying that

$$\sum_{j=0}^{\ell-1} \left| \#A_j - \frac{|A|}{\ell} \right| \ge \left| \sum_{j=0}^{\ell-1} \left(\#A_j - \frac{|A|}{\ell} \right) e\left(\frac{j}{\ell}\right) \right| \ge \hat{A}(m) - O\left(\frac{|A|}{\ell}\right) \gg \delta |A|.$$

Adding this to $\sum_j (\#A_j - \frac{|A|}{\ell}) = 0$, we find that there exists j for which

$$\left(\#A_j - \frac{|A|}{\ell} \right) \gg \delta \frac{|A|}{\ell}.$$

What we would like to do now is to define $A' := \{i : \ [jN/\ell] + i \in A_j\}$, a subset of $\{1, 2, \dots, N'\}$ where $N' = [N/\ell]$, with $|A'| \ge (1 + c\delta)\delta N'$ and then assert that A' contains no non-trivial 3-term arithmetic progressions. To prove this last remark, we proceed by noting that if $u, v, w \in A'$ for which $u + w = 2v$ then there exist $a, b, c \in A$ such that $ma \equiv [jN/\ell] + u \pmod N$, $mb \equiv [jN/\ell] + v \pmod N$, $mc \equiv [jN/\ell] + w \pmod N$ so that $m(a + c - 2b) \equiv u + w - 2v \equiv 0 \pmod N$, and therefore $a + c \equiv 2b \pmod N$. *However* there is no guarantee that this implies that $a + c = 2b$ (as above), since there may be "wraparound" (that is, $a + c$ might equal $2b \pm N$ or $2b \pm 2N$ or ...), and so we need to refine our construction to be able to make this final step.

The trick is to use the well-known result that if $RS = N$ where R and S are real numbers > 1 then there exist integers r and s, with $0 < r < R$ and $0 < s < S$, such that $\pm m \equiv s/r \pmod N$. (Proof: There are more than N integers of the form $j + im$ with $0 \le i < R$ and $0 \le j < S$, so two must be congruent mod N. Thus their difference $s \pm rm \equiv 0 \pmod N$.) For convenience we will assume $m \equiv s/r \pmod N$ where $R = \sqrt{N/\delta^3}, S = \sqrt{N\delta^3}$, with $x = [jN/\ell]$ and $y = [N/\ell]$ and $\ell \asymp 1/\delta$, so that

$$\#\{a \in A : x < (ma)_N \le x + y\} \ge (1 + c\delta)\delta y.$$

We begin by partitioning this set depending only on the value of $(ma)_N \pmod s$: For $1 \le i \le s$ let $\alpha_i = ((x+i)/m)_N$, and then define

$$A_i := \left\{ a \in A : \ a \equiv \alpha_i + jr \pmod N \text{ and } 0 \le j \le \left[\frac{y-i}{s} \right] \right\}.$$

Note that $ma \equiv m(\alpha_i + jr) \equiv x + (i+js)$ so that $x < (ma)_N \le x+y$ for $a \in A$. Therefore there exists some value of i for which $\#A_i \ge (1+c\delta)\delta y/s$. Even within A_i we still have the possibility of the "wraparound problem"; so we deal with this by partitioning A_i:

Let $K = [(\alpha_i + ry/s)/N]$ so that $\alpha_i \le \alpha_i + jr \le \alpha_i + ry/s < (K+1)N$. For each $0 \le k \le K$ define

$$A_{i,k} := \{a \in A_i : \ kN < \alpha_i + jr \le (k+1)N\}.$$

Let $\alpha_{i,0} = \alpha_i - r$, and let $\alpha_{i,k}$ be the largest integer $\le kN$ which is $\equiv \alpha_i \pmod r$ for $1 \le k \le K$. Then $A_{i,k} = \{a \in A_i : \ a = \alpha_{i,k} + jr, \ 1 \le j \le J_k + O(1)\}$ where $J_0 = N/r - \alpha_i/r$, $J_k = N/r$ for $1 \le k \le K-1$, and $J_K = y/s - KN/r + \alpha_i/r$. We let T be the set of indices $k, 1 \le k \le K-1$ together with $k=0$ provided $J_0 > c\delta^2 y/4s$, and with $k=K$ provided $J_K > c\delta^2 y/4s$. Note that

$$\sum_{k \in T} \#A_{i,k} \ge \#A_i - c\delta^2 y/2s \ge (1+c\delta/2)\delta y/s \ge (1+c\delta/2)\delta \sum_{k \in T} J_k.$$

Thus there exists $k \in K$ such that $\#A_{i,k} \ge (1+c\delta/2)\delta J_k$. Now define $N' = [J_k]$ and $A' = \{j : 1 \le j \le N', \ \alpha_{i,k} + jr - kN \in A\}$, a subset of $\{1, 2, \dots, N'\}$, so that $\#A' = \#A_{i,k} \ge (1+c\delta/2)\delta N'$. We claim that A' does not contain any non-trivial 3-term arithmetic progressions; else if $u+v=2w$ with $u,v,w \in A'$ then $a = \alpha_{i,k} + ur - kN$, $b = \alpha_{i,k} + vr - kN$, $c = \alpha_{i,k} + wr - kN \in A$ and $a+b=2c$, contradicting the fact that A does not contain any non-trivial 3-term arithmetic progressions. Note that $N' \ge \min\{N/r, c\delta^2 y/4s\} \gg \min\{N/R, \delta^2 N/\ell S\} \gg \sqrt{\delta^3 N}$.

We have obtained the induction hypothesis that we wanted. If we iterate we find that we increase the constant $\delta = 2^{-n}$ to $2\delta = 2^{-(n-1)}$ in $\asymp 1/\delta$ iterations by which time the size of our set is roughly 2^{-3n} times N to the power $(1/2)^{2^{n+O(1)}}$. Thus when we get all the way up to $\delta = 1$ the size of our set is N to the power $(1/2)^{2^{n+O(1)}}$. To ensure that this is not negligible we must have $2^{2^{n+O(1)}} = o(\log N)$; that is $2^n \ll \log\log N$ and so $\delta = 2^{-n} \gg 1/\log\log N$.

In the other direction we have

Behrend's Theorem. *There exists a subset $A \subset \{1, \dots, N\}$ with $\#A \ge \frac{N}{\exp(c\sqrt{\log N})}$, such that A has no non-trivial 3-term arithmetic progression.*

Proof. Let $T := \{(x_0, \dots, x_{n-1}) \in \mathbb{Z}^n : \ 0 \le x_i < d\}$ and $T_k := \{\mathbf{x} \in T : |\mathbf{x}|^2 = k\}$. We have $|T| = d^n$, and $|\mathbf{x}|^2 < nd^2$ for every $\mathbf{x} \in T$, so there exists a positive integer k for which T_k has $\ge d^{n-2}/n$ elements. Let

$$A := \{x_0 + x_1(2d) + \cdots + x_{n-1}(2d)^{n-1} : \ \mathbf{x} \in T_k\}.$$

If $a + b = 2c$ with $a, b, c \in A$ then $a_0 + b_0 \equiv 2c_0 \pmod{2d}$ and $-2d < a_0 + b_0 - 2c_0 < 2d$ so that $a_0 + b_0 = 2c_0$; similarly one proves that $a_1 + b_1 = 2c_1$, and indeed $a_i + b_i = 2c_i$ for each $i \ge 0$. But then $\mathbf{a}+\mathbf{b}=2\mathbf{c}$, that is $\mathbf{a},\mathbf{b},\mathbf{c} \in T_k$ are collinear, which is impossible as T_k is a sphere! Therefore A contains no non-trivial 3-term arithmetic progressions.

The elements of A are all $\le (d-1)(1 + 2d + \cdots + (2d)^{n-1}) < N := 2^{n-1}d^n$. The result follows by taking $n \approx \sqrt{\log N}$ and $d = [(2N)^{1/N}/2]$.

For each integer $N \ge 1$, define $R(N)$ to be the size of the largest subset A of $\{1, \dots, N\}$ which does not contain any non-trivial 3-term arithmetic progressions. We know, after Behrend and Bourgain, that

$$N\sqrt{\frac{\log\log N}{\log N}} \gg R(N) \gg \frac{N}{\exp(c\sqrt{\log N})};$$

the question is whether $R(N)$ is really near to one of these bounds, or somewhere in-between. There does not seem to be any convincing heuristic to predict the truth; at the school we asked the lecturers to all venture a guess – it seemed that people's intuitions varied substantially! It would be most exciting if one could prove that $R(N) \le (1 - \epsilon)N/\log N$ for sufficiently large N since this would give an "automatic proof" that there are infinitely many three term arithmetic progressions of primes.

Finally we prove the following slight strengthening of Roth's theorem

Varnavides' Theorem. *Fix $1 \ge \delta > 0$. There exist constants $C(\delta) > 0$ and $N(\delta)$ such that if $N \ge N(\delta)$ and $A \subset \{1, \dots, N\}$ with $\#A \ge \delta N$, then A has at least $C(\delta)N^2$ 3-term arithmetic progressions.*

Proof. By Roth's theorem we know that there exists an integer M such that any set of $\delta M/2$ integers from an arithmetic progression of length M contains a non-trivial three term arithmetic progression. We will apply this result to the subset of A lying in each arithmetic progression of length M taken from the integers in $\{1, \dots, N\}$. Let $\mathcal{P}(b, d)$ be the arithmetic progression $b, b + d, \dots, b + (M-1)d$, for $1 \le b \le N - (M-1)d$, and let $A(b, d)$ be the number of elements of A in $\mathcal{P}(b, d)$. Since every element of A from the interval $((M-1)d, N-(M-1)d]$ is counted in exactly M of these arithmetic progressions, we deduce that $\sum_{1 \le b \le N-(M-1)d} A(b, d) \ge M(\delta N - 2(M-1)d)$. Since each $A(b, d) \le M$, we can deduce that there are $\ge \delta N/2$ values of b for which $A(b, d) \ge \delta M/2$, provided $N \ge 6(M-1)d/\delta^2$. Now, each of these contains a non-trivial three term arithmetic

progression, making for a total of $\geq \delta^3 N^2/12(M-1)$ non-trivial three term arithmetic progressions, when we consider all $d \leq \delta^2 N/6(M-1)$, though many of these may have been counted more than once. Now if $a, a+D, a+2D$ is counted in some $A(b,d)$ then d divides D and $2D/d \leq M-1$; and it is counted in $A(b,d)$ for no more than $M-2D/d$ values of b. Writing $D/d = h$, we find that $a, a+D, a+2D$ has been counted no more than $\leq \sum_{1\leq h\leq M/2}(M-2h) \leq (M/2)^2$ times. Therefore A contains $\geq \delta^3 N^2/3M^3$ distinct non-trivial three term arithmetic progressions.

By Bourgain's result we may take $M = (1/\delta)^{c/2\delta^2}$ for some constant $c > 0$, and therefore $C(\delta) \geq \delta^{c/\delta^2}$ in Varnavides' theorem. A small modification of the proof of Behrend's theorem implies that $C(\delta) \leq \delta^{c' \log(1/\delta)}$, for some constant $c' > 0$.

III.4. Large Fourier coefficients

We saw in the previous section that proving Roth's theorem is difficult only in the case that there are large Fourier coefficients, $\hat{A}(m)$, with $m \not\equiv 0 \pmod N$. It is worth noting a few other results which reflect consequences of having large Fourier coefficients:

An easy one to prove is that for any $\eta > 0$ there exists $\delta > 0$ such that

$$r_{A-A}(n) > (1-\eta)|A| \quad \textit{if and only if} \quad \sum_{m:|(mn)_N|\leq \epsilon N} |\hat{A}(m)|^2 \geq (1-\delta)\sum_m |\hat{A}(m)|^2.$$

(See Lecture I for further discussion of $r_{A-A}(n)$, the number of representations of n as $a-a'$ with $a, a' \in A$.)

A manifestation of the *uncertainty principle* (which roughly states that a non-trivial function and its Fourier transforms cannot all be too small) is given by: If $A \subset \mathbb{Z}/N\mathbb{Z}$ has no elements in $(x-L, x+L)$ then there exists m, $0 < m < (N/L)^2$ such that $|\hat{A}(m)| \geq (L/2N)|A|$.

In many proofs it is important to know how often $|\hat{A}(m)|$ can be large? Let $R := \{r \pmod N : \ |\hat{A}(r)| > \rho|A|\}$. From Parseval's identity we see that

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

so that $|R| \leq \rho^{-2}N/|A|$. Note that if $r, s \in R$ then this says that the numbers $(ra)_N, a \in A$ and $(sa)_N, a \in A$ have a bias towards being close to certain values x and y respectively. In that case we might expect that the numbers $((r+s)a)_N, a \in A$ have a bias towards $(x+y)_N$ so that $r+s \in R$. Therefore we might expect that R has some lattice structure, an intuition that is verified by Chang's result [C1] that R is contained in a cube of dimension $\leq 2\rho^{-2}\log(N/|A|)$ (cubes, that is sets of numbers $\{\sum_{s\in S}\epsilon_s s : \ \epsilon_s \in \{-1,0,1\}\}$ for given S, were discussed in the previous two lectures.)

III.5. Four term arithmetic progressions

One can prove (using the proof of Theorem 2) that if $A \gg N$ and $\hat{A}(m) = o(N)$ for all $m \not\equiv 0 \pmod N$ then A has $\sim |A|^3/N$ 3-term arithmetic progressions. This leads one to ask: *What about 4-term arithmetic progressions?*
Does $\hat{A}(m) = o(N)$ imply that A has $\sim |A|^4/N^2$ 4-term arithmetic progressions (that is, the expected number)? As an example consider the set

$$A_\delta := \left\{ n \pmod N : \left\| \frac{n^2}{N} \right\| < \frac{\delta}{2} \right\}$$

for N prime. For $J = \delta N/2$ we have

$$\hat{A}_\delta(m) = \sum_{n \pmod N} e\left(\frac{mn}{N}\right) \sum_{-J<j<J} \frac{1}{N} \sum_{r \pmod N} e\left(r\frac{(j-n^2)}{N}\right)$$

$$\text{so that} \qquad |\hat{A}_\delta(m)| \le \frac{1}{N} \sum_{r \pmod N} \sum_{-J<j<J} e\left(\frac{rj}{N}\right) \sum_{n \pmod N} e\left(\frac{mn-rn^2}{N}\right).$$

Now $\sum_n e(\frac{mn}{N}) = 0$ if $m \neq 0$, and $= N$ if $m = 0$; and if $r \neq 0$ then $\sum_n e(\frac{mn-rn^2}{N})$ is a Gauss sum and so has absolute value $\sqrt{N}$. Moreover $|\sum_{-J\le j\le J} e(\frac{rj}{N})| \ll N/|r|$ for $1 \le |r| \le N/2$. Inputting all this into the equation above we obtain $|\hat{A}_\delta(m)| \ll \sqrt{N}\log N$ for each $m \not\equiv 0 \pmod N$ and $\#A_\delta = |\hat{A}_\delta(0)| = \delta N + O(\sqrt{N}\log N)$. It follows from the proof of Theorem 2 that A_δ has $\sim \delta^3 N^2$ 3-term arithmetic progressions $a, a+d, a+2d$. Now

$$(a+3d)^2 = 3(a+2d)^2 - 3(a+d)^2 + a^2$$

so if $a, a+d, a+2d \in A_\delta$ then $\|\frac{(a+3d)^2}{N}\| < \frac{7\delta}{2}$, and hence $a, a+d, a+2d, a+3d \in A_{7\delta}$. But this implies that $A_{7\delta}$ has $\ge \{1+o(1)\}\delta^3 N^2$ 4-term arithmetic progressions far more than the expected, $\sim (7\delta)^4 N^2$, once δ is sufficiently small.

Thus we have shown, from this example, that in order to prove that a set of residues of positive density has the expected number of 4-term arithmetic progressions it is insufficient to simply assume that all of the Fourier transforms are small. What else we need to assume is at the heart of the subject of additive combinatorics – see Ben Green's article in these proceedings [G4].

Acknowledgements: Thanks to Jason Lucier for his careful reading of these notes.

References

[Ba] A. Balog, *Many additive quadruples*, Herein.
[BS] A. Balog and E. Szemerédi, *A statistical theory of set addition*, Combinatorica **14** (1994), 263-268.
[Bi] Yuri Bilu, *Structure of sets with small sumsets*, Astérisque **258** (1999), 77-108.

[B1] J. Bourgain, *On triples in arithmetic progressions*, GAFA **9** (1999), 968–984.
[B2] J. Bourgain, *Mordell's exponential sum estimate revisited*, J. Amer. Math. Soc **18** (2005), 477–499.
[BC] J. Bourgain and Mei-Chu Chang, *Exponential sum estimates over subgroups and almost subgroups of $\mathbb{Z}_Q^*$, where Q is composite with few prime factors.*, GAFA **16** (2006), 327–366.
[BG] J. Bourgain, A.A. Glibichuk and S.V. Konyagin, *Estimates for the number of sums and products and for exponential sums in fields of prime order*, J. London Math. Soc **73** (2006), 380–398..
[BK] J. Bourgain, N. Katz and T. Tao, *A sum-product estimate in finite fields, and applications*, GAFA **14** (2004), 27–57.
[C1] Mei-Chu Chang, *A polynomial bound in Freiman's theorem*, Duke Math J **113** (2002), 399-419.
[C2] Mei-Chu Chang, *Some problems related to sum-product theorems*, Herein.
[Fr] G.A. Freiman, *Foundations of a structural theory of set addition*, vol. 37, (Translations of Mathematical Mongraphs) AMS, Providence, R.I., 1973.
[Fu] H. Furstenberg, *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math **31** (1977), 204–256.
[G1] W.T. Gowers, *A new proof of Szemerdi's theorem for arithmetic progressions of length four*, GAFA **8** (1998), 529–551.
[G2] W.T. Gowers, *A new proof of Szemerédi's theorem*, GAFA **11** (2001), 465–588.
[G3] Ben Green, *Structure theory of set addition*, `http://www-math.mit.edu/~green/icmsnotes.pdf` .
[G4] Ben Green, *Quadratic Fourier analysis*, Herein.
[HR] H. Halberstam and K. Roth, *Sequences*, Springer-Verlag, London, 1966.
[HB] D.R. Heath-Brown, *Integer sets containing no arithmetic progressions*, JLMS **35** (1987), 385–394.
[He] H. Helfgott, *Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$* (to appear).
[Kr] B. Kra, *Ergodic methods in combinatorial number theory*, Herein.
[Ku] P. Kurlberg, *Bounds on exponential sums over small multiplicative subgroups*, Herein.
[Pl] H. Plünnecke, *Eigenschaften und Abschätzungen von Wirkingsfunktionen*, Gesellschaft für Mathematik und Datenverarbeitung (1969), Bonn.
[Ro] K.F. Roth, *On certain sets of integers*, JLMS **28** (1953), 104–109.
[R1] Imre Ruzsa, *An application of graph theory to additive number theory*, Sci. Ser. A **3** (1989), 97-109.
[R2] Imre Ruzsa, *Generalized arithmetic progressions and sumsets*, Acta Math Hungar **65** (1994), 379-388.
[R3] Imre Ruzsa, *Cardinality questions about sumsets*, Herein.
[S1] Jozsef Solymosi, *On the number of sums and products*, Bull. London Math. Soc **37** (2005), 491–494.
[S2] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progressions*, Acta Arithmetica **27** (1975), 299-345.
[S3] E. Szemerédi, *Integer sets containing no arithmetic progressions*, Acta Math Hungar **56** (1990), 155–158.
[S4] E. Szemerédi, *An old new proof of Roth's theorem*, Herein.
[TV] T. Tao and V.H. Vu, *Additive Combinatorics*, Cambridge studies in advanced math, vol. 105, Cambridge University Press, Cambridge, 2006.
[Ta] T. Tao, *The ergodic and combinatorial approaches to Szemerédi's theorem*, Herein.
[W] H. Weyl, *Über ein Problem aus dem Gebeit der diophantischen Approximationen*, Nachr. Ges. Wiss. Göttingen (math.-phys. Kl.) (1914), 234–244.