Arithmetic progressions in certain sumsets

Norbert Hegyvári
Eötvös University and Rényi Institute, Budapest

Atelier de Combinatoire Additive

2014, May 5
The simplest and most structured objects of the set of integers are the arithmetic progressions.
Introduction

One of the first combinatorial question of this type is the

Theorem (Van der Waerden)

For any given positive integers $r$ and $k$, there is some least number $N = W(k, r)$ such that if the integers $\{1, 2, \ldots, N\}$ are colored, each with one of $r$ different colors, then there is a monochromatic $k$–term arithmetic progression.

There are no strict bounds for $W(k, r)$. The best upper bound currently known is

Theorem (T. Gowers)

$$W(r, k) < 2^{2^{r2^k+9}}.$$
The density version of this theorem is the celebrated

**Theorem (Szemerédi)**

If $A \subseteq \mathbb{N}$ and $\overline{d}(A) > 0$, then for every $k$ there exists a $k$–term arithmetic progression containing in $A$, where $\overline{d}(A)$ is the upper density of $A$ defined by

$$\overline{d}(A) := \limsup_{n \to \infty} \frac{|A \cap [1, n]|}{n}$$
Introduction

Theorem (Green - Tao)

For every $k$ there exists a $k$–term arithmetic progression containing in sequence of primes

Many generalization:

Theorem (J. Pintz)

There exists $K \in \mathbb{N}$, such that for every $k$ there exists a $k$–tuple of prime-pairs $\{(p_i, p'_i)\}_{i=1}^k$, where $\{p_i\}_{i=1}^k$, forms a $k$–term arithmetic progression and for $i = 1, 2, \ldots k$ $p_i - p'_i = K$.

Note: Pintz’s theorem remains true if we assume just

$$p_i - p'_i < K' \ i = 1, 2, \ldots k.$$
Arithmetic progressions in $k$–fold sumsets

• The case $k = 2$

Theorem (Bourgain)

Let $A, B \subseteq \{1, 2, \ldots, N\}$ be sets with $|A| = \alpha N$, $|B| = \beta N$. Then there is a constant $c = c(\alpha, \beta)$ such that $A + B$ contains an arithmetic progression of length $e^{c(\log N)^{1/3}}$.

Green and others improved $1/3$ to $1/2$ which is the best possible thanks to Ruzsa.
The case $k \geq 3$

**Theorem (Freiman, Halberstam and Ruzsa)**

Let $A \subseteq \mathbb{Z}_p$ with $|A| = \gamma p$. Then $kA$ contains $\gamma p/2$ many arithmetic progressions with common differences, and length $c_1 p^{c_2}$ where $c_1 = c_1(k)$ and $c_2 = c_2(k, \gamma)$. 
Definition:

• Let $A \subseteq \mathbb{N}$, define the subset sums $P(A)$ of $A$ as follows:

$$P(A) := \left\{ \sum_{b \in B} b : B \subseteq A; \ |B| < \infty \right\},$$

and when $B = \{\emptyset\}$, we mean $\sum_{b \in B} b = 0$.

• If $P(A)$ contains an infinite arithmetic progression, we say that $A$ is subcomplete,

• If the difference of this infinite arithmetic progression is 1 we say that $A$ is complete.
In 1962 Erdős conjectured that for $A \subseteq \mathbb{N}$, the condition $A(n) > c\sqrt{n}$, $(c > 0)$ yields that $A$ is subcomplete. (it is easy to see that we could not give a better bound than $c\sqrt{n}$). He proved a weaker result :

**Theorem (Erdős)**

Assume that $A \subseteq \mathbb{N}$, and for some $c > 0$ $A(n) > cn^{(\sqrt{5}-1)/2}$. Then $A$ is subcomplete.

One year later Folkman improved :

**Theorem (Folkman)**

Assume that $A \subseteq \mathbb{N}$, and for some $c > 0$ $A(n) > cn^{1/2+\varepsilon}$. Then $A$ is subcomplete.
After a comparatively long break I improved (and independently in the same year T. Łuczak and T. Schoen) it to

Theorem (H, indep. Ł-S)

Assume that $A \subseteq \mathbb{N}$, and $A(n) > 300 \sqrt{n \log n}$. Then $A$ is subcomplete

Finally Endre Szemerédi and Van Vu (and independently Y. Gao-Chen) could prove the conjecture of Erdős:

Theorem (Szemerédi- Vu, indep. Gao-Chen)

*If for some $c > 0$ $A(n) > c \sqrt{n}$, then $A$ is subcomplete*
A simple example for complete sets is the 2-powers, and clearly \( Y_0 = \{ p^n : n = 0, 1, \ldots \} \), is complete if and only if \( p = 2 \).

Erdős asked the following: Take \( Y = \{ p^n q^m : n, m = 0, 1, \ldots \} \), where \( 1 < p, q \in \mathbb{N} \)

A plausible conjecture (of Erdős again) that \( Y \) is complete if and only if \( \gcd(p, q) = 1 \).
Arithmetic progressions in subset sums of exponential type sets

In 1959 Birch proved this conjecture. Later Cassels proved a more general result which implies Birch’s result:

**Theorem (Cassels)**

Let $A \subseteq \mathbb{N}$, and assume that $\lim_{n \to \infty} \frac{A(2n) - A(n)}{\log \log n} = \infty$. Furthermore assume that for all $\theta$ real number, $(0 < \theta < 1)$ $\sum_{i=1}^{\infty} \|a_i \theta\| = \infty$. Then $A$ is complete.
Then Davenport and Erdős made a stronger conjecture:
For all $p, q > 1$, $\gcd(p, q) = 1$ there exists a $K = K(p, q)$ such that the set $Y_K = \{p^n q^m : n = 0, 1, \ldots \ 0 \leq m \leq K\}$ is complete.

Erdős wrote:
"Of course the exact value of $K(p, q)$ is not known and no doubt will be very difficult to determine."

Unfortunately there are no good (lower-upper) bound for $K = K(p, q)$. 
I proved

**Theorem (H)**

*With conditions above*

\[ K(p, q) \leq 2p^{2c^2q^{4p+3}}. \]

Later Gao-Chen and J.-Fang could reduce one step of mine, nevertheless their bound is also ”tower-exponential” (with one less height).
Perturbated Graham sequences

R. Graham asked the following:

For which pairs of positive reals \((\alpha, \beta)\) is the sequence
\[
\{[2^n \alpha], [2^m \beta] : n, m \in \mathbb{N}\}
\]
complete?

(I obtained some result, say: when \(\alpha\) is a finite and \(\beta\) is an infinite diadal fraction then it is complete).

With Gerard Rauzy we proved the completeness of an Erdős-Birch-Graham type set by proving

**Theorem (H-Rauzy)**

Let \(B\) be an arbitrary infinite sequence of positive integers. Then the set
\[
\{b_m[2^n \alpha], : n, m \in \mathbb{N}; b_m \in B\}
\]
is complete.
Complete sequences in the integer lattice

For $A \subseteq \mathbb{N}^2$, we define the subset sums $P(A)$ of $A$ in the same way:

$$P(A) := \left\{ \sum_{b \in B} b : B \subseteq A; |B| < \infty \right\},$$

We can define two types of completeness:

• $A$ is $L$-complete (line complete)
  If there is a line $L(x_0, m) := \{ x_0 + m \cdot t : t \in \mathbb{N} \} \subseteq P(A) \subseteq \mathbb{N}^2$
  ($x_0, m \in \mathbb{N}^2$)

• $A$ is $p$-complete (angle-plane complete)
  If there is an angle-region
  $S(x_0, u, v) := \{ x_0 + \alpha \cdot u + \beta \cdot v : \alpha, \beta \in \mathbb{R}^+ \} \cap \mathbb{N}^2 \subseteq P(A)$, where
  ($x_0, u, v \in \mathbb{N}^2$)
Complete sequences in the integer lattice

$\mathbb{N} \leftarrow \text{BIG DIFFERENCE} \rightarrow \mathbb{N}^2$

Recall: $A(N) \gg \sqrt{N}$ implies $A$ is subcomplete

Theorem (H)

There exists an $A \subseteq \mathbb{N}^2$ for which

$A(N) \gg N^2,$

and $A$ is not $L$-complete
A SUFFICIENT CONDITION FOR COMPLETENESS IN \( \mathbb{N} \): If \( A \subseteq \mathbb{N} \) and \( A = A_1 \sqcup A_2 \) for which \( P(A_1) \) has bounded gaps, and \( P(A_2) \) is thick (contains sufficiently large intervals) then \( A \) is complete.

Interval in \( \mathbb{N} \) \( \rightarrow \) Lattice points a rectangle \( R(a, b) \) with sizes \( a \) and \( b \)

Bounded gaps in \( \mathbb{N} \) \( \rightarrow \) Walk in the lattice \( \mathbb{N}^2 \)
i.e. a sequence \( A = \{a_n\} \subseteq \mathbb{N}^2 \) and

\[
a_{n+1} - a_n = (0, 1) \text{ or } (1, 0).
\]
Walk in the lattice; a question of Sárközy

Theorem (H)

There exists a walk $A \subseteq \mathbb{N}^2$ for which $A = A_1 \sqcup A_2$ and

1. for every $r, s \in \mathbb{N}$ $P(A_1)$ contains a discrete rectangle $R(s, r)$

and

2. $P(A_2)$ has bounded gaps

and $A$ is not $L$-complete

Sárközy: Modify the question; what is the max $\varepsilon > 0$ for which the following true

If $A \subseteq \mathbb{N}^2$ and for every lattice point $n \in \mathbb{N}^2$ there exists a point $a \in A$ for which

$$\|a - n\| < \|n\|^{\varepsilon},$$

then $A$ is $p$-complete? ($\| \cdot \|$ is the distance from the origin)
I proved

Theorem (H)

\[ \frac{1}{8} \leq \max \varepsilon < \frac{\sqrt{13} + 1}{6} \approx 0.7675 \]
Another type of completeness

Completeness in groups

Theorem (Olson)

If $p$ is a prime and $A$ is a subset of $\mathbb{Z}_p$ with cardinality larger than $2\sqrt{p}$, then $A$ is complete.

It is extend by Van Vu

Theorem (Vu)

There is a constant $C$ such that the following holds. Let $n$ be a sufficiently large positive integer and $A$ be a subset of $\mathbb{Z}_n$, where $|A| > C\sqrt{n}$, and the elements of $A$ are co-primes with $n$. Then $A$ is complete.
Another type of completeness

Completeness in groups and fields

Reacher structure: $\mathbb{F}_p$

Using two operations:

**Theorem (Sárközy)**

For $A, B, C, D \subseteq \mathbb{F}_p$, the equation $a + b = cd$, $(a, b, c, d) \in A \times B \times C \times D$ has a solution, provided $|A||B||C||D| > p^3$.

From this

**Theorem**

For every $n \in \mathbb{F}_p$ there exist $a_1, a_2, a_3, a_4 \in A$, such that $n = a_1 + a_2 + a_3 \cdot a_4$, provided $|A| > p^{3/4}$. 
Completeness in groups and fields

Reacher structure: $\mathbb{F}_p$

Other generalizations

Define the following four maps:

$G_u(x, y) = x^{1+u}y + x^{2-u}h(y)$ where $u \in \{0, 1\}$, $h(y) \in \mathbb{Z}[y]$ is a non constant polynomial.

$F_{p,v}(x, y) = x^{1+u}y + x^{2-u}g_p^y$ for any $p$ where $g_p$ generates $\mathbb{F}_p^\times$ and $v \in \{0, 1\}$ is fixed.
Another type of completeness

Completeness in groups and fields

Reacher structure: $\mathbb{F}_p$

We proved

**Theorem (H-Hennecart)**

There exist real numbers $0 < \delta, \delta' < 1$ s.t. for any $p$ and for any sets $A, B, C, D \subseteq \mathbb{F}_p$ with

$$|C| > p^{1/2-\delta}, \quad |D| > p^{1/2-\delta} \quad |A||B| > p^{2-\delta'},$$

for every $n, m \in \mathbb{F}_p$ there exist $a, a' \in A, b, b' \in B, c, c' \in C, d, d' \in D$ for which

$$n = a + b + F_{p,v}(c, d)$$

and

$$m = a' + b' + G_u(c', d').$$
Application of (almost) completeness

On structure theorems in Heisenberg group over primefield

With elements

\[ [x, y, z] = \begin{pmatrix} 1 & x & z \\ 0 & l_n & ty \\ 0 & 0 & 1 \end{pmatrix}, \]

where \( x = (x_1, x_2, \ldots, x_n) \), \( y = (y_1, y_2, \ldots, y_n) \), \( x_i, y_i, z \in \mathbb{F} \), \( i = 1, 2, \ldots, n \), and \( l_n \) is the \( n \times n \) identity matrix.

and operations

\[ [x, y, z][x', y', z'] = [x + x', y + y', \langle x, y' \rangle + z + z'], \]

where \( \langle \cdot, \cdot \rangle \) is the inner product

Two results related to (almost) completeness
Application of (almost) completeness

For the third coordinate of a special sets of Heisenberg group we have

**Theorem (H-Hennecart)**

Let $n, m \in \mathbb{N}$, $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots Y_n \subseteq \mathbb{F}_p^* = \mathbb{F}_p \setminus \{0\}$, $Z \subseteq \mathbb{F}_p$. We have

$$mZ + \sum_{j=1}^{n} X_j \cdot Y_j := \left\{ z_1 + \cdots + z_m + \sum_{j=1}^{n} x_j y_j, \ z_i \in Z, \ x_j \in X_j, \ y_j \in Y_j \right\} = \mathbb{F}_p,$$

provided

$$|Z|^2 \prod_{i=1}^{n} |X_i|^n \prod_{i=1}^{n} |Y_i|^n > p^{n(n+1)+2}.$$
Application of (almost) completeness

Freiman model in Heisenberg groups

A key step at the proof of Freiman’s theorem is the following: if

\[ |A + A| < K|A| \]

then \( A \) can be embedded into a group \( G \) by a \( t \)-Freiman isomorphism such that

\[ |G| \leq C(K, t)|A|. \]

With F. Hennecart we prove that at the Heisenberg groups the picture is completely different.

One of the tool is related to ”almost completeness”
Application of (almost) completeness

We start a ”big” subset $A$ of

$$A_0 = \{[x, y, z] : 0 \leq x < p^\alpha, \ y, z \in \mathbb{F}_p\},$$

Taking projections we obtain $x_0, y_0, z_0, z'_0, u, v \in \mathbb{F}_p$ and $X, Y, Z \subset \mathbb{F}$ such that :

$$[X, y_0, z_0] \cup [x_0, Y, z'_0] \cup [u, v, Z] \subset A$$

$$|X| \geq \frac{|A|}{p^2}, \quad |Y| \geq \frac{|A|}{p^{1+\alpha}}, \quad |Z| \geq \frac{|A|}{p^{1+\alpha}}.$$ 

For $(x, y, z) \in X \times Y \times Z$, one has

$$[x, y_0, z_0][x_0, y, z'_0][x, y_0, z_0]^{-1}[x_0, y, z'_0]^{-1}[u, v, z] = [u, v, xy + z - x_0y_0].$$
Application of (almost) completeness

The third coordinate of

\[ [u, v, xy + z - x_0y_0] \]

is a "sum-product" expression

Let \( R(t) \) be the number of triples \((x, y, z) \in X \times Y \times Z\) such that

\[
t = xy + z - x_0y_0, \quad C := \{ t : R(t) > 0 \}.
\]

By Cauchy one has

\[
|C| \geq \frac{(|X||Y||Z|)^2}{\sum_t R(t)^2}.
\]
∑ \in \mathbb{R}^t \ counts the incidence of solutions of
\[ xy + z = x'y' + z', \quad x, x' \in X, \quad y, y' \in Y, \quad z, z' \in Z. \]

Fixing
\[ x = x_1, \quad x' = x_1', \quad z' = z_1', \]

it is an equation for a hyperplane \( D_{x_1, x'_1, z'_1} \) in \( \mathbb{F}_p^3 \):
\[ x_1 y - x'_1 y' + z - z'_1 = 0. \]

These hyperplanes are different and there are \( |X|^2 |Z| \) such hyperplanes.
The possible number of points \((y, y', z) \in Y \times Y \times Z\) is \( |Y|^2 |Z| \).
Application of (almost) completeness

A useful result is

**Lemma (Vinh)**

Let $d \geq 2$. Let $\mathcal{P}$ be a set of points in $\mathbb{F}_p^d$ and $\mathcal{H}$ be a set of hyperplanes in $\mathbb{F}_p^d$. Then

$$\left| \{(P, D) \in \mathcal{P} \times \mathcal{H} : P \in D\}\right| \leq \frac{|\mathcal{P}||\mathcal{H}|}{p} + (1 + o(1))p^{(d-1)/2}(|\mathcal{P}||\mathcal{H}|)^{1/2}.$$ 

With $d = 3$, we get for any large $p$

$$\sum_t R(t)^2 \leq \frac{(||X||Y||Z||)^2}{p} + 2p||X||Y||Z||,$$

and hence

$$|C| \geq p - \frac{2p^3}{||X||Y||Z||}.$$
When

$$|X||Y||Z| = o(p^3)$$

the set of the third coordinate is almost everything.

Merci pour votre attention