Additive Structure in Sumsets (An Introduction & Easy Results)

Many familiar theorems in mathematics have the following common feature:

"The set of differences from a suff. large set contains non-trivial structure."

<u>(Infinite) examples</u> that "taking differences is a smoothing operation":

1. (Steinhaus) If $A \subseteq \mathbb{R}$ and $m(A) > 0$, then $A-A$ contains interval centred at origin

2. (Bourgain) If $A \subseteq \mathbb{R}^n$ $(n \geq 2)$ measurable with positive upper density,

   i.e. $\displaystyle \lim_{R \to \infty} \frac{m(A \cap B_R(0))}{R^n} \neq 0$, then

   (centred at origin)

   $A-A$ intersects <u>all</u> sufficiently large spheres.

3. (Sárközy) If $A \subseteq \mathbb{N}$ with positive upper density, i.e. $\displaystyle \lim_{N \to \infty} \frac{|A \cap \{1, \dots, N\}|}{N} \neq 0$

   then $A-A$ contains infinitely many integers of the form n^2 & $p-1$

   prime.

<u>Exercise 1</u>: Show that if $A \subseteq \mathbb{N}$ with positive upper density, then $A-A$ is <u>syndetic</u> (i.e. has bounded gaps).

- Central theme in arithmetic combinatorics:

  "If $A, B \subseteq G$ (finite subsets of group G) are large, what can one say about the structure of $A+B$?"

In this lecture we focus on $A, B \subseteq \mathbb{Z}_N$, or $\{1, \dots, N\}$. $\longrightarrow$

A simple structure theorem for $A-A$ when $A \subseteq \mathbb{Z}_N$

Ruzsa's Covering lemma

Theorem 1: If $A \subseteq \mathbb{Z}_N$ with $|A| = \delta N$, then if $\delta \geq \frac{1}{m}$ we can cover $\mathbb{Z}_N$ with no more than m translates of $A-A$.

* Note that we are covering $\mathbb{Z}_N$ with $A-A$ (<u>not</u> A). This reflects the fact that $A-A$ is smoother than A and tends to contain less "holes" which would render it unsuitable for covering $\mathbb{Z}_N$.

<u>Proof</u>: Consider the family

$$A_t = \{A + t : t \in \mathbb{Z}_N\}.$$

of translates of A by elements in $\mathbb{Z}_N$. Note that $|A_t| = |A|$. for all $t \in \mathbb{Z}_N$.

Let T be the <u>largest</u> collection of shifts such that

$$\{A+t : t \in T\} \text{ are pairwise disjoint.}$$

Since

$$|T||A| = \left| \bigcup_{t \in T} A_t \right| \leq N$$

it follows that $|T| \leq 1/\delta$. <u>We now show that $\mathbb{Z}_N \subseteq T + (A-A)$:</u>

Let $x \in \mathbb{Z}_N$, $\exists\, t \in T$ such that $(t+A) \cap (x+A) \neq \emptyset$ (by maximality assumption.)

$$\Rightarrow \quad x - t \in A - A \Leftrightarrow x \in t + (A-A)$$

as required. $\square$

We note that a simple averaging argument allows us to deduce from Theorem 1 the following structural result for $A+B$ with $A,B \subseteq \mathbb{Z}_N$.

<u>Corollary 1</u>: Let $A,B \subseteq \mathbb{Z}_N$ & $m \in \mathbb{N}$.

If $\dfrac{|A||B|}{N^2} \geq \dfrac{1}{m}$, then $\mathbb{Z}_N$ can be covered by m translates of $A+B$

<u>Proof</u>: Since

$$\sum_{t \in \mathbb{Z}_N} |B \cap (t-A)| = |A||B|$$

it follows $\exists\ t \in \mathbb{Z}_N$ such that, if we set $D = B \cap (t-A)$, then

$$|D| \geq \frac{|A||B|}{N}.$$

The result follows from Theorem 1 since $D - D \subseteq (A+B) - t$. $\square$

<u>Arithmetic Progressions in Sumsets</u>

A good measure of the amount of additive structure in a given subset of $\{1,\dots,N\}$ (or $\mathbb{Z}_N$) is provided by the size of the longest arithmetic progression that the set contains.

Using only simple combinatorial arguments, Croot, Ruzsa & Schoen establish the following result along these lines for $A+B$ with $A,B \subseteq \{1,\dots,N\}$.

<u>Theorem 2</u> (Croot, Ruzsa, Schoen)

Let $A, B \subseteq \{1, \ldots, N\}$ with densities α & β and $\ell \in \mathbb{N}$, then

$A+B$ contains arith. prog. of length $2\ell+1$, provided $\alpha\beta \geq \frac{C}{N^{1/\ell}}$.

<u>Proof</u> As with Corollary 1, it suffices to show that if $A \subseteq \{1, \ldots, N\}$ with density α and $\ell \in \mathbb{N}$, then $A-A$ will contain an arith. prog. of length $2\ell+1$ provided $\alpha \geq CN^{-1/\ell}$.

(Change needed in Remark below)

For each $w = (w_1, \ldots, w_\ell) \in \mathbb{Z}^\ell$ we define

$$R_w = \{ r \in \{1, \ldots, N/\ell\} : jr + w_j \in A \ (1 \leq j \leq \ell) \}$$

$N/2\ell$ $\quad$ $2j$

and note that if, for some $w \in \mathbb{Z}^\ell$, $\exists\, r', r'' \in R_w$ with $r' \neq r''$, then it will follow immediately that

$$jr', jr'' \in A - w_j \iff j(r'-r'') \in A-A$$

for all $1 \leq j \leq \ell$ and hence, utilizing fact that $A-A$ is symmetric, that $A-A$ will contain an arith. prog. of length $2\ell+1$.

In order to show that $|R_w| \geq 2$ for some $w \in \mathbb{Z}^\ell$ we will naturally restrict our attention to those w for which R_w has at least a chance of being non-empty, namely $\rightarrow$ Could just do $1-N$?

$$W = \{ w \in \mathbb{Z}^\ell : 1 - \frac{jN}{\ell} \leq w_j \leq N-1 \ (1 \leq j \leq \ell) \}.$$

<u>Note</u>: $$|W| \leq N^\ell \prod_{j=1}^{\ell} (1+ j/\ell) \leq 2^\ell N^\ell.$$

Since
$$\frac{1}{|W|}\sum_{w\in W} |R_w| = \frac{1}{|W|}\sum_{w\in W}\sum_{r=1}^{N/\ell}\prod_{j=1}^{\ell} 1_A(jr+w_j) = \frac{1}{|W|}|A|^\ell \frac{N}{\ell}$$

it follows that $\exists\, w\in W$ s.t.
$$|R_w| \geq \left(\frac{|A|}{N}\right)^\ell \frac{N}{\ell 2^\ell}$$

and consequently, for this choice of w, the set R_w will satisfy $|R_w|\geq 2$ provided
$$\frac{|A|}{N} \geq C_\ell \frac{1}{N^{1/\ell}}$$

where $C_\ell = 2(2\ell)^{1/\ell}$. [Note $2\leq C_\ell \leq 4$] $\square$.

<u>Remark</u>: Same argument as given above shows:

If $A\subseteq\{1,\dots,N\}$ with density α & $\ell\in\mathbb{N}$, then
$$A-A \supseteq \pm\{r, 2r, 4r, \dots, 2^{\ell-1}r\} \text{ for some } r\neq 0.$$

Consequently, using the fact that $\ell T-\ell T \supseteq r\cdot\{-2^\ell,\dots,2^\ell\}$, it follows that
$$\ell T-\ell T \supseteq P$$
where P is some arithmetic progression with $|P|\geq 2^{\ell+1}$.

Further Remarks

If $\alpha\beta \geq (\log N)^{-1+\varepsilon}$, for any $\varepsilon > 0$, then the conclusion of Theorem 2 can be strengthened significantly. Using Fourier analytic techniques Green, improving on earlier work of Bourgain, prove the following

Theorem 3 (Green, 2002) If $A, B \subseteq \{1, \ldots, N\}$ with densities α & β, then $A+B$ contains an arith. prog. of length at least

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

* In the next set of notes we give a simple new proof of this result, using random sampling in frequency space (but, surprisingly, very little Fourier analysis) due to Croot, Laba & Sisask.

* In the same paper, Croot, Laba & Sisask, prove the following strengthening of Theorem 3.

Theorem 4 (Croot, Laba, Sisask, 2011) If $A, B \subseteq \{1, \ldots, N\}$ with densities α & β, then $A+B$ contains an arith. prog. of length at least

$$\exp\left(c\left(\frac{\alpha\log N}{(\log(2\beta^{-1}))^3}\right)^{1/2} - \log(\beta^{-1}\log N)\right).$$

The improvement of Theorem 4 over Theorem 3 enters when one of the sets has density decreasing with N.

* Theorem 3 only yields non-trivial results if at least one of the sets has density at least $\frac{\log\log N}{\sqrt{\log N}}$ and both sets have density at least $\frac{(\log\log N)^2}{\log N}$.

* Theorem 4 allows for both sets to have density about

$$\frac{(\log\log N)^c}{\log N}$$

and allows one of the sets to have density as low as

$$\exp(-(\log N)^c).$$