Long Arithmetic Progressions in Sumsets via Random Sampling in Frequency Space

In this lecture we give a simple proof (due to Croot, Laba & Sisask) of the following result.

<u>Theorem</u> (Green, 2002) If $A, B \subseteq \{1, \ldots, N\}$ with densities α & β, then $A+B$ contains an arithmetic progression of length at least

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

provided $\alpha\beta \geqslant C(\log\log N)^2/\log N$.

We will deduce this from:

<u>Proposition</u> Let $p \geqslant 2$ & $\varepsilon > 0$. If $A, B \subseteq \mathbb{Z}_N$ with densities α & β, then $\exists$ symm. arith. prog. Q with $|Q| \geqslant \varepsilon N^{c\varepsilon^2/p} = \exp\left(\frac{c\varepsilon^2\log N}{p} - \log \varepsilon^{-1}\right)$ such that

$$\mathbb{E}_x \left| 1_A * 1_B(x+t) - 1_A * 1_B(x)\right|^p \leqslant \varepsilon^p (\alpha\beta)^{p/2}$$

for all $t \in Q$.

<u>Proof of Theorem:</u> By embedding the interval $\{1, \ldots, N\}$ in the cyclic group $\mathbb{Z}_{N'}$, with $4N \leqslant N' \leqslant 8N$ prime, it suffices to establish the result for cyclic gps instead of intervals. We shall therefore prove the result for $A, B \subseteq \mathbb{Z}_N$ with densities α & β.

For values of $p \geq 2$ and $\varepsilon > 0$ to be determined, let Q be the arithmetic progression from the proposition. Our goal is to show that $\exists\, x$ such that

$$1_A * 1_B (x+t) > 0 \quad \forall\, t \in Q,$$

since from this it would follow that $x + Q \subseteq A + B$.

Note that

$$\begin{aligned} -\min_{t \in Q} 1_A * 1_B(x+t) &= \max_{t \in Q} \left(-1_A * 1_B(x+t)\right) \\ &\leq \underbrace{\max_{t \in Q} \left| 1_A * 1_B(x+t) - 1_A * 1_B(x) \right| - 1_A * 1_B(x)}_{(*)} \end{aligned}$$

and that it suffices to show that $\mathbb{E}_x (*) < 0$.

Since $\mathbb{E}_x 1_A * 1_B(x) = \alpha\beta$, we are thus trying to show:

$$\max_{t \in Q} \left| 1_A * 1_B(x+t) - 1_A * 1_B(x) \right| < \alpha\beta.$$

Since

$$\begin{aligned} \max_{t \in Q} \left| 1_A * 1_B(x+t) - 1_A * 1_B(x) \right| &\leq \mathbb{E}_x \Big(\sum_{t \in Q} \left| \text{---}\,''\,\text{---} \right|^p \Big)^{1/p} \\ \text{Hölder} \quad &\leq \Big(\mathbb{E}_x \sum_{t \in Q} \left| \text{---}\,''\,\text{---} \right|^p \Big)^{1/p} \\ &\leq |Q|^{1/p} \varepsilon (\alpha\beta)^{1/2} \quad \text{by Proposition} \\ &\leq \frac{|Q|^{1/p}}{e} (\alpha\beta). \end{aligned}$$

$\boxed{\text{Now choose } \varepsilon = \frac{(\alpha\beta)^{1/2}}{e}}$

So done if $|Q| < e^p$, which follows for all $p \geq C\sqrt{\alpha\beta \log N}$. $\square$

Proof of Proposition

Let $f = 1_A * 1_B$. Our goal is to show that $\exists\, \Gamma \subseteq \mathbb{Z}_N$ with $|\Gamma| \geq Cp/\varepsilon^2$ such that

$$\mathbb{E}_x\, |1_A * 1_B(x+t) - 1_A * 1_B(x)|^p \leq \varepsilon^p (\alpha\beta)^{p/2}$$

for all $t \in B(\Gamma, c\varepsilon)$.

- Recall: The Bohr set $B(\Gamma, c\varepsilon) := \{ t \in \mathbb{Z}_N : |e^{2\pi i t\xi/N} - 1| \leq c\varepsilon \}$ contains a symm. arith. prog. of length $\varepsilon N^{1/|\Gamma|}$.

Recall the Fourier inversion formula:

$$f(x) = \sum_{\xi \in \mathbb{Z}_N} \hat{f}(\xi) e^{2\pi i x\xi/N}$$

$$\Updownarrow$$

$$\frac{f(x)}{\|\hat{f}\|_{\ell^1}} = \sum_{\xi \in \mathbb{Z}_N} \frac{|\hat{f}(\xi)|}{\|\hat{f}\|_{\ell^1}} \left(\frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i x\xi/N} \right)$$

where $\|\hat{f}\|_{\ell^1} = \sum_{\xi \in \mathbb{Z}_N} |\hat{f}(\xi)|$.

"We can thus view $\frac{f(x)}{\|\hat{f}\|_{\ell^1}}$ as the expectation of a complex-valued random variable $X(x)$, where $X(x)$ takes the value $\frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i x\xi/N}$ with probability $\frac{|\hat{f}(\xi)|}{\|\hat{f}\|_{\ell^1}}$."

Thus if we choose $\Gamma \subseteq \mathbb{Z}_N$ with $|\Gamma| = k$ randomly with each $\xi \in \mathbb{Z}_N$ being chosen independently with probability $|\hat{f}(\xi)| / \|\hat{f}\|_{\ell^1}$, the average

$$\frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i x \xi / N}$$

ought to approximate its expectation, namely $\dfrac{f(x)}{\|\hat{f}\|_{\ell^1}}$, provided k is not too small; this is indeed the case (& follows from the Marcinkiewicz – Zygmund Inequality).

<u>Claim:</u> $\exists\ \Gamma \subseteq \mathbb{Z}_N$ with $|\Gamma| \leq Cp/\varepsilon^2$ such that

$$\left(\mathbb{E}_x \left| \frac{f(x)}{\|\hat{f}\|_{\ell^1}} - \frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i x \xi / N} \right|^p \right)^{1/p} \leq \varepsilon/3 .$$

Assuming this "quantitative law of large number result" for now, lets finish the proof of the proposition. Since

$$\widehat{f(\cdot + t)}(\xi) = \hat{f}(\xi) e^{2\pi i x \xi / N}$$

we also have
$$\left(\mathbb{E}_x \left| \frac{f(x+t)}{\|\hat{f}\|_{\ell^1}} - \frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i (x+t) \xi / N} \right|^p \right)^{1/p} \leq \varepsilon/3$$

for all $t \in \mathbb{Z}_N$. Thus
$$\left(\mathbb{E}_x \left| \frac{f(x+t)}{\|\hat{f}\|_{\ell^1}} - \frac{f(x)}{\|\hat{f}\|_{\ell^1}} \right|^p \right)^{1/p} \leq \frac{2\varepsilon}{3} + \underbrace{\left(\mathbb{E}_x \left| e^{2\pi i t \xi / N} - 1 \right|^p \right)^{1/p}}_{\leq \varepsilon/3 \ \forall\ t \in B(\Gamma, \varepsilon/3)}$$

$$\Rightarrow \left(\mathbb{E}_x |f(x+t) - f(x)|^p \right)^{1/p} \leq \varepsilon \|\hat{f}\|_{\ell^1} = \varepsilon \sum_{\xi \in \mathbb{Z}_N} |\hat{1}_A(\xi)| |\hat{1}_B(\xi)| \leq \varepsilon (\alpha\beta)^{1/2} \quad \forall\ t \in B(\Gamma, \varepsilon/3)$$

By Plancherel & Cauchy-Schwarz.

Proof of Claim

This relies on the following

Lemma (Marcinkiewicz - Zygmund Inequality)

Suppose $X_1, \dots, X_k$ are independent, mean-zero complex-valued random variables with $\mathbb{E}|X_j|^p < \infty$. Then

$$\mathbb{E} \left| \frac{1}{k} \sum_{j=1}^{k} X_j \right|^p \leq \left(\frac{Cp}{k} \right)^{p/2} \mathbb{E} \left(\frac{1}{k} \sum_{j=1}^{k} |X_j|^2 \right)^{p/2}.$$

* For a proof of this result see supplementary note.

Choose $\Gamma \subseteq \mathbb{Z}_N$ with $|\Gamma| = k$ randomly with each $\xi \in \mathbb{Z}_N$ being chosen independently with probability $|\hat{f}(\xi)| / \|\hat{f}\|_{\ell^1}$. Then,

$$\mathbb{E} \left(\mathbb{E}_x \left| \frac{f(x)}{\|\hat{f}\|_{\ell^1}} - \frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i x \xi / N} \right|^p \right)$$

$$= \mathbb{E}_x \, \mathbb{E} \left| \frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \left(\frac{f(x)}{\|\hat{f}\|_{\ell^1}} - \frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i x \xi / N} \right) \right|^p$$

M-Z inequality

$$\leq \mathbb{E}_x \left(\frac{Cp}{k} \right)^{p/2} \mathbb{E} \left(\frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \left| \frac{\hat{f}(\xi)}{|\hat{f}(\xi)|} e^{2\pi i x \xi / N} - \frac{f(x)}{\|\hat{f}\|_{\ell^1}} \right|^2 \right)^{p/2}$$

$$= \left(\frac{Cp}{k} \right)^{p/2} \mathbb{E} \, \mathbb{E}_x \left(\frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \left| \text{—''—} \right|^2 \right)^{p/2}$$

Holder on $\sum_\xi$

$$\leq \left(\frac{Cp}{k} \right)^{p/2} \mathbb{E} \, \mathbb{E}_x \frac{1}{|\Gamma|} \sum_{\xi \in \Gamma} \left| \text{—''—} \right|^p$$

$$\leq \left(4Cp / |\Gamma| \right)^{p/2} \quad \text{since } \mathbb{E}_x \left| \text{—''—} \right|^p \leq 2^p.$$

$$\leq \varepsilon^p, \text{ provided } |\Gamma| \geq 4Cp/\varepsilon^2 \text{ as required.} \qquad \square$$