

# HA: Roth's Theorem

22/1

Let  $f: \mathbb{Z} \rightarrow \mathbb{C}$ , finitely supported. For  $\alpha \in \mathbb{R}$

$$\text{let } \hat{f}(\alpha) = \sum_{m \in \mathbb{Z}} f(m) e^{-2\pi i m \alpha}$$

Note  $\hat{f}(\alpha) = \hat{f}(\alpha+1)$  and  $k \in \mathbb{N} \exists f^{(k)}(\alpha) \Rightarrow \hat{f} \in C^\infty(\mathbb{T}^1)$ .

Fourier inversion:  $f(n) = \int_0^1 \hat{f}(\alpha) e^{2\pi i n \alpha} d\alpha$

Plancherel,  $\int_0^1 \hat{f}(\alpha) \overline{\hat{g}(\alpha)} d\alpha = \sum_{n \in \mathbb{Z}} f(n) \overline{g(n)}$

Convolution;  $(\text{as } \int_0^1 e^{2\pi i(n-m)\alpha} d\alpha = \delta(n-m))$

$$f * g(n) = \sum_{m \in \mathbb{Z}} f(n-m) g(m) = \sum_{m \in \mathbb{Z}} f(m) g(n-m)$$

It is easy to see that:  $\widehat{f * g}(\alpha) = \hat{f}(\alpha) \hat{g}(\alpha)$ .

Van der Waerden's thm let  $r \in \mathbb{N}$ . Then  $\exists N = N(r)$

s.t.  $\forall r$ -coloring of  $[1, N]$   $\exists$  a progression  $P = \{a, a+d, a+2d\} \subseteq [1, N]$  which is monochromatic.

Q: If  $C_1, \dots, C_r$  are the color classes, then

$$N = |C_1| + \dots + |C_r| \Rightarrow \exists i \text{ s.t. } |C_i| \geq \frac{1}{r} N$$

Is it true that  $C_i$  already contains

a 3-term progression?

Roth's theorem Let  $\delta > 0$ . Then  $\exists N = N(\delta)$ ,

such that if  $A \subseteq [1, N]$ ,  $|A| = \delta N$  then  $A$  necessarily contains a 3-term progression (a 3AP).

Note.  $\{m_1, m_2, m_3\}$  is a 3-AP  $\iff m_1 + m_2 - 2m_3 = 0$

For any  $m_1, m_2, m_3$  we have

$$\int_0^1 e^{2\pi i(m_1 + m_2 - 2m_3)\alpha} d\alpha = \begin{cases} 1, & \text{if } m_1 + m_2 - 2m_3 = 0 \\ 0, & \text{otherwise} \end{cases}$$

Let  $N(A)$  be the number of 3-AP's in  $A$ .

$$\text{Then } N(A) = \sum_{m_1 \in A} \sum_{m_2 \in A} \sum_{m_3 \in A} \int_0^1 e^{-2\pi i(m_1 + m_2 - 2m_3)\alpha} d\alpha$$

$$= \int_0^1 \sum_{m_1, m_2, m_3} |A(m_1)| |A(m_2)| |A(m_3)| e^{-2\pi i(m_1 + m_2 - 2m_3)\alpha} d\alpha$$

$$= \int_0^1 \hat{1}_A(\alpha)^2 \hat{1}_A(-2\alpha) d\alpha \quad (2.1)$$

Let  $f_A := |A(m) - \delta 1_{[N]}(m)$  be the balanced function of  $A$ ,

Definition Let  $\epsilon > 0, L > 1$ . We say that  $A$  is  $(\epsilon, L)$ -uniformly distributed, if for every progression  $P$  of length  $|P| \geq L$ , one has

$$|A \cap P| \leq (\delta + \epsilon) |P|$$

Note This means that the relative density of  $A$  on  $P$ ,  $\delta(A|P) = \frac{|A \cap P|}{|P|} \leq \delta + \epsilon$

Lemma If  $A$  is  $(\epsilon, L)$ -uniform with  $\epsilon \leq \delta^2/10$  and  $L := \frac{\epsilon}{2} N^{\frac{1}{2}}$

then  $|\hat{f}_A(\alpha)| \leq \delta \epsilon N, \forall \alpha$  (22.2)

Assuming Lemma 1,

Proof of Roth's theorem Write  $1_A = \delta 1_{[N]} + f_A$ , then

$$N(A) = \delta \underbrace{\int_0^1 \hat{1}_A(\alpha)^2 \hat{1}_{[N]}(-2\alpha) d\alpha}_M + \underbrace{\int_0^1 \hat{1}_A(\alpha)^2 \hat{f}_A(-2\alpha) d\alpha}_E$$

Main term

$$M = \delta \sum_{m_1, m_2, m_3} 1_A(m_1) 1_A(m_2) 1_{[N]}(m_3) \times \int_0^1 e^{2\pi i(m_1 + m_2 - 2m_3)\alpha} d\alpha \geq \frac{1}{4} \delta^3 N^2$$

Error term

$$|E| \leq \left( \sup_{\alpha} |\hat{f}_A(-2\alpha)| \right) \int_0^1 |\hat{1}_A(\alpha)|^2 d\alpha = \left( \sup_{\alpha} |\hat{f}_A(-2\alpha)| \right) \cdot \delta N$$

• Now, if  $A$  is  $(\epsilon, L)$ -unif. distr. with  $\epsilon = \frac{\delta^2}{10}$  and  $L := \epsilon N^{1/2}$ , then we have

$$N(A) \geq \frac{1}{4} \delta^3 N^2 - \epsilon \delta N^2 \geq \frac{1}{8} \delta^3 N^2 \stackrel{N > N_0}{\Rightarrow} A \text{ contains a } \overbrace{3\text{-AP}}^{\text{non-triv}}$$

• If  $A$  is not  $(\epsilon, L)$ -unif. distributed then

$$\exists P := x + q[-L/2, L/2] \text{ s.t. that } |A \cap P| \geq (\delta + \epsilon) |P|$$

Let  $P \mapsto [1, L]$  (with  $\phi: n \mapsto (n-x)/q$ ) and  $A \mapsto A_1$ .

Now we have  $|A_1| \geq (\delta + \delta^2/10) L$ .

Suppose indirect that  $A$  does not contain a  $\overbrace{3\text{-AP}}^{\text{non-triv}}$ .  $N > 100 \delta^{-6}$

Then  $\exists A_1 \subseteq [1, N_1]$  s.t.  $|A_1| \geq (\delta + \delta^2/10) N_1$  and  $A_1$

$A_1$  also does not contain a 3-AP.

Continuing with this procedure  $m = \lceil \frac{10}{\delta} \rceil$  times

we get  $A_m \subseteq [1, N_m]$ , s.t.  $|A_m| \geq 2\delta \cdot N_m$

After  $\lceil \frac{10}{\delta} + \frac{10}{2\delta} + \dots \rceil \leq \frac{20}{\delta}$  steps, the density becomes larger than one which is not possible.

Note

$$N_1 = \frac{\delta^2 N}{10}^{\frac{1}{2}} \Rightarrow \log N_1 \geq \frac{1}{2} \log N - 2 \log \frac{1}{\delta}$$

$$\Rightarrow \log \log N_1 \geq \log \log N - \log 10 \quad (\leftarrow \text{say})$$

$$\Rightarrow \log \log N_m \geq \log \log N - m \log 10 \quad ($$

Thus the procedure works, if  $\log \log N \geq C/\delta$  ( $\leftarrow m$ )  
 $\Leftrightarrow N \geq \exp \exp(C/\delta)$

Proof of Lemma 1

Let  $x \in [0, 1]$ . Then,  $\exists \square$

$\exists q \leq 4L$  s.t.  $\|q\alpha\| \leq \frac{1}{4L}$  by Dirichlet's principle  
 (where  $\| \beta \| = \min_{n \in \mathbb{Z}} |\beta - n|$ ).

Let  $P_0 := \{lq \mid -L/2 \leq l \leq L/2\}$ , then

$$\hat{1}_{P_0}(\alpha) = \sum_{|l| \leq L/2} e^{2\pi i l q \alpha} = 1 + 2 \sum_{l \leq L/2} \cos(2\pi l q \alpha) \geq L/2, \quad (*)$$

indeed  $\cos(2\pi l q \alpha) = \cos(2\pi l \|q\alpha\|) = \cos(\beta_l) \geq \frac{1}{2}$   
 as  $|\beta_l| \leq \pi L \times \frac{1}{4L} = \frac{\pi}{4}$

thus

$$\sum_m |f_A * 1_{P_0}(m)| \geq |\hat{f}_A(\alpha)| |\hat{1}_{P_0}(\alpha)| \geq |\hat{f}_A(\alpha)| \cdot \frac{L}{2}$$

But

$$f_A * 1_{P_0}(m) = \sum_{|l| \leq \frac{L}{2}} f_A(m - lq) = |A \cap P_m| - \delta |P_m \cap [1, N]| \leq \varepsilon L \quad (**)$$

where  $P_m = m + P_0$ , if  $P_m \subseteq [1, N]$

Note that  $P_0 \in [-2L^2, 2L^2]$  as  $q \leq 4L, |l| \leq \frac{L}{2}$ .

We have

$$\sum_m f_A * l_{P_0}(m) = \left( \sum_m f_A(m) \right) \left( \sum l_{P_0}(m) \right) = 0 \quad (*3)$$

Let  $g_+(m) := \begin{cases} g(m), & \text{if } g(m) \geq 0 \\ = 0, & \text{otherwise} \end{cases}$

We have  $(f_A * l_{P_0})_+(m) \leq \varepsilon L$ , if  $P_m = m + P_0 \in [1, N]$   
 $\leq L$ , if  $P_m \notin [1, N]$   
 $= 0$ ; if  $P_m \cap [1, N] = \emptyset$

however  $\# \{m; P_m \notin [1, N]; P_m \cap [1, N] \neq \emptyset\} \leq 4L^2$

$$\Rightarrow \sum_m (f_A * l_{P_0}(m))_+ \leq \varepsilon NL + 4L^3 \leq 2\varepsilon NL$$

$\frac{4L^2 P_m}{m}$

$$\text{as } \varepsilon N \geq 4L^2 \Rightarrow L \leq \frac{\varepsilon}{2} N^{\frac{1}{2}}$$

Since  $|g| = 2g_+ - g_-$  ( $g = g_+ - g_-$ ,  $|g| = g_+ + g_-$ )

We get

$$\sum_m |f_A * l_{P_0}(m)| \leq 4\varepsilon NL$$

$$\Rightarrow |\hat{f}_A(\alpha)| \cdot \frac{L}{2} \leq 4\varepsilon NL \Rightarrow |\hat{f}_A(\alpha)| \leq 8\varepsilon N$$

□