Szemerédi's proof of Roth's theorem

In this exercise sheet, we outline Szemerédi's proof of Roth's theorem on the existence of 3-term arithmetic progressions in large sets. Szemerédi's proof appeared some 10 years after Roth's original proof and contain elements of his general result on the existence of k-term arithmetic progressions.

As in the Fourier analytic proof the key will again be to show that if a set $A \subseteq [1, N]$ with $|A| = \delta N$ does not contain any 3-term arithmetic progressions, then there exists a "long" arithmetic progression on which the (relative) density of A increases, to say $\delta + \delta^2/20$. This can then be iterated a large number of time as before and, provided N is sufficiently large, the density of A will eventually exceed 1 on some progression, which is clearly impossible.

Throughout this sheet we will use of the convenient notation $\log N := \log_2 N$.

1. (Cube Lemma) Let $a, d_1, \ldots, d_k \in \mathbb{N}$. We define a k-dimensional cube to be a set of the form

$$Q(a, d_1, \ldots, d_k) = \{a + \varepsilon_1 d_1 + \ldots + \varepsilon_k d_k : \varepsilon_i = 0 \text{ or } 1 \text{ for all } 1 \leq i \leq k\}.$$

   (a) Let $\delta > 0$ and $k \in \mathbb{N}$. Use induction to show that if $N \geq (2/\delta)^{2^k}$ and $A \subseteq [1, N]$ with $|A| = \delta N$ then A must contain a k-dimensional cube.

   (b) Note in particular that if $\delta > 0$, $N \geq 2^{(\log 2/\delta)^2}$ and $A \subseteq [1, N]$ with $|A| = \delta N$ then A must contain a k-dimensional cube with $k \geq \frac{1}{2} \log \log N$.

2. Let $A, P \subseteq [1, N]$ and $P = P_1 \cup \cdots \cup P_k$ be a partition.

   (a) Verify that from the simple observation that

$$\frac{|A \cap P|}{|P|} = \sum_{j=1}^{k} \frac{|P_j|}{|P|} \frac{|A \cap P_j|}{|P_j|}$$

   it follows that there must exist $1 \leq j \leq k$ such that

$$\frac{|A \cap P_j|}{|P_j|} \geq \frac{|A \cap P|}{|P|}.$$

   (b) Let $\delta > 0$. Verify that if $A \subseteq P$ with $|A| = \delta N$ and $|P| \leq (1 - \delta/10)N$, then in fact there must exist $1 \leq j \leq k$ such that

$$\frac{|A \cap P_j|}{|P_j|} \geq \delta + \delta^2/20 \quad \text{and} \quad |P_j| \geq \frac{\delta^3}{20} \frac{N}{k}.$$

   (c) Note in particular that if $\delta > 0$, $|A| = \delta N$ and we know that A can be covered by a union of disjoint progressions $P_1, \ldots, P_k$ with $k \leq N/\log \log N$ whose union P satisfies $|P| \leq (1 - \delta/10)N$, then there must exist $1 \leq j \leq k$ such that

$$\frac{|A \cap P_j|}{|P_j|} \geq \delta + \delta^2/20 \quad \text{and} \quad |P_j| \geq \frac{\delta^3}{20} \log \log N.$$

3. Let $A \subseteq [1, N]$ and $d \in \mathbb{N}$. We will say that $a, b \in A$ are equivalent, and write $a \sim b$, if there exists $L \in \mathbb{N}$ such that either

$$\{a, a+d, \ldots, a+(L-1)d = b\} \subseteq A \quad \text{or} \quad \{b, b+d, \ldots, b+(L-1)d = a\} \subseteq A.$$

   (a) Verify that the relation $\sim$ is an equivalence relation whose equivalence classes are (maximal) progressions of step size d.

   (b) Show that the number of equivalence classes k, satisfies $k = |(A+d) \setminus A|$.

   (c) Verify that the complement of A can be partitioned into at most $k+d$ progressions with step size d, that is

$$[1, N] \setminus A = P_1 \cup \cdots \cup P_\ell$$

   where each P_j is a progressions with step size d and $\ell \leq k + d$.

4. Let $\delta > 0$, $N \geq 2^{2(\log 2/\delta)^2}$ and $A \subseteq [1, N]$ with $|A| = \delta N$.
   Set $B := A \cap [N/3, 2N/3]$ and suppose $|B| \geq \delta N/6$.

   (a) Show that B must contain a k-dimensional cube $Q(a, d_1, \ldots, d_k)$ with $k \geq \frac{1}{2}\log\log N$ and $1 \leq d_1 + \cdots + d_k \leq \sqrt{N}$.
   Hint: Decompose $[N/3, 2N/3]$ *in to intervals of length* $\sqrt{N}$.

   (b) For each $1 \leq j \leq k$ we set $Q_j := Q(a, d_1, \ldots, d_j)$, $Q_0 := \{a\}$ and introduce the sets

$$B_j := \{x : x = 2z - y,\ z \in Q_j \text{ and } y \in B\}.$$

   Verify that for each $0 \leq j \leq k$ the sets B_j satisfy the following:

   i. $B_j \cap A = \emptyset$ if A contains no 3-term arithmetic progressions

   ii. $|B_j| \geq |B| \geq \delta N/6$

   iii. $B_j \cup (B_j + 2d_{j+1}) \subset B_{j+1}$ (for $0 \leq j \leq k-1$)

   (c) i. Verify that there must exist $0 \leq j \leq k-1$ such that

$$|B_{j+1} \setminus B_j| \leq \frac{2N}{\log\log N}.$$

   ii. Conclude from this that the complement of the set B_j can be partitioned into at most ℓ progressions with step size $d = 2d_{j+1}$ where

$$\ell \leq d + |B_{j+1} \setminus B_j| \leq 2\sqrt{N} + \frac{2N}{\log\log N} \leq \frac{3N}{\log\log N}.$$

   Hint: The set B_j *can be partitioned into* $|(B_j + 2d_{j+1}) \setminus B_j|$ *progressions with step size* $2d_{j+1}$ *and it follows from 4(b)iii that* $|(B_j + 2d_{j+1}) \setminus B_j| \leq |B_{j+1} \setminus B_j|$.

5. Let $A \subseteq [1, N]$ and set $\delta = |A|/N$.

   (a) Verify that if A contains <u>no</u> non-trivial 3-term arithmetic progressions, then either $N < 2^{2(\log 2/\delta)^2}$ or there exists an arithmetic progression P with $|P| \geq c\delta^3 \log\log N$, such that

$$\frac{|A \cap P|}{|P|} \geq \delta + \frac{\delta^2}{20}.$$

   (b) By iterating this argument (in exactly the same way as we did in the Fourier analytic proof) we can prove Roth's Theorem. What bounds do you obtain from this combinatorial approach?