Roth's Theore

William Gasarch

September 9, 2024

0.1 Every Set of Positive Upper Density has a 3-AP

0.1.1 Combinatorial Proof

Consider the following statement:

If $A \subseteq [n]$ and $|A|$ is 'big' then A must have a 3-AP.

This statement, made rigorous, is true. In particular, the following is true and easy:

Let $n \geq 3$. If $A \subseteq [n]$ and $|A| \geq 0.7n$ then A must have a 3-AP.

Can we lower the constant 0.7? We can lower it as far as we like if we allow n to start later:

Roth [**?**, **?**, **?**] proved the following using analytic means.

$(\forall \lambda > 0)(\exists n_0 \in \mathbb{N})(\forall n \geq n_0)(\forall A \subseteq [n])[|A| \geq \lambda n \implies A \text{ has a 3-AP}]$.

The analogous theorem for 4-APs was later proven by Szemeredi [**?**, **?**] by a combinatorial proof. Szemeredi [**?**] later (with a much harder proof) generalized from 4 to any k.

We prove the $k = 3$ case using the combinatorial techniques of Szemeredi. Our proof is essentially the same as in the book *Ramsey Theory* by Graham, Rothschild, and Spencer [**?**].

More is known. A summary of what else is known will be presented in the next section.

Def 0.1.1 Let $sz(n)$ be the least number such that, for all $A \subseteq [n]$, if $|A| \geq sz(n)$ then A has a 3-AP. Note that if $A \subseteq [a, a+n-1]$ and $|A| \geq sz(n)$ then A has a 3-AP. Note also that if $A \subseteq \{a, 2a, 3a, \ldots, na\}$ and $|A| \geq sz(n)$ then A has a 3-AP. More generally, if A is a subset of any equally spaced set of size n, and $|A| \geq sz(n)$, then A has a 3-AP.

We will need the following Definition and Lemma.

Def 0.1.2 Let $k, e, d_1, \ldots, d_k \in \mathbb{N}$. The *cube on* $(e, d_1, \ldots, d_k)$, denoted $C(e, d_1, \ldots, d_k)$, is the set $\{e + b_1 d_1 + \cdots + b_k d_k \mid b_1, \ldots, b_k \in \{0, 1\}\}$. A *$k$-cube* is a cube with k d's.

Lemma 0.1.3 *Let I be an interval of $[1,n]$ of length L. If $|B| \subseteq I$ then there is a cube $(e, d_1, \ldots, d_k)$ contained in B with $k = \Omega(\log\log|B|)$ and $(\forall i)[d_i \leq L]$.*

Proof:

The following procedure produces the desired cube.

1. Let $B_1 = B$ and $\beta_1 = |B_1|$.

2. Let D_1 be all $\binom{\beta_1}{2}$ positive differences of elements of B_1. Since $B_1 \subseteq [n]$ all of the differences are in $[n]$. Hence some difference must occur $\binom{\beta_1}{2}/n \sim \beta_1^2/2n$ times. Let that difference be d_1. Note that $d_1 \leq L$.

3. Let $B_2 = \{x \in B_1 : x + d_1 \in B_1\}$. Note that $|B_2| \geq \beta_1^2/2n$. Let $|B_2| = \beta_2$. Note the trivial fact that

   $x \in B_1 \implies x + d_1 \in B$.

4. Let D_2 be all $\binom{\beta_2}{2}$ positive differences of elements of B_2. Since $B_2 \subseteq [n]$ all of the differences are in $[n]$. Hence some difference must occur $\binom{\beta_1}{2}/n \sim \beta_2^2/2n$ times. Let that difference be d_2. Note that $d_2 \leq L$.

5. Let $B_3 = \{x \in B_2 : x + d_2 \in B_2\}$. Note that $|B_3| \geq \beta_2^2/2n$. Let $|B_3| = \beta_3$. Note that

   $x \in B_3 \implies x + d_2 \in B$

   $x \in B_3 \implies x \in B_2 \implies x + d_1 \in B$

   $x \in B_3 \implies x + d_2 \in B_2 \implies x + d_1 + d_2 \in B$

6. Keep repeating this procedure until $B_{k+2} = \emptyset$. (We leave the details of the definition to the reader.) Note that if $i \leq k+1$ then

   $x \in B_i \implies x + b_1 d_1 + \cdots + b_{i-1} d_{i-1} \in B$ for any $b_1, \ldots, b_{i-1} \in \{0,1\}$.

7. Let e be any element of B_{k+1}. Note that we have $e + b_1 d_1 + \cdots + b_k d_k \in B$ for any $b_1, \ldots, b_k \in \{0,1\}$.

We leave it as an exercise to formally show that $C(e, d_1, \ldots, d_k)$ is contained in B and that $k = \Omega(\log\log|B|)$. ∎

The next lemma states that if A is 'big' and 3-free then it is somewhat uniform. There cannot be sparse intervals of A. The intuition is that if A has a sparse interval then the rest of A has to be dense to make up for it, and it might have to be so dense that it has a 3-AP.

Lemma 0.1.4 *Let $n, n_0 \in \mathbb{N}; \lambda, \lambda_0 \in (0,1)$. Assume $\lambda < \lambda_0$ and $(\forall m \geq n_0)[sz(m) \leq \lambda_0 m]$. Let $A \subseteq [n]$ be a 3-free set such that $|A| \geq \lambda n$.*

1. *Let a, b be such that $a < b$, $a > n_0$, and $n - b > n_0$. Then $\lambda_0(b-a) - n(\lambda_0 - \lambda) \leq |A \cap [a,b]|$.*
2. *Let a be such that $n - a > n_0$. Then $\lambda_0 a - n(\lambda_0 - \lambda) \leq |A \cap [1,a]|$.*

Proof:
1) Since A is 3-free and $a \geq n_0$ and $n - b \geq n_0$ we have $|A \cap [1, a-1]| < \lambda_0(a-1) < \lambda_0 a$ and $|A \cap [b+1, n]| < \lambda_0(n-b)$. Hence

$$\begin{aligned} \lambda n \leq |A| = & \ |A \cap [1,a-1]| + |A \cap [a,b]| + |A \cap [b+1,n]| \\ \lambda n \leq & \ \lambda_0 a + |A \cap [a,b]| + \lambda_0(n-b) \\ \lambda n - \lambda_0 n + \lambda_0 b - \lambda_0 a \leq & \ |A \cap [a,b]| \\ \lambda_0(b-a) - n(\lambda_0 - \lambda) \leq & \ |A \cap [a,b]|. \end{aligned}$$

2) Since A is 3-free and $n - a > n_0$ we have $|A \cap [a+1, n]| \leq \lambda_0(n-a)$. Hence

$$\begin{aligned} \lambda n \leq |A| = & \ |A \cap [1,a]| + |A \cap [a+1,n]| \\ \lambda n \leq & \ |A \cap [1,a]| + \lambda_0(n-a) \\ \lambda n - \lambda_0 n + \lambda_0 a \leq & \ |A \cap [1,a]| \\ \lambda_0 a - (\lambda_0 - \lambda)n \leq & \ |A \cap [1,a]|. \end{aligned}$$

∎

Lemma 0.1.5 *Let $n, n_0 \in \mathbb{N}$ and $\lambda, \lambda_0 \in (0,1)$. Assume that $\lambda < \lambda_0$ and that $(\forall m \geq n_0)[sz(m) \leq \lambda_0 m]$. Assume that $\frac{n}{2} \geq n_0$. Let $a, L \in \mathbb{N}$ such that $a \leq \frac{n}{2}$, $L < \frac{n}{2} - a$, and $a \geq n_0$. Let $A \subseteq [n]$ be a 3-free set such that $|A| \geq \lambda n$.*

1. *There is an interval $I \subseteq [a, \frac{n}{2}]$ of length $\leq L$ such that*

$$|A \cap I| \geq \left\lfloor \frac{2L}{n-2a}(\lambda_0(\frac{n}{2} - a) - n(\lambda_0 - \lambda)) \right\rfloor.$$

2. *Let α be such that $0 < \alpha < \frac{1}{2}$. If $a = \alpha n$ and $\sqrt{n} << \frac{n}{2} - \alpha n$ then there is an interval $I \subseteq [a, \frac{n}{2}]$ of length $\leq O(\sqrt{n})$ such that*

$$|A \cap I| \geq \left\lfloor \frac{2\sqrt{n}}{(1-2\alpha)}(\lambda_0(\frac{1}{2} - (\lambda_0 - \lambda) - \alpha)) \right\rfloor = \Omega(\sqrt{n}).$$

Proof: By Lemma 0.1.4 with $b = \frac{n}{2}$, $|A \cap [a, \frac{n}{2}]| \geq \lambda_0(\frac{n}{2} - a - n(\lambda_0 - \lambda)$. Divide $[a, \frac{n}{2}]$ into $\lceil \frac{n-2a}{2L} \rceil$ intervals of size $\leq L$. There must exist an interval I such that

$$|A \cap I| \geq \left\lfloor \frac{2L}{n-2a}(\lambda_0(\frac{n}{2} - a) - n(\lambda_0 - \lambda)) \right\rfloor.$$

If $L = \lceil \sqrt{n} \rceil$ and $a = \alpha n$ then

$$\begin{aligned} |A \cap I| &\geq \left\lfloor \tfrac{2L}{n-2a}(\lambda_0(\tfrac{n}{2} - a) - n(\lambda_0 - \lambda)) \right\rfloor \\ &\geq \left\lfloor \tfrac{2\sqrt{n}}{n(1-2\alpha)}(\lambda_0(\tfrac{n}{2} - \alpha n) - n(\lambda_0 - \lambda))) \right\rfloor \\ &\geq \left\lfloor \tfrac{2\sqrt{n}}{(1-2\alpha)}(\lambda_0(\tfrac{1}{2} - \alpha) - (\lambda_0 - \lambda)) \right\rfloor = \Omega(\sqrt{n}). \end{aligned}$$

∎

Theorem 0.1.6 *For all λ, $0 < \lambda < 1$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $sz(n) \leq \lambda n$.*

Proof:

Let $S(\lambda)$ be the statement

there exists n_0 such that, for all $n \geq n_0$, $sz(n) \leq \lambda n$.

It is a trivial exercise to show that $S(0.7)$ is true.

Let

$$C = \{\lambda \mid S(\lambda)\}.$$

C is closed upwards. Since $0.7 \in C$ we know $C \neq \emptyset$. Assume, by way of contradiction, that $C \neq (0, 1)$. Then there exists $\lambda < \lambda_0$ such that $\lambda \notin C$ and $\lambda_0 \in C$. We can take $\lambda_0 - \lambda$ to be as small as we like. Let n_0 be such that $S(\lambda_0)$ is true via n_0. Let $n \geq n_0$ and let $A \subseteq [n]$ such that $|A| \geq \lambda n$ but A is 3-free. At the end we will fix values for the parameters that (a) allow the proof to go through, and (b) imply $|A| < \lambda n$, a contradiction.

PLAN : We will obtain a $T \subseteq \overline{A}$ that will help us. We will soon see what properties T needs to help us. Consider the bit string in $\{0,1\}^n$ that represents $T \subseteq [n]$. Say its first 30 bits looks like this:

$$T(0)T(1)T(2)T(3)\cdots T(29) = 00011111110000111001011110000 0$$

The set A lives in the blocks of 0's of T (henceforth 0-blocks). We will bound $|A|$ by looking at A on the 'small' and on the 'large' 0-blocks of T. Assume there are t 1-blocks. Then there are $t+1$ 0-blocks. We call a 0-block *small* if it has $< n_0$ elements, and *big* otherwise. Assume there are t^{small} small 0-blocks and t^{big} big 0-blocks. Note that $t^{\text{small}} + t^{\text{big}} = t+1$ so $t^{\text{small}}, t^{\text{big}} \leq t+1$. Let the small 0-blocks be $B_1^{\text{small}}, \ldots, B_{t^{\text{small}}}^{\text{small}}$, let their union be B^{small}, let the big 0-blocks be $B_1^{\text{big}}, \ldots, B_{t^{\text{big}}}^{\text{big}}$, and let their union be B^{big}. It is easy to see that

$$|A \cap B^{\text{small}}| \leq t^{\text{small}} n_0 \leq (t+1)n_0.$$

Since each B_i^{big} is bigger than n_0 we must have, for all i, $|A \cap B_i^{\text{big}}| < \lambda_0 |B_i^{\text{big}}|$ (else $A \cap B_i^{\text{big}}$ has a 3-AP and hence A does). It is easy to see that

$$|A \cap B^{\text{big}}| = \sum_{i=1}^{t^{\text{big}}} |A \cap B_i^{\text{big}}| \leq \sum_{i=1}^{t^{\text{big}}} \lambda_0 |B_i^{\text{big}}| \leq \lambda_0 \sum_{i=1}^{t^{\text{big}}} |B_i^{\text{big}}| \leq \lambda_0 (n - |T|).$$

Since A can only live in the (big and small) 0-blocks of T we have

$$|A| = |A \cap B^{\text{small}}| + |A \cap B^{\text{big}}| \leq (t+1)n_0 + \lambda_0(n - |T|).$$

In order to use this inequality to bound $|A|$ we will need T to be big and t to be small, so we want T to be a big set that has few blocks.

If only it was that simple. Actually we can now reveal the

REAL PLAN: The real plan is similar to the easy version given above. We obtain a set $T \subseteq \overline{A}$ and *a parameter* d. A *1-block* is a maximal AP with difference d that is contained in T (that is, if $FIRST$ and $LAST$ are the first and last elements of the 1-block then $FIRST - d \notin T$ and $LAST + d \notin T$). A *0-block* is a maximal AP with difference d that is contained in $\overline{T}$. Partition T into 1-blocks. Assume there are t of them.

Let $[n]$ be partitioned into $N^0 \cup \cdots \cup N^{d-1}$ where $N_j = \{x \mid x \leq n \wedge x \equiv j \pmod{d}\}$.

Fix j, $0 \leq j \leq d-1$. Consider the bit string in $\{0,1\}^{\lfloor n/d \rfloor}$ that represents $T \cap N_j$ Say the first 30 bits of $T \cap N_j$ look like

$$T(j)T(d+j)T(2d+j)T(3d+j)\cdots T(29d+j) = 000111111100001110010111111100$$

During PLAN we had an intuitive notion of what a 0-block or 1-block was. Note that if we restrict to N_j then that intuitive notion is still valid. For example the first block of 1's in the above example represents $T(3d+j)$, $T(4d+j)$, $T(5d+j)$, ..., $T(9d+j)$ which is a 1-block as defined formally.

Each 1-block is contained in a particular N_j. Let t_j be the number of 1-blocks that are contained in N_j. Note that $\sum_{j=0}^{d-1} t_j = t$. The number of 0-blocks that are in N_j is at most $t_j + 1$.

Let j be such that $0 \leq j \leq d-1$. By reasoning similar to that in the above PLAN we obtain

$$|A \cap N_j| \leq (t_j + 1)n_0 + \lambda_0(N_j - |T|).$$

We sum both sides over all $j = 0$ to $d-1$ to obtain

$$|A| \leq (t+d)n_0 + \lambda_0(n - |T|)$$

In order to use this inequality to bound $|A|$ we need T to be big and t, d to be small. Hence we want a big set T which when looked at mod d, for some small d, decomposes into a small number of blocks.

What is a 1-block within N_j? For example, lets look at $d = 3$ and the bits sequence for T is

$$\begin{array}{cccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17; \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0. \end{array}$$

Note that T looked at on $N_2 \cup T$ has bit sequence

$$\begin{array}{cccccc} 2 & 5 & 8 & 11 & 14 & 17; \\ 0 & 1 & 1 & 1 & 1 & 0. \end{array}$$

The numbers $5, 8, 11, 14$ are all in T and form a 1-block in the N_2 part. Note that they also from an arithmetic progression with spacing $d = 3$. Also note that this is a maximal arithmetic progression with spacing $d = 3$ since $0 \notin T$ and $17 \notin T$. More generally *1-blocks of T within N_j are maximal arithmetic progressions with spacing d.* With that in mind we can restate the kind of set T that we want.

We want a set $T \subseteq \overline{A}$ and a parameter d such that

1. T is big (so that $\lambda_0(n - |T|)$ is small),

2. d is small (see next item), and

3. the number of maximal arithmetic progressions of length d within T, which is the parameter t above, is small (so that $(t + d)n_0$ is small).

How do we obtain a big subset of $\overline{A}$? We will obtain many pairs $x, y \in A$ such that $2y - x \leq n$. Note that since $x, y, 2y - x$ is a 3-AP and $x, y \in A$ we must have $2y - x \in \overline{A}$.

Let α, $0 < \alpha < \frac{1}{2}$, be a parameter to be determined later. (For those keeping track, the parameters to be determined later are now λ_0, λ, n, and α. The parameter n_0 depends on λ_0 so is not included in this list.)

We want to apply Lemma 0.1.5.2.b to $n, n_0, a = \alpha n$. Hence we need the following conditions.

$$\begin{aligned} \alpha n &\geq n_0 \\ \tfrac{n}{2} &\geq n_0 \\ \tfrac{n}{2} - \alpha n &\geq \sqrt{n} \end{aligned}$$

Assuming these conditions hold, we proceed. By Lemma 0.1.5.b there is an interval $I \subseteq [\alpha n, \frac{n}{2}]$ of length $O(\sqrt{n})$ such that

$$|A \cap I| \geq \left\lfloor \frac{2\sqrt{n}}{(1-2\alpha)}(\lambda_0(\frac{1}{2} - \alpha) - (\lambda_0 - \lambda)) \right\rfloor = \Omega(\sqrt{n}).$$

By Lemma 0.1.3 there is a cube $C(e, d_1, \ldots, d_k)$ contained in $|A \cap I|$ with $k = \Omega(\log\log|A \cap I|) = \Omega(\log\log\sqrt{n}) = \Omega(\log\log n)$ and $d \geq \sqrt{n}$.

For i such that $1 \leq i \leq k$ we define the following.

1. Define $C_0 = \{e\}$ and, for $1 \leq i \leq k$, define $C_i = C(e, d_1, \ldots, d_i)$.

2. T_i is the third terms of AP's with the first term in $A \cap [1, e-1]$ and the second term in C_i. Formally $T_i = \{2m - x \mid x \in A \cap [1, e-1] \wedge m \in C_i\}$.

Note that, for all i, $T_i \cap A = \emptyset$. Hence we look for a large T_i that can be decomposed into a small number of blocks. We will end up using $d = 2d_{i+1}$.

Note that $T_0 \subseteq T_1 \subseteq T_2 \subseteq \cdots \subseteq T_k$. Hence to obtain a large T_i it suffices to show that T_0 is large and then any of the T_i will be large (though not necessarily consist of a small number of blocks).

Since $C_0 = \{e\}$ we have

$T_0 = \{2m - x \mid x \in A \cap [1, e-1] \wedge m \in C_0\} = \{2e - x \mid x \in A \cap [1, e-1]\}$.

Clearly there is a bijection from $A \cap [1, e-1]$ to T_0, hence $|T_0| = |A \cap [1, e-1]|$. Since $e \in [\alpha n, \frac{n}{2}]$ we have $|A \cap [1, e]| \geq |A \cap [1, \alpha n]|$.

We want to use Lemma 0.1.4.2 on $A \cap [1, \alpha n]$. Hence we need the condition

$$n - \alpha n \geq n_0.$$

By Lemma 0.1.4

$$|T_0| \geq |A \cap [1, \alpha n]| \geq \lambda_0 \alpha n - n(\lambda_0 - \lambda) = n(\lambda_0 \alpha - (\lambda_0 - \lambda)).$$

In order for this to be useful we need the following condition

$$\begin{aligned} \lambda - \lambda_0 + \lambda_0 \alpha &> 0 \\ \lambda_0 \alpha &> \lambda_0 - \lambda \end{aligned}$$

We now show that some T_i has a small number of blocks. Since $|T_k| \leq n$ (a rather generous estimate) there must exist an i such that $|T_{i+1} - T_i| \leq \frac{n}{k}$. Let $t = \frac{n}{k}$ (t will end up bounding the number of 1-blocks).

Partition T_i into maximal AP's with difference $2d_{i+1}$. We call these maximal AP's 1-blocks. We will show that there are $\leq t$ 1-blocks by showing a bijection between the blocks and $T_{i+1} - T_i$.

If $z \in T_i$ then $z = 2m - x$ where $x \in A \cap [1, \alpha n - 1]$ and $m \in C_i$. By the definitions of C_i and C_{i+1} we know $m + d_{i+1} \in C_{i+1}$. Hence $2(m + d_{i+1}) - x \in T_{i+1}$. Note that $2(m + d_{i+1}) - x = z + 2d_{i+1}$. In short we have

$$z \in T_i \implies z + 2d_{i+1} \in T_{i+1}.$$

NEED PICTURE

We can now state the bijection. Let $z_1, \ldots, z_m$ be a block in T_i. We know that $z_m + 2d_{i+1} \notin T_i$ since if it was the block would have been extended to include it. However, since $z_m \in T_i$ we know $z_m + 2d_{i+1} \in T_{i+1}$. Hence $z_m + 2d_{i+1} \in T_{i+1} - T_i$. This is the bijection: map a block to what would be the next element if it was extended. This is clearly a bijection. Hence the number of 1-blocks is at most $t = |T_{i+1} - T_i| \leq n/k$.

To recap, we have

$$|A| \leq (t + d)n_0 + \lambda_0(n - |T|)$$

with $t \leq \frac{n}{k} = O(\frac{n}{\log \log n})$, $d = O(\sqrt{n})$, and $|T| \geq n(\lambda_0 \alpha - (\lambda_0 - \lambda))$. Hence we have

$$|A| \leq O((\frac{n}{\log\log n} + \sqrt{n})n_0) + n\lambda_0(1 - \lambda + \lambda_0 - \lambda_0\alpha).$$

We want this to be $< \lambda n$. The term $O((\frac{n}{\log\log n} + \sqrt{n})n_0)$ can be ignored since for n large enough this is less than any fraction of n. For the second term we need

$$\lambda_0(1 - \lambda + \lambda_0 - \lambda_0\alpha) < \lambda$$

We now gather together all of the conditions and see how to satisfy them all at the same time.

$$\begin{aligned}
\alpha n &\geq n_0 \\
\tfrac{n}{2} &\geq n_0 \\
\tfrac{n}{2} - \alpha n &\geq \sqrt{n} \\
n - \alpha n &\geq n_0 \\
\lambda_0\alpha &> \lambda_0 - \lambda \\
\lambda_0(1 - \lambda + \lambda_0 - \lambda_0\alpha) &< \lambda
\end{aligned}$$

We first choose λ and λ_0 such that $\lambda_0 - \lambda < 10^{-1}\lambda_0^2$. This is possible by first picking an initial (λ', λ_0') pair and then picking (λ, λ_0) such that $\lambda' < \lambda < \lambda_0 < \lambda_0'$ and $\lambda_0 - \lambda < 10^{-1}(\lambda')^2 < 10^{-1}\lambda_0^2$. The choice of λ_0 determines n_0. We then chose $\alpha = 10^{-1}$. The last two conditions are satisfied:

$\lambda_0\alpha > \lambda_0 - \lambda$ becomes

$$\begin{aligned}
10^{-1}\lambda_0 &> 10^{-1}\lambda_0^2 \\
1 &> \lambda_0
\end{aligned}$$

which is clearly true.

$\lambda_0(1 - \lambda + \lambda_0 - \lambda_0\alpha) < \lambda$ becomes

$$\begin{aligned}
\lambda_0(1 - 10^{-1}\lambda_0^2 - 10^{-1}\lambda_0) &< \lambda \\
\lambda_0 - 10^{-1}\lambda_0^3 - 10^{-1}\lambda_0^2 &< \lambda \\
\lambda_0 - \lambda - 10^{-1}\lambda_0^3 - 10^{-1}\lambda_0^2 &< 0 \\
10^{-1}\lambda_0^2 - 10^{-1}\lambda_0^3 - 10^{-1}\lambda_0^2 &< 0 \\
-10^{-1}\lambda_0^3 &< 0
\end{aligned}$$

which is clearly true.

Once λ, λ_0, n_0 are picked, you can easily pick n large enough to make the other inequalities hold. ∎

0.1.2 Analytic Proof

Consider the following statement:

If $A \subseteq [n]$ and $\#(A)$ is 'big' then A must have a 3-AP.

This statement, made rigorous, is true. In particular, the following is true and easy:

Let $n \geq 3$. If $A \subseteq [n]$ and $\#(A) \geq 0.7n$ then A must have a 3-AP.

Can we lower the constant 0.7? We can lower it as far as we like if we allow n to start later:

Roth [**?**, **?**, **?**] proved the following using analytic means.

$(\forall \lambda > 0)(\exists n_0 \in \mathbb{N})(\forall n \geq n_0)(\forall A \subseteq [n])[\#(A) \geq \lambda n \implies A \text{ has a 3-AP}]$.

The analogous theorem for 4-APs was later proven by Szemeredi [**?**, **?**] by a combinatorial proof. Szemeredi [**?**] later (with a much harder proof) generalized from 4 to any k.

We prove the $k = 3$ case using the analytic techniques of Roth; however, we rely heavily on Gowers [**?**, **?**]

Def 0.1.7 Let $sz(n)$ be the least number such that, for all $A \subseteq [n]$, if $\#(A) \geq sz(n)$ then A has a 3-AP. Note that if $A \subseteq [a, a+n-1]$ and $\#(A) \geq sz(n)$ then A has a 3-AP. Note also that if $A \subseteq \{a, 2a, 3a, \ldots, na\}$ and $\#(A) \geq sz(n)$ then A has a 3-AP. More generally, if A is a subset of any equally spaced set of size n, and $\#(A) \geq sz(n)$, then A has a 3-AP.

Throughout this section the following hold.

1. $n \in \mathbb{N}$ is a fixed large prime.
2. $\mathbb{Z}_n = \{1, \ldots, n\}$ with modular arithmetic.
3. $\omega = e^{2\pi i/n}$.
4. If a is a complex number then $|a|$ is its length.
5. If A is a set then $|A|$ is its cardinality.

Counting 3-AP's

Lemma 0.1.8 *Let $A, B, C \subseteq [n]$. The number of $(x, y, z) \in A \times B \times C$ such that $x + z \equiv 2y \pmod{n}$ is*

$$\frac{1}{n} \sum_{x,y,z \in [n]} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)}.$$

Proof:

We break the sum into two parts:
Part 1:

$$\frac{1}{n} \sum_{x,y,z \in [n], x+z \equiv 2y \pmod{n}} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)}.$$

Note that we can replace $\omega^{-r(x-2y+z)}$ with $\omega^0 = 1$. We can then replace $\sum_{r=1}^{n} 1$ with n. Hence we have

$$\frac{1}{n} \sum_{x,y,z \in [n], x+z \equiv 2y \pmod{n}} A(x)B(y)C(z)n = \sum_{x,y,z \in [n], x+z \equiv 2y \pmod{n}} A(x)B(y)C(z)$$

This is the number of $(x, y, z) \in A \times B \times C$ such that $x + z \equiv 2y \pmod{n}$.
Part 2:

$$\frac{1}{n} \sum_{x,y,z \in [n], x+z \not\equiv 2y \pmod{n}} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)}.$$

We break this sum up depending on what the (nonzero) value of $w = x + z - 2y \pmod{n}$. Let

$$S_u = \sum_{x,y,z \in [n], x-2y+z=2} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-ru}.$$

Since $u \neq 0$, $\sum_{r=1}^{n} \omega^{-ru} = \sum_{r=1}^{n} \omega^{-r} = 0$. Hence $S_u = 0$.
Note that

$$\frac{1}{n} \sum_{x,y,z \in [n], x+z \not\equiv 2y \pmod{n}} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)} = \frac{1}{n} \sum_{u=1}^{n-1} S_u = 0$$

The lemma follows from Part 1 and Part 2. ∎

Lemma 0.1.9 *Let $A \subseteq [n]$. Let $B = C = A \cap [n/3, 2n/3]$. The number of $(x, y, z) \in A \times B \times C$ such that x, y, z forms a 3-AP is at least*

$$\frac{1}{2n} \sum_{x,y,z \in [n]} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)} - O(n).$$

Proof: By Lemma 0.1.8

$$\frac{1}{n} \sum_{x,y,z \in [n]} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)}$$

is the number of $(x, y, z) \in A \times B \times C$ such that $x + z \equiv 2y \pmod{n}$. This counts three types of triples:

- Those that have $x = y = z$. There are $n/3$ of them.
- Those that have $x + z = 2y + n$. There are $O(1)$ of them.
- Those that have $x \neq y$, $y \neq z$, $x \neq z$, and $x + z = 2y$.

Hence

$$\#(\{(x, y, z) : (x+z = 2y) \wedge x \neq y \wedge y \neq z \wedge x \neq z\}) = \frac{1}{n} \sum_{x,y,z \in [n]} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)} - O(n).$$

We are not done yet. Note that $(5, 10, 15)$ may show up as $(15, 10, 5)$. Every triple appears at most twice. Hence

$\#(\{(x, y, z) : (x + z = 2y) \wedge x \neq y \wedge y \neq z \wedge x \neq z\})$

$\leq \quad 2\#(\{(x, y, z) : (x < y < z) \wedge (x+z = 2y) \wedge x \neq y \wedge y \neq z \wedge x \neq z\})$.

Therefore

$$\frac{1}{2n} \sum_{x,y,z \in [n]} A(x)B(y)C(z) \sum_{r=1}^{n} \omega^{-r(x-2y+z)} - O(n) \leq \text{ the number of 3-AP's with } x \in A,\ y \in B,\ z \in C\ .$$

∎

We will need to re-express this sum. For that we will use Fourier Analysis.

Fourier Analysis

Def 0.1.10 If f:$\mathbb{Z}_n \to \mathbb{N}$ then $\hat{f}$:$\mathbb{Z}_n \to \mathbb{C}$ is

$$\hat{f}(r) = \sum_{s\in[n]} f(s)\omega^{-rs}.$$

$\hat{f}$ is called the *Fourier Transform* of f.

What does $\hat{f}$ tell us? We look at the case where f is the characteristic function of a set $A \subseteq [n]$. Henceforth we will use $A(x)$ instead of $f(x)$.

We will need the following facts.

Lemma 0.1.11 *Let $A \subseteq \{1,\ldots,n\}$.*

1. $\hat{A}(n) = \#(A)$.
2. $\max_{r\in[n]} |\hat{A}(r)| = \#(A)$.
3. $A(s) = \frac{1}{n}\sum_{r=1}^{n} \hat{A}(r)\omega^{-rs}$. *DO WE NEED THIS?*
4. $\sum_{r=1}^{n} |\hat{A}(r)|^2 = n\#(A)$.
5. $\sum_{s=1}^{n} A(s) = \frac{1}{n}\sum_{r=1}^{n} \hat{A}(r)$.

Proof:

Note that $\omega^n = 1$. Hence

$$\hat{A}(n) = \sum_{s\in[n]} A(s)\omega^{-ns} = \sum_{s\in[n]} A(s) = \#(A).$$

Also note that

$$|\hat{A}(r)| = |\sum_{s\in[n]} A(s)\omega^{-rs}| \leq \sum_{s\in[n]} |A(s)\omega^{-rs}| \leq \sum_{s\in[n]} |A(s)||\omega^{-rs}| \leq \sum_{s\in[n]} |A(s)| = \#(A).$$

∎

Informal Claim: If $\hat{A}(r)$ is large then there is an arithmetic progression P with difference $r^{-1} \pmod n$ such that $\#(A \cap P)$ is large.

We need a lemma before we can proof the claim.

Lemma 0.1.12 *Let $n, m \in \mathbb{N}$, $s_1, \ldots, s_m$, and $0 < \lambda, \alpha, \epsilon < 1$ be given (no order on $\lambda, \alpha, \epsilon$ is implied). Assume that $(\lambda - \frac{m-1}{m}(\lambda + \epsilon)) \geq 0$. Let $f(x_1, \ldots, x_m) = |\sum_{j=1}^m x_j \omega^{s_j}|$. The maximum value that $f(x_1, \ldots, x_m)$ can achieve subject to the following two constraints (1) $\sum_{j=1}^m x_j \geq \lambda n$, and (2) $(\forall j)[0 \leq x_i \leq (\lambda + \epsilon)\frac{n}{m}]$ is bounded above by $\epsilon mn + (\lambda + \epsilon)\frac{n}{m}|\sum_{j=1}^m \omega^{s_j}|$*

Proof:

Assume that the maximum value of f, subject to the constraints, is achieved at $(x_1, \ldots, x_m)$. Let MIN be the minimum value that any variable x_i takes on (there may be several variables that take this value). What is the smallest that MIN could be? By the constraints this would occur when all but one of the variables is $(\lambda + \epsilon)\frac{n}{m}$ and the remaining variable has value MIN. Since $\sum_{x_i} \geq \lambda n$ we have

$MIN + (m-1)(\lambda + \epsilon)\frac{n}{m} \geq \lambda n$
$MIN + \frac{m-1}{m}(\lambda + \epsilon)n \geq \lambda n$
$MIN \geq \lambda n - \frac{m-1}{m}(\lambda + \epsilon)n$
$MIN \geq (\lambda - \frac{m-1}{m}(\lambda + \epsilon))n$
Hence note that, for all j,
$x_j - MIN \leq x_j - (\lambda - \frac{m-1}{m}(\lambda + \epsilon))n$
Using the bound on x_j from constraint (2) we obtain

$$\begin{aligned} x_j - MIN &\leq (\lambda + \epsilon)\tfrac{n}{m} - (\lambda - \tfrac{m-1}{m}(\lambda + \epsilon))n \\ &\leq ((\lambda + \epsilon)\tfrac{1}{m} - (\lambda - \tfrac{m-1}{m}(\lambda + \epsilon)))n \\ &\leq ((\lambda + \epsilon)\tfrac{1}{m} - \lambda + \tfrac{m-1}{m}(\lambda + \epsilon))n \\ &\leq \epsilon n \end{aligned}$$

Note that

$$\begin{aligned} |\sum_{j=1}^m x_j \omega^{s_j}| = & \ |\sum_{j=1}^m (x_j - MIN)\omega^{s_j} + \sum_{j=1}^m MIN\omega^{s_j}| \\ \leq & \ |\sum_{j=1}^m (x_j - MIN)\omega^{s_j}| + |\sum_{j=1}^m MIN\omega^{s_j}| \\ \leq & \ \sum_{j=1}^m |(x_j - MIN)||\omega^{s_j}| + MIN|\sum_{j=1}^m \omega^{s_j}| \\ \leq & \ \sum_{j=1}^m \epsilon n + MIN|\sum_{j=1}^m \omega^{s_j}| \\ \leq & \ \epsilon mn + MIN|\sum_{j=1}^m \omega^{s_j}| \\ \leq & \ \epsilon mn + (\lambda + \epsilon)\tfrac{n}{m}|\sum_{j=1}^m \omega^{s_j}| \end{aligned}$$

∎

Lemma 0.1.13 *Let $A \subseteq [n]$, $r \in [n]$, and $0 < \alpha < 1$. If $|\hat{A}(r)| \geq \alpha n$ and $|A| \geq \lambda n$ then there exists $m \in \mathbb{N}$, $0 < \epsilon < 1$, and an arithmetic progression P within $\mathbb{Z}_n$, of length $\frac{n}{m} \pm O(1)$ such that $\#(A \cap P) \geq (\lambda + \epsilon)\frac{n}{m}$. The parameters ϵ and m will depend on λ and α but not n.*

Proof: Let m and ϵ be parameters to be picked later. We will note constraints on them as we go along. (Note that ϵ will not be used for a while.)

Let $1 = a_1 < a_2 < \cdots < a_{m+1} = n$ be picked so that

$a_2 - a_1 = a_3 - a_2 = \cdots = a_m - a_{m-1}$ and $a_{m+1} - a_m$ is as close to $a_2 - a_1$ as possible.

For $1 \leq j \leq m$ let

$$P_j = \{s \in [n] : a_j \leq rs \pmod n < a_{j+1}\}.$$

Let us look at the elements of P_j. Let r^{-1} be the inverse of r mod n.

1. s such that $a_j \equiv rs \pmod n$, that is, $s \equiv a_j r^{-1} \pmod n$.
2. s such that $a_j + 1 \equiv rs \pmod n$, that is $s \equiv (a_j + 1)r^{-1} \equiv a_j r^{-1} + r^{-1} \pmod n$.
3. s such that $a_j + 2 \equiv rs \pmod n$, that is $s \equiv (a_j + 2)r^{-1} \equiv a_j r^{-1} + 2r^{-1} \pmod n$.
4. $\vdots$.

Hence P_j is an arithmetic progression within $\mathbb{Z}_n$ which has difference r^{-1}. Also note that $P_1, \ldots, P_m$ form a partition of $\mathbb{Z}_n$ into m parts of size $\frac{n}{m} + O(1)$ each.

Recall that

$$\hat{A}(r) = \sum_{s \in [n]} A(s)\omega^{-rs}.$$

Lets look at $s \in P_j$. We have that $a_j \leq rs \pmod n < a_{j+1}$. Therefore the values of $\{\omega^{rs} : s \in P_j\}$ are all very close together. We will pick $s_j \in P_j$ carefully. In particular we will constrain m so that it is possible to pick $s_j \in P_j$ such that $\sum_{j=1}^{m} \omega^{-rs_j} = 0$. For $s \in P_j$ we will approximate ω^{-rs} by ω^{-rs_j}. We skip the details of how good the approximation is.

We break up the sum over s via P_j.

$$\begin{aligned}
\hat{A}(r) &= \sum_{s\in[n]} A(s)\omega^{-rs} \\
&= \sum_{j=1}^{m}\sum_{s\in P_j} A(s)\omega^{-rs} \\
&\sim \sum_{j=1}^{m}\sum_{s\in P_j} A(s)\omega^{-rs_j} \\
&= \sum_{j=1}^{m}\omega^{-rs_j}\sum_{s\in P_j} A(s) \\
&= \sum_{j=1}^{m}\omega^{-rs_j}\#(A\cap P_j) \\
&= \sum_{j=1}^{m}\#(A\cap P_j)\omega^{-rs_j} \\
\alpha n \le |\hat{A}(r)| &= |\sum_{j=1}^{m}\#(A\cap P_j)\omega^{-rs_j}|
\end{aligned}$$

We will not use ϵ. We intend to use Lemma 0.1.12; therefore we have the constraint $(\lambda - \frac{m-1}{m}(\lambda+\epsilon)) \ge 0$.

Assume, by way of contradiction, that $(\forall j)[|A\cap P_j| \le (\lambda+\epsilon)\frac{n}{m}$. Applying Lemma 0.1.12 we obtain

$$|\sum_{j=1}^{m}\#(A\cap P_j)\omega^{-rs_j}| \le \epsilon mn + (\lambda+\epsilon)\frac{n}{m}|\sum_{j=1}^{m}\omega^{-rs_j}| = \epsilon mn.$$

Hence we have
$\alpha n \le \epsilon mn$
$\alpha \le \epsilon m$.
In order to get a contradiction we pick ϵ and m such that $\alpha > \epsilon m$.
Having done that we now have that $(\exists j)[|A\cap P_j| \ge (\lambda+\epsilon)\frac{n}{m}]$.
We now list all of the constraints introduced and say how to satisfy them.

1. m is such that there exists $s_1 \in P_1, \ldots, s_m \in P_m$ such that $\sum_{j=1}^{m}\omega^{-rs_j} = 0$, and
2. $(\lambda - \frac{m-1}{m}(\lambda+\epsilon)) \ge 0$.
3. $\epsilon m < \alpha$.

First pick m to satisfy item 1. Then pick ϵ small enough to satisfy items 2,3. ∎

Lemma 0.1.14 *Let $A, B, C \subseteq [n]$. The number of 3-AP's $(x, y, z) \in A \times B \times C$ is bounded below by*

$$\frac{1}{2n}\sum_{r=1}^{n}\hat{A}(r)\hat{B}(-2r)\hat{C}(r) - O(n).$$

Proof:

The number of 3-AP's is bounded below by

$$\frac{1}{2n}\sum_{x,y,z\in[n]} A(x)B(y)C(z)\sum_{r=1}^{n}\omega^{-r(x-2y+z)} - O(n) =$$

We look at the inner sum.

$$\sum_{x,y,z\in[n]} A(x)B(y)C(z)\sum_{r=1}^{n}\omega^{-r(x-2y+z)} =$$

$$\sum_{r=1}^{n}\sum_{x,y,z\in[n]} A(x)\omega^{-rx}B(y)\omega^{2yr}C(z)\omega^{-rz} =$$

$$\sum_{r=1}^{n}\sum_{x\in[n]} A(x)\omega^{-rx}\sum_{y\in[n]} B(y)\omega^{2yr}\sum_{z\in\mathbb{Z}_r} C(z)\omega^{-rz} =$$

$$\sum_{r=1}^{n}\hat{A}(r)\hat{B}(-2r)\hat{C}(r).$$

The Lemma follows. ∎

Main Theorem

Theorem 0.1.15 *For all λ, $0 < \lambda < 1$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $sz(n) \leq \lambda n$.*

Proof:

Let $S(\lambda)$ be the statement

there exists n_0 such that, for all $n \geq n_0$, $sz(n) \leq \lambda n$.

It is a trivial exercise to show that $S(0.7)$ is true.

Let

$$C = \{\lambda : S(\lambda)\}.$$

C is closed upwards. Since $0.7 \in C$ we know $C \neq \emptyset$. Assume, by way of contradiction, that $C \neq (0,1)$. Then there exists $\lambda < \lambda_0$ such that $\lambda \notin C$ and $\lambda_0 \in C$. We can take $\lambda_0 - \lambda$ to be as small as we like. Let n_0 be such

that $S(\lambda_0)$ is true via n_0. Let $n \geq n_0$ and let $A \subseteq [n]$ such that $\#(A) \geq \lambda n$ but A is 3-free.

Let $B = C = A \cap [n/3, 2n/3]$.

By Lemma 0.1.14 the number of 3-AP's of A is bounded below by

$$\frac{1}{2n} \sum_{r=1}^{n} \hat{A}(r)\hat{B}(-2r)\hat{C}(r) - O(n).$$

We will show that either this is positive or there exists a set $P \subseteq [n]$ that is an AP of length XXX and has density larger than λ. Hence P will have a 3-AP.

By Lemma 0.1.11 we have $\hat{A}(n) = \#(A)$, $\hat{B}(n) = \#(B)$, and $\hat{C}(n) = \#(C)$. Hence

$$\frac{1}{2n}\hat{A}(n)\hat{B}(n)\hat{C}(n) + \frac{1}{2n} \sum_{r=1}^{n-1} \hat{A}(r)\hat{B}(-2r)\hat{C}(r) - O(n) =$$

$$\frac{1}{2n}\#(A)\#(B)\#(C) + \frac{1}{2n} \sum_{r=1}^{n-1} \hat{A}(r)\hat{B}(-2r)\hat{C}(r) - O(n).$$

By Lemma 0.1.5 we can take $\#(B), \#(C) \geq n\lambda/4$. We already have $\#(A) \geq \lambda n$. This makes the lead term $\Omega(n^3)$; hence we can omit the $O(n)$ term. More precisely we have that the number of 3-AP's in A is bounded below by

$$\frac{\lambda^3 n^2}{32} + \frac{1}{2n} \sum_{r=1}^{n-1} \hat{A}(r)\hat{B}(-2r)\hat{C}(r)).$$

We are assuming that this quantity is ≤ 0.

$$\frac{\lambda^3 n^2}{32} + \frac{1}{2n} \sum_{r=1}^{n-1} \hat{A}(r)\hat{B}(-2r)\hat{C}(r)) < 0.$$

$$\frac{\lambda^3 n^2}{16} + \frac{1}{n} \sum_{r=1}^{n-1} \hat{A}(r)\hat{B}(-2r)\hat{C}(r)) < 0.$$

$$\frac{\lambda^3 n^2}{16} < -\frac{1}{n} \sum_{r=1}^{n-1} \hat{A}(r)\hat{B}(-2r)\hat{C}(r)).$$

Since the left hand side is positive we have

$$\begin{aligned}\frac{\lambda^3 n^2}{16} &< |\frac{1}{n}\sum_{r=1}^{n-1}\hat{A}(r)\hat{B}(-2r)\hat{C}(r)| \\ &< \frac{1}{n}(\max r\hat{A}(r))\sum_{r=1}^{n-1}|\hat{B}(-2r)||\hat{C}(r)|\end{aligned}$$

By the Cauchy Schwartz inequality we know that

$$\sum_{i=1}^{n-1}|\hat{B}(-2r)||\hat{C}(r)| \leq (\sum_{i=1}^{n-1}|\hat{B}(-2r)|^2)^{1/2})(\sum_{i=1}^{n-1}|\hat{C}(r)|^2)^{1/2}).$$

Hence

$$\frac{\lambda^3 n^2}{16} < |\frac{1}{n}\max_{1\leq r\leq n-1}|\hat{A}(r)|(\sum_{i=1}^{n-1}|\hat{B}(-2r)|^2)^{1/2})(\sum_{i=1}^{n-1}|\hat{C}(r)|^2)^{1/2}).$$

By Parsaval's inequality and the definition of B and C we have

$$\sum_{i=1}^{n-1}|\hat{B}(-2r)|^2)^{1/2} \leq n\#(B) = \frac{\lambda n^2}{3}$$

and

$$\sum_{i=1}^{n-1}|\hat{C}(r)|^2)^{1/2} \leq n\#(C) = \frac{\lambda n^2}{3}$$

Hence

$$\frac{\lambda^3 n^2}{16} < (\max_{1\leq r\leq n-1}|\hat{A}(r)|)\frac{1}{n}\frac{\lambda n^2}{3} = (\max_{1\leq r\leq n-1}|\hat{A}(r)|)\frac{\lambda n}{3}.$$

Therefore
$|\hat{A}(r) \geq \frac{3\lambda^2 n}{16}$. ∎

0.1.3 What more is known?

The following is known.

Theorem 0.1.16 *For every $\lambda > 0$ there exists n_0 such that for all $n \geq n_0$, $sz(n) \leq \lambda n$.*

This has been improved by Heath-Brown [**?**] and Szemeredi [**?**]

Theorem 0.1.17 *There exists c such that* $sz(n) = \Omega(n\frac{1}{(\log n)^c})$. *(Szemeredi estimates* $c \leq 1/20$*).*

Bourgain [**?**] improved this further to obtain the following.

Theorem 0.1.18 $sz(n) = \Omega(n\sqrt{\frac{\log\log n}{\log n}})$.