Roth’s Theorem: If \( A \subseteq [n] \) is large then it has a 3-AP

Szemeredi’s Proof

by William Gasarch (gasarch@cs.umd.edu)

1 Roth’s Theorem

**Notation 1.1** Let \( [n] = \{1, \ldots, n\} \). If \( k \in \mathbb{N} \) then \( k \)-AP means an arithmetic progression of size \( k \).

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 \([2, 4, 5]\) proved the following using analytic means.

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

The analogous theorem for 4-APs was later proven by Szemeredi \([2, 6]\) by a combinatorial proof. Szemeredi \([7]\) 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, Rothchild, and Spencer \([2]\).

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

**Definition 1.2** 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.

**Definition 1.3** 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 \( \{e + b_1d_1 + \cdots + b_kd_k \mid b_1, \ldots, b_k \in \{0, 1\}\} \). A \( k \)-cube is a cube with \( k \) \( d \)'s.

**Lemma 1.4** 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 
\( x \in B_1 \Rightarrow x + d_1 \in B \).

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 \( \binom{\beta_2}{2}/n \sim \beta_2^2/2n \) 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 
\( x \in B_3 \Rightarrow x + d_2 \in B \)
\( x \in B_3 \Rightarrow x \in B_2 \Rightarrow x + d_1 \in B \)
\( x \in B_3 \Rightarrow x + d_2 \in B_2 \Rightarrow x + d_1 + d_2 \in B \)

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 
\( x \in B_i \Rightarrow x + b_1d_1 + \cdots + b_{i-1}d_{i-1} \in B \) 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 \( e + b_1d_1 + \cdots + b_dd_k \in B \) 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 \( k = \Omega(\log \log |B|) \).

We now note that the above gives a good upper bound on the Hilbert Cube Numbers.

**Corollary 1.5** For \( k, c \) let \( H(k, c) \) be the least \( H \) such that for any \( c \)-coloring of \( \{1, \ldots, H\} \) there is a monochromatic \( k \)-cube. Then \( H(k, c) \leq c2^{O(k)} \).

**Proof:** Let \( H = c2^{2^k} \) where we define \( A \) later. Let \( COL \) be a \( c \)-coloring of \( \{1, \ldots, H\} \). Some color appears \( H/c = 2^{2^k} \) times. Let \( B \) be the set of integers with that color, so \( |B| = 2^{2^k} \). By Lemma 1.4 there is a monochromatic cube of size \( \Omega(\log_2(\log_2(|B|))) = \Omega(2^k) \). Pick \( A \) big enough so that this term is \( \geq k \).

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 1.6** 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 \( \lambda_0 a - n(\lambda_0 - \lambda) \leq |A \cap [a, 1]| \).
Proof:
1) Since $A$ is 3-free and $a \geq n_0$ and $n - b \geq n_0$ we have $|A \cap [1, a - 1]| < \lambda_0(a - 1) < \lambda_0a$ and $|A \cap [b + 1, n]| < \lambda_0(n - b)$. Hence
\[
\lambda n \leq |A| = |A \cap [1, a - 1]| + |A \cap [a, b]| + |A \cap [b + 1, n]|,
\]
\[
\lambda n \leq \lambda_0a + |A \cap [a, b]| + \lambda_0(n - b),
\]
\[
\lambda n - \lambda_0n + \lambda_0b - \lambda_0a \leq |A \cap [a, b]|,
\]
\[
\lambda_0(b - a) - n(\lambda_0 - \lambda) \leq |A \cap [a, b]|.
\]

2) Since $A$ is 3-free and $n - a > n_0$ we have $|A \cap [a + 1, n]| \leq \lambda_0(n - a)$. Hence
\[
\lambda n \leq |A| = |A \cap [1, a + 1]| + |A \cap [a + 1, n]|,
\]
\[
\lambda n \leq |A \cap [1, a + 1]| + \lambda_0(n - a),
\]
\[
\lambda n - \lambda_0n + \lambda_0a \leq |A \cap [1, a]|,
\]
\[
\lambda_0a - (\lambda_0 - \lambda)n \leq |A \cap [1, a]|.
\]

Lemma 1.7 Let $n, n_0 \in \mathbb{N}$ and $\lambda, \lambda_0 \in (0, 1)$. Assume that $\lambda < \lambda_0$ and that $(\forall m \geq n_0) [sz(m) \leq \lambda_0m]$. 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
\[
|A \cap I| \geq \frac{2L}{n - 2a} (\lambda_0(\frac{n}{2} - a) - n(\lambda_0 - \lambda)) \, .
\]

2. Let $\alpha$ 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
\[
|A \cap I| \geq \frac{2\sqrt{n}}{(1 - 2\alpha)} (\lambda_0(\frac{1}{2} - (\lambda_0 - \lambda) - \alpha)) = \Omega(\sqrt{n}).
\]

Proof: By Lemma 1.6 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 $\left[\frac{n - 2a}{2L}\right]$ intervals of size $\leq L$. There must exist an interval $I$ such that
\[
|A \cap I| \geq \frac{2L}{n - 2a} (\lambda_0(\frac{n}{2} - a) - n(\lambda_0 - \lambda)) \, .
\]

If $L = \left\lfloor\sqrt{n}\right\rfloor$ and $a = \alpha n$ then
\[
|A \cap I| \geq \frac{2L}{n - 2a} (\lambda_0(\frac{n}{2} - a) - n(\lambda_0 - \lambda)) \geq \frac{2\sqrt{n}}{(1 - 2\alpha)} (\lambda_0(\frac{1}{2} - \alpha n) - n(\lambda_0 - \lambda))) \geq \frac{2\sqrt{n}}{(1 - 2\alpha)} (\lambda_0(\frac{1}{2} - (\lambda_0 - \lambda)) = \Omega(\sqrt{n}).
\]
Theorem 1.8 For all $\lambda$, $0 < \lambda < 1$, 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 $|A| \leq \lambda n$ 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:

$$T(0)T(1)T(2)T(3) \cdots T(29) = 00011111100001110010111100000$$

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 $t_{\text{small}}$ small 0-blocks and $t_{\text{big}}$ big 0-blocks. Note that $t + 1 = t_{\text{small}} + t_{\text{big}} + 1$ since $t_{\text{small}}, t_{\text{big}} \leq t + 1$. Let the small 0-blocks be $B_{1}^{\text{small}}, \ldots, B_{t_{\text{small}}}^{\text{small}}$, let their union be $B^{\text{small}}$, let the big 0-blocks be $B_{1}^{\text{big}}, \ldots, B_{t_{\text{big}}}^{\text{big}}$, and let their union be $B^{\text{big}}$. If $A$ lives in the (big and small) 0-blocks of $T$ we have

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

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

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

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
A 0-block is a maximal AP with difference $d$ that is contained in $T$. 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 $N_j = \{ x \mid x \leq n \land x \equiv j \pmod{d} \}$.

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

$$T(j)T(d+j)T(2d+j)T(3d+j) \cdots T(29d+j) = 00011111110000111001011111100$$

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 $T(3d+j), T(4d+j), T(5d+j), \ldots, T(9d+j)$ 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

$$1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12 \ 13 \ 14 \ 15 \ 16 \ 17;$$

$$0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0.$$

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

$$2 \ 5 \ 8 \ 11 \ 14 \ 17;$$

$$0 \ 1 \ 1 \ 1 \ 1 \ 0.$$

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 $T \subseteq A$ 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 \( A \)? 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 \( 2y - x \in A \).

Let \( \alpha, 0 < \alpha < \frac{1}{2} \), be a parameter to be determined later. (For those keeping track, the parameters to be determined later are now \( \lambda_0, \lambda, n, \) and \( \alpha \). The parameter \( n_0 \) depends on \( \lambda_0 \) so is not included in this list.)

We want to apply Lemma 1.7.2.b to \( n, n_0, a = \alpha n \). Hence we need the following conditions.

\[
\frac{\alpha n}{2} \geq n_0 \quad \frac{n}{2} \geq n_0 \quad \frac{n}{2} - \alpha n \geq \sqrt{n}
\]

Assuming these conditions hold, we proceed. By Lemma 1.7.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 1.4 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 \( T_i = \{2m - x \mid x \in A \cap [1, e - 1] \land m \in C_i\} \).

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

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 1.6.2 on \( A \cap [1, \alpha n] \). Hence we need the condition

\[
n - \alpha n \geq n_0.
\]

By Lemma 1.6

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

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

\[
\lambda - \lambda_0 + \lambda_0 \alpha > 0 \quad \lambda_0 \alpha > \lambda_0 - \lambda
\]
We know 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 $|T_{i+1} - T_i| \leq \frac{n}{2}$. Let $t = \frac{n}{2}$ (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 $x \in A \cap [1, an - 1]$ 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

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

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 $t \leq \frac{n}{2} = O\left(\frac{n}{\log \log n}\right)$, $d = O(\sqrt{n})$, and $|T| \geq n(\lambda_0 \alpha - (\lambda_0 - \lambda))$. Hence we have

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

We want this to be $< \lambda n$. The term $O\left(\left(\frac{n}{\log \log n}\right) + \sqrt{n}\right)n_0$ 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.

$$\begin{align*}
an & \geq n_0 \\
\frac{n}{2} & \geq n_0 \\
\frac{n}{2} - an & \geq \sqrt{n} \\
n - an & \geq n_0 \\
\lambda_0 \alpha & > \lambda_0 - \lambda \\
\lambda_0(1 - \lambda + \lambda_0 - \lambda_0 \alpha) & < \lambda
\end{align*}$$

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

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

which is clearly true.
\[ \lambda_0(1 - \lambda + \lambda_0 - \lambda_0\alpha) < \lambda \]

becomes

\[
\begin{align*}
\lambda_0(1 - 10^{-1}\lambda_0^2 - 10^{-1}\lambda_0) &< \lambda \\
\lambda_0 - 10^{-1}\lambda_0^3 - 10^{-1}\lambda_0^2 &< \lambda \\
\lambda_0 - \lambda - 10^{-1}\lambda_0^3 - 10^{-1}\lambda_0^2 &< 0 \\
10^{-1}\lambda_0^2 - 10^{-1}\lambda_0^3 - 10^{-1}\lambda_0^2 &< 0 \\
-10^{-1}\lambda_0^2 &< 0
\end{align*}
\]

which is clearly true.

Once \( \lambda, \lambda_0, n_0 \) are picked, you can easily pick \( n \) large enough to make the other inequalities hold. \( \blacksquare \)

2 What more is known?

The following is known.

**Theorem 2.1** 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 [3] and Szemeredi [8]

**Theorem 2.2** There exists \( c \) such that \( sz(n) = \Omega(n^{\frac{1}{\log n}}) \). (Szemeredi estimates \( c \leq 1/20 \)).

Bourgain [1] improved this further to obtain the following.

**Theorem 2.3** \( sz(n) = \Omega(n^{\sqrt{\frac{\log \log n}{\log n}}} \).

References


