EXTENSIONS OF VARNAVIDES' PROOF ON 3-TERM ARITHMETIC PROGRESSIONS

LUCIA PETITO

Abstract

Varnavides proved, as a result of Roth's theorem, a lower bound on the number of three term arithmetic progressions in a set $[1, N]$. We will extend it to find a lower bound on the number of K-term arithmetic progressions and the number of solutions to the equation $c_1x_1 + c_2x_2 + c_3x_3 = 0$ such that $c_1 + c_2 + c_3 = 0$ in a subset of $[1, N]$.

Date: July 2010.

1. Introduction

In 1927, dutch mathematician Baartel Leendert van der Waerden (1903-1996) proved the following theorem about arithmetic progressions:

Given r *and* K*, there exists a positive integer* $N = N(r, K)$ *such that if the integers in* $[1, N]$ *are* r*-colored, there will exist at least one* K*-term arithmetic progression in* $[1, N]$

However, the bound he got on the size of N needed was not very strong. Hungarian mathematicians Paul Erdös (1913-1996) and Paul Turán (1910-1976) posed the following conjecture in 1936 with hopes to improve on the bounds:

Given δ *and* K*, there exists a positive integer* $N = N(\delta, K)$ *such that any subset* A *of* $[1, N]$ *with* $|A| \geq \delta N$ *contains a non-trivial* K*-term arithmetic progression.*

No progress was made on this conjecture for more than 20 years. Finally, in 1953, Klaus Roth (b. 1925) proved it for the case $K = 3$. Szemerédi (b. 1940) proved the $K = 4$ case in 1969, with Roth presenting an alternative proof in 1972, but it was not until 1975 that Szemerédi was able to finally prove the Erdös/Turán conjecture.

During the course of this program we were able to prove the following theorem using the method used in Roth's proof of the $K = 3$ case.

Given an equation $c_1x_1 + \cdots + c_kx_k = 0$ *such that* $\sum_{i=1}^{k} c_i = 0$ *in* $A \subseteq [1, N]$*,* $\delta > 0$*, and* $|A| \geq \delta N$ *for* $N \geq N(\delta, K)$*, a solution* $(x_1, x_2, \ldots, x_k)$ *to the given equation exists in* A*.*

The constant δ, as in all the theorems above, is also known as the density of A in $[1, N]$. While Roth and Szemerédi were thinking about proving the existence of one K-term arithmetic progression, Varnavides was proving the existence of multiple K-term arithmetic progressions in $[1, N]$. In 1955, Varnavides was able to prove the existence of multiple 3-term arithmetic progressions in [1, N], assuming that one existed. In 1958, he published a second paper with a better bound on the number of 3-term arithmetic progressions in $[1, N]$:

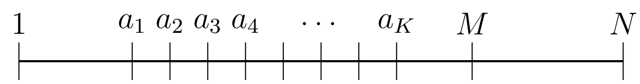
Given $\delta > 0$*, there exists* $N \geq N(\delta)$ *such that any subset* A *of* $[1, N]$ *with* $|A| \geq \delta N$ *contains at least* CN^2 *3-term arithmetic progressions, where* C *is a constant depending on* δ*.*

In this paper we will extend the argument Varnavides presented in 1958 to K-term arithmetic progressions, as well as solutions to linear polynomials in 3 variables, as in $c_1x_1 + c_2x_2 + c_3x_3 = 0$.

2. Varnavides' Bound for K-term Arithmetic Progressions

Theorem: *Given $\delta > 0$, $K \geq 3$ and a set $A \subseteq [1, N]$ of distinct integers with $|A| \geq \delta N$ and $N > N(\delta)$, there exist at least CN^2 distinct K-term arithmetic progressions in A.*

Proof: First, we must create a situation to which we can apply Szemerédi's theorem. Given $\frac{1}{2}\delta > 0$, look at the set of integers $[1, M]$, where $M > N(\frac{1}{2}\delta)$. In a subset $A \subseteq [1, M]$, with $|A| \geq \frac{\delta M}{2}$ there exists at least one K-term arithmetic progression according to Szemerédi's theorem.



Now consider all M-term arithmetic progressions in $[1, N]$ of the form:

$$1 \leq u \leq u + d \leq \ldots \leq u + (M-1)d \leq N$$

We will refer to these progressions as $P_{u,d}$, where u is the start point of the progression, and d is its step size. We are interested in the size of the intersection of $P_{u,d}$ and A, i.e. $|A \cap P_{u,d}|$. We define a *good progression* to be a $P_{u,d}$ such that $|A \cap P_{u,d}| > \frac{\delta M}{2}$. We assume $d < \frac{\delta}{M^2}N$.

Claim: The number of good progressions in $[1, N]$ with a fixed step size d is at least $\frac{1}{4}\delta N$, that is

$$|\{P_{u,d} : |A \cap P_{u,d}| > \tfrac{1}{2}\delta M\}| > \tfrac{1}{4}\delta N$$

Let GP be the total number of good progressions in A for any step size. Let GP_d be the total number of good progressions in A with a fixed step size d.

Look at all $a \in A$ such that $Md \leq a \leq N - Md$. We claim each of those a is contained in exactly M M-term arithmetic progressions with a fixed step size d.

Each a is of the form $a = u + ld$ where $l = 0, 1, 2, \ldots, M-1$. We know $Md \leq a \leq N - Md$. We must consider the end points, that is $a = Md$ and $b = N - Md$. It is clear that if a M-term arithmetic progression begins at $a = Md$, the M-term progression will fit in the interval [1, N]; the same is clear for a M-term arithmetic progression that ends at $a = N - Md$. It suffices to check that a M-term arithmetic progression ending at $a = Md$ fits in $[1, N]$ and a M-term arithmetic progression starting at $b = N - Md$ fits in $[1, N]$. Let $a = Md = d + (M-1)d$.

$$\begin{aligned} a - (M-1)d &= a - (d + (M-2)d) \\ &= d + (M-1)d - d - (M-2)d \\ &= d > 1 \end{aligned}$$

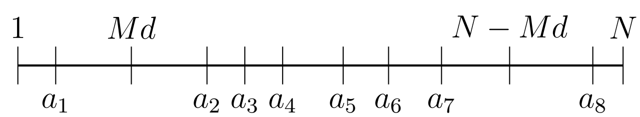
Now let $b = N - Md$.

$$\begin{aligned} b + (M-1)d &= N - Md + (M-1)d \\ &= N - d < N \end{aligned}$$

Since each of the endpoints of our interval can be the beginning and the ending of an M-term arithmetic progression, they can also occur at any point in the M-term arithmetic progression. Therefore, they can be contained in$M$ M-term arithmetic progressions. Since the endpoints can be contained in M M-term arithmetic progressions, each a inside our interval can be contained in M M-term arithmetic progressions with fixed step size d.

Since we know $Md < \frac{\delta N}{M}$ (from $d < \frac{\delta N}{M^2}$), the number of a's that fall in the interval $[Md, N - Md]$ is at least

$$\delta N - 2Md < \delta N - 2\tfrac{\delta N}{M} = \delta(1 - \tfrac{2}{M})N$$



Keep d fixed. Now find a lower bound on the sum of the sizes of the intersections of A and $P_{u,d}$

$$\sum_{u=1}^{N-(M-1)d} |A \cap P_{u,d}| \geq M\delta(1 - \tfrac{2}{M})N > \tfrac{3}{4}\delta MN$$

Now find an upper bound on the sum of the sizes of the intersections of A and $P_{u,d}$. The number of M-term arithmetic progressions with fixed d such that $1 \leq u \leq \cdots \leq u + (M-1)d \leq N$ is clearly less than N (this is a very rough estimate), so

$$\sum_{u=1}^{N-(M-1)d} |A \cap P_{u,d}| < \tfrac{1}{2}\delta MN + (GP_d)M$$

$$\tfrac{1}{4}\delta MN < (GP_d)M \Rightarrow (GP_d) > \tfrac{1}{4}\delta N$$

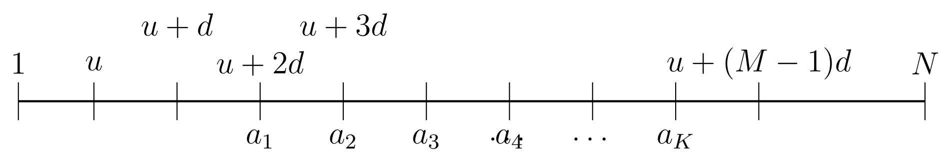
Since we have a lower bound on the number of good progressions for a given step size d, we can find a lower bound on the number of good progressions in A.

$$GP > \tfrac{1}{4}\delta N \tfrac{\delta N}{M^2} = \tfrac{\delta^2 N^2}{4M^2} = CN^2$$

Each $P_{u,d}$ in the set of good progressions contains a K-term arithmetic progression, by Szemerédi's theorem applied to the set:

$$\{1 + \tfrac{a_i - u}{d} : a_i \in A \cap P_{u,d}\}$$

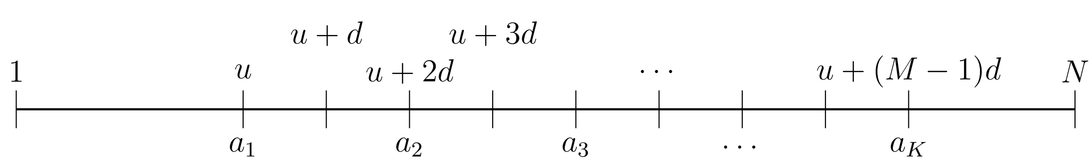
However, this approximation double counts some K-term arithmetic progressions.



Given $P_{u,d}$, an M-term arithmetic progression, our solution with $d = d'$ can be found in at most $M - K + 1$ progressions.

a_1	a_K
$u + 0d$	$u + (K-1)d$
$u + d$	$u + Kd$
$\vdots$	$\vdots$
$u + (M-K)d$	$u + (M-1)d$

The K-term arithmetic progression does not fit in $P_{u,d}$ if a_1 is any of the last $M-K+1$ terms in the M-term arithmetic progression; ie if $l = \{M-K, M-K+1, \ldots, M-1\}$. Thus, we have $M - K + 1$ options for the beginning placement of the K-term arithmetic progression, $l = \{0, 1, 2, \ldots, M-K\}$



Now consider our K-term arithmetic progression iwth step size $\frac{d'}{t} = d$. We have a progression $P_{u,d}$ with step size d. We need $Md > (K-1)d'$ to guarantee that the

K-term arithmetic progression fits inside a M-term arithmetic progression with step size d. This implies

$$M\frac{d'}{t} > (K-1)d' \rightarrow t < \frac{M}{K-1} \quad (d' < \frac{dM}{K-1})$$

We have $\frac{M}{(K-1)}$ options for step sizes, and a very rough bound on the number of M-term arithmetic progressions our K-term arithmetic progression fits in, $M-K+1$. Thus, our K-term arithmetic progression can be found in at most $\frac{M}{K-1}(M-K+1)$ M-term arithmetic progressions.

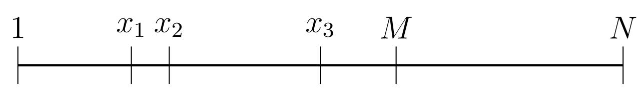
Thus, the same K-term arithmetic progression cannot occur more than C_2 times (where C_2 is a constant depending on M, which in turn depends on δ and K). We can say the number of distinct K-term arithmetic progressions is at least $\frac{C_1}{C_2}N^2 = CN^2$.

3. Varnavides' Bound for Solutions to Linear Polynomials in 3 Variables

As stated in the introduction, it is known that given $\delta > 0$, a set $A \subseteq [1, N]$ of distinct integers with $|A| \geq \delta N$ with $N > N(\delta)$, and an equation $c_1x_1 + \cdots + c_kx_k = 0$ such that $\sum_{i=1}^{k} c_i = 0$, we are guaranteed to have at least one solution in A.

Theorem: *Given $\delta > 0$, a set $A \subseteq [1, N]$ of distinct integers with $|A| \geq \delta N$ with $N > N(\delta)$, and an equation $c_1x_1 + c_2x_2 + c_3x_3 = 0$ with $c_1 + c_2 + c_3 = 0$, there exist at least CN^2 solutions to this equation in A.*

Proof: First, we must create a situation that fits our theorem. Look at $[1, M]$, where $M \geq N(\frac{1}{2}\delta)$. Given $c_1x_1 + c_2x_2 + c_3x_3 = 0$ with $c_1 + c_2 + c_3 = 0$, in a subset $A \subseteq [1, M]$, with $|A| \geq \frac{1}{2}\delta M$, there exists at least one solution to this equation, according to our theorem.



Now consider all M-term arithmetic progressions that satisfy

$$1 \leq u \leq u + d \leq \ldots \leq u + (M-1)d \leq N$$

We assume $d < \frac{\delta}{M^2}N$.

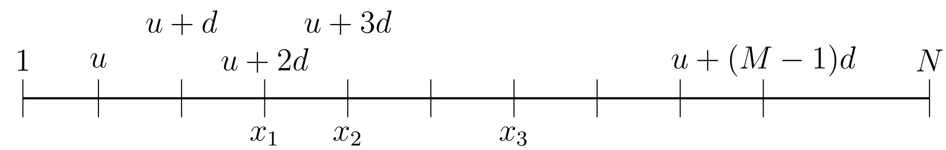
Fix d. The number of good progressions in $[1, N]$ is at least $\frac{1}{4}\delta N$, as shown in section 2. As before, let GP stand for the total number of good progressions for all step sizes.

$$\text{GP} > \frac{1}{4}\delta N\frac{\delta N}{M^2} = \frac{\delta^2 N^2}{4M^2} = CN^2$$

Each $P_{u,d}$ in the set of good progressions contains a solution to our monomial in 3 variables, by the theorem stated earlier applied to the set:

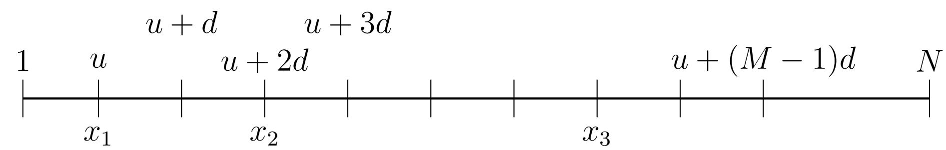
$$\{1 + \tfrac{a_i - u}{d} : a_i \in A \cap P_{u,d}\}$$

However, this approximation double counts a some solutions.



Given $P_{u,d}$, an M-term arithmetic progression, our solution with $d = gcd|x_i - x_j|$ for $i, j \in \{1, 2, 3\}$ can be found in at most $M - 2$ M-term arithmetic progressions.

Worst case scenario: Assuming without loss of generality $x_1 < x_2 < x_3$, if $d = 1 = |x_2 - x_1| = |x_3 - x_2|$, then we have a 3-term arithmetic progression, so we can place this solution in $M - 2$ M-term progressions. The solution does not fit in $P_{u,d} = \{u, u + d, \ldots, u + ld, \ldots, u + (M-1)d\}$ if x_1 is any of the last two terms in the M-term arithmetic progression; ie if $l = \{M - 2, M - 1\}$. Thus, we have M-2 options for the placement of $x_1 = u + ld$, $l = \{0, 1, 2, \ldots, M - 3\}$



Now consider our solution with $\frac{gcd|x_i - x_j|}{t} = d$ for $i, j \in [1, 3]$. We have a progression $P_{u,d}$ with step size d. We know $Md > \max|x_i - x_j|(i, j \in \{1, 3\})$ to ensure that the solution fits in the M-term arithmetic progression. By definition, $\max|x_i - x_j| = wgcd|x_i - x_j|(i, j \in \{1, 2, 3\})$, where $w \in [1, N]$ This means

$$M\frac{gcd|x_i - x_j|}{t} > \max|x_i - x_j| = wgcd|x_i - x_j|$$
$$t < \frac{M}{w} \quad (d' < \frac{dM}{2})$$

We have $\frac{M}{2}$ options for step sizes, and a very rough bound on the number of M-term arithmetic progressions our solution fits in, $M - 2$. Thus, our solution can be found in at most $\frac{M}{2}(M - 2)$ M-term arithmetic progressions.

Thus, the same solution cannot occur more than C_2 times (where C_2 is a constant depending on M, which in turn depends on δ). Therefore, we can say the number of distinct solutions is at least $\frac{C_1}{C_2}N^2 = CN^2$.

References

[1] Roth, Klaus. *On Certain Sets of Integers* Journal London Math. Soc., 28 (1953), 104-109.

[2] Varnavides, P. *On Certain Sets of Positive Density* Journal London Math. Soc. 34 (1959) 358-360.

I would like to thank Neil Lyall, Mariah Hamel, and Alex Rice for their knowledge and support, as well as the other participants in the 2010 Mathematics REU at the University of Georgia: Stephanie Bell, Cliff Blakestad, Bryan Gillespie, Will Grodzicki, Catherine Ha Ta, Hans Parshall, and Frank Xiao.