Distinct Distances and Arithmetic Progressions

Adrian Dumitrescu*

October 9, 2017

Abstract

According to a theorem of Szemerédi (1975), for every positive integer k and every $\delta > 0$, there exists $S = S(k, \delta)$ such that every subset X of $\{1, 2, \ldots, S\}$ of size at least δS contains an arithmetic progression with k terms. Here we generalize the above result to any set of real numbers. For every positive integer k and every $c > 0$, there exists $N = N(k, c)$ such that every set of $n \geq N$ points on the line that determine at most cn pairwise distances contains k points that form an arithmetic progression. A quantitative expression for $N(k, c)$ is derived.

Keywords: Ramsey theory, arithmetic progression, Szemerédi's theorem.

1 Introduction

As remarked by Erdős et al. [5] in 1973, perhaps the first Ramsey type result of a geometric nature is van der Waerden's theorem on arithmetic progressions (see also [9]):

Theorem 1. (van der Waerden [23]). *For every positive integers k and r, there exists a positive integer $W = W(k, r)$ with the following property: For every r-coloring of the integers $1, 2, \ldots, W$ there is a monochromatic arithmetic progression of k terms.*

As early as 1936, Erdős and Turán had suggested that a stronger *density* statement must hold; this was confirmed about 40 years later:

Theorem 2. (Szemerédi [20]). *For every positive integer k and every $\delta > 0$, there exists $S = S(k, \delta)$ such that every subset X of $\{1, 2, \ldots, S\}$ of size at least δS contains an arithmetic progression with k terms.*

It is easy to see that this result is indeed a strengthening of van der Waerden's theorem and that $W(k, r)$ can be chosen to be $S(k, 1/r)$.

Alternative proofs of this result were subsequently obtained by Furstenberg [6], Furstenberg, Katznelson and Ornstein [7], Gowers [8], Rödl et al. [14], Tao [21], Green and Tao [10]. See also the survey by Shkredov [19] for an account on existent proofs. Most notably, some of the alternative proofs also produced much desired quantitative bounds that were previously unavailable (throughout this note all logarithms are in base 2):

Theorem 3. (Gowers [8]). *For every positive integer k there exists a constant $\alpha_k > 0$ such that every subset of $\{1, 2, \ldots, N\}$ of size at least $N(\log \log N)^{-\alpha_k}$ contains an arithmetic progression with k terms. Moreover, α_k can be taken to be $2^{-2^{k+9}}$.*

*Department of Computer Science, University of Wisconsin–Milwaukee, USA. Email: dumitres@uwm.edu

An alternative form of Theorem 3 is the following. We shall use the notation $a \uparrow b$ for a^b; for example, $a \uparrow b \uparrow c = a \uparrow (b \uparrow c)$.

Theorem 4. (Gowers [8]). *Let $0 < \delta \leq 1/2$, let k be a positive integer, let $N \geq 2 \uparrow 2 \uparrow \delta^{-1} \uparrow 2 \uparrow 2 \uparrow (k+9)$, and let A be a subset of the set $\{1, 2 \ldots, N\}$ of size at least δN. Then A contains an arithmetic progression with k terms.*

By combining a result on integer set difference due to Ruzsa [15] with a classic result on simultaneous approximation [11, Theorem 200] (and implicitly assuming Theorems 2 and 3), here we obtain the following generalization of Szemerédi's theorem:

Theorem 5. *For every positive integer k and every $c > 0$, there exists $N = N(k, c)$ such that every set of $n \geq N$ points on the line that determine at most cn pairwise distances contains k points that form an arithmetic progression. Moreover, one can take $N(k, c) = 2 \uparrow 2 \uparrow (5^4 c^4) \uparrow 2 \uparrow 2 \uparrow (k+9)$.*

Let A be a set of n points on the line, and write $g(A)$ for the number of distinct distances determined by A. It is clear that $g(A) = \Theta(|A - A|)$, where $A - A = \{a' - a'' \mid a', a'' \in A\}$; more precisely, $|A - A| = 2g(A) + 1$. It is also clear that $g(A) \geq |A| - 1$, with equality for equidistant points; so if $g(A) \leq c|A|$, one can always assume that $c \geq 1$.

It is plain that Theorem 5 implies Theorem 2: indeed, if $X \subset \{1, 2, \ldots, S\}$ has size at least δS, then $g(X) \leq g(\{1, 2, \ldots, S\}) = S - 1 \leq \delta^{-1}|X|$, and the result follows from Theorem 5. In Section 2 we show that Theorem 2 implies Theorem 5.

Problems on distances determined by n points have been first considered by Erdős in the 1940s [4] and the area is still very active today. See for instance a recent survey by Sheffer [17]. Here we show that the problem of distinct distances and that of arithmetic progressions are connected and establish a quantitative relation.

Preliminaries. Following [15], let $r_k(n)$ denote the maximum number of integers that can be selected from $\{1, 2, \ldots, n\}$ without including an arithmetic progression of k terms and write $w_k(n) = n/r_k(n)$. It is known from Szemerédi's theorem [20] that $w_k(n) \to \infty$ for every fixed k; see also [9, 22]. According to a result of Ruzsa [15], if A is a set of n integers that does not contain a k-term arithmetic progression, then

$$|A - A| \geq \frac{1}{2} w_k(n)^{1/4} \, n, \tag{1}$$

and thus $|A - A|/|A| \to \infty$ for every fixed k. In particular, $g(A)/|A| \to \infty$ for every fixed k.

The real numbers $x_1 < x_2 < \cdots < x_k$ form a *k-term arithmetic progression* if $x_i = x_1 + (i-1)r$, for $i = 2, \ldots, k$, and some $r > 0$. Equivalently, $x_1 < x_2 < \cdots < x_k$ form a k-term arithmetic progression if and only if they satisfy the system of $k - 2$ equations:

$$\begin{aligned} x_1 + x_3 &= 2x_2, \\ x_2 + x_4 &= 2x_3, \\ &\vdots \\ x_{k-2} + x_k &= 2x_{k-1}. \end{aligned} \tag{2}$$

2 Proof of Theorem 5

The argument is analogous with earlier proofs in [1, 2, 13] relying on simultaneous approximation. Let $A = \{a_1 < \ldots < a_n\}$ be a set of n points on the line with $g(A) \leq c|A|$ and no k-term arithmetic progression. We reduce this case to that when A consists of n integers. Using simultaneous approximation [11, Theorem 200], for any positive integer m, there exist n rational points $a_i' = r_i/m$, $i = 1, \ldots, n$ where $r_i, m \in \mathbb{N}$, and

$$\left| a_i - \frac{r_i}{m} \right| \leq \frac{1}{m^{1+1/n}}$$

holds for all $1 \leq i \leq n$. By choosing m large enough, we also have $a_1' < \ldots < a_n'$ and the following two conditions hold: (i) $A' = \{a_1', \ldots, a_n'\}$ has no k-term arithmetic progression, and (ii) $g(A') \leq g(A)$.

Indeed, assume that A' contains a k-term arithmetic progression, as in (2):

$$\begin{aligned} a_1' + a_3' &= 2a_2', \\ a_2' + a_4' &= 2a_3', \\ &\vdots \\ a_{k-2}' + a_k' &= 2a_{k-1}'. \end{aligned} \tag{3}$$

Observe that

$$a_i + a_{i+2} - 2a_{i+1} = (a_i - a_i') + (a_{i+2} - a_{i+2}') - 2(a_{i+1} - a_{i+1}'), \text{ for } i = 1, 2, \ldots, k-2.$$

Applying the triangle inequality gives the following system of inequalities:

$$\begin{aligned} |a_i + a_{i+2} - 2a_{i+1}| &\leq |a_i - a_i'| + |a_{i+2} - a_{i+2}'| + 2|a_{i+1} - a_{i+1}'| \\ &\leq \frac{4}{m^{1+1/n}}, \text{ for } i = 1, 2, \ldots, k-2. \end{aligned}$$

Since this holds for arbitrary large m, we consequently have

$$\begin{aligned} a_1 + a_3 &= 2a_2, \\ a_2 + a_4 &= 2a_3, \\ &\vdots \\ a_{k-2} + a_k &= 2a_{k-1}. \end{aligned} \tag{4}$$

This contradicts the absence of k-term AP's in A thus (i) holds for m large enough.

Assume now that a pair of equal distances in A yields a pair of distinct distances for the corresponding points in A'. Since any pair of equal distances yields a pair whose corresponding intervals are non-overlapping in their interiors, we may assume that

$$|a_j - a_i| = |a_l - a_k| \text{ and } r_j - r_i \neq r_l - r_k, \text{ for some } i < j \leq k < l. \tag{5}$$

Recall that $r_j - r_i \geq 1$ and $r_l - r_k \geq 1$ (since $a_1' < \ldots < a_n'$). We have

$$\frac{r_l - r_k}{m} = \left| \frac{r_l}{m} - a_l + a_l - \frac{r_k}{m} + a_k - a_k \right| \leq |a_l - a_k| + \frac{2}{m^{1+1/n}}, \tag{6}$$

and

$$\frac{r_j - r_i}{m} = \left| \frac{r_j}{m} - a_j + a_j - \frac{r_i}{m} + a_i - a_i \right| \geq |a_j - a_i| - \frac{2}{m^{1+1/n}}. \tag{7}$$

The two inequalities above imply

$$1 \leq |(r_l - r_k) - (r_j - r_i)| \leq m|a_l - a_k| + \frac{2}{m^{1/n}} - m|a_j - a_i| + \frac{2}{m^{1/n}} = \frac{4}{m^{1/n}}. \tag{8}$$

As m tends to infinity, this leads to $(r_l - r_k) = (r_j - r_i)$, in contradiction with (5). Thus (ii) also holds for m large enough.

We next finalize the argument. Multiplying the numbers in A' by m yields a set A'' of n integers, such that A'' has no k-term arithmetic progression and $g(A'')/|A''| = g(A')/|A'| \leq g(A)/|A|$. As such, $g(A'') \leq c|A''|$; however, this is in contradiction to the result of Rusza mentioned in the preliminaries, according to which $g(A'')/|A''| \to \infty$ for every fixed k. This completes the proof of the first part.

To obtain the quantitative bound in Theorem 5, we combine the previous inequalities concerning A'' with (1):

$$\frac{1}{2}w_k(n)^{1/4}\, n \leq |A'' - A''| = 2g(A'') + 1 < \frac{5}{2}\, g(A'') \leq \frac{5}{2}\, c|A''| = \frac{5}{2}\, cn.$$

It follows that $w_k(n) < (5c)^4$. On the other hand, by Theorem 3,

$$r_k(n) \leq \frac{n}{(\log\log n)^{\alpha_k}}, \text{ where } \alpha_k = \frac{1}{2 \uparrow 2 \uparrow (k+9)}, \tag{9}$$

and so

$$(5c)^4 > w_k(n) = \frac{n}{r_k(n)} \geq (\log\log n)^{\alpha_k}. \tag{10}$$

Consequently, the assumption that A has no k-term arithmetic progression implies that

$$n < 2 \uparrow 2 \uparrow (5^4 c^4) \uparrow 2 \uparrow 2 \uparrow (k+9).$$

The proof of Theorem 5 is now complete. $\square$

Remark. It is worth noting that the qualitative statement in Theorem 5 can be also deduced via a different argument based on applying Freiman's theorem, as pointed out to us by Sheffer [18]; a sketch of the argument is included as a final token. On the other hand, the issue of whether a quantitative bound similar (or sharper) to that in Theorem 5 can be obtained via other means remains open for further investigation.

Let d and $n_1, \ldots, n_d$ be positive integers and $x_1, \ldots, x_d$ real numbers. A set G is a *generalized arithmetic progression* of dimension d and size $|G| = n_1 \cdot n_2 \cdot \ldots \cdot n_d$ if

$$G = \left\{ \sum_{i=1}^{d} k_i x_i \mid 0 \leq k_i < n_i \text{ for } i = 1, \ldots, d \right\},$$

and all these elements are distinct; cf. [3, Ch. 1].

In the following, let $c_i \geq 1$ for $i = 1, 2, \ldots$, denote suitable positive constants. Assume that A determines at most $c|A|$ distinct distances, and so $|A - A| \leq c_1|A|$, where $c_1 = 5c/2$. Since $|A - A| \leq c_1|A|$, we also have that $|A + A| \leq c_2|A|$, where $c_2 = c_2(c_1)$; see, e.g., [22, Proposition 2.27]. By Freiman's theorem (see, e.g., [3, Ch. 1], [12, Theorem 8.1], [16], [22, Theorem 5.33]), A is contained in a generalized arithmetic progression G of a bounded dimension $d = d(c_2)$, where $|G| \leq c_3|A|$ and $c_3 = c_3(c_2)$. It follows that A contains $\delta_1|P|$ elements of some standard arithmetic progression P, where $P = \{p_1 + (i-1)r \ : \ i = 1, \ldots, m\}$, $r > 0$, $m \geq |A|^{1/d}$, and $\delta_1 = 1/c_3$; see for instance [12, Theorem 9.1]. Let $B = A \cap P$. A simple bijective transformation, call it f, turns P into $P' = \{1, 2, ..., m\}$; then by Theorem 2, $f(A)$ contains a k-term arithmetic progression provided that $n = |A|$ is large enough. Consequently, B (and thus also A) contains a k-term arithmetic progression, as required.

References

[1] A. Dumitrescu, On distinct distances and λ-free point sets, *Discrete Mathematics* **308** (2008), 6533–6538.

[2] A. Dumitrescu, On distinct distances among points in general position and other related problems, *Periodica Mathematica Hungarica* **57(2)** (2008), 165–176.

[3] G. Elekes, SUMS versus PRODUCTS in number theory, algebra and Erdős geometry, in *Paul Erdős and his Mathematics* vol. II, G. Halász, L. Lovász, M. Simonovits, V. Sós (editors), pp. 241-290, Bolyai Society Mathematical Studies, vol. 11, Springer, 2002.

[4] P. Erdős, On sets of distances of n points, *American Mathematical Monthly* **53** (1946), 248–250.

[5] P. Erdős, R. Graham, P. Montgomery, B. Rothschild, J. Spencer, and E. Straus, Euclidean Ramsey theorems, I, *J. Combinatorial Theory Ser. A* **14** (1973), 341–363.

[6] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.*, **31** (1977), 204–256.

[7] H. Furstenberg, Y. Katznelson, and D. Ornstein, The ergodic theoretical proof of Szemerédi's theorem, *Bulletin of the AMS* **7(3)** (1982), 527–552.

[8] W. T. Gowers, A new proof of Szemerédi's theorem, *Geometric and Functional Analysis* **11(3)** (2001), 465–588.

[9] R. Graham, B. Rothschild, and J. Spencer, *Ramsey Theory*, second edition, John Wiley, New York, 1990.

[10] B. Green and T. Tao, Yet another proof of Szemerédi's theorem, in *An irregular mind: Szemerédi is 70*, I. Bárány, J. Solymosi (editors), pp. 335-342, Bolyai Society Mathematical Studies, vol. 21, Springer, 2010.

[11] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, fifth edition, Oxford University Press, 1979.

[12] M. B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, Springer, 1996.

[13] J. Pach, Midpoints of segments induced by a point set, *Geombinatorics* **13(2)** (2003), 98–105.

[14] V. Rödl, M. Schacht, E . Tengan, and N. Tokushige, Density theorems and extremal hypergraph problems, *Israel Journal of Mathematics* **152** (2006), 371–380.

[15] I. Z. Ruzsa, Arithmetical progressions and the number of sums, *Periodica Mathematica Hungarica* **25(1)** (1992), 105–111.

[16] I. Z. Ruzsa, Sumsets and structure, in *Combinatorial Number Theory and Additive Group Theory*, pp. 87–210, 2009.

[17] A. Sheffer, Distinct distances: open problems and current bounds, preprint, May 2015; `arXiv: 1406.1949`.

[18] A. Sheffer, personal communication, Nov. 2016.

[19] I. D. Shkredov, Szemerédi's theorem and problems on arithmetic progressions, *Russian Mathematical Surveys* **61(6)** (2006), 1101–1166.

[20] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arithmetica* **27** (1975), 199–245.

[21] T. Tao, A quantitative ergodic theory proof of Szemerédi's theorem, *Electronic Journal of Combinatorics* **13** (2006), #R99.

[22] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge University Press, 2006.

[23] B. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wiskunde* **15** (1927), 212–216.