Approximate Euclidean Ramsey theorems

Adrian Dumitrescu*

October 27, 2010

Abstract

According to a classical result of Szemerédi, every dense subset of $1, 2, \ldots, N$ contains an arbitrary long arithmetic progression, if N is large enough. Its analogue in higher dimensions due to Fürstenberg and Katznelson says that every dense subset of $\{1, 2, \ldots, N\}^d$ contains an arbitrary large grid, if N is large enough. Here we present geometric variants of these results for separated point sets on the line and respectively in the Euclidean space: (i) every dense separated set of points in some interval $[0, L]$ on the line contains an arbitrary long approximate arithmetic progression, if L is large enough. (ii) every dense separated set of points in the d-dimensional cube $[0, L]^d$ in $\mathbb{R}^d$ contains an arbitrary large approximate grid, if L is large enough. A further generalization for any finite pattern in $\mathbb{R}^d$ is also established. The separation condition is shown to be necessary for such results to hold. In the end we show that every sufficiently large point set in $\mathbb{R}^d$ contains an arbitrarily large subset of almost collinear points. No separation condition is needed in this case.

Keywords: Euclidean Ramsey theory, approximate arithmetic progression, approximate homothetic copy, almost collinear points.

1 Introduction

Let us start by recalling the classical result of Ramsey from 1930:

Theorem 1 (Ramsey [24]). *Let $p \leq q$, and r be positive integers. Then there exists a positive integer $N = N(p, q, r)$ with the following property: If X is a set with N elements, for any r-coloring of the p-element subsets of X, there exists a subset Y of X with at least q elements such that all p-element subsets of Y have the same color.*

As noted in [5], perhaps the first Ramsey type result of a geometric nature is Van der Waerden's theorem on arithmetic progressions:

Theorem 2 (Van der Waerden [31]). *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 have suggested that a stronger *density* statement must hold. Only in 1975, Szemerédi succeeded to confirm this belief with his celebrated result:

*Department of Computer Science, University of Wisconsin–Milwaukee, Email: `dumitres@uwm.edu`. Supported in part by NSF CAREER grant CCF-0444188.

Theorem 3 (Szemerédi [29]). *For every positive integer k and every $c > 0$, there exists $N = N(k,c)$ such that every subset X of $\{1,2,\dots,N\}$ of size at least cN contains an arithmetic progression with k terms.*

This is a deep result with relations to many areas in mathematics. Szemerédi's proof is very complicated and is regarded as a mathematical tour de force in combinatorial reasoning [19, 23]. Another proof of this result was obtained by means of ergodic theory by Fürstenberg [9] in 1977. More recently, alternative proofs have been also given by Gowers [11], Rödl et al. [25], and Tao [30]. See also the survey by Shkredov [26] for an account on the various proofs.

A homothetic copy of $\{0,1,\dots,k-1\}^d$ is also called a *k-grid* in $\mathbb{R}^d$. The following generalization of Van der Waerden's theorem to higher dimensions is given by the Gallai–Witt theorem [19, 23]:

Theorem 4 (Gallai–Witt [23]). *For every positive integers d, k and r, there exists a positive integer $N = N(d,k,r)$ with the following property: For every r-coloring of the integer lattice points in $\{1,2,\dots,N\}^d$, there exists a monochromatic homothetic copy of $\{0,1,\dots,k-1\}^d$. More precisely, there exist $(a_1,a_2,\dots,a_d) \in \{1,2,\dots,N\}^d$, and a positive integer x such that all points of the form*

$$(a_1 + i_1 x, a_2 + i_2 x, \dots, a_d + i_d x), \quad i_1, i_2, \dots, i_d \in \{0,1,\dots,k-1\}$$

are of the same color.

A higher dimensional generalization of Szemerédi's density theorem was obtained by Fürstenberg and Katznelson [10]; see also [23].

Theorem 5 (Fürstenberg–Katznelson [10]). *For every positive integers d, k and every $c > 0$, there exists a positive integer $N = N(d,k,c)$ with the following property: every subset X of $\{1,2,\dots,N\}^d$ of size at least cN^d contains a homothetic copy of $\{0,1,\dots,k-1\}^d$.*

The proof of Fürstenberg and Katznelson uses infinitary methods in ergodic theory. Although some later proofs might lead to quantitative bounds on the numbers $N(d,k,c)$, we are not aware of such explicit bounds in the literature; see [12, 27].

In Section 2 we present analogues of Theorems 2, 3, 4, and 5, for point sets in the Euclidean space. Specifically, we obtain (restricted) Ramsey theorems for *separated* point sets, for finding approximate homothetic copies of an arithmetic progression on the line and respectively of a grid in $\mathbb{R}^d$. The latter result carries over for any finite pattern point set and every dense and sufficiently large separated point set in $\mathbb{R}^d$. It is worth noting that the separation condition is necessary for such results to hold (Proposition 1 in Section 2). While for Theorems 2, 3, 4, and 5, the separation condition comes for free for any set of integers, it has to be explicitly enforced for point sets.

The exact statements of our results (Theorems 6, 7 and 8) are to be found in Section 2 following the definitions. Fortunately, the proofs of these theorems are much simpler than of their exact counterparts previously mentioned. Moreover, the resulting upper bounds are much better than those one would get from the integer theorems. The proofs are constructive and yield very simple algorithms for computing the respective approximate homothetic copies given input point sets satisfying the requirements.

In Section 3 we present an unrestricted theorem (Theorem 9) which shows the existence of an arbitrary large subset of almost collinear points in every sufficiently large point set in $\mathbb{R}^d$. No separation condition is needed in this result.

Applications. Many other Ramsey type problems in the Euclidean space have been investigated in a series of papers by Erdős et al. [5, 6, 7] in the early 1970s, and later by Graham [13, 14, 15, 16]. See also Ch. 6.3 in the problem collection by Braß, Moser and Pach [1]. Van der Waerden's theorem on arithmetic progressions has inspired new connections and numerous results in number theory, combinatorics, and combinatorial geometry [2, 3, 8, 13, 17, 18, 19, 20, 21, 23, 27, 28], where we only named a few here.

Our analogues of Theorems 2, 3, 4, and 5, for point sets in the Euclidean space may find fruitful applications in combinatorial and computational geometry. It is obvious that general point sets are much more common in these areas than the rather special integer or lattice point sets that occur in number theory and integer combinatorics. A first application needs to be mentioned: a result similar to our Theorem 6 has been proved instrumental in settling a conjecture of Mitchell [22] on illumination for maximal unit disk packings. It is shown [4] that any dense (circular) forest with congruent unit trees that is deep enough has a hidden point (i.e., a point within the forest that cannot be seen from outside the forest). The result that is needed there is an approximate equidistribution lemma for separated points on the line, which is a relaxed version of our Theorem 6.

2 Approximate homothetic copies of any pattern

Definitions. Let $\delta > 0$. A point set S in $\mathbb{R}^d$ is said to be *δ-separated* if the minimum pairwise distance among points in S is at least δ. For two points $p, q \in \mathbb{R}^d$, let $d(p, q)$ denote the Euclidean distance between them. The closed ball of radius r in $\mathbb{R}^d$ centered at point $z = (z_1, \ldots, z_d)$ is

$$B_d(z, r) = \{x \in \mathbb{R}^d \mid d(z, x) \leq r\} = \{(x_1, \ldots, x_d) \mid \sum_{i=1}^{d} (x_i - z_i)^2 \leq r^2\}.$$

Given a point set (or "pattern") $P = \{p_1, \ldots, p_k\}$ of k points in $\mathbb{R}^d$ and another point set Q with k points:

- Q is *similar* to P, if it is a magnified/shrunk and possibly rotated copy[1] of P.
- Q is *homothetic* to P, if it is a magnified/shrunk copy of P in the same position (with no rotations).

Approximate similar copies and approximate homothetic copies are defined as follows. See also Fig. 1 for an illustration. Given point sets P and Q as above and $0 < \varepsilon \leq 1/3$:

- Q is an *ε-approximate similar* copy of of P, if there exists Q' so that Q' is similar to P, and each point $q'_i \in Q'$ contains a (distinct) point $q_i \in Q$ in the ball of radius εd centered at q'_i, where d is the minimum pairwise distance among points in Q'.
- Q is an *ε-approximate homothetic* copy of of P, if there exists Q' so that Q' is homothetic to P, and each point $q'_i \in Q'$ contains a (distinct) point $q_i \in Q$ in the ball of radius εd centered at q'_i, where d is the minimum pairwise distance among points in Q'.

The condition $\varepsilon \leq 1/3$ is imposed to ensure that any two balls of radius εd around points in Q' are disjoint, and moreover, that any two distinct points of Q are separated by a constant times d, in this case by at least $d/3$.[2] In our theorems, ε-approximate means ε-approximate homothetic

[1] A copy of P is a translate of P.

[2] The choice of the constant 1/3 in this definition is rather arbitrary. One could relax this inequality and require $\varepsilon < 1/2$ instead, however this would allow two points in Q be close to each other, which may defeat the intent.

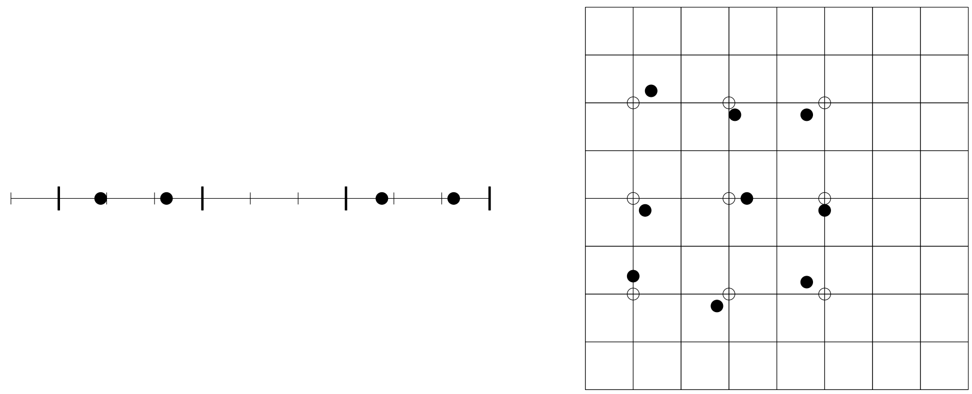

Figure 1: Left: a 4-term arithmetic progression (thick vertical bars) and a 1/3-approximate 4-term arithmetic progression (filled circles) on the line. Right: a 3-grid (empty circles) and a 1/4-approximate 3-grid (filled circles) in $\mathbb{R}^2$.

copy. We start with ε-approximate arithmetic progressions on the line by proving the following analogue of Theorem 3 for points on the line:

Theorem 6 *For every positive integer k, $c, \delta > 0$, and $0 < \varepsilon \leq 1/3$, there exists a positive number $Z_0 = Z_0(k, c, \delta, \varepsilon)$ with the following property: Let S be a δ-separated point set in an interval I of length $|I| = L$ with at least cL points, where $L \geq Z_0$. Then S contains a k-point subset that forms an ε-approximate arithmetic progression of k terms. Moreover, one can set*

$$Z_0(k, c, \delta, \varepsilon) = 2\delta \cdot (ks)^j, \quad \text{where} \quad s = \left\lceil \frac{1}{\varepsilon} \right\rceil, \quad r = \frac{k}{k-1}, \quad j = \left\lceil \frac{\log \frac{2}{c\delta}}{\log r} \right\rceil.$$

Proof. Without loss of generality, $I = [0, L]$. Put $s = \lceil \frac{1}{\varepsilon} \rceil$. Conduct an iterative process as follows. In step 0: Let $I_0 = I$, and subdivide the interval I_0 into ks half-closed intervals[3] of equal length. Let $x = L/(ks)$ be the common length of the sub-intervals. For $t = 0, \ldots, s-1$ consider the system $\mathcal{I}_t$ of k disjoint sub-intervals with left endpoints of coordinates $tx, (t+s)x, (t+2s)x, \ldots, (t+(k-1)s)x$. Observe that the s systems of intervals $\mathcal{I}_t$ partition the interval I_0. If for some t , $0 \leq t \leq s-1$, each of the k intervals contains at least one point in S, stop. Otherwise in each of the s systems of k intervals, at least one of the k intervals is empty, so all the points are contained in at most $ks - s$ intervals from the total of ks. Now pick one of the remaining $(k-1)s$ intervals, which contains the most points of S, say I_1. In step i, $i \geq 1$: Subdivide I_i into ks half-closed intervals of equal length and proceed as before.

In the current step i, the process either (i) terminates successfully by finding an interval I_i subdivided into ks sub-intervals making s systems of intervals, and in at least one of the systems, each sub-interval contains at least one point in S, or (ii) it continues with another subdivision in step $i+1$. We show that if L is large enough, and the number of subdivision steps is large enough, the iterative process terminates successfully.

Let $L_0 = |I| = L$ be the initial interval length, and $m_0 \geq c|I|$ be the (initial) number of points in I_0. At step i, $i \geq 0$, let m_i be the number of points in I_i, and let $L_i = |I_i|$ be the length of interval I_i. For convenience, reset the origin of the x-axis to the left endpoint of I_i, that is, $I_i = [0, L_i]$.

[3] When subdividing a closed interval, the first $k-1$ resulting sub-intervals are half-closed, and the kth sub-interval is closed. When subdividing a half-closed interval, all resulting sub-intervals are half-closed.

Redefine x for the current step as $x = L_i/(ks)$. Clearly

$$L_i = \frac{L}{(ks)^i}, \quad \text{and} \quad m_i \geq \frac{m_0}{(k-1)^i s^i} \geq \frac{cL}{(k-1)^i s^i}. \tag{1}$$

Let j be a positive integer so that

$$c \cdot \delta \cdot \left(\frac{k}{k-1}\right)^j \geq 2, \quad \text{e.g., set} \quad j = \left\lceil \frac{\log \frac{2}{c\delta}}{\log r} \right\rceil, \quad \text{where} \quad r = \frac{k}{k-1}. \tag{2}$$

Now set $Z_0(k, c, \delta, \varepsilon) = 2\delta \cdot (ks)^j$. If $L \geq Z_0$, as assumed, then by (1) and (2) we have

$$L_j = \frac{L}{(ks)^j} \geq \frac{Z_0}{(ks)^j} = \frac{2\delta \cdot (ks)^j}{(ks)^j} = 2\delta, \tag{3}$$

and

$$m_j \cdot \delta \geq \frac{cL \cdot \delta}{(k-1)^j s^j} = c \cdot \delta \cdot \left(\frac{k}{k-1}\right)^j \frac{L}{(ks)^j} \geq \frac{2L}{(ks)^j} = 2L_j. \tag{4}$$

Since the point set is δ-separated, an interval packing argument on the line using (3) gives

$$m_j \delta \leq L_j + \frac{\delta}{2} + \frac{\delta}{2} = L_j + \delta \leq \frac{3}{2} L_j. \tag{5}$$

Observe that (5) is in contradiction to (4), which means that the iterative process cannot reach step j. We conclude that for some $0 \leq i \leq j-1$, step i is successful: we found a system of k intervals of length x with left endpoints at coordinates $a_0 = tx, a_1 = (t+s)x, a_2 = (t+2s)x, \ldots, a_{k-1} = (t + (k-1)s)x$, each containing a distinct point, say $b_p \in S$, $p = 0, 1, \ldots, k-1$. Observe that the k points $\{a_p : p = 0, 1, \ldots, k-1\}$ form an (exact) arithmetic progression of k terms with common difference equal to sx. Note that the minimum pairwise distance among these points is sx. It is now easy to verify that the k points b_p form an ε-approximate arithmetic progression of k terms, since for $p = 0, 1, \ldots, k-1$

$$a_p \leq b_p \leq a_p + x \quad \text{and} \quad \varepsilon s \geq 1, \quad \text{thus} \quad x \leq \varepsilon s x \quad \text{and} \quad b_p \in [a_p, a_p + \varepsilon s x].$$

This completes the proof. □

The next proposition shows that the separation condition in the theorem is necessary, for otherwise, even a 3-term approximate arithmetic progression cannot be guaranteed, irrespective of the size of the point set.

Proposition 1 *For any n and $0 \leq \varepsilon < 1/3$, there exists a set of n points in $[0, 1]$, without an ε-approximate arithmetic progression of 3 terms.*

Proof. Let $\xi = \frac{1}{3} - \varepsilon$. Let $S = \{\xi^i \mid i = 0, \ldots, n-1\}$. Assume for contradiction that $\{q_1, q_2, q_3\}$ is an ε-approximate arithmetic progression of 3 terms, where $q_1 < q_2 < q_3$, and $q_1, q_2, q_3 \in S$. Then there exist a and $r > 0$, so that $a - r$, a and $a + r$ form a 3-term arithmetic progression, and we have:

$$a - r - \varepsilon r \leq q_1 \leq a - r + \varepsilon r,$$
$$a - \varepsilon r \leq q_2 \leq a + \varepsilon r,$$
$$a + r - \varepsilon r \leq q_3 \leq a + r + \varepsilon r.$$

From the first and the third inequalities we obtain

$$a - \varepsilon r \le \frac{q_1 + q_3}{2} \le a + \varepsilon r,$$

therefore

$$\left| \frac{q_1 + q_3}{2} - q_2 \right| \le 2\varepsilon r. \tag{6}$$

Further note that

$$q_3 - q_1 \ge a + r - \varepsilon r - (a - r + \varepsilon r) = 2(1 - \varepsilon) r,$$

hence

$$r \le \frac{q_3 - q_1}{2(1 - \varepsilon)}.$$

By substituting this bound into (6), we have

$$\left| \frac{q_1 + q_3}{2} - q_2 \right| \le \frac{\varepsilon}{1 - \varepsilon} \cdot (q_3 - q_1) \le \frac{\varepsilon}{1 - \varepsilon} \cdot q_3. \tag{7}$$

On the other hand

$$\left| \frac{q_1 + q_3}{2} - q_2 \right| \ge \frac{q_1 + q_3}{2} - q_2 \ge \frac{q_3}{2} - q_2. \tag{8}$$

Putting inequalities (7) and (8) together and dividing by q_3 yields

$$\frac{1}{2} - \frac{q_2}{q_3} \le \frac{\varepsilon}{1 - \varepsilon}.$$

Obviously, $\frac{q_2}{q_3} \le \xi$, hence

$$\frac{1}{6} + \varepsilon = \frac{1}{2} - \frac{1}{3} + \varepsilon = \frac{1}{2} - \xi \le \frac{1}{2} - \frac{q_2}{q_3} \le \frac{\varepsilon}{1 - \varepsilon}.$$

Equivalently,

$$\frac{1}{6} \le \frac{\varepsilon^2}{1 - \varepsilon},$$

which is impossible for $\varepsilon < 1/3$. Indeed, the quadratic function $f(x) = 6x^2 + x - 1$ is strictly negative for $0 < x < 1/3$. We have reached a contradiction and thereby conclude that S has no ε-approximate arithmetic progression of 3 terms. □

Remark. The following slightly different form of Proposition 1 may be convenient: For any n there exists a set of n points in $[0, 1]$, without an ε-approximate arithmetic progression of 3 terms, for any $0 \le \varepsilon \le 1/4$. For the proof, take $S = \{1/8^i \mid i = 0, \dots, n-1\}$, and proceed in the same way.

For a d-dimensional cube $\Pi_{i=1}^d [a_i, b_i]$, let us refer to $(a_1, \dots, a_d)$ as the *first vertex* of the d-dimensional cube. We now continue with ε-approximate grids in $\mathbb{R}^d$ by proving the following analogue of Theorem 5 for points in $\mathbb{R}^d$:

Theorem 7 *For every positive integers d, k, and $c, \delta > 0$, and $0 < \varepsilon \leq 1/3$, there exists a positive number $Z_0 = Z_0(d, k, c, \delta, \varepsilon)$ with the following property: Let S be a δ-separated point set in the d-dimensional cube $Q = [0, L]^d$, with at least cL^d points, where $L \geq Z_0$. Then S contains a subset that forms an ε-approximate k-grid in $\mathbb{R}^d$. Moreover, one can set*

$$Z_0(d, k, c, \delta, \varepsilon) = 2\delta \cdot (ks)^j, \quad \text{where} \quad s = \left\lceil \frac{\sqrt{d}}{\varepsilon} \right\rceil, \quad r = \frac{k^d}{k^d - 1}, \quad j = \left\lceil \frac{\log \frac{\kappa_d}{c\delta}}{\log r} \right\rceil.$$

Here κ_d (in the expression of j) is a constant depending on d:

$$\kappa_d = \left\lceil \frac{3^d \cdot (d/2)!}{\pi^{d/2}} \right\rceil, \text{ if } d \text{ is even, and } \kappa_d = \left\lceil \frac{3^d \cdot (1 \cdot 3 \cdots d)}{2 \cdot (2\pi)^{(d-1)/2}} \right\rceil, \text{ if } d \text{ is odd.} \tag{9}$$

Proof. For simplicity of calculations, we first present the proof for $d = 2$ by outlining the differences from the one-dimensional case; the argument for $d \geq 3$ is analogous, with the specific calculations in the second part of the proof.

Recall that we have set $\kappa_2 = \lceil \frac{9}{\pi} \rceil = 3$. Put $s = \lceil \frac{\sqrt{2}}{\varepsilon} \rceil$. Conduct an iterative process as follows. In step 0: Let $Q_0 = Q$, and subdivide the square Q_0 into $(ks)^2$ smaller congruent squares. Let $x = L/(ks)$ be the common side length of these squares. For $t_1, t_2 \in \{0, \ldots, s-1\}$ consider the system $\mathcal{Q}_{t_1,t_2}$ of k^2 disjoint squares with first vertices of coordinates $(t_1 + i_1 s, t_2 + i_2 s)x$, where $i_1, i_2 \in \{0, 1, \ldots, k-1\}$. Observe that the s^2 systems of squares $\mathcal{Q}_{t_1,t_2}$ partition the square Q_0. If for some (t_1, t_2) , $0 \leq t_1, t_2 \leq s-1$, each of the k^2 squares in the respective system contains at least one point in S, stop. Otherwise in each of the s^2 systems of k^2 squares, at least one of the k^2 squares is empty, so all the points are contained in at most $k^2 s^2 - s^2$ squares from the total of $k^2 s^2$. Now pick one of the remaining $s^2(k^2 - 1)$ squares, which contains the most points of S, say Q_1. In step i, $i \geq 1$: Subdivide Q_i into $(ks)^2$ smaller congruent squares and proceed as before.

In the current step i, the process either (i) terminates successfully by finding a square Q_i subdivided into $(ks)^2$ smaller squares making s^2 systems of squares, and in at least one of the systems, each smaller square contains at least one point in S, or (ii) it continues with another subdivision in step $i + 1$. We show that similar to the one-dimensional case, if L is large enough, and the number of subdivision steps is large enough, the iterative process terminates successfully.

Let $L_0 = L$ be the initial square side of Q_0, and $m_0 \geq cL^2$ be the (initial) number of points in Q_0. At step i, $i \geq 0$, let m_i be the number of points in Q_i, and let L_i be the side length of Q_i. For convenience, reset the origin of the coordinate system to the lower left vertex of Q_i, that is, $Q_i = [0, L_i]^2$. Redefine x for the current step as $x = L_i/(ks)$. Clearly

$$L_i = \frac{L}{(ks)^i}, \quad \text{and} \quad m_i \geq \frac{m_0}{(k^2-1)^i s^{2i}} \geq \frac{cL^2}{(k^2-1)^i s^{2i}}.$$

Let j be a positive integer so that

$$c \cdot \delta^2 \cdot \left(\frac{k^2}{k^2 - 1} \right)^j \geq \kappa_2 = 3, \quad \text{e.g., set} \quad j = \left\lceil \frac{\log \frac{3}{c\delta^2}}{\log r} \right\rceil, \quad \text{where} \quad r = \frac{k^2}{k^2 - 1}. \tag{10}$$

Now set $Z_0(2, k, c, \delta, \varepsilon) = 2\delta \cdot (ks)^j$. If $L \geq Z_0$, as assumed, then by our choice of parameters we have

$$L_j = \frac{L}{(ks)^j} \geq \frac{Z_0}{(ks)^j} = \frac{2\delta \cdot (ks)^j}{(ks)^j} = 2\delta, \tag{11}$$

and

$$m_j \cdot \delta^2 \geq \frac{cL^2 \cdot \delta^2}{(k^2-1)^j s^{2j}} = c \cdot \delta^2 \cdot \left(\frac{k^2}{k^2-1}\right)^j \frac{L^2}{(ks)^{2j}} \geq \frac{3L^2}{(ks)^{2j}} = 3L_j^2. \tag{12}$$

Note that (11) is identical with (3) from the one-dimensional case. Since S is δ-separated, the disks of radius $\delta/2$ centered at the points of S are interior-disjoint. This set of disks is contained in the square $[-\delta/2, L_j + \delta/2]^2$ A straightforward packing argument yields

$$m_j \frac{\pi\delta^2}{4} \leq (L_j + \delta)^2 \leq \left(\frac{3}{2}L_j\right)^2 = \frac{9}{4}L_j^2, \tag{13}$$

where the last inequality is implied by (11). Inequality (13) is equivalent to

$$m_j \cdot \delta^2 \leq \frac{9}{\pi} L_j^2. \tag{14}$$

However this is contradiction with inequality (12) (by the setting $\kappa_2 = \lceil \frac{9}{\pi} \rceil = 3$). This means that the iterative process cannot reach step j.

We conclude that for some $0 \leq i \leq j-1$, step i is successful: we found a system of k^2 disjoint squares of side x with first vertices $a_{i_1,i_2} = (t_1 + i_1 s, t_2 + i_2 s)x$, where $i_1, i_2 \in \{0, 1, \ldots, k-1\}$, each containing a distinct point, say $b_{i_1,i_2} \in S$, for $i_1, i_2 \in \{0, 1, \ldots, k-1\}$. Observe that the k^2 points a_{i_1,i_2} form an (exact) grid Q' of k^2 points with side length equal to sx. As in the one-dimensional case, it is now easy to verify that the k^2 points b_{i_1,i_2} form an ε-approximate grid of k^2 points, since for $i_1, i_2 \in \{0, 1, \ldots, k-1\}$

$$\varepsilon s \geq \sqrt{2}, \quad \text{thus} \quad d(a_{i_1,i_2}, b_{i_1,i_2}) \leq x\sqrt{2} \leq \varepsilon x s. \tag{15}$$

Note that the minimum distance among the points in Q' is sx, and this completes the proof for the planar case ($d = 2$).

The argument for the general case $d \geq 3$ is analogous and the calculations in deriving the upper bound are as follows. The inequality (11) remains valid. By the choice of parameters r and j, we have

$$c \cdot \delta^d \cdot \left(\frac{k^d}{k^d - 1}\right)^j \geq \kappa_d. \tag{16}$$

The analogue of (12) is

$$m_j \cdot \delta^d \geq \frac{cL^d \cdot \delta^d}{(k^d-1)^j s^{dj}} = c \cdot \delta^d \cdot \left(\frac{k^d}{k^d-1}\right)^j \frac{L^d}{(ks)^{dj}} \geq \kappa_d \cdot \frac{L^d}{(ks)^{dj}} = \kappa_d \cdot \left(\frac{L}{(ks)^j}\right)^d = \kappa_d \cdot L_j^d. \tag{17}$$

The packing argument in $\mathbb{R}^d$ yields

$$m_j \cdot \mathrm{Vol}_d\left(\frac{\delta}{2}\right) \leq \left(\frac{3}{2}\right)^d L_j^d, \tag{18}$$

where $\mathrm{Vol}_d(r)$ is the volume of the sphere of radius r in $\mathbb{R}^d$. It is well-known that

$$\mathrm{Vol}_d(r) = \begin{cases} \dfrac{\pi^{d/2}}{(d/2)!} \cdot r^d & \text{if } d \text{ is even,} \\[2ex] \dfrac{2 \cdot (2\pi)^{(d-1)/2}}{1 \cdot 3 \cdots d} \cdot r^d & \text{if } d \text{ is odd.} \end{cases} \tag{19}$$

To obtain a contradiction in the argument, as in the previous cases, one sets κ_d as in (9) taking into account (19). The setting of s is such that the analogue of (15) is ensured. This completes the proof of Theorem 7. $\square$

By selecting a sufficiently fine grid in Theorem 7, one obtains by similar means the following general statement for any pattern in $\mathbb{R}^d$:

Theorem 8 *For every positive integer d, finite pattern $P \subset \mathbb{R}^d$, $|P| = k$, and $c, \delta > 0$, and $0 < \varepsilon \leq 1/3$, there exists a positive number $Z_0 = Z_0(d, P, c, \delta, \varepsilon)$ with the following property: Let S be a δ-separated point set in the d-dimensional cube $Q = [0, L]^d$, with at least cL^d points, where $L \geq Z_0$. Then S contains a subset that is an ε-approximate homothetic copy of P.*

Proof. (Sketch.) Given P, consider its smallest axis-aligned bounding box B. Let ϱ be the minimum pairwise distance among the points in P. Subdivide B into a grid of side length $\leq \frac{\varepsilon\varrho}{3\sqrt{d}}$ by drawing axis-parallel hyperplanes. Let G be the set of grid points (in B); write $|G| = g^d$. Map each point in P to the closest grid point. Observe that this mapping is injective, i. e., no two points of P are mapped to the same grid point. Let P' be the resulting set of grid points; $|P| = |P'| = k$. By Theorem 7 for the grid pattern G, and $\varepsilon/3$, if S is a set of points satisfying the requirements it contains a subset that forms an $\varepsilon/3$-approximate g-grid in $\mathbb{R}^d$. Its subset of k points corresponding to the grid points in P' forms an ε-approximate homothetic copy of P. $\square$

Observe that the iterative procedures used in the proofs of Theorems 6, 7 and 8, yield very simple algorithms for computing the respective approximate homothetic copies, given input point sets satisfying the imposed requirements. For instance in Theorems 6 and 7, the number of iterations, j, is given by (2) and respectively (10), and each iteration takes linear time (in the number of points). On the other hand, these requirements are probably too high, and it is likely that such copies exist under much weaker conditions.

Remarks. The following connection between Theorem 6 and Szemerédi's Theorem 3 is worth making. If one makes abstraction of the bounds obtained, the qualitative statement in Theorem 6 can be obtained as a corollary from Theorem 3. Here is a proof. For simplicity let $n = L$ be integer. Take any set of cn points. Since the set is δ-separated every interval $[i, i+1]$ has at most $1/\delta$ points. Therefore, there are at least $c\delta n$ intervals with at least one point. By Theorem 3, we know that if n is large enough then we can find k/ε intervals which form an arithmetic progression of length k/ε (just think of each interval $[i, i+1]$ as the integer i). To be more precise, if Theorem 3 works for $n \geq N_0(k, c)$ then we apply it with $N_0(k/\varepsilon, c\delta)$. Let $i_0, \ldots, i_{k/\varepsilon}$ be the intervals of this arithmetic progression. Then, by definition each of these k/ε intervals has a point from the set. Pick an arbitrary element of the set from the k intervals $i_0, i_{1/\varepsilon}, i_{2/\varepsilon}, \ldots, i_{k/\varepsilon}$. Then we get an ε-approximate k-term arithmetic progression since the distance between these intervals is at least $1/\varepsilon$, so the error from picking an arbitrary point in each interval is at most ε relative to the distance between the points.

It is also worth noting that our proof of Theorem 6 is self contained and much simpler (from first principles) than the proof one gets from Szemerédi's theorem as described above. Moreover, the upper bound resulting from our proof is much better than that one gets from the integer theorem. That is, with the quantitative bounds included, the two theorems (6 and 7) cannot be derived as corollaries of the classical integer theorems. Also, as mentioned in the introduction, no explicit quantitative bounds seem to be available for the higher dimensional generalization of Szemerédi's theorem.

3 Almost collinear points

Let $0 < \varepsilon < \pi/3$, and let S be a finite point set in $\mathbb{R}^d$. S is said to be ε-*collinear*, if in every triangle determined by S, two of its (interior) angles are at most ε. Note that in particular, this condition implies that an ε-collinear point set is contained in a section of a cylinder whose axis is a diameter pair of the point set, and with radius εD, where D is the diameter; the cylinder radius is at most $\frac{D}{2} \tan \varepsilon \leq \varepsilon D$, for $\varepsilon < \pi/3$.

Theorem 9 *For any dimension d, positive integer k, and $0 < \varepsilon < \pi/3$, there exists $N = N(d, k, \varepsilon)$, such that any point set S in $\mathbb{R}^d$ with at least N points has a subset of k points that is ε-collinear.*

Proof. For simplicity, we present the proof for $d = 2$; the argument for $d \geq 3$ is analogous (sketched at the end). Finitely color all the segments determined by S as follows. Choose a coordinate system, so that no two points have the same x-coordinate. Put $r = \lceil \pi/\varepsilon \rceil + 1$, and let $\mathcal{I}$ be a uniform subdivision of the interval $[-\pi/2, \pi/2]$ into r half-closed subintervals of length at most ε.

Let pq be any segment, where $x(p) < x(q)$. Color pq by i if the angle made by pq with the x-axis belongs to the ith subinterval. Obviously this is an r-coloring of the segments determined by S. Let $N = N(2, k, r)$, where $N(\cdot)$ is as in Theorem 1. By Ramsey's theorem (Theorem 1), for every r-coloring of the segments of an N-element point set, there exists a monochromatic set K of k points, that is, all segments have the same color, say i. Let Δpqr be any triangle determined by K, and assume that $x(p) < x(q) < x(r)$. Then by construction, we have $\angle qpr, \angle prq \leq \varepsilon$. This means that K is ε-collinear, as required.

For $d \geq 3$, consider the unit ball B centered at the origin o in $\mathbb{R}^d$. Partition the solid angle subtended by the upper hemisphere ($x_d \geq 0$) of B into a finite number of solid angles with apex o, such that any two rays from o in the same solid angle make an angle of at most ε. Number (label) these solid angles arbitrarily. Finitely color all segments determined by S, each by the label of the solid angle containing the line through the origin parallel to the corresponding supporting line of the segment. The argument is completed in the same way, by noting that in any triangle determined by K, exactly two of its interior angles are at most ε. □

Remarks. We are not aware of any previous consideration of the above angle condition and its connection with almost collinear point sets. While such angle conditions are preserved under similarity, this is not the case for other measures, such as area. For instance, the classical Heilbronn triangle problem asks for the smallest $a(n)$, such that any set of n points in the unit square spans a triangle whose area is at most $a(n)$. Its generalization asks for the smallest $a_k(n)$, such that any set of n points in the unit square contains k elements whose convex hull has area at most $a_k(n)$ [1, pp. 443–450]. It is easy to see that the small area condition does not imply the small angle condition, even for $k = 3$ (and even in the unit square).

Acknowledgments. The author thanks the anonymous reviewer of an earlier version for the observation and the proof in the remark at the end of Section 2.

References

[1] P. Braß, W. Moser, and J. Pach: *Research Problems in Discrete Geometry*, Springer Verlag, New York, 2005.

[2] P. Braß and J. Pach: Problems and results on geometric patterns, in *Graph Theory and Combinatorial Optimization* (D. Avis et al., editors), Springer Verlag, 2005, pp. 17–36.

[3] A. Dumitrescu and M. Jiang: Monochromatic simplices of any volume, *Discrete Mathematics*, **310** (2010), 956–960.

[4] A. Dumitrescu and M. Jiang: The forest hiding problem, *Discrete & Computational Geometry*, to appear. Online First, May 8, 2010.

[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] P. Erdős, R. Graham, P. Montgomery, B. Rothschild, J. Spencer, and E. Straus: Euclidean Ramsey theorems, II, in *Infinite and Finite Sets*, vol. I; Colloq. Math. Soc. János Bolyai, vol. 10, North-Holland, Amsterdam, 1975, pp. 529–557.

[7] P. Erdős, R. Graham, P. Montgomery, B. Rothschild, J. Spencer, and E. Straus: Euclidean Ramsey theorems, III, in *Infinite and Finite Sets*, vol. I; Colloq. Math. Soc. János Bolyai, vol. 10, North-Holland, Amsterdam, 1975, pp. 559–583.

[8] P. Erdős and R. Graham: Old and New Problems and Results in Combinatorial Number Theory: van der Waerden's Theorem and Related Topics, *L'Enseignement Math.*, **25** (1979), 325–344.

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

[10] H. Fürstenberg and Y. Katznelson: An ergodic Szemerédi theorem for commuting transformations, *J. Analyse Math.*, **34** (1978), 275–291.

[11] W. T. Gowers: A new proof of Szemerédi's theorem, *Geom. Funct. Anal.*, **11(3)** (2001), 465–588.

[12] W. T. Gowers: Some unsolved problems in additive and combinatorial number theory, preprint, 2001; `http://www.dpmms.cam.ac.uk/~wtg10/papers.html#surveys`.

[13] R. Graham: On Partitions of $\mathbb{E}^n$, *J. Combinatorial Theory Ser. A*, **28** (1980), 89–97.

[14] R. Graham: Old and New Euclidean Ramsey Theorems, *Discrete Geometry and Convexity, Annals of the N.Y. Acad. Sci.* **440** (1985), 20–30.

[15] R. Graham: Topics in Euclidean Ramsey Theory, in *Mathematics of Ramsey Theory* (J. Nesetril and V. Rodl, editors), Springer Verlag, New York, (1990), pp. 200–213.

[16] R. Graham: Recent trends in Euclidean Ramsey Theory, *Discrete Mathematics*, **136** (1994), 119–127.

[17] R. Graham: Euclidean Ramsey Theory, in *Handbook of Discrete and Computational Geometry* (E. Goodman and J. O'Rourke, editors), Chap. 11, CRC Press, 2nd ed., 2004, pp. 239–254.

[18] R. Graham: Old and new problems and results in Ramsey theory, in *Horizons of Combinatorics (Bolyai Society Mathematical Studies)*, (E. Gyori, G. Katona and L. Lovász, editors), 2008, 105–118.

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

[20] R. Graham and J. Solymosi: Monochromatic equilateral right triangles on the integer grid, *Topics in Discrete Mathematics*, Algorithms and Combinatorics **26** (2006), 129–132.

[21] A. W. Hales and R. I. Jewett: Regularity and positional games, *Transactions of American Mathematical Society*, **106** (1963), 222–229.

[22] J. Mitchell: Dark points among disks, in *Open Problems from the 2007 Fall Workshop in Computational Geometry*, `http://www.research.ibm.com/people/l/lenchner/fwcg2007/fwcg_open_problems.pdf`.

[23] J. Nešetřil: Ramsey Theory, in *Handbook of Combinatorics* (R. Graham, M. Grötschel and L. Lovász, editors), Vol. II, 1331–1403, Elsevier, Amsterdam, 1995.

[24] E. P. Ramsey: On a problem of formal logic, *Proc. London. Math. Soc.*, **30** (1930), 264–286.

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

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

[27] J. Solymosi: A note on a question of Erdős and Graham, *Combinatorics, Probability and Computing*, **13(2)** (2004), 263–267.

[28] J. Spencer: *Ten Lectures on the Probabilistic Method*, 2nd edition, SIAM, Philadelphia, 1994.

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

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

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