LARGE SETS CONTAINING NO COPIES OF A GIVEN INFINITE SEQUENCE

MIHAIL N. KOLOUNTZAKIS AND EFFIE PAPAGEORGIOU

ABSTRACT. Suppose a_n is a real, nonnegative sequence that does not increase exponentially. For any $p < 1$ we construct a Lebesgue measurable set $E \subseteq \mathbb{R}$ which has measure at least p in any unit interval and which contains no affine copy $\{x + ta_n : n \in \mathbb{N}\}$ of the given sequence (for any $x \in \mathbb{R}, t > 0$). We generalize this to higher dimensions and also for some "non-linear" copies of the sequence. Our method is probabilistic.

CONTENTS

1. INTRODUCTION

In Euclidean Ramsey Theory one is interested in assuming some kind of largeness for sets E in Euclidean space $\mathbb{R}^d$, or, sometimes in $\mathbb{Z}^d$, and concluding that E then contains a "copy" of a pattern. The most famous such example is perhaps Szemeredi's Theorem [Sze75] which states that any subset of the integers with positive density contains arbitrarily long arithmetic progressions. Another well known example is the theorem of Falconer and Marstrand [FM86], Furstenberg, Katznelson and Weiss [FKW90] and Bourgain [Bou86] (see also [Kol04]) that if the set $E \subseteq \mathbb{R}^d$ has positive Lebesgue density (this means that there are arbitrarily large cubes where E takes up at least a constant fraction of the measure) then its points implement all sufficiently large distances (conjecture by Székely [Szé83]).

Another well known problem, very much related to the contents of this paper, is the so-called *Erdős similarity problem*: A set $\mathbb{A} \subseteq \mathbb{R}$ is called *universal in measure* if whenever $E \subseteq \mathbb{R}$ has positive Lebesgue measure we can find an affine copy of A contained in E. In other words $x + t\mathbb{A} \subseteq E$ for some $x \in \mathbb{R}, t > 0$. It is easy to see that every finite set $\mathbb{A}$ is universal (just look close enough to some point of density of E, shrink $\mathbb{A}$ enough and

average the number of points of the copy of $\mathbb{A}$ that belong to E over translates of $\mathbb{A}$ nearby) but it has been conjectured [Erd15] (see also [CFG12, p. 183]) that no infinite set $\mathbb{A}$ can be universal in measure. This is known for many classes of infinite sets but not for all [Fal84, GLW23, HL98, Kom83, Chl15]. Clearly it would suffice to prove this for $\mathbb{A}$ being a positive sequence a_n decreasing to 0 but if a_n decays fast to 0 (so it is in some sense sparse, hence not that hard to contain) this is still unknown. On the contrary this is known when $\log \frac{1}{a_n} = o(n)$. This is not known if $a_n = 2^{-n}$, for example.

In this paper we consider an analogue of the Erdős similarity problem "in the large". Let $\mathbb{A} \subseteq \mathbb{R}$ be a discrete, unbounded, infinite set in $\mathbb{R}$. Can we find a "large" measurable set $E \subseteq \mathbb{R}$ which does not contain any affine copy $x + t\mathbb{A}$ of $\mathbb{A}$ (for any $x \in \mathbb{R}, t > 0$)? Our attention to this problem was drawn by a recent paper by Bradford, Kohut and Mooroogen [BKM23] in which the authors prove that if $\mathbb{A}$ is an infinite arithmetic progression then this is indeed possible: for any $p \in [0, 1)$ they construct a Lebesgue measurable set E, with measure at least p in any interval of length 1, which does not contain any affine copy of $\mathbb{A}$. This is clearly equivalent to being able to obtain, for any $p \in [0, 1)$ a set E avoiding all infinite arithmetic progressions and having measure $\geq p$ in any interval of length 1 *whose endpoints are integers*. (Indeed, if the set E has measure at least p in every interval of the form $[n, n + 1]$, $n \in \mathbb{Z}$, then, since for any x the interval $[x, x + 1]$ is contained in the union of two such unit-length intervals with integer endpoints, we obtain that $[x, x + 1] \setminus E$ has measure at most $2(1 - p)$. Since p can be as close to 1 as we want, this implies that $[x, x + 1] \setminus E$ has measure as close to 0 as we want.) From now on we follow this simplification and we deal only with intervals with integer endpoints (in any dimension).

We generalize the result of [BKM23] to sequences of nonnegative numbers $\mathbb{A}$ which do not grow too fast. To state our result, we introduce the following class of sequences.

Definition 1.1. We say that a real sequence $\mathbb{A} = \{a_n,\ n \in \mathbb{N}\}$ is in the *class (A)* if

    (1) $a_0 = 0$,
    (2) $a_{n+1} - a_n \geq 1$, for every $n \in \mathbb{N}$.
    (3) $\log a_n = o(n)$

Remark. Since the problem we are studying is translation invariant condition 1.1.(1) in Definition 1.1 is unnecessary, but we keep it as it simplifies the proofs somewhat.

Writing

$$(1.1) \qquad\qquad\qquad A(t) = |\mathbb{A} \cap [0, t]|$$

for the *counting function* of the set $\mathbb{A}$, notice that the growth condition 1.1.(3) is equivalent to the limit, as $t \to +\infty$,

$$(1.2) \qquad\qquad\qquad \frac{A(t)}{\log t} \to +\infty.$$

Our main result is the following.

Theorem 1.1. *Consider the sequence $\mathbb{A} = \{a_n :\ n \in \mathbb{N}\}$ which belongs to the class (A). Then, for each $0 \leq p < 1$, there exists a Lebesgue measurable set $E \subseteq \mathbb{R}$ such that*

$$|E \cap [m, m + 1]| \geq p, \quad \text{for all } m \in \mathbb{Z},$$

but E does not contain any affine copy of $\mathbb{A}$.

As in the case of the Erdős similarity problem described above, the sparser the set $\mathbb{A}$ is the easier it should be to be contained in large sets, so it is not surprising that we had to impose a growth condition (to belong to the class (A)). It remains an open question if a similar set E can be constructed when $\mathbb{A}$ grows exponentially or faster.

Question 1. *Is there a sequence $0 < a_n \to +\infty$ and a number $p \in [0, 1)$ such that one can find an affine copy of $\mathbb{A} = \{a_n : n \in \mathbb{N}\}$ in any set $E \subseteq \mathbb{R}$ which has measure more than p in any interval of length 1?*

Unlike the approach taken in [BKM23] our method of proof is probabilistic. We construct a family of random sets and we show that, with high probability, such a random set will have all the properties we want. This method turns out to be extremely flexible, and this allows us to generalize. Not only can we deal with essentially arbitrary and unstructured sequences $\mathbb{A}$ but we can also relax the sense in which we seek copies of $\mathbb{A}$ in the large set E. Instead of scaling the elements of $\mathbb{A}$ and translating them

$$x + ta_n, \quad x \in \mathbb{R}, t > 0,$$

we can allow for more general transformations

$$(1.3) \qquad\qquad x + \phi(n, t) \cdot a_n, \quad x \in \mathbb{R}, t > 0.$$

Theorem 1.2. *Consider the set $\mathbb{A} = \{a_n : n \in \mathbb{N}\}$, which belongs to the class (A), and let $\phi(n, t) : \mathbb{N} \times (0, +\infty) \to (0, +\infty)$ be such that for each n the function $\phi(n, t)$ is increasing in t and is such that for all $n \in \mathbb{N}$ we have*

$$(1.4) \qquad\qquad C_1 t \leq \phi(n + 1, t)a_{n+1} - \phi(n, t)a_n$$

and

$$(1.5) \qquad\qquad \phi(n, t) \leq C_2 t, \quad \text{for all } t > 0,$$

for some $C_1, C_2 > 0$. Then, for each $0 \leq p < 1$, there exists a Lebesgue measurable set $E \subseteq \mathbb{R}$ such that E intersects every interval of unit length in a set of measure at least p, but E does not contain the set

$$\left\{ x + \phi(n, t) \cdot a_n : n \in \mathbb{N} \right\}$$

for any choice of $x \in \mathbb{R}, t > 0$.

We adopt certain arguments from [Kol97, Section 3] where it is proved, on the Erdős similarity problem, that sequences with a finite limit, say 0, which are not decaying very fast (e.g. they decay polynomially or subexponentially but not, for instance, exponentially fast – compare to our growth condition (1.1.(3))), cannot be universal in measure, by showing the existence of a randomly constructed set $E \subseteq [0, 1]$, avoiding all affine copies of the sequence.

The measure assumption makes this problem different than other "avoidance" problems, where the avoiding set is often taken to have zero Lebesgue measure but to have large Hausdorff dimension or Fourier dimension. For example, in [Kel08], a compact subset of $\mathbb{R}$ is constructed that has full Hausdorff dimension but does not contain any 3-term arithmetic progression. See also [CLP22, DPZ21, FP18, Mag11, Mát17, Shm17, Yav21].

We can also prove the following result in higher dimension. We phrase it as avoiding linear images of a set in Euclidean space into another Euclidean space. In this manner we obtain easily some corollaries, Theorem 1.1 one of them, and its proof is rather simpler

than that of Theorem 1.1 given in §3. But it does not extend easily to more complicated transformations such as those in Theorem 1.2, so we choose to stay with linear maps.

Theorem 1.3. *Let $d_1, d \geq 1$, $b, f > 0$, $p \in [0, 1)$. Let also $\alpha(R)$ be a function satisfying $\dfrac{\alpha(R)}{\log R} \to +\infty$ as $R \to +\infty$.*

Then if $\mathbb{A} \subseteq \mathbb{R}^{d_1}$ is a discrete point set such that

$$(1.6) \qquad |\mathbb{A} \cap B_R(0)| \leq C_2 R^b, \quad (R > 0)$$

there is a set $E \subseteq \mathbb{R}^d$ such that

  i. $\left| E \cap (m + [0, 1]^d) \right| \geq p$ for all $m \in \mathbb{Z}^d$,
  ii. *For any linear map $T : \mathbb{R}^{d_1} \to \mathbb{R}^d$ if for arbitrarily large values of R*

$$(1.7) \qquad T(\mathbb{A}) \cap B_R(0)$$

  contains at least $\alpha(R)$ points with separation R^{-f} then

$$(1.8) \qquad T(\mathbb{A}) \text{ is not contained in } E.$$

Proof of Theorem 1.1 using Theorem 1.3. Apply Theorem 1.3 with $d_1 = 2, d = 1, b = 1, \alpha(x) = A(x^{1/2})$ (where $A(x)$ is the counting function of $\mathbb{A}$), $f = 1$ (there is great flexibility in choosing $\alpha(x), b, f$) and the set

$$P = \mathbb{A} \times \{1\} \subseteq \mathbb{R}^2$$

to obtain a set $E \subseteq \mathbb{R}$ satisfying $|E \cap [m, m+1]| \geq p$ for all $m \in \mathbb{Z}$. We see that (1.6) is satisfied. Let now $T : \mathbb{R}^2 \to \mathbb{R}$ be given by the 1×2 matrix $T = (t, x)$ so that

$$T(P) = x + t\mathbb{A}.$$

For any $x \in \mathbb{R}, t > 0$, the set $(x + t\mathbb{A}) \cap [-R, R]$ contains at least $A(R/t)$ points of separation t, so, if R is large enough, it contains $\alpha(R) = A(R^{1/2})$ points with separation R^{-1}. It follows that $x + t\mathbb{A}$ is not contained in E. $\square$

Corollary 1.4 (Avoiding linear images of general sets in high dimension)**.** *Let $p \in [0, 1)$, $d \geq 1$, $a_n \in \mathbb{R}^d$, for $n \in \mathbb{N}$, with $\log |a_n| = o(n)$ and $|a_n - a_{n+1}| \geq 1$ for all $n \in \mathbb{N}$. Then there is a set $E \subseteq \mathbb{R}^d$ such that for all $m \in \mathbb{Z}^d$ we have $\left| E \cap (m + [0, 1]^d) \right| \geq p$ and such that for all $x \in \mathbb{R}^d$ and for all non-singular linear $T : \mathbb{R}^d \to \mathbb{R}^d$ the set $\{x + Ta_n : n \in \mathbb{N}\}$ is not contained in E.*

Proof. Take $\mathbb{A} \subseteq \mathbb{R}^{2d}$ to be the set $A \times \{\underbrace{(1, 0, \dots, 0)}_{d}\}$, where $A = \{a_n : n \in \mathbb{N}\}$. Writing $A(s) = \#(A \cap B_s(0))$ for the counting function of A we have $\dfrac{A(R)}{\log R} \to +\infty$. Use Theorem 1.3 with $d_1 = 2d, b = 1, \alpha(R) = A(R^{1/2}), f = 1$. Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be non-singular, $x \in \mathbb{R}^d$, and define the linear map $S : \mathbb{R}^{2d} \to \mathbb{R}^d$ by

$$S(u, v) = S(u, v_1, v_2, \dots, v_d) = Tu + v_1 x.$$

In other words the $d \times (2d)$ matrix of S is $(T \mid x \mid 0)$ in block form. It follows that

$$S(\mathbb{A}) = \{Ta_n + x : n \in \mathbb{N}\}.$$

Since T is non-singular it follows that if $R > 0$ is sufficiently large the set $S(\mathbb{A}) \cap B_R(0)$ contains at least $\alpha(R)$ points with separation $\geq R^{-1}$ so the set $E \subseteq \mathbb{R}^d$ furnished by Theorem 1.3 does not contain $S(\mathbb{A})$, as we had to prove. □

Corollary 1.5 (Corollary 6 from [BKM23]). *If $p \in [0, 1)$ then there exists a set $E \subseteq \mathbb{R}^d$ such that $\left| E \cap (m + [0, 1]^d) \right| \geq p$ for all $m \in \mathbb{Z}^d$ and it does not contain any set of the form $x + \mathbb{N}\Delta$, with $x \in \mathbb{R}^d$, $\Delta \in \mathbb{R}^d \setminus \{0\}$ (an arithmetic progression in $\mathbb{R}^d$).*

Proof. We use Corollary 1.4 with the sequence $a_n = (n, 0, \ldots, 0) \in \mathbb{R}^d$, $x \in \mathbb{R}^d$ and any non-singular $d \times d$ matrix T that maps $(1, 0, \ldots, 0)$ to Δ. □

The outline of this note is as follows. In §3 we give the proof of Theorem 1.1 without using Theorem 1.3, and we indicate how the same proof also works for Theorem 1.2. In §4 we extend our technique to cover linear transformations of given sequences from one Euclidean space to another and prove Theorem 1.3 and some corollaries.

Added in revision: The results in [BGK+22], which came after this paper was submitted, are very relevant to the results in this paper and contain some improvements.

2. WARM-UP AND SOME BASIC TOOLS: NO TRANSLATIONAL COPIES

In this section we introduce the basic probabilistic method by proving the more restricted Theorem 2.1: we can avoid all translations of a given infinite sequence $0 \leq a_n \to +\infty$ with a set which is arbitrarily large everywhere. This is considerably easier than avoiding all affine copies of the sequence, when scaling the sequence as well as translating it is allowed. For translations we have only one degree of freedom while for affine copies we have two. Still, some important ingredients of the method will be evident in the proof of Theorem 2.1 below. In §3 we will introduce the extra discretization in scaling space that will be required.

Theorem 2.1. *Let $\mathbb{A} = \{a_0 = 0 < a_1 < a_2 < \cdots\} \subseteq \mathbb{R}$ be a sequence with $a_n \to +\infty$, and $p \in [0, 1)$. Then we can find a Lebesgue measurable set $E \subseteq \mathbb{R}$ such that no translate of $\mathbb{A}$*

$$x + \mathbb{A}, \quad x \in \mathbb{R},$$

is contained in E, and such that for each $m \in \mathbb{Z}$ we have

$$|E \cap [m, m + 1]| \geq p.$$

Proof. Let $q < 1$ be defined by $1 - q = \frac{1}{2}(1 - p)$ (or $q = \frac{1}{2}(1 + p)$). Passing to a subsequence we can assume that $a_{n+1} - a_n \geq 1$ for all n. We construct a random set E by breaking up each unit interval $[m, m + 1]$, $m \in \mathbb{Z}$, into a number N_m of equal intervals and keeping each of these subintervals with probability q, independently, into our set E. As $|m|$ increases the number N_m will also have to increase, so let us take $N_m = \max\{K, |m|\}$ say, where the large positive integer K will be determined later.
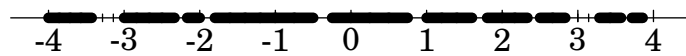

FIGURE 1. How the random set E looks like

Define now for $x \in \mathbb{R}$ the random function

$$\phi(x) = \mathbb{1} \left(x + \mathbb{A} \subseteq E \right).$$

Since all points of $x + \mathbb{A}$ are in different random intervals it follows, by independence, that $\mathbb{E}\phi(x) = \mathbb{P}\left[x + \mathbb{A} \subseteq E\right] = 0$. Let the set of "bad" x be

$$B = \{x \in \mathbb{R} : \ x + \mathbb{A} \subseteq E\}.$$

We have

$$\mathbb{E}|B| = \int \mathbb{E}\phi(x)\, dx = 0,$$

hence $|B|$ is almost surely 0.

It remains to make sure that $|E \cap [m, m+1]| \geq p$ for all $m \in \mathbb{Z}$. Fix m and let $X_1, \ldots, X_{N_m}$ be 0/1 random variables such that X_i is 0 if we included the i-th subinterval of $[m, m+1]$ into the set E and is 1 otherwise. In other words, X_i denotes the absence of the i-the subinterval from the set E. Clearly $\mathbb{E}X_i = 1 - q$ and the random variable

$$X = \sum_{i=1}^{N_m} X_i \text{ (the number of missing subintervals)}$$

is a sum of independent indicator random variables with $\mathbb{E}X = (1-q)N_m$ and we can use the very versatile large deviation *Chernoff inequality* (to be used repeatedly in §§3.1,4 below)

$$(2.1) \qquad\qquad \mathbb{P}\left[|X - \mathbb{E}X| \geq \epsilon \mathbb{E}X\right] \leq 2e^{-c_\epsilon \mathbb{E}X}$$

(see [Che52, AS16]) with $\epsilon = 1$ to obtain

$$\begin{aligned} \mathbb{P}\left[|E \cap [m, m+1]| < p\right] &= \mathbb{P}\left[X > (1-p)N_m\right] \\ &= \mathbb{P}\left[X - \mathbb{E}X > \mathbb{E}X\right] \\ (2.2) \qquad\qquad &\leq 2\exp(-c_1(1-q)\max\{K, |m|\}). \end{aligned}$$

Define now the bad events $B_m = \{|E \cap [m, m+1]| < p\}$ which we want not to hold, for all $m \in \mathbb{Z}$, and observe that the above inequality means that we can choose K large enough to achieve

$$\sum_{m \in \mathbb{Z}} \mathbb{P}\left[B_m\right] < \frac{1}{2}.$$

This means that with probability at least $1/2$ none of the bad events B_m hold and, with the same probability, the set B has measure 0. We now amend our random set E by removing from it the set B (the set of first terms of those $x + \mathbb{A}$ which are contained in E). Thus arises a set E', which differs from E by a set of measure 0, and which contains no translate of $\mathbb{A}$. $\square$

Remark 2.1. It is not necessary to assume that $a_n \to +\infty$ in Theorem 2.1. It suffices to assume that the set $\mathbb{A}$ is infinite. If $\mathbb{A}$ does not contain a sequence tending to infinity (for Theorem 2.1 to apply to it) then it will have a finite accumulation point, so a result of Komjáth [Kom83] guarantees the existence of a set $\widetilde{E} \subseteq [0, 1]$, of measure arbitrarily close to 1, which contains no translate of $\mathbb{A}$. Repeating $\widetilde{E}$ 1-periodically

$$E = \bigcup_{n \in \mathbb{Z}} \widetilde{E} + n$$

we obtain a set E with the required properties. For a probabilistic proof of this result in the spirit of the present paper see [Kol97].

Remark 2.2. The Chernoff inequality (2.1) is extremely useful when one needs to control a random variable X (this means that one wants to ensure, with high probability, that X is near its mean $\mathbb{E}X$) which is a sum of indicator, independent random variables. The key is that the mean $\mathbb{E}X$ cannot be very small, as it appears in the exponent in the right hand side of (2.1). Since one usually wants to do so *simultanesouly* for a large number of random variables X, one key situation to keep in mind is the following: if the number of random variables to be controlled is polynomial in N (a parameter) it is enough that their mean are at least a large multiple of $\log N$.

With minor modifications of the proof we can get a progressively denser set E avoiding all translates. We throw in the whole negative half line (as we could have done in Theorem 1.1 too).

Theorem 2.2. *Let* $\mathbb{A} = \{a_0 = 0 < a_1 < a_2 < \cdots\} \subseteq \mathbb{R}$ *be a sequence with* $a_n \to +\infty$. *Then we can find a Lebesgue measurable set* $E \subseteq \mathbb{R}$ *such that no translate of* $\mathbb{A}$

$$x + \mathbb{A}, \quad x \in \mathbb{R},$$

is contained in E, *and such that*

$$(-\infty, 0] \subseteq E \quad and \quad |E \cap [m, m+1]| \to 1^- \ as \ m \to +\infty.$$

Proof. We indicate the differences with the proof of Theorem 2.1 and omit some details.

Our random set E now will be of the same type as in the proof of Theorem 2.1 but with the probability of including the small subintervals tending slowly to 1 as we go out to $+\infty$ and with the negative half line contained in E to begin with.

Let us view the probability of keeping an interval as a function $p(s)$ defined on the real line. In the proof of Theorem 2.1 this function was constant. Here it will be constant on all intervals of the form $[m, m+1]$, $m \in \mathbb{Z}$.

With $\phi(x) = \mathbb{1}\,(x + \mathbb{A} \subseteq E)$ we need again to ensure that $\mathbb{E}\phi(x) = 0$ for all $x \in \mathbb{R}$. After assuming, as in the previous proof, that the points of $\mathbb{A}$ differ by at least 1, we again have independence of all events $x + a \in E$ for $a \in \mathbb{A}$ so that $\mathbb{E}\phi(x) = 0$ becomes equivalent to

$$\prod_{a \in \mathbb{A}} p(x + a) = 0,$$

which, writing $q(s) = 1 - p(s)$, is equivalent to

$$(2.3) \qquad\qquad \sum_{a \in \mathbb{A}} q(x + a) = +\infty.$$

Let $0 = k_1 < k_2 < \cdots$ be those positive integers for which

$$[k, k+1) \cap \mathbb{A} \neq \emptyset.$$

Define then $q(x)$ to be $1/i$ in the interval $[k_i, k_{i+1})$, $i = 1, 2, \ldots$. It follows easily that for all $x \in \mathbb{R}$ we have (2.3): since the function $q(\cdot)$ is decreasing we have $q(x + a_n) \geq q(a_n)$ if $x \leq 0$ and if $x \geq 0$ we have $q(x + a_n) \geq q(a_{\lceil x \rceil + n})$ since $a_{k+1} - a_k \geq 1$ for all $k \in \mathbb{N}$. In both cases the series (2.3) contains a tail of the series $\sum_{a \in \mathbb{A}} q(a)$ which is divergent.

It remains to ensure that the random variables $|[m, m+1] \setminus E|$ tend to 0 with $m \to +\infty$. These random variables are $\frac{1}{N_m}$ times a sum of independent indicator random variables

(one for each of the N_m subintervals into which we break up $[m, m + 1]$) of mean $q(m)N_m$ so we can use the Chernoff bound (2.1) to obtain

$$\mathbb{P}\left[\,|[m, m + 1] \setminus E| > 2q(m)\right] \leq 2\exp(-c_1 q(m)N_m).$$

To ensure that the sum, over all $m \in \mathbb{Z}$ of the left hand side is < 1 we can of course pick the integers N_m to be very large, say $N_m = K\frac{1}{q(m)}|m|$, with a sufficiently large constant $K > 0$.

$\square$

3. No affine copies for slowly increasing sequences

In this section we prove Theorem 1.1 and explain why the proof also gives the more general Theorem 1.2.

Lemma 3.1. *Let $\mathbb{A} \in (A)$. For all $0 < a < b$, $0 \leq p < 1$ and $\epsilon > 0$, there is $N_0 \in \mathbb{N}$, such that for all $N \geq N_0$, there is a set $E \subseteq [-N, N]$ such that*

(i) for all $m \in \{-N, -N + 1, ..., N - 1\}$, we have $|E \cap [m, m + 1]| \geq p$, and
(ii) if the set B consists of all $x \in [-N, N]$ for which there is $t \in [a, b]$ such that
  (a) $(x + t\mathbb{A}) \cap [-N, N] \subseteq E$ and
  (b) $\#((x + t\mathbb{A}) \cap [-N, N]) \geq A\left(\frac{N}{10b}\right)$,

then $|B| < \epsilon$. Here, $A(\cdot)$ is the counting function (1.1) of the set $\mathbb{A}$ and $A\left(\frac{N}{10b}\right) = |\mathbb{A} \cap [0, N/(10b)]|$.

Let us first show how one derives Theorem 1.1 from Lemma 3.1. We give the proof of Theorem 1.1 in two steps: the first one verifies the result for a restricted scale, that is, for scales in a compact interval, and the second one concludes for all positive scales, by writing the whole scaling interval $(0, +\infty)$ as a countable union of intervals of the above type.

Step 1. *For all $0 < a < b$ and for each $0 \leq p < 1$, there exists a set $E \subseteq \mathbb{R}$, such that $|E \cap [m, m + 1]| \geq p$ for all $m \in \mathbb{Z}$, but E does not contain any affine copies of $\mathbb{A}$ with scale in $[a, b]$.*

Consider $0 \leq p < 1$ and a positive increasing sequence $\{p_n\}$, $n = 1, 2, ...$ such that $p_n \to 1^-$ and, moreover,

$$(3.1) \qquad \sum_{n=0}^{\infty}(1 - p_n) < 1 - p.$$

Take also any positive sequence $\epsilon_n \to 0$. According to Lemma 3.1, for $0 < a < b$, we can choose an increasing sequence of natural numbers $N_n = N_n(p_n, \epsilon_n, a, b) \to \infty$, for which there exist sets $E_n \subseteq [-N_n, N_n]$ with the following properties:

(i) for all $m = -N_n, ..., N_n - 1$, we have $|E_n \cap [m, m + 1]| \geq p_n$,
(ii) if

$$\mathbb{A}_n(x, t) = (x + t\mathbb{A}) \cap [-N_n, N_n]$$

and

$$B_n = \{x \in [-N_n, N_n] : \exists t \in [a, b] \text{ s.t. } \mathbb{A}_n(x, t) \subseteq E_n$$
$$\text{and } \#\mathbb{A}_n(x, t) \geq A\left(\frac{N_n}{10b}\right)\},$$

then $|B_n| < \epsilon_n$.

Now take

$$\widetilde{E_n} = (-\infty, -N_n] \cup E_n \cup [N_n, +\infty)$$

and

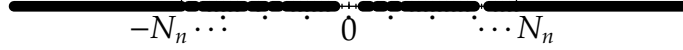$$E = \bigcap_{n=1}^{\infty} \widetilde{E_n}.$$

$$-N_n \cdots \qquad 0 \qquad \cdots N_n$$

FIGURE 2. The set $\widetilde{E_n}$.

Then, since $\left|\widetilde{E_n} \cap [m, m+1]\right| \geq p_n$ for all $m \in \mathbb{Z}$, we get from (3.1) that the set E has measure at least p at every unit interval with integer endpoints. Also, if there exist x, t such that $x + t\mathbb{A} \subseteq E$, then $x + t\mathbb{A}$ is also contained in each $\widetilde{E_n}$. Having fixed x and t we can then find n_0 large enough such that for all $n \geq n_0$, we have $\#((x + t\mathbb{A}) \cap [-N_n, N_n]) \geq A\left(\frac{N_n}{10b}\right)$. This implies that for every $n \geq n_0$, $x \in B_n$. It follows that for every $n \geq n_0$, $|B_n| < \epsilon_n$. Since $\epsilon_n \to 0$, setting

$$B = \{x : \ \exists t \in [a, b] \text{ s.t. } x + t\mathbb{A} \subseteq E\},$$

we get $|B| = 0$. The null set of "bad" translates B is contained in E (since we assumed that $0 \in \mathbb{A}$), thus removing it from E results in a set E', which still has measure $|E' \cap [m, m+1]| \geq p$ for all $m \in \mathbb{Z}$, but contains no affine copy of $\mathbb{A}$ with scale in $[a, b]$.

Step 2. *Completion of the proof of Theorem 1.1.*

Take a positive sequence $p'_n \in [0, 1)$, $n \in \mathbb{Z}$, such that

$$(3.2) \qquad \sum_{n \in \mathbb{Z}} (1 - p'_n) < 1 - p.$$

Consider the intervals $[a_n, b_n] = [2^{n-1}, 2^n]$, $n \in \mathbb{Z}$. Then, according to Step 1, for each p'_n, there exists a set E_n such that $|E_n \cap [m, m+1]| \geq p'_n$, for all $m \in \mathbb{Z}$, but for all $x \in \mathbb{R}$ and for all $t \in [a_n, b_n]$, the set $x + t\mathbb{A}$ is not contained in E_n.

Take

$$E = \bigcap_{n \in \mathbb{Z}} E_n.$$

Assume that for some $x \in \mathbb{R}$ and some $t > 0$, $x + t\mathbb{A} \subseteq E$. Then, $x + t\mathbb{A} \subseteq E_n$, for all $n \in \mathbb{Z}$. However, since there is $n_0 \in \mathbb{Z}$ such that $t \in [2^{n_0 - 1}, 2^{n_0}]$, the inclusion $x + t\mathbb{A} \subseteq E_{n_0}$ cannot be true. Thus, E does not contain any affine copy of $\mathbb{A}$ with positive scale. Finally, due to (3.2) we have $|[m, m+1] \setminus E| < 1 - p$, or $|E \cap [m, m+1]| \geq p$.

3.1. **Proof of Lemma 3.1.** Fix the scale $t \in [a, b]$ and let $0 \leq p < 1$. Consider the positive sequence given by

$$(3.3) \qquad p_N = 1 - \sqrt{\frac{\log\left(\frac{N}{10b}\right)}{A\left(\frac{N}{10b}\right)}}.$$

From (1.2) this implies $p_N \to 1^-$.

Partition $[-N, N]$ into unit intervals $[m, m + 1]$, $m = -N, -N + 1, ..., N - 1$. Divide each $[m, m + 1]$ further, into k_N equal subintervals

$$I_{i,m} = m + \left[\frac{i - 1}{k_N}, \frac{i}{k_N}\right], \quad i = 1, ..., k_N,$$

where

(3.4)
$$k_N = \left\lceil \frac{10}{a} \right\rceil \frac{N}{1 - p_N}.$$

Notice that $k_N / N \to +\infty$.

Construct a random set $E = E_N$ as follows: keep each $I_{i,m}$ in E independently of the other intervals and with probability p_N as in (3.3). Then, $\mathbb{P}(x \in E) = p_N$ for each $x \in [-N, N]$.
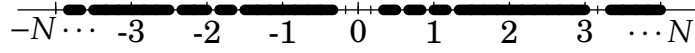

$$-N \cdots \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad \cdots N$$

FIGURE 3. The random set E.

Let $M_N(x, t)$ be the number of elements of $(x + t\mathbb{A}) \cap [-N, N]$ and observe that

(3.5)
$$M_N(x, t) \leq A(2N/a), \quad \text{for } x \in [-N, N].$$

For a given set $E \subseteq [-N, N]$, consider the set of "bad" translates

$$B = \left\{ x \in [-N, N] : \exists t \in [a, b] \text{ s.t. } (x + t\mathbb{A}) \cap [-N, N] \subseteq E \right.$$

$$\left. \text{and } M_N(x, t) \geq A\left(\frac{N}{10b}\right) \right\}.$$

We first deal with the measure of B. We have

$$\mathbb{E}|B| = \mathbb{E} \int_{-N}^{N} \mathbb{1}_B(x) dx$$

(3.6)
$$= \int_{-N}^{N} \mathbb{P}\left[\exists t \in [a, b] : (x + t\mathbb{A}) \cap [-N, N] \subseteq E \right.$$

(3.7)
$$\left. \text{and } M_N(x, t) \geq A\left(\frac{N}{10b}\right) \right] dx.$$

In what follows, we estimate from above the probability in (3.7), uniformly in $x \in [-N, N]$.

Fix $x \in [-N, N]$. To check whether there exists $t \in [a, b]$ such that $(x + t\mathbb{A}) \cap [-N, N] \subseteq E$, it is sufficient to check whether such a t exists in a finite set

(3.8)
$$S = S(x) = \{t_1, t_2, ..., t_u\} \subseteq [a, b].$$

Write $\alpha'_0 < \alpha'_1 < ... < \alpha'_{M_N(x,t)-1}$ for the elements of $(x + t\mathbb{A}) \cap [-N, N]$. Then, the set S consists exactly of those $t \in [a, b]$ for which some $\alpha'_j = x + ta_j$, $j = 0, ..., M_N(x, t) - 1$, is in the set $m + \left\{0, \frac{1}{k_N}, \frac{2}{k_N}, ..., \frac{k_N - 1}{k_N}, 1\right\}$, for some $m \in \{-N, -N + 1, ..., N - 1\}$. Each of the points $\alpha'_j = x + ta_j$ traverses, as t moves from a to b, and as long as the point α'_j remains in $[-N, N]$, an interval
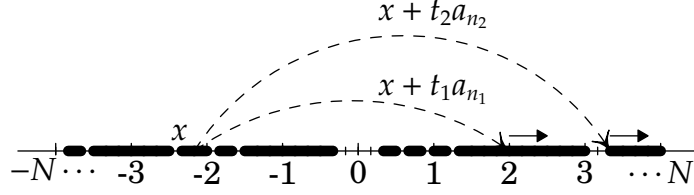
FIGURE 4. As x is held fixed and t grows the points $x + ta_n$ cross over interval endpoints creating events that need to be checked.

of length at most $2N$, therefore it meets at most $2Nk_N$ interval endpoints of the intervals $I_{i,m}$. Altogether, we have

$$(3.9) \qquad u \leq 2Nk_N \sup_{a \leq t \leq b} M_N(x,t) \leq c(a)N^2(1 - p_N)^{-1}A\left(\frac{2N}{a}\right),$$

where for the last inequality, we used (3.4) and (3.5).

Since $k_N \to +\infty$, we can take N large enough, say $N \geq N_0$, so that $k_N > 1/a$, for every $N \geq N_0$. Then, the length of each $I_{i,m}$ is small enough, $\leq a$, to ensure that, for each $t \in [a,b]$, the points α'_j, $j = 0, \dots, M_N - 1$, all belong to different intervals $I_{i,m}$. Therefore, for any fixed x and t,

$$\mathbb{P}\left[(x + t\mathbb{A}) \cap [-N,N] \subseteq E \text{ and } M_N(x,t) \geq A\left(\frac{N}{10b}\right)\right]$$

$$\leq \mathbb{P}\left[(x + t\mathbb{A}) \cap [-N,N] \subseteq E \mid M_N(x,t) \geq A\left(\frac{N}{10b}\right)\right]$$

$$(3.10) \qquad \leq p_N^{A\left(\frac{N}{10b}\right)}.$$

Thus, using the bound (3.9),

$$\mathbb{P}\left[\exists t \in S : (x + t\mathbb{A}) \cap [-N,N]) \subseteq E\right]$$

$$\leq c(a)N^2(1 - p_N)^{-1}A\left(\frac{2N}{a}\right)p_N^{A\left(\frac{N}{10b}\right)}.$$

Thus, (3.7) yields

$$\mathbb{E}|B| \leq 2c(a)N^3(1 - p_N)^{-1}A\left(\frac{2N}{a}\right)p_N^{A\left(\frac{N}{10b}\right)}.$$

We want to have

$$N^3(1 - p_N)^{-1}A\left(\frac{2N}{a}\right)p_N^{A\left(\frac{N}{10b}\right)} \to 0,$$

while $p_N \to 1^-$, as $N \to \infty$. Since $A(\cdot)$ grows at most linearly at infinity, it suffices to show that

$$(3.11) \qquad A\left(\frac{N}{10b}\right)\log p_N\left(4\frac{\log N}{A\left(\frac{N}{10b}\right)\log p_N} - \frac{\log(1 - p_N)}{A\left(\frac{N}{10b}\right)\log p_N} + 1\right)$$

$$\to -\infty.$$

To show (3.11), observe first that since $\lim_{x\to+\infty} x\log\left(1 - x^{-1/2}\right) = -\infty$, we have

$$(3.12) \qquad \frac{A\left(\frac{N}{10b}\right)\log p_N}{\log N} \to -\infty,$$

due to (3.3). Therefore, we also have $A\left(\frac{N}{10b}\right)\log p_N \to -\infty$. Finally, by (3.3) and (3.12) we get

$$\frac{\log(1 - p_N)}{A\left(\frac{N}{10b}\right)\log p_N} = -\frac{1}{2}\frac{\log A\left(\frac{N}{10b}\right)}{A\left(\frac{N}{10b}\right)\log p_N}\left\{1 - \frac{\log\log\frac{N}{10b}}{\log A\left(\frac{N}{10b}\right)}\right\} \to 0.$$

In other words, we have shown that for every $\epsilon > 0$, there is $N_1 \geq N_0$ such that for all $N \geq N_1$, $\mathbb{E}|B| < \epsilon/2$, which implies that

$$(3.13) \qquad \mathbb{P}(|B| \geq \epsilon) < 1/2, \quad \forall N \geq N_1.$$

We now turn to the measure of E in every unit interval with integer endpoints. Fix $m \in [-N, N]$. Let $X_1^m, X_2^m, ..., X_{k_N}^m$ be independent indicator random variables, with $X_i^m = 1$ if and only if $I_{i,m} \subseteq E$. Let $Y_i^m = 1 - X_i^m$ and denote by $X^m = \sum_{i=1}^{k_N} X_i^m$, $Y^m = \sum_{i=1}^{k_N} Y_i^m$ their sums. Then, $\mathbb{E}Y^m = (1 - p_N)k_N$. Notice also that the total measure kept in $[m, m+1] \cap E$ is equal to X^m/k_N.

For any $\delta > 0$ we define the "bad" events

$$A_m = \{|Y^m - \mathbb{E}Y^m| > \delta\,\mathbb{E}Y^m\}, \quad m = -N, -N + 1, ..., N - 1.$$

To control $\mathbb{P}[A_m]$, we use Chernoff's inequality, [AS16, Che52]: for all $\delta > 0$,

$$\mathbb{P}[A_m] \leq 2e^{-c_\delta \mathbb{E}Y^m},$$

where $c_\delta = \min\left\{(1 + \delta)\log(1 + \delta) - \delta\log\delta,\ \delta^2/2\right\}$. Take $\delta = 1/2$. It follows that

$$\mathbb{P}\left[|Y^m - (1 - p_N)k_N| > \frac{1}{2}(1 - p_N)k_N\right]$$
$$\leq 2\exp\left(-\frac{1}{2}(1 - p_N)k_N\right).$$

Thus, the probability that there is some $[m, m + 1] \subseteq [-N, N]$, such that A_m holds, is at most

$$4N\exp\left(-\frac{1}{2}(1 - p_N)k_N\right)$$

and the right hand side tends to zero as $N \to +\infty$, by our choice of k_N in (3.4). Thus, there is $N_2 \geq N_1$ such that

$$(3.14) \qquad \mathbb{P}\left[\exists m \in \{-N, -N + 1, ..., N - 1\} : A_m \text{ holds}\right] < \frac{1}{2},$$

for all $N \geq N_2$. Then, (3.13) and (3.14) imply the existence of a set $E \subseteq \mathbb{R}$ such that, on the one hand, it satisfies

$$|B| < \epsilon$$

and on the other hand,

$$X^m - p_N k_N \geq -\frac{1}{2}(1 - p_N)k_N,$$

for all $m = -N, -N + 1, ..., N - 1$, for all $N \geq N_2$. Thus the measure of E in each unit interval $[m, m + 1]$, is at least $p_N - \frac{1}{2}(1 - p_N) \to 1$, as $p_N \to 1^-$. In other words, for all $0 \leq p < 1$, there is $N_3 \geq N_2$ such that for all $N \geq N_3$, we have $|E \cap [m, m + 1]| \geq p$. The proof of Lemma 3.1 is now complete.

Remark 3.1. Let us indicate here why the proof of Theorem 1.1 just completed also applies to Theorem 1.2 without any essential changes. First of all, the implication from Lemma 3.1 to Theorem 1.3 (finite to infinite) remains true almost verbatim. So it suffices to ensure that Lemma 3.1 is true in this case. The main ingredients of the proof of Lemma 3.1 are the following. Having fixed x and varying t we have to make sure that the following conditions hold.

C.1 All points of the the (x, t)-copy of the set remain well separated, so that independence applies and we can multiply the probabilities that they belong to our random set. This is ensured by (1.4).

C.2 The number of points in the (x, t)-copy of the set in the interval $[-N, N]$ has to be large as this is the exponent in the upper bound (3.10). Condition (1.5) guarantees this.

C.3 The number of events that need to be checked so that we are certain that for all t no (x, t)-copy is contained in our random set is small. This is the number u in (3.8). What we are doing in the proof is to count how many times each of the points of our set (as x is held fixed and t increases from a to b) crosses over an interval boundary. Since the $\phi(n, t)$ are assumed increasing in t this remains as before.

It should be clear that the conditions imposed on the scaling functions $\phi(n, t)$ in Theorem 1.2 are far from optimal. They are rather indicative of what can be accomplished with the method and it is clear that the method could work under different sorts of conditions.

4. THE PROBLEM IN HIGHER DIMENSION

We will derive Theorem 1.3 as a consequence of the more finitary theorem below.

Theorem 4.1. *Let $d_1, d \geq 1$, $\beta, \zeta > 0$, $p \in (0, 1)$. Let also $\alpha(N)$ be a function satisfying*
$$\frac{\alpha(N)}{\log N} \to +\infty.$$

Then if N is sufficiently large and $P \subseteq \mathbb{R}^{d_1}$ is a point set with at most N^ζ points there is a set $E_N \subseteq [-N, N]^d$ such that

(1) $\left| E_N \cap \left(m + [0, 1]^d \right) \right| \geq p$ *for all $m = (m_1, \ldots, m_d) \in \mathbb{Z}^d$, with $-N \leq m_j < N$,*

(2) For any linear map $T : \mathbb{R}^{d_1} \to \mathbb{R}^d$ if

$$(4.1) \qquad\qquad T(P) \cap [-N, N]^d$$

contains at least $\alpha(N)$ points with separation$\geq N^{-\beta}$ then

$$(4.2) \qquad\qquad \left(T(P) \cap [-N, N]^d \right) \subsetneq E_N.$$

Proof. Let $\gamma > \beta$ and split the cube $[-N, N]^d$ with a $N^{-\gamma} \times \cdots \times N^{-\gamma}$-spaced grid of $O(dN^{1+\gamma})$ hyperplanes perpendicular to the d coordinate axes. Define the random set E to contain

each of the $N^{-\gamma} \times \cdots \times N^{-\gamma}$-sized cubes independently with probability $p' \in (p, 1)$. We show that with positive probability one can take $E_N = E$.

The first property of E is a simple consequence of Chernoff bounds and we can assume it holds with probability $> \frac{1}{2}$ working as in the proof of Theorem 1.1.

Let $T = (T_{i,j})$ be a linear map $\mathbb{R}^{d_1} \to \mathbb{R}^d$. This depends on $d \cdot d_1$ real variables $T_{i,j}$, so we view T as an element of $\mathbb{R}^{d \cdot d_1}$. Instead of checking condition (2) for all $T \in \mathbb{R}^{d \cdot d_1}$ we first show that there is a small number (polynomial in N) of T's that need to be checked.

Indeed, the set of $N^{-\gamma} \times \cdots \times N^{-\gamma}$-sized cubes that contain $T(P)$ does not change when T varies except when one or more of the points in $T(P)$ cross a dividing hyperplane of those that subdivide $[-N, N]^d$. Let H be one of those $O(dN^{1+\gamma})$ hyperplanes and fix an arbitrary point $h \in H$. Let also u be a unit vector orthogonal to H. For a point $x \in \mathbb{R}^d$ to belong to H it must satisfy the linear equation

$$E(H, x): \quad u \cdot x = u \cdot h.$$

Let $q \in P$. For the point $T(q)$ to belong to H we must have

$$E(H, T(q)): \quad u \cdot T(q) = u \cdot h,$$

which is a linear equation in $T \in \mathbb{R}^{d \cdot d_1}$. Taking all such equations in T, over all dividing hyperplanes H and all $q \in P$ we obtain a subdivision of $\mathbb{R}^{d \cdot d_1}$ by

$$n = O(d \cdot N^{1+\gamma} \cdot |P|)$$

hyperplanes. These n hyperplanes subdivide $\mathbb{R}^{d \cdot d_1}$ into $m = O(n^{d \cdot d_1})$ connected regions (this is easily proved by induction on the dimension, or see [Buc43]). For any two points T_1, T_2 in the same region condition (4.2) is either true for both or false for both since we can move continuously from T_1 to T_2 without leaving the region and, therefore, without any of the point $T(q)$ touching any of the dividing hyperplanes H.

It suffices therefore to check condition (4.2) for one point per region. Let us call these points $T_1, \ldots, T_m$. To guarantee that (4.2) holds for all T it is enough for it to be true for all $T_j, j = 1, 2, \ldots, m$. Define the bad events

$$B_j = \bigcap_{q \in P} \left\{ T_j(q) \in E \right\}.$$

We need to ensure that none of the B_j holds, but we only need to check those B_j for which there is a T in the cell of T_j for which (4.1) holds. For such a j the number of different $N^{-\gamma} \times \cdots \times N^{-\gamma}$-sized cubes touched by $T_j(P)$ is the same as the number touched by $T(P)$ which is at least $\alpha(N)$ so

$$\mathbb{P}\left[B_j \right] \le p'^{\alpha(N)},$$

and it is therefore enough to make sure that

$$n^{d \cdot d_1} p'^{\alpha(N)} = O\left(N^{\zeta \cdot d \cdot d_1} N^{(1+\gamma)d \cdot d_1} p'^{\alpha(N)} \right)$$

can be made arbitrarily small by choosing N large. This is clearly possible since the term $p'^{\alpha(N)}$ decays faster than any power of N. $\qquad \square$
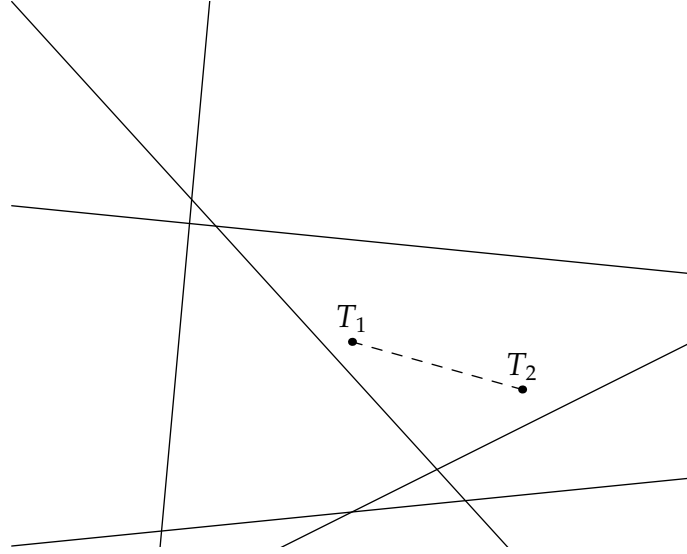
FIGURE 5. The regions defined in T-space by the equations $E(H, T(q))$ for all H, q. Only one of the transformations T_1, T_2 needs to be checked.

Proof of Theorem 1.3. Let $p_n \in (0, 1)$ be such that

$$(4.3) \qquad \sum_{n=1}^{\infty} (1 - p_n) < 1 - p.$$

Apply Theorem 4.1 successively for $N = n$, p_n, $\zeta = b$, $\alpha(N) = \alpha(R)$, $\beta = f$ and the set $P = \mathbb{A} \cap [-n, n]^{d_1}$ to obtain sets $E_n \subseteq [-n, n]^d$. Define

$$E = \bigcap_{n=1}^{\infty} \left(E_n \cup (\mathbb{R}^d \setminus [-n, n]^d) \right).$$

It is easy to see because of (4.3) that for any $m \in \mathbb{Z}^d$ we have $\left| E \cap m + [0, 1]^d \right| \geq p$. Let $T : \mathbb{R}^{d_1} \to \mathbb{R}^d$ and let R be such that $T(\mathbb{A}) \cap B_R(0)$ contains $\alpha(R)$ points which are R^{-f} separated. Let $n = \lceil R \rceil$. It follows from Theorem 4.1 that $T(\mathbb{A}) \cap [-n, n]^d$ is not contained in $E_n \cup (\mathbb{R}^d \setminus [-n, n]^d)$ and therefore not contained in E, as we had to show. $\square$

REFERENCES

[AS16] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2016. 6, 12

[BGK+22] Alex Burgin, Samuel Goldberg, Tamás Keleti, Connor MacMahon, and Xianzhi Wang. Large sets avoiding infinite arithmetic/geometric progressions. *arXiv preprint arXiv:2210.09284*, 2022. 5

[BKM23] Laurestine Bradford, Hannah Kohut, and Yuveshen Mooroogen. Large subsets of euclidean space avoiding infinite arithmetic progressions. *Proceedings of the American Mathematical Society*, 151(08):3535–3545, 2023. 2, 3, 5

[Bou86] Jean Bourgain. A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$. *Israel Journal of Mathematics*, 54(3):307–316, 1986. 1

[Buc43] Robert Creighton Buck. Partition of space. *The American Mathematical Monthly*, 50(9):541–544, 1943. 14

[CFG12] Hallard T Croft, Kenneth Falconer, and Richard K Guy. *Unsolved problems in geometry*, volume 2. Springer Science & Business Media, 2012. 2

[Che52] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, pages 493–507, 1952. 6, 12

[Chl15] Miroslav Chlebik. On the Erdős similarity problem. *arXiv preprint arXiv:1512.05607*, 2015. 2

[CLP22] Angel Cruz, Chun-Kit Lai, and Malabika Pramanik. Large sets avoiding affine copies of infinite sequences. *arXiv preprint arXiv:2204.12720*, 2022. 3

[DPZ21] Jacob Denson, Malabika Pramanik, and Joshua Zahl. Large sets avoiding rough patterns. In *Harmonic analysis and applications*, pages 59–75. Springer, 2021. 3

[Erd15] Paul Erdös. My Scottish Book problems. In *The Scottish Book*, pages 27–33. Springer, 2015. 2

[Fal84] Kenneth J Falconer. On a problem of Erdős on sequences and measurable sets. *Proceedings of the American Mathematical Society*, 90(1):77–78, 1984. 2

[FKW90] Hillel Fürstenberg, Yitzchak Katznelson, and Benjamin Weiss. Ergodic theory and configurations in sets of positive density. In *Mathematics of Ramsey theory*, pages 184–198. Springer, 1990. 1

[FM86] Kenneth J Falconer and John M Marstrand. Plane sets with positive density at infinity contain all large distances. *Bulletin of the London Mathematical Society*, 18(5):471–474, 1986. 1

[FP18] Robert Fraser and Malabika Pramanik. Large sets avoiding patterns. *Analysis & PDE*, 11(5):1083–1111, 2018. 3

[GLW23] John Gallagher, Chun-Kit Lai, and Eric Weber. On a topological erdős similarity problem. *Bulletin of the London Mathematical Society*, 2023. 2

[HL98] Paul Humke and Miklós Laczkovich. A visit to the Erdős problem. *Proceedings of the American Mathematical Society*, 126(3):819–822, 1998. 2

[Kel08] Tamás Keleti. Construction of one-dimensional subsets of the reals not containing similar copies of given patterns. *Analysis & PDE*, 1(1):29–33, 2008. 3

[Kol97] Mihail N Kolountzakis. Infinite patterns that can be avoided by measure. *Bulletin of the London Mathematical Society*, 29(4):415–424, 1997. 3, 6

[Kol04] Mihail N Kolountzakis. Distance sets corresponding to convex bodies. *Geometric & Functional Analysis GAFA*, 14(4):734–744, 2004. 1

[Kom83] Péter Komjáth. Large sets not containing images of a given sequence. *Canadian Mathematical Bulletin*, 26(1):41–43, 1983. 2, 6

[Mag11] Péter Maga. Full dimensional sets without given patterns. *Real Analysis Exchange*, 36(1):79–90, 2011. 3

[Mát17] András Máthé. Sets of large dimension not containing polynomial configurations. *Advances in Mathematics*, 316:691–709, 2017. 3

[Shm17] Pablo Shmerkin. Salem sets with no arithmetic progressions. *International Mathematics Research Notices*, 2017(7):1929–1941, 2017. 3

[Sze75] Endre Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith*, 27(299-345):21, 1975. 1

[Szé83] László A Székely. *Remarks on the chromatic number of geometric graphs*, pages 312–315. Teubner, Leipzig, 1983. 1

[Yav21] Alexia Yavicoli. Large sets avoiding linear patterns. *Proceedings of the American Mathematical Society*, 149(10):4057–4066, 2021. 3

Department of Mathematics and Applied Mathematics, University of Crete, Voutes Campus, 70013 Heraklion, Crete, Greece.

Email address: kolount@gmail.com

Department of Mathematics and Applied Mathematics, University of Crete, Voutes Campus, 70013 Heraklion, Crete, Greece.

Email address: papageoeffie@gmail.com