Ergod. Th. & Dynam. Sys., (2025), **45**, 1370–1402

doi:10.1017/etds.2024.73

Intersective sets for sparse sets of integers

PIERRE-YVES BIENVENU†, JOHN T. GRIESMER‡, ANH N. LE§ and THÁI HOÀNG LÊ¶

† *Institute of Discrete Mathematics and Geometry, TU Wien, Wiedner Hauptstr. 8–10, A-1040 Wien, Austria (e-mail: pierre.bienvenu@tuwien.ac.at)*

‡ *Department of Applied Mathematics and Statistics, Colorado School of Mines, 1005 14th Street, Golden, CO 80401, USA (e-mail: jtgriesmer@gmail.com)*

§ *Department of Mathematics, University of Denver, 2390 S. York St, Denver, CO 80210, USA (e-mail: anh.n.le@du.edu)*

¶ *Department of Mathematics, University of Mississippi, University, MS 38677, USA (e-mail: leth@olemiss.edu)*

(*Received* 7 *March* 2024 *and accepted in revised form* 4 *September* 2024)

Abstract. For $E \subset \mathbb{N}$, a subset $R \subset \mathbb{N}$ is *E-intersective* if for every $A \subset E$ having positive relative density, $R \cap (A - A) \neq \varnothing$. We say that R is *chromatically E-intersective* if for every finite partition $E = \bigcup_{i=1}^{k} E_i$, there exists i such that $R \cap (E_i - E_i) \neq \varnothing$. When $E = \mathbb{N}$, we recover the usual notions of intersectivity and chromatic intersectivity. We investigate to what extent the known intersectivity results hold in the relative setting when $E = \mathbb{P}$, the set of primes, or other sparse subsets of $\mathbb{N}$. Among other things, we prove the following: (1) the set of shifted Chen primes $\mathbb{P}_{\text{Chen}} + 1$ is both intersective and $\mathbb{P}$-intersective; (2) there exists an intersective set that is not $\mathbb{P}$-intersective; (3) every $\mathbb{P}$-intersective set is intersective; (4) there exists a chromatically $\mathbb{P}$-intersective set which is not intersective (and therefore not $\mathbb{P}$-intersective).

Key words: recurrence, intersectivity, difference set
2020 Mathematics Subject Classification: 37B20, 11B05 (Primary); 37A44, 11B30, 11B13 (Secondary)

Contents

1. *Introduction*

1.1. *Combinatorial theorems in dense sets of integers and transference to sparse sets.*

Let $\mathbb{N}$ be the set of positive integers $\{1, 2, 3, \ldots\}$ and for $N \in \mathbb{N}$, define $[N] = \{1, 2, \ldots, N\}$. If $A, E \subset \mathbb{N}$, the *upper density of A relative to E* is defined as

$$\overline{d}_E(A) := \limsup_{N\to\infty} \frac{|A \cap E \cap [N]|}{|E \cap [N]|}.$$

Similarly, the *lower density of A relative to E* is

$$\underline{d}_E(A) := \liminf_{N\to\infty} \frac{|A \cap E \cap [N]|}{|E \cap [N]|}.$$

When the ambient set E is unambiguous from context, we simply say 'the upper relative density' and the 'the lower relative density' of A without mentioning E.

Note that $\underline{d}_E(A) \le \overline{d}_E(A)$. If equality holds, we denote by $d_E(A)$ the common value and call it the *density of A relative to E*. If $E = \mathbb{N}$, we omit E and simply write $\overline{d}(A), \underline{d}(A), d(A)$ and call them the upper density, lower density, and density of A, respectively. We say a set A of integers is *dense* if $\overline{d}(A) > 0$ and *sparse* if $\overline{d}(A) = 0$. More generally, we say that A is *dense relative to E* if $\overline{d}_E(A) > 0$, and that A is *sparse relative to E* otherwise.

Dense subsets of $\mathbb{N}$ are known to inherit many combinatorial properties of $\mathbb{N}$. For example, Roth [44] proved that every dense set contains infinitely many three-term arithmetic progressions. Szemerédi [48] showed such a set contains arbitrarily long arithmetic progressions. That being said, these properties are not exclusive to dense sets. For instance, despite being a sparse set, the set of primes $\mathbb{P}$ also enjoys the same properties. In [20], Green devised a transference principle to deduce from Roth's theorem that every set which is dense relative to $\mathbb{P}$ contains three-term arithmetic progressions. This transference principle was a precursor to another one which enabled Green and Tao [21] to prove that a dense subset of primes contains arbitrarily long arithmetic progressions. Since then, many variants of the transference principle were devised to prove combinatorial theorems in sparse sets of integers such as the squares [10], the sums of two squares [37], and various relatively sparse subsets of the primes.

Against this backdrop, the goal of our paper is to investigate whether other combinatorial properties of $\mathbb{N}$ may be transferred to $\mathbb{P}$ and other sparse sets. The properties we will study are the so-called intersective properties, which we now define.

1.2. *Intersectivity.*

Given $A, B \subset \mathbb{Z}$, we define their *sumset* and *difference set* to be $A + B := \{a + b : a \in A, b \in B\}$ and $A - B := \{a - b : a \in A, b \in B\}$, respectively.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Definition 1.1. For an infinite set $E \subset \mathbb{N}$, a subset $R \subset \mathbb{N}$ is said to be *E-intersective* if for every $A \subset E$ with $\overline{d}_E(A) > 0$, we have $R \cap (A - A) \neq \varnothing$.

Thus, R is E-intersective if every A satisfying $\overline{d}_E(A) > 0$ contains two distinct elements differing by an element of R. We will refer to a $\mathbb{P}$-intersective set as *prime intersective*, and we will refer to an $\mathbb{N}$-intersective set as *intersective*.

An intersective set is also called a *set of recurrence*. This terminology is motivated by the following connection to dynamical systems: a set $R \subset \mathbb{N}$ is intersective if and only if for every measure-preserving system $(X, \mathcal{B}, \mu, T)$ and $B \in \mathcal{B}$ satisfying $\mu(B) > 0$, there exists $r \in R$ such that

$$\mu(B \cap T^{-r} B) > 0. \tag{1}$$

(A measure-preserving system is a quadruple $(X, \mathcal{B}, \mu, T)$, where $(X, \mathcal{B}, \mu)$ is a probability space and $T : X \to X$ is a $\mathcal{B}$-measurable map satisfying $\mu(T^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}$.) One direction of the equivalence is provided by Furstenberg's correspondence principle [15, Theorem 1.1]; see [5, Théorème 1] for a proof of the equivalence.

In the late 1970s, Sárközy [46] and Furstenberg [15, 17] proved independently that $\{n^2 : n \in \mathbb{N}\}$ is intersective. Furstenberg used ergodic theory, while Sárközy's proof is inspired by the original proof of Roth's theorem [44] that employs the circle method. Sárközy [47] proved subsequently that $\{n^2 - 1 : n > 1\}$, $\mathbb{P} - 1$ and $\mathbb{P} + 1$ are also intersective, confirming conjectures of Erdős. Kamae and Mendès France's criterion [29] provides a generalization of Sárközy's results to arbitrary polynomials of integer coefficients, which we now state.

THEOREM 1.2. [29] *Suppose $Q \in \mathbb{Z}[x]$ has positive leading coefficient.*

(1) *The set $Q(\mathbb{N}) \cap \mathbb{N}$ is intersective if and only if for every $m \in \mathbb{N}$, there is $n \in \mathbb{Z}$ such that $Q(n) \equiv 0 (\mathrm{mod}\ m)$. If this condition holds, we say Q is an intersective polynomial.*

(2) *The set $Q(\mathbb{P}) \cap \mathbb{N}$ is intersective if and only if for every $m \in \mathbb{N}$, there is $n \in \mathbb{Z}$ such that $\gcd(m, n) = 1$ and $Q(n) \equiv 0 (\mathrm{mod}\ m)$.*

All intersective sets mentioned above are also prime intersective. Indeed, using Green's transference principle [20], the fourth author [33] proved that if Q is an intersective polynomial, then $Q(\mathbb{N})$ is prime intersective. This result is now superseded by Rice [43], who proved that if $Q(\mathbb{N})$ is not E-intersective, then $|E \cap [N]| \leq c_1(N/(\log N)^{c_2 \log\log\log\log N})$, where $c_1 > 0$ is a constant depending on Q and $c_2 > 0$ depends only on the degree of Q. Li and Pan [36] showed that if $Q(1) = 0$, then $Q(\mathbb{P})$ is also prime intersective. For the general case (when Q satisfies item (2)), $Q(\mathbb{P})$ is proved to be prime intersective by Rice [42].

1.3. *Shifted Chen primes.* Our first result contributes to the previously mentioned pool of sets that are both intersective and prime intersective. The sets that we study come from almost twin primes. A prime p is a *Chen prime* if $p + 2$ is a product of at most two primes. Chen [11] proved that the set of Chen primes, denoted by $\mathbb{P}_{\mathrm{Chen}}$, is infinite. Another type of almost twin primes is *bounded gap primes*. For a fixed natural number h, let $\mathbb{P}_{\mathrm{bdd},h}$ be the

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

set of primes p such that $p + h$ is a prime. The celebrated theorem of Zhang [51] shows that there exists $h \in \mathbb{N}$ such that $\mathbb{P}_{\mathrm{bdd},h}$ is infinite.

As previously mentioned, the results of Sárközy [47] and Li–Pan [36] say that $\mathbb{P} - 1$ and $\mathbb{P} + 1$ are both intersective and prime intersective. Therefore, a natural question is whether $\mathbb{P}_{\mathrm{Chen}} - 1$, $\mathbb{P}_{\mathrm{Chen}} + 1$, $\mathbb{P}_{\mathrm{bdd},h} - 1$ and $\mathbb{P}_{\mathrm{bdd},h} + 1$ are intersective (and prime intersective) for some $h \in \mathbb{N}$.

An intersective set must contain a non-zero multiple of every natural number. If $p \in \mathbb{P}_{\mathrm{bdd},h}$, then $p + h \in \mathbb{P}$. Therefore, $\mathbb{P}_{\mathrm{bdd},h} - 1$ is a subset of $\mathbb{P} - h - 1$ which does not contain a non-zero multiple of $h + 1$. Thus, $\mathbb{P}_{\mathrm{bdd},h} - 1$ cannot be intersective. For a similar reason, $\mathbb{P}_{\mathrm{bdd},h} + 1$ is not intersective unless $h = 2$. This leads to a question: Is $\mathbb{P}_{\mathrm{bdd},2} + 1$ intersective? Even though this question is interesting, it is out of the scope of our investigation because a positive answer will imply that there are infinitely many twin primes. The matter is more tractable regarding Chen primes, and we are able to prove the following theorem.

THEOREM A. $\mathbb{P}_{\mathrm{Chen}} + 1$ *is both intersective and prime intersective.*

To prove Theorem A, we use a transference principle developed by the first author, Shao, and Teräväinen [7]. Due to a local obstruction, we cannot use the same method for $\mathbb{P}_{\mathrm{Chen}} - 1$ and so the question as to whether $\mathbb{P}_{\mathrm{Chen}} - 1$ is intersective is still open.

1.4. *Separating intersective sets and prime intersective sets.* The similarities in known examples of intersective sets and prime intersective sets raise the question of the existence of intersective sets which are not prime intersective. This question was also asked by the third author in his survey [34].

Question 1.3. [34, Problem 6] Does there exist an intersective set which is not prime intersective?

In this paper, we give a positive answer to this question in a rather strong way. To explain our results in detail, we first introduce some important classes of subsets of integers: thick sets and syndetic sets.

A set $S \subset \mathbb{Z}$ is *thick* if S contains arbitrarily long intervals of the form $\{m, m + 1, \ldots, m + n\}$. We say S is *syndetic* if $S - F \supset \mathbb{Z}$ for some finite set F; equivalently, S is syndetic if the gaps between consecutive elements of S are bounded. Note that these two classes of sets are dual to each other in the following sense: given a family X of subsets of $\mathbb{Z}$, its dual is $X^* = \{E \subset \mathbb{Z} : E \cap A \neq \varnothing \text{ for all } A \in X\}$. When X is upward closed, i.e. any superset of any member of X is again a member of X, it is easy to see that $X^{**} = X$. Therefore, a set is thick if, and only if, it intersects every syndetic set; and a set is syndetic if, and only if, it intersects every thick set.

A folklore result says that every thick subset of $\mathbb{N}$ is intersective (see [16]). Our next result says that there are thick sets which are not prime intersective, and in fact, a thick set can fail very badly to be prime intersective.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

THEOREM B. *There exists $A \subset \mathbb{P}$ such that $d_{\mathbb{P}}(A) = 1$ and $A - A$ is not syndetic. Consequently, $R = \mathbb{N} \setminus (A - A)$ is a thick set (and so an intersective set) but not a prime intersective set.*

This naturally raises the question of whether $\mathbb{P} - \mathbb{P}$ itself is syndetic; a theorem of Pintz [41] implies that it is indeed syndetic.

The main ingredient in the proof of Theorem B is a classical sieve-theoretic bound for the number $\pi_m(x)$ of primes p smaller than x such that $p + m$ is also a prime. In fact, the theorem applies not only to the primes but to a broad range of sparse sets that satisfy similar bounds, for example, the images of nonlinear polynomials with integer coefficients and the images of those polynomials evaluated at primes. (See Proposition 3.1.)

To state our next result, we need to introduce the notions of piecewise syndetic sets and thickly syndetic sets. A set $S \subset \mathbb{Z}$ is *piecewise syndetic* if $S = T \cap R$, where T is a thick set and R is syndetic. A set $S \subset \mathbb{Z}$ is *thickly syndetic* if S intersects every piecewise syndetic set. Equivalently, S is thickly syndetic if for every N, there is a syndetic set J such that $[N] + J \subset S$, i.e. S contains syndetically many copies of intervals of arbitrary length. In particular, a thickly syndetic set is both thick and syndetic. We say that S is *thickly syndetic in* $\mathbb{N}$ if $S = R \cap \mathbb{N}$ for some thickly syndetic set R.

We do not know if the conclusion of Theorem B can be upgraded to 'R is thickly syndetic in $\mathbb{N}$ but not prime intersective'. However, this upgrade is possible if we slightly weaken the hypothesis on the largeness of A.

THEOREM C. *For every $\epsilon > 0$, there is a set $A \subset \mathbb{P}$ of relative density at least $1 - \epsilon$ such that $\mathbb{P} - A$ is not piecewise syndetic. In particular, $R := \mathbb{Z} \setminus (\mathbb{P} - A)$ is thickly syndetic, while $R \cap \mathbb{N}$ is not prime intersective.*

While the proof of Theorem B uses a quantitative input from number theory regarding the number of bounded gap primes, the proof of Theorem C uses a softer approach based on dynamics.

At the cost of replacing 'relative density' with 'upper relative density', Theorem C can be extended from $\mathbb{P}$ to any set E whose closure in $b\mathbb{Z}$, the Bohr compactification of $\mathbb{Z}$, has Haar measure zero (see §4 for definition of $b\mathbb{Z}$). This applies to the image of an integer polynomial $\{P(n) : n \in \mathbb{N}\}$ where $\deg P \geq 2$, the set of sums of two squares $\{x^2 + y^2 : x, y \in \mathbb{N}\}$, and the set of integers is represented by a norm form, for example, $\{x^3 + 2y^3 + 4z^3 - 6xyz : x, y, z \in \mathbb{Z}\}$.

It is easy to see that if $\overline{d}(A) > 0$, then $A - A$ is syndetic. (Here is a simple proof of this fact: let k be the largest integer such that there are distinct integers $n_1 < \cdots < n_k$ for which $A + n_i$ are pairwise disjoint. Thus, for any integer n, the shift $A + n$ intersects one of the shifts $A + n_i$ and so $n - n_i \in A - A$. It follows that $A - A + \{n_1, \ldots, n_k\} \supset \mathbb{Z}$.) Now if $\underline{d}(E) > 0$ and $A \subset E$ satisfies $\overline{d}_E(A) > 0$, then $\overline{d}(A) > 0$ and so $A - A$ is syndetic. As a result, it is crucial in Theorems B and C that $\underline{d}(\mathbb{P}) = 0$. Given this fact, a natural question is whether having zero lower density is all that is needed to satisfy these two theorems. While we do not know the answer to this question for Theorem C (see Conjecture 7.5), we show in Proposition 3.2 that the ambient set merely having zero lower density does not suffice for Theorem B. In fact, we prove something stronger regarding upper Banach

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

density. For $E \subset \mathbb{Z}$, its *upper Banach density* is defined by

$$d^*(E) = \lim_{N\to\infty} \sup_{M\in\mathbb{N}} \frac{|E \cap [M, M+N)|}{N}.$$

Note that we always have $d^*(E) \geq \overline{d}(E) \geq \underline{d}(E)$. Proposition 3.2 below states that there exists a set E with $d^*(E) = 0$ such that if $\overline{d}_E(A) = 1$, then $A - A = \mathbb{Z}$ (in particular, $A - A$ is syndetic).

1.5. *Prime intersective sets must be intersective.* Theorems B and C say that there exists an intersective set that is not prime intersective. Interestingly, the converse is false: every prime intersective set must be intersective. Moreover, the next theorem shows that the same is true if one replaces $\mathbb{P}$ by any infinite subset of $\mathbb{N}$.

THEOREM D. *For any infinite $E \subset \mathbb{N}$, every E-intersective set is intersective.*

Theorem D follows from a more general result regarding sets of multiple recurrence (see Theorem 5.2). The idea is to use Furstenberg's correspondence principle [**15**] to recast the problem into a question about sets of recurrence. We then use Fatou's lemma to show that a set which is not a set of recurrence (for $\mathbb{N}$) cannot be a set of recurrence for E.

1.6. *Chromatic intersectivity versus density intersectivity.* Additive combinatorics is often concerned with the contrast between a structure arising from density and a structure arising from partitions. In our setting, this leads to the following definition.

Definition 1.4. Given $E \subset \mathbb{N}$, a set $R \subset \mathbb{N}$ is said to be *chromatically E-intersective* if for every finite partition $E = \bigcup_{i=1}^{k} E_i$, there exists i such that $R \cap (E_i - E_i) \neq \varnothing$.

Equivalently, R is chromatically E-intersective if for any finite coloring of E, there are distinct m, n of the same color such that $m - n \in R$.

If $E = \mathbb{N}$, we simply say that R is *chromatically intersective*. In dynamical systems language, R is chromatically intersective if and only if for any minimal topological dynamical system (X, T) and any non-empty open set $U \subset X$, there exists $n \in R$ such that $U \cap T^{-n}U \neq \varnothing$. (A *topological dynamical system* is a pair (X, T), where X is a compact Hausdorff space and $T : X \to X$ is a continuous map. Here, (X, T) is *minimal* if for every $x \in X$, the orbit $\{T^n x : n \in \mathbb{N}\}$ is dense in X.) Due to this characterization, a chromatically intersective set is also called a *set of topological recurrence* (in contrast to measurable recurrence as defined in equation (1)).

In any partition $\mathbb{N} = \bigcup_{i=1}^{\ell} A_i$, one of the A_i has positive upper density, and so an intersective set is always chromatically intersective. Therefore, $\mathbb{N}$, $\{n^2 : n \in \mathbb{N}\}$, $\{n^2 - 1 : n \in \mathbb{N}\}$, $\mathbb{P} - 1$, and $\mathbb{P} + 1$ are each chromatically intersective. However, Kříž [**30**] proved that there exists a chromatically intersective set which is not intersective.

For a similar reason, for any $E \subset \mathbb{N}$, an E-intersective set is chromatically E-intersective. From Kříž's example, it is natural to ask whether there exists a chromatically E-intersective set which is not E-intersective. Our next result confirms this is the case. In fact, Theorem E below strengthens Kříž's example by showing that for

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

any subset $E \subset \mathbb{N}$, there exists a chromatically E-intersective set which is not intersective. Calling this set R, Theorem D implies that R is not E-intersective and as a result, R is an example of a chromatically E-intersective set which is not E-intersective.

THEOREM E. *For any infinite $E \subset \mathbb{N}$, there exists a chromatically E-intersective set which is not intersective (and thus not E-intersective).*

The proof of Theorem E is based on a recent refinement of the third author on Kříž's theorem [30]. Using Theorem E, we can find a set R that separates chromatic intersectivity and density intersectivity of any infinite E. Nevertheless, it is hard to extract from the construction any combinatorial properties of R. In contrast, in the special case when $E = \mathbb{P}$, we can take R to be thickly syndetic. Indeed, Theorem C says that there is a thickly syndetic set R which is not prime intersective; however, our next theorem states that every thick set is chromatically prime intersective. Therefore, the thickly syndetic set R found in Theorem C is chromatically prime intersective but not prime intersective.

THEOREM F. *For any finite partition $\mathbb{P} = \bigcup_{i=1}^{k} E_i$, the union $\bigcup_{i=1}^{k}(E_i - E_i)$ is syndetic. Equivalently, every thick set is chromatically prime intersective.*

Piecewise syndeticity is a partition regular property. That is, if $k \in \mathbb{N}$ and $A_1, \ldots, A_k \subset \mathbb{N}$ have the property that $\bigcup_{i=1}^{k} A_i$ is piecewise syndetic, then one of the A_i is piecewise syndetic. A proof of this standard fact can be found in [9, Lemma 1] or [17, Theorem 1.24]. Thus, for any partition $\mathbb{P} = \bigcup_{i=1}^{k} E_i$, there exists $i \in \{1, \ldots, k\}$ such that $E_i - E_i$ is piecewise syndetic. It remains unclear whether there is i such that $E_i - E_i$ is syndetic. The proof of Theorem F relies on Maynard and Tao's famous results [38] on the Hardy–Littlewood prime tuple conjecture. In fact, Theorem F follows from the more general Theorem 6.3 below which says that for any set E which satisfies some 'finite-tuple' property, every thick set is chromatically E-intersective. This applies to a broad range of sets, such as random sets, various subsets of the primes, and the set of sums of two squares.

A crucial property of $\mathbb{P}$ used in Theorem F is that there are infinitely many bounded gap primes. The situation is different for an ambient set whose gaps tend to infinity.

THEOREM G. *If $E = \{n_1 < n_2 < \cdots\} \subset \mathbb{N}$ satisfies $\lim_{i\to\infty}(n_{i+1} - n_i) = \infty$, then there is a partition $E = E_1 \cup E_2$ such that $(E_1 - E_1) \cup (E_2 - E_2)$ is not syndetic. In particular, there exists a thick set which is not chromatically E-intersective.*

Theorem G applies, for instance, to $E = \{P(n) : n \in \mathbb{N}\}$, where P is a polynomial of degree at least 2 and $E = \{\lfloor n^{1+\epsilon} \rfloor : n \in \mathbb{N}\}$ for any $\epsilon > 0$.

1.7. *Diagrams.* Figure 1 is a diagram of relations among thick sets, intersective sets, prime intersective sets and their chromatic counterparts. The implications we prove are marked with thick arrows; strike-out arrows mean 'does not imply'.

Figure 2 displays the relations between intersective sets and E-intersective sets for arbitrary subset $E \subset \mathbb{N}$. Recall that 'intersective' is the same as '$\mathbb{N}$-intersective'.

1.8. *Outline of the paper.* Theorems A, B, C, and D are proved in §§2, 3, 4, and 5, respectively. We prove Theorem E in §6 by verifying that the examples constructed in [24,

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

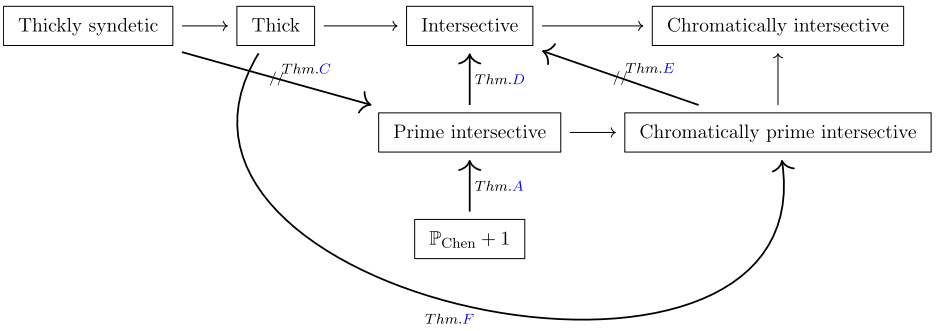

FIGURE 1. Relations among thick sets, (chromatically) intersective sets, and (chromatically) prime intersective sets.

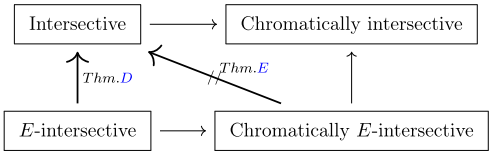

FIGURE 2. Relations between (chromatically) intersective sets and (chromatically) E-intersective sets for arbitrary $E \subset \mathbb{N}$.

Theorem 1.2] are actually chromatically E-intersective. The rest of §6 is devoted to the proofs of Theorems F and G. Section 7 collects questions suggested by our study. Lastly, Appendix A supplements Theorem C by providing a list of subsets of $\mathbb{N}$ whose closures have zero Haar measure in $b\mathbb{Z}$.

2. *Shifted Chen primes*

2.1. *Preliminaries.* The goal of §2 is to prove Theorem A. First, we recall some terminology and results from a recent paper by Bienvenu, Shao, and Teräväinen [7].

Let $\mathbb{P}'_{\text{Chen}}$ denote the set

$$\{p \in \mathbb{P} : p+2 \text{ is a prime or a product of two primes } p_1 p_2 \text{ with } p_1, p_2 \geq p^{1/10}\}.$$

Note that $\mathbb{P}'_{\text{Chen}}$ is a subset of $\mathbb{P}_{\text{Chen}} = \{p \in \mathbb{P} : p+2 \text{ is a prime or a product of two primes}\}$. We define the following weighted indicator function of $\mathbb{P}'_{\text{Chen}}$:

$$\theta(n) := (\log n)^2 1_{\mathbb{P}'_{\text{Chen}}}(n) 1_{p|n(n+2) \Longrightarrow p \geq n^{1/10}}.$$

Chen's theorem [11] says

$$\sum_{n \in [N]} \theta(n) \gg N.$$

Here, for two functions $f, g : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, $f(N) \gg g(N)$ means that there exists a positive constant C such that $f(N) \geq Cg(N)$ for all sufficiently large N. Alternatively, we also use the notation $g(N) \ll f(N)$ for the same meaning.

For any $W \in \mathbb{N}$, the set of Chen primes is heavily biased toward the congruence classes that are coprime to W. To remove this local obstruction, we use a standard procedure called the 'W-trick'. For any $w \in \mathbb{N}$, let $W = W(w) = \prod_{p \leq w, p \in \mathbb{P}} p$. For $b \in \mathbb{Z}$ such that

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

$(b, W) = 1$, define

$$\theta_{W,b}(n) = \left(\frac{\phi(W)}{W}\right)^2 \theta(Wn + b). \tag{2}$$

Since our focus will be on $\mathbb{P}_{\text{Chen}} + 1$, we will consider only $b = -1$.

For a function $f : S \to \mathbb{C}$ on a finite set S, let $\mathbb{E}_{n\in S} f(n)$ denote the average

$$\frac{1}{|S|} \sum_{n\in S} f(n).$$

For a function $f : \mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ and $k \in \mathbb{N}$, the *Gowers U^k-norm* of f is defined by

$$\|f\|_{U^k(\mathbb{Z}_N)} = \left(\mathbb{E}_{x\in\mathbb{Z}_N} \mathbb{E}_{\underline{h}\in(\mathbb{Z}_N)^k} \prod_{\underline{\omega}\in\{0,1\}^k} \mathcal{C}^{|\underline{\omega}|} f(x + \underline{\omega} \cdot \underline{h})\right)^{1/2^k},$$

where $\mathcal{C}$ means taking complex conjugates, $|\underline{\omega}| = \sum_{i\in[k]} \omega_i$, and $\underline{\omega} \cdot \underline{h}$ is the dot product of $\underline{w}$ and $\underline{h}$.

We have the following proposition from [7].

PROPOSITION 2.1. *There exists a constant $\delta > 0$ such that the following holds: for any $k \in \mathbb{N}$, $\epsilon > 0$, sufficiently large $N = N(k, \epsilon)$ and $w = w(k, \epsilon)$, letting $W = \prod_{p<w, p\in\mathbb{P}} p$, and $\theta_{W,-1}$ be defined as in (2), we can decompose*

$$\theta_{W,-1} = f_1 + f_2 \text{ on } [N]$$

such that $\delta \le f_1 \le 2$ pointwise and $\|f_2\|_{U^{k+1}(\mathbb{Z}_N)} \le \epsilon$.

Proof. This proposition follows from the transference principle stated in [7, Proposition 3.9]. Propositions 4.2 and 5.2 in the same paper show that for large enough w, depending only on the degree $k + 1$ and the quality ϵ of the required uniformity, the function $f = \theta_{W,-1}$ satisfies the conditions in the hypothesis of [7, Proposition 3.9], and so proves Theorem 2.1. □

2.2. *Intersectivity of shifted Chen primes.* The fact that $\mathbb{P}_{\text{Chen}} + 1$ is intersective follows from a much more general result below, which says that $\mathbb{P}_{\text{Chen}} + 1$ is a 'set of multiple polynomial recurrence for measure-preserving systems of commuting transformations'. In other words, by Furstenberg's correspondence principle [15, Theorem 1.1], the intersectivity asserted in Theorem A corresponds to the case $\ell = m = 1$ and $q_{1,1}(n) = n$ of the following proposition.

PROPOSITION 2.2. *Let $\ell \in \mathbb{N}$, $(X, \mathcal{B}, \mu)$ be a probability space and let $T_1, \ldots, T_\ell : X \to X$ be commuting measure-preserving transformations (i.e. $\mu(T_i^{-1}B) = \mu(B)$ for all $i \in [\ell]$ and $B \in \mathcal{B}$). Let $m \in \mathbb{N}$ and $q_{i,j} : \mathbb{Z} \to \mathbb{Z}$ be polynomials with $q_{i,j}(0) = 0$ for $i \in [\ell]$ and $j \in [m]$. For any $A \in \mathcal{B}$ with $\mu(A) > 0$, the set*

$$\left\{n \in \mathbb{N} : \mu\left(A \cap \left(\prod_{i=1}^{\ell} T_i^{-q_{i,1}(n)} A\right) \cap \cdots \cap \left(\prod_{i=1}^{\ell} T_i^{-q_{i,m}(n)} A\right)\right) > 0\right\}$$

has non-empty intersection with $\mathbb{P}_{\text{Chen}} + 1$.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Proof. Define

$$\alpha(n) = \mu\Big(A \cap \Big(\prod_{i=1}^{\ell} T_i^{-q_{i,1}(n)} A\Big) \cap \cdots \cap \Big(\prod_{i=1}^{\ell} T_i^{-q_{i,m}(n)} A\Big)\Big).$$

By [14, Corollary 4.2] (see also [3, Theorem 3.2]), there is a constant $c > 0$ depending only on $\mu(A)$, m and the polynomials $q_{i,j}$ such that for all W in $\mathbb{N}$ and all sufficiently large N,

$$\mathbb{E}_{n\in[N]}\alpha(Wn) \geq c. \tag{3}$$

Let δ be the constant found in Theorem 2.1 and let $\epsilon > 0$ be very small compared with $c\delta$. By Theorem 2.1, for sufficiently large $N = N(k, \epsilon)$ and $w = w(k, \epsilon)$, we have

$$\mathbb{E}_{n\in[N]}\theta_{W,-1}(n)\alpha(Wn) = \mathbb{E}_{n\in[N]} f_1(n)\alpha(Wn) + \mathbb{E}_{n\in[N]} f_2(n)\alpha(Wn), \tag{4}$$

where f_1, f_2 satisfy the conclusion of Theorem 2.1 with respect to the aforementioned δ and ϵ.

Since f_1 is bounded below pointwise by δ, equation (3) implies that

$$\mathbb{E}_{n\in[N]} f_1(n)\alpha(Wn) \geq c\delta. \tag{5}$$

However, by [14, Lemma 3.5], there exists an integer k that depends only on the maximum degree of the polynomials $q_{i,j}$ and the integers ℓ, m such that

$$\mathbb{E}_{n\in[N]} f_2(n)\alpha(Wn) \ll \|f_2 \cdot 1_{[N]}\|_{U^k(\mathbb{Z}_{kN})} + o_{N\to\infty}(1) \ll \|f_2\|_{U^k(\mathbb{Z}_N)} + o_{N\to\infty}(1),$$

where the term $o_{N\to\infty}(1)$ is independent of W. Therefore, for sufficiently large N,

$$\mathbb{E}_{n\in[N]} f_2(n)\alpha(Wn) \leq \epsilon. \tag{6}$$

Combining equations (4), (5), and (6), we deduce that for sufficiently large N and w,

$$\mathbb{E}_{n\in[N]}\theta_{W,-1}(n)\alpha(Wn) \gg 1,$$

where the implicit constant depends only on $\mu(A)$, ℓ, m, and the $q_{i,j}$. Thus, there exists $n \in \mathbb{N}$ such that $\alpha(Wn) > 0$ and $\theta_{W,-1}(n) > 0$.

Observe that $\theta_{W,-1}(n) > 0$ implies $\theta(Wn - 1) > 0$, which in turn implies $Wn - 1 \in \mathbb{P}_{\text{Chen}}$. Therefore, the fact that $\alpha(Wn) > 0$ and $\theta_{W,-1}(n) > 0$ implies $Wn \in \{m \in \mathbb{N} : \alpha(m) > 0\} \cap (\mathbb{P}_{\text{Chen}} + 1)$. In particular, this intersection is non-empty and our theorem follows. □

2.3. *Prime-intersectivity of shifted Chen primes.* As in the preceding subsection, we will derive the prime intersectivity of $\mathbb{P}_{\text{Chen}} + 1$ (the second part of Theorem A) from a more general result concerning multiple recurrence.

PROPOSITION 2.3. *Let $A \subset \mathbb{P}$ be such that $\overline{d}_{\mathbb{P}}(A) > 0$. Then, for every $k \in \mathbb{N}$, there exists $p \in \mathbb{P}_{\text{Chen}}$ such that*

$$a, a + (p + 1), \ldots, a + k(p + 1) \in A.$$

Proof. The idea is similar to the proof of Theorem 2.2. The main difference is that we will use the Green–Tao theorem [21] on arithmetic progressions in primes instead of the

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

uniform Bergelson–Leibman polynomial multiple recurrence theorem ([14, Corollary 4.2], [3, Theorem 3.2]).

For this part of the proof, w and N will be parameters to be chosen later, with N very large compared with $W = \prod_{p<w, p\in\mathbb{P}} p$.

Suppose $\overline{d}_{\mathbb{P}}(A) = \rho > 0$. For $b \in [W]$ coprime to W, define $f_{W,b}(n) := \phi(W)/W \log(Wn+b)1_A(Wn+b)$. Note that

$$\begin{aligned}\overline{d}_{\mathbb{P}}(A) = \limsup_{N\to\infty} \frac{|A\cap\mathbb{P}\cap[N]|}{|\mathbb{P}\cap[N]|} &= \limsup_{N\to\infty} \mathbb{E}_{n\in[N]} \log(n) 1_A(n) \\ &= \limsup_{N\to\infty} \mathbb{E}_{\substack{b\in[W]\\(b,W)=1}} \mathbb{E}_{n\in[N/W]} f_{W,b}(n),\end{aligned}$$

where the second equality follows from the prime number theorem and the fact that $\log n = (1+O(\eta))\log N$ for all $n \in [N^{1-\eta}, N]$, where η tends to 0 sufficiently slowly with N, e.g. $\eta = (\log\log N)^{-1}$. Thus, there is a $b \in [W]$ coprime to W such that $\limsup_{N\to\infty} \mathbb{E}_{n\in[N]} f_{W,b}(n) \geq \rho$. Define

$$\alpha_{W,N}(n) = \mathbb{E}_{a\in[N]} f_{W,b}(a) f_{W,b}(a+n)\cdots f_{W,b}(a+kn).$$

By [21, Theorem 3.5 and Proposition 9.1] (or see p. 524 in that paper), there exists a positive constant $c = c(k,\rho)$ such that

$$\mathbb{E}_{n\in[N]} \alpha_{W,N}(n) \geq c - o_{w,N\to\infty}(1), \tag{7}$$

where $o_{w,N\to\infty}(1)$ means a quantity that approaches 0 as w and N approach infinity. Define $\theta_{W,-1}$ as in equation (2). Let δ be the constant found in Theorem 2.1 and let ϵ be very small compared with $c\delta$. In view of Theorem 2.1 and by increasing w and N if necessary, $\theta_{W,-1}$ can be decomposed on $[N]$ as

$$\theta_{W,-1} = f_1 + f_2,$$

where $\delta \leq f_1 \leq 2$ pointwise and $\|f_2\|_{U^{k+1}(\mathbb{Z}_N)} \leq \epsilon$. It follows from equation (7) that

$$\mathbb{E}_{n\in[N]} f_1(n)\alpha_{W,N}(n) \geq c\delta - o_{w,N\to\infty}(1). \tag{8}$$

Now we want to show

$$\mathbb{E}_{n\in[N]} f_2(n)\alpha_{W,N}(n) = \mathbb{E}_{n\in[N]} f_2(n)\mathbb{E}_{a\in[N]} \prod_{j=0}^{k} f_{W,b}(a+jn) = O_k(\epsilon) + o_{w,N\to\infty}(1), \tag{9}$$

where $O_k(\epsilon)$ means a quantity bounded by $C\epsilon$ in which C only depends on k. To do this, we will use the generalized von Neumann theorem (van der Corput lemma) (see [21, Proposition 5.3], [22, Proposition 7.1], or [7, Proposition 3.8]), which says that if the integers a_j, b_j satisfy $a_i b_j \neq a_j b_i$ for all $i \neq j$ and if $|f_j| \leq \nu + 2$ for all j, where ν is a 'pseudorandom measure' on $[N]$, then

$$\mathbb{E}_{m,n\in[N]} \prod_{j=0}^{k} f_j(a_j m + b_j n) = O\Big(\inf_{0\leq j\leq k} \|f_j\|_{U^k(\mathbb{Z}_N)}\Big) + o_{w,N\to\infty}(1),$$

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

where the implicit constant depends only on k, a_j, b_j. By [21, Theorem 9.1], $|f_{W,b}|$ is bounded by a pseudorandom measure ν. It remains to check that $|f_2| \leq \nu + 2$. Since $f_2 = \theta_{W,-1} - f_1$ and $0 \leq f_1 \leq 2$, it suffices to check that $0 \leq \theta_{W,-1} \leq \nu$. This in turn was shown in [7, Proposition 4.2].

Putting equations (8) and (9) together, we obtain

$$\mathbb{E}_{n\in[N]}\theta_{W,-1}(n)\alpha_{W,N}(n) \geq c\delta - O_k(\epsilon) - o_{w,N\to\infty}(1).$$

Thus, by choosing ϵ very small compared with $c\delta$, there exist large w and N such that

$$\mathbb{E}_{n\in[N]}\theta_{W,-1}(n)\alpha_{W,N}(n) > 0.$$

As a result, there exist $a, n \in [N]$ such that

$$Wa - 1, W(a+n) - 1, \ldots, W(a+kn) - 1 \in A$$

and $Wn \in \mathbb{P}_{\text{Chen}} + 1$. Thus, $Wa - 1, Wa - 1 + Wn, \ldots, Wa - 1 + k(Wn) \in A$, where $Wn \in \mathbb{P}_{\text{Chen}} + 1$. □

Remark 1. In this remark, we explain why the proofs of Propositions 2.2 and 2.3 do not work for $\mathbb{P}_{\text{Chen}} - 1$. A main ingredient in these proofs is Theorem 2.1, which in turn uses [7, Proposition 5.2]. The hypothesis of [7, Proposition 5.2] requires both b and $b + 2$ to be coprime to W, where $W = \prod_{p\in\mathbb{P},p<w} p$ comes from the W-tricked weighted indicator function of Chen primes $\theta_{W,b}$. (This requirement is due ultimately to the fact that $\theta_{W,b}$ is supported on $\{n : Wn + b \text{ is a Chen prime}\}$.) If we work with $\mathbb{P}_{\text{Chen}} + 1$, then $b = -1$ and $b + 2 = 1$, both of which are coprime to W. However, if we work with $\mathbb{P}_{\text{Chen}} - 1$, then $b = 1$ and so $b + 2 = 3$, which is not coprime to W. At the moment, we do not know how to remove this local obstruction.

3. *A quantitative approach and Proof of Theorem B*

3.1. *Proof of Theorem B.*

For two functions $f, g : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, the expression $f(x) = o(g(x))$ means $\lim_{x\to\infty} f(x)/g(x) = 0$. Theorem B will be a simple consequence of the following proposition.

PROPOSITION 3.1. *Let $E \subset \mathbb{N}$ be an infinite set and $E(x) = |E \cap [1, x]|$ its counting function. Let $E_m(x) = |\{n \in E \cap [1, x] : n + m \in E\}|$. Assume that there is a thick set $T \subset \mathbb{N}$ and an increasing bijection $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ such that the following hold as x tends to infinity:*

(1) $f^{-1}(x) = o(E(x))$;

(2) $\max\{E_m(x) : m \in T, m \leq f(x)\} = o(E(x))$.

Then, there exists $A \subset E$ of relative density 1 such that $A - A$ is not syndetic.

If both conditions are satisfied as x tends to infinity along a common subsequence, then there exists $A \subset E$ of upper relative density 1 such that $A - A$ is not syndetic.

Remark 2. The hypothesis of Proposition 3.1 is cumbersome, but certainly some condition on $E_m(x)$ is necessary. We discuss this issue in §3.2.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Proof of Proposition 3.1. By definition, the thick set T contains a set of the form $R = \bigcup_{k\in\mathbb{N}} I_k$, where $I_k = \mathbb{N} \cap [g(k) - k, g(k)]$ and $g : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is an increasing bijection. We may choose g to grow as fast as we like, and we will determine a suitable rate of growth later.

We shall construct a set $C \subset E$ of relative density 1 in E such that $B = (C + R) \cap E$ is relatively sparse in E. Then, $A := C \setminus B$ satisfies $A \cap (A + R) \subset (C + R) \cap (E \setminus B) = \varnothing$, so $(A - A) \cap R = \varnothing$, in particular, $A - A$ is not syndetic.

Observe that for any $C \subset E$, we have

$$C + R = \bigcup_{\substack{s\in C,k\in\mathbb{N}\\ g(k)\le f(s)}} (s + I_k) \cup \bigcup_{\substack{s\in C,k\in\mathbb{N}\\ g(k)>f(s)}} (s + I_k).$$

Now we try to construct C such that $\bigcup_{\substack{s\in C,k\in\mathbb{N}\\ g(k)\le f(s)}} (s + I_k)$ does not meet E.

Define

$$C := \{s \in E : (s + I_k) \cap E = \varnothing \text{ for all } k \text{ such that } g(k) \le f(s)\}.$$

Let $E'(x) := |(E \setminus C) \cap [x]|$ be the counting function of $E \setminus C$. If $s \in E \setminus C$, then by definition of C, there exists $k \in \mathbb{N}$ such that $g(k) \le f(s)$ (so $k \le g^{-1}(f(s))$) and $m \in I_k \subset T$ such that $s + m \in E$. Therefore,

$$\begin{aligned} E'(x) &\le \sum_{k:g(k)\le f(x)} \sum_{m\in I_k} E_m(x) \\ &< g^{-1}(f(x))^2 \cdot \max\{E_m(x) : m < f(x), m \in T\} = o(E(x)) \end{aligned}$$

if g grows sufficiently quickly, using hypothesis (2).

Then, $(C + R) \cap E \subset \bigcup_{\substack{s\in E,k\in\mathbb{N}\\ g(k)>f(s)}} (s + I_k) =: D$. We show that this set is sparse. Observe that if $g(k) > f(s)$, then $s < f^{-1}(g(k))$ and $s + I_k \subset [g(k) - k, g(k) + f^{-1}(g(k))]$. Therefore,

$$D \subset \bigcup_k [g(k) - k, g(k) + f^{-1}(g(k))].$$

In particular, $D \cap [1, x] \subset \bigcup_{g(k)<x} [g(k) - g^{-1}(x), g(k) + f^{-1}(x)]$, whose cardinality is at most

$$g^{-1}(x) \cdot ((f^{-1}(x) + g^{-1}(x)) = o(E(x))$$

if g grows sufficiently quickly (in terms of f and E), in view of hypothesis (1).

It follows that the counting function $B(x)$ of $B = (C + R) \cap E$ satisfies $B(x) = o(E(x))$, as desired, and we are done.

If the hypotheses only hold for x in an increasing sequence $(x_n)_{n\in\mathbb{N}}$ of integers, we still have $E'(x_n) = o(E(x_n))$ and also $B(x_n) = o(E(x_n))$ as n tends to infinity, which concludes the proof. □

Proof of Theorem B. To prove Theorem B, it suffices to check the hypotheses of Proposition 3.1 for the set of primes. In this case, one may take $T = \mathbb{N}$ and $f : x \mapsto x^2$.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

The counting function of the primes, $E(x)$, is asymptotic to $x/\log x$ by the prime number theorem, so the first hypothesis is satisfied.

By definition, $E_m(x)$ is the number of primes $p \le x$ such that $p+m$ is prime. We know, by Selberg's sieve, that

$$E_m(x) \ll \frac{x}{\log^2 x} \prod_{p|m}(1+1/p),$$

where the implied constant is absolute; see for instance [26]. Furthermore, denoting the number of distinct prime factors of m by $\omega(m)$ and observing that $\omega(m) \le \log_2 m$ and $\prod_{p\le n}(1+1/p) \ll \log n$, we have

$$\prod_{p|m}(1+1/p) \le \prod_{p\le\omega(m)}(1+1/p) \ll \log\log m.$$

Thus, for any $m \le x^2$, we have $E_m(x) \ll (x \log\log x/\log^2 x) = o(E(x))$, confirming the second hypothesis. □

Proposition 3.1 also applies to $E = \{P(n) : n \in \mathbb{N}\}$, where $P \in \mathbb{Z}[x]$ with a positive leading coefficient and $d = \deg(P) \ge 2$. In this case, we take $T = \mathbb{N}$ and $f : x \to x^{d+1}$. Since $P(x) = o(x^{d+1})$, we have $f^{-1}(x) = x^{1/(d+1)} = o(E(x))$. For $m \in \mathbb{N}$, $E_m(x) \ll d(m)$, where $d(m)$ is the number of divisors of m. Applying the divisor to the bound $d(m) = m^{o(1)}$, it is easy to see the second hypothesis is true in this case. For the same reason, we can show that Proposition 3.1 also applies to $E = \{P(p) : p \in \mathbb{P}\}$, where $P \in \mathbb{Z}[x]$ having a positive leading coefficient and degree ≥ 2.

3.2. *A discussion on Proposition 3.1.* Here, we discuss the necessity of certain hypotheses of Proposition 3.1.

As mentioned in the introduction, if $\underline{d}(E) > 0$ and $A \subset E$ satisfies $\overline{d}_E(A) > 0$, then $A - A$ is syndetic. The hypothesis $\underline{d}(E) = 0$ is therefore necessary in Proposition 3.1. However, $\underline{d}(E) = 0$ is not sufficient to guarantee the conclusion of Proposition 3.1, as we shall see. In fact, even $d^*(E) = 0$ is not sufficient, as we shall see.

First, recall the notation $E(x)$ and $E_m(x)$ from Proposition 3.1. If for some $m \in \mathbb{N}$ and constant $c_m > 0$ and every $x > 0$ we have $E_m(x) \ge c_m E(x)$, then every $A \subset E$ such that $m \notin A - A$ must have $\overline{d}_E(A) < 1 - c_m/2$. Indeed, for every pair $\{n, n+m\} \subset E$, at least one element of the pair must be outside of A; and any element of E may be in at most two such pairs. This observation already makes the condition on $E_m(x)$ in Proposition 3.1 less surprising and will be implicitly at the core of the next proposition.

PROPOSITION 3.2. *Let E be the set of all positive integers which have the same number of zeros and ones in their binary expansions; that is,*

$$E = \bigcup_{k=0}^{\infty} \left\{ 2^{2k+1} + \sum_{i\in I} 2^i : I \subset [0, 2k], |I| = k \right\}.$$

Then, E has the following properties:

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

(i) $d^*(E) = 0$;
(ii) *if* $E' \subset E$ *is such that* $\overline{d}_E(E') = 1$, *then* $E' - E' = \mathbb{Z}$.

Proof. First we prove property (i). Note that

$$|E \cap [1, 2^{2n+2}]| = \sum_{k=0}^{n} \binom{2k}{k} = O(2^{2n}/\sqrt{n}), \tag{10}$$

so E has upper density 0.

We will now prove the stronger statement that E has upper Banach density 0. It suffices to show that as $n \to \infty$, for every $u \in \mathbb{N}$, we have $|E \cap [u, u + 2^n)| = o(2^n)$. Also, we may assume that u is divisible by 2^n. We consider two cases.

Case 1: $u = 0$. In this case, the claim follows from the estimate in equation (10).

Case 2: $u \geq 2^n$. We write

$$u = \sum_{s=1}^{\ell} 2^{j_s},$$

where $n \leq j_1 < \cdots < j_\ell$. Suppose $v \in E \cap [u, u + 2^n)$. By the definition of E,

$$v = 2^{2k+1} + \sum_{r=1}^{k} 2^{i_r}$$

for some $0 \leq i_1 < \cdots < i_k \leq 2k$. Therefore,

$$v = 2^{2k+1} + \sum_{r=1}^{k} 2^{i_r} = \sum_{s=1}^{\ell} 2^{j_s} + y$$

for some $0 \leq y < 2^n$. We necessarily have $2k + 1 = j_\ell$ (in particular, k is uniquely determined in terms of u). Furthermore, $i_k = j_{\ell-1}, \ldots, i_{k-\ell+2} = j_1$ and these exponents are also uniquely determined. Hence,

$$y = \sum_{r=1}^{k-\ell+1} 2^{i_r}.$$

Therefore,

$$|E \cap [u, u + 2^n)| \leq \binom{n}{k - \ell + 1} = o(2^n).$$

We now proceed to prove property (ii). Let $a \in \mathbb{N}$ be arbitrary. Suppose $a = \sum_{i \in A} 2^i$ is the binary representation of a. Then, we have

$$a = \sum_{i \in A+1} 2^i - \sum_{i \in A} 2^i.$$

Let $|A| = m$ and $\ell = \max A$. By equation (10), there exists a constant $c_{\ell,m} > 0$ such that for all n sufficiently large, we have

$$\binom{2n - \ell - 1}{n - m} > c_{\ell,m} |E \cap [1, 2^{2n+4}]|.$$

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Since $\overline{d}_E(E') = 1$, we have $|E' \cap [1, x]| \geq (1 - c_{\ell,m}/2)|E \cap [1, x]|$ for infinitely many $x \in \mathbb{N}$. Let x be any such number and let n be such that $2^{2n+2} < x \leq 2^{2n+4}$.

Note that for any subset $J \subset [\ell + 2, 2n]$ with $|J| = n - m$, we have

$$a = s_1 - s_2 = \left(2^{2n+1} + \sum_{i \in J \cup (A+1)} 2^i\right) - \left(2^{2n+1} + \sum_{i \in J \cup A} 2^i\right)$$

is a difference of two elements $s_1, s_2 \in E \cap [1, 2^{2n+2}] \subset E \cap [1, x]$. The number of such representations is

$$\binom{2n - \ell - 1}{n - m} > c_{\ell,m}|E \cap [1, 2^{2n+4}]| \geq c_{\ell,m}|E \cap [1, x]|$$

if n is sufficiently large.

Since $|E' \cap [1, x]| \geq (1 - c_{\ell,m}/2)|E \cap [1, x]|$, the number of representations $a = s_1 - s_2$, with $s_1, s_2 \in E \cap [1, x]$ and at least one of them not belonging in $E' \cap [1, x]$, is at most $c_{\ell,m}|E \cap [1, x]|$. Thus, there exists a representation $a = s_1 - s_2$, where $s_1, s_2 \in E' \cap [1, x]$. This shows that $E' - E' = \mathbb{Z}$, and we are done. □

With the notation from Proposition 3.1, in Proposition 3.2, we just proved $E_a(x) \geq c_a E(x)$ for some $c_a > 0$, for any a, and for x large enough in terms of a. Yet for many a, the constant c_a may be expected to be very small, in view of Cusick's conjecture and partial results toward it such as [**13**]. When c_a is positive but indeed very small on a sufficiently large set of integers a, we have the following relaxation of Proposition 3.1 (in which both hypothesis and conclusion are slightly weaker compared with Proposition 3.1).

PROPOSITION 3.3. *Let E be a set of positive integers and $E(x) = |E \cap [1, x]|$ its counting function. Let $E_m(x) = |\{n \in E \cap [1, x] : n + m \in E\}|$. Assume that there is a sequence $(c_m)_{m \in \mathbb{N}}$ and an increasing bijection $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ such that the following hold:*

(1) *$f^{-1}(x) = o(E(x))$ as x tends to infinity;*

(2) *for all $m \in \mathbb{N}$, for all $x > f^{-1}(m)$, one has $E_m(x) \leq c_m \cdot E(x)$.*

Assume additionally that there is a thick set $T \subset \mathbb{N}$ such that $\lim_{m \in T, m \to \infty} c_m = 0$. Then, for every $\epsilon > 0$, there exists $A \subset E$ of lower relative density at least $1 - \epsilon$ such that $A - A$ is not syndetic.

We omit the proof since it is extremely close to that of Proposition 3.1. Note that given a sequence (c_m) of positive real numbers, the negation of the statement that there exists a thick set T such that $\lim_{m \in T, m \to \infty} c_m = 0$ is the statement that there exists $\eta > 0$ such that $\{m \in \mathbb{N} : c_m > \eta\}$ is syndetic.

In practice, Proposition 3.3 is unwieldy because hypothesis (2) is difficult to prove. A more manageable hypothesis would be $\lim_{x \to \infty} E_m(x)/E(x) \leq c_m$, but it is not clear whether such a hypothesis is sufficient. However, the negation of an hypothesis of this form leads to a conclusion of a positive upper Banach density, as stated in the proposition below.

PROPOSITION 3.4. *Let $E \subset \mathbb{N}$. Suppose $c_m := \liminf_{x \to \infty} E_m(x)/E(x)$ is such that $\{m : c_m > \eta\}$ is syndetic (or even has positive lower density) for some $\eta > 0$. Then, $d^*(E) > 0$.*

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Proof. Let $\eta > 0$ be such that $M := \{c_m > \eta\}$ has $\underline{d}(M) > 0$. Enumerating M as $m_1 < m_2 < m_3 < \cdots$, we have $\underline{d}_S(S - m_i) \geq \eta$ for every i. Let ν be a finitely additive probability measure on $\mathbb{N}$ such that $\nu(S - m) \geq \underline{d}_S(S - m)$ for every m. Such a ν can be obtained as a weak*-limit of the probability measures

$$\nu_N := \frac{1}{|S \cap [N]|} \sum_{x \in S \cap [N]} \delta_x,$$

where δ_x is the Dirac point mass at x. We then have $\nu(S - m) \geq \eta$ for all $m \in M$. By Bergelson's intersectivity lemma [2, Theorem 1.1], there is a subset $I \subset \mathbb{N}$ with $\bar{d}(I) > 0$ such that by letting $M' := \{m_i : i \in I\}$, we have

$$\nu\Big(\bigcap_{m \in F} (S - m)\Big) > 0$$

for any finite set $F \subset M'$. (Bergelson's intersectivity lemma was originally stated for countably additive measures. However, the proof of [2, Theorem 2.1] shows that it applies to finitely additive measures, as well.) In particular,

$$\bigcap_{m \in F} (S - m) \neq \varnothing \quad \text{for all finite } F \subset M'. \tag{11}$$

The assumption that $\underline{d}(M) > 0$ and $\bar{d}(I) > 0$ imply that $\bar{d}(M') > 0$. Equation (11) implies S contains a translate of every finite subset of M', so $d^*(S) > 0$. □

4. *Bohr compactification and proof of Theorem C*

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the one-dimensional torus and $\mathbb{T}_d$ be $\mathbb{T}$ endowed with the discrete topology. The *Bohr compactification* of $\mathbb{Z}$ is the Pontryagin dual of $\mathbb{T}_d$ and denoted by $b\mathbb{Z}$. Then, $b\mathbb{Z}$ is a compact abelian group. For every $n \in \mathbb{Z}$, let $\tau(n) \in b\mathbb{Z}$ be the character on $\mathbb{T}_d$ defined by $\tau(n)(\chi) = \chi(n)$, for every $\chi \in \mathbb{T} \cong \hat{\mathbb{Z}}$. Then, $\tau(\mathbb{Z})$ is dense in $b\mathbb{Z}$ (for a proof, see [45, Theorem 1.8.2]). We use $m_{b\mathbb{Z}}$ to denote the normalized Haar measure on $b\mathbb{Z}$. We remark that $(b\mathbb{Z}, \tau)$ has the following universal property: if K is any compact Hausdorff topological group and $\phi : \mathbb{Z} \to K$ is a homomorphism, then there is a unique continuous homomorphism $\tilde{\phi} : b\mathbb{Z} \to K$ such that $\phi = \tilde{\phi} \circ \tau$.

We say a sequence $E = \{c_n : n \in \mathbb{N}\} \subset \mathbb{N}$ with $c_1 < c_2 < \cdots$ is *good for the pointwise ergodic theorem* if for any measure-preserving system $(X, \mathcal{B}, \mu, T)$ and any $f \in L^\infty(X)$, the pointwise limit $\lim_{N \to \infty} (1/N) \sum_{n=1}^{N} f(T^{c_n} x)$ exists for almost all $x \in X$.

Work of Bourgain [8], Wierdl [50], and Nair [40] provides examples of sequences which are good for the pointwise ergodic theorem: $\{P(n) : n \in \mathbb{N}\}$, $\{\lfloor Q(n) \rfloor : n \in \mathbb{N}\}$, $\{p_n : n \in \mathbb{N}\}$, $\{P(p_n) : n \in \mathbb{N}\}$. Here, P is any polynomial in $\mathbb{Z}[x]$, Q is any polynomial in $\mathbb{R}[x]$, p_n is the nth prime, and $\lfloor \cdot \rfloor$ denotes the integer part.

Theorem C follows from a more general result.

PROPOSITION 4.1. *Let $E \subset \mathbb{N}$ be such that its closure in $b\mathbb{Z}$ has measure 0, i.e. $m_{b\mathbb{Z}}(\overline{\tau(E)}) = 0$. Then, for every $\epsilon > 0$, there exist $A \subset E$ with $\overline{d}_E(A) > 1 - \epsilon$ such that $E - A$ is not piecewise syndetic. In particular, $R := \mathbb{Z} \setminus (E - A)$ is thickly syndetic, while $R \cap \mathbb{N}$ is not E-intersective.*

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Furthermore, if the natural enumeration of E is good for the pointwise ergodic theorem, then $\overline{d}_E(A)$ can be replaced by $d_E(A)$.

In [12], Dressler and Pigno showed that the closures of the following sets in $b\mathbb{Z}$ have measure zero.

(1) The set of prime powers $\{p^n : p \in \mathbb{P}, n \in \mathbb{N}\}$. This explains why Theorem C follows from Theorem 4.1.
(2) The set of sums of two squares $\{x^2 + y^2 : x, y \in \mathbb{Z}\}$.
(3) The set of square-full numbers, that is, the set of numbers n so that every exponent in the prime factorization of n is at least two.

In Appendix A, we give more examples of such sets.

(4) The set of values of a polynomial of degree ¿1, i.e. $\{P(n) : n \in \mathbb{Z}\}$, where $P \in \mathbb{Z}[x]$, deg $P > 1$.
(5) The set of values of a binary quadratic form, i.e. $\{ax^2 + bxy + cy^2 : x, y \in \mathbb{Z}\}$, whose discriminant $D = b^2 - 4ac$ is not a perfect square.
(6) More generally, the set of integers represented by a norm form, e.g. $\{x^3 + 2y^3 + 4z^3 - 6xyz : x, y, z \in \mathbb{Z}\}$. (A norm form is a homogeneous form $F(x_1, \ldots, x_d) = N_{K/\mathbb{Q}}(x_1\omega_1 + \cdots + x_d\omega_d)$, where K is an algebraic number field of degree $d \geq 2$, $\{\omega_1, \ldots, \omega_d\}$ is a basis of the ring of integers of K as a $\mathbb{Z}$-module, and $N_{K/\mathbb{Q}}$ denotes the norm.)

All the examples presented above have zero Banach density, and this is not a coincidence: if $E \subset \mathbb{Z}$, we always have $m_{b\mathbb{Z}}(\overline{\tau(E)}) \geq d^*(E)$, and so if E has positive upper Banach density, it will not satisfy the hypothesis of Theorem 4.1.

As mentioned above, Theorem 4.1 applies to $E = \{P(n) : n \in \mathbb{N}\}$, where $P \in \mathbb{Z}[x]$ has degree ≥ 2. Note that for such sets, we also have Theorem G, which says that there exists a thick set (however, not thickly syndetic) that is not chromatically E-intersective.

In the proof of Theorem 4.1, we make use of the following two lemmas from [23]. For completeness, we include the short proofs. Note that the compact abelian groups appearing in this section are not assumed to be metrizable.

The first lemma says that we can create sumsets with large measure and empty interior in a separable compact abelian group.

LEMMA 4.2. [23, Lemma 2.7] *Let K be a separable compact abelian group with the normalized Haar measure m_K and let $E \subset K$ be a compact set with $m_K(E) = 0$. For all $\epsilon > 0$, there exists a compact set $F \subset K$ with $m_K(F) > 1 - \epsilon$ such that $E + F$ has empty interior.*

Proof. Let $X = \{x_n\}_{n=1}^{\infty}$ be a dense, countable subset of K. By taking $G = K \setminus \bigcup_{n=1}^{\infty}(x_n - E)$, we have $x_n \notin E + G$ for all n and since $m_K(E) = 0$, $m_K(G) = 1$. By regularity of the Haar measure on a compact group, G contains a compact set F of measure more than $1 - \epsilon$, and $(E + F) \cap X \subset (E + G) \cap X = \varnothing$. Thus, $E + F$ has empty interior. □

The next lemma says that the set of return times to a closed nowhere dense set is not piecewise syndetic.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

LEMMA 4.3. [23, Lemma 4.1] *Let K be a compact abelian group and $\tau : \mathbb{Z} \to K$ be a homomorphism such that $\tau(\mathbb{Z})$ is dense in K. Let E be a compact subset of K with empty interior. Then, the set $R := \{n \in \mathbb{Z} : \tau(n) \in E\}$ is not piecewise syndetic.*

Proof. Suppose for a contradiction that the set R defined above is piecewise syndetic. Then there is a finite set $A \subset \mathbb{N}$ such that $R' := \bigcup_{a\in A}(R + a)$ is thick. Note that $R' = \{n \in \mathbb{Z} : \tau(n) \in \bigcup_{a\in A}(E + \tau(a))\}$.

We claim that $K = \bigcup_{a\in A}(E + \tau(a))$. Suppose this is not true. Then, $V := K \setminus \bigcup_{a\in A}(E + \tau(a))$ is non-empty and open (since E is compact). Since $\tau(\mathbb{Z})$ is dense in K, the action of $\mathbb{Z}$ on K given by $T_n(x) = x + \tau(n)$ for all $x \in K$ defines a minimal topological dynamical system. By the uniform recurrent property of minimal dynamical systems (for example, see [17, Theorem 1.15]), the set $\mathbb{Z} \setminus R' = \{n \in \mathbb{Z} : \tau(n) \in V\}$ is syndetic. However, this contradicts the fact that R' is thick.

Hence, $K = \bigcup_{a\in A}(E + \tau(a))$, so one of the $E + \tau(a)$ has a non-empty interior and E has a non-empty interior. This is a contradiction. □

To prove Theorem 4.1, we need a new lemma.

LEMMA 4.4. *Let K be a compact abelian group and $\tau : \mathbb{Z} \to K$ an arbitrary map. Then, for every $E \subset \mathbb{N}$ and every measurable set $D \subset K$, there exists $z \in K$ such that*

$$\overline{d}_E(\{n \in \mathbb{N} : z - \tau(n) \in D\}) \geq m_K(D).$$

Proof. Enumerate E as the sequence $1 \leq c_1 < c_2 < \cdots$. Let $f = 1_D$, the indicator function of D, and for $N \in \mathbb{N}$, define the function $f_N : K \to [0, 1]$ by

$$f_N(z) := \frac{1}{N}\sum_{n=1}^{N} f(z - \tau(c_n)).$$

Since m_K is translation invariant, $\int_K f_N \, dm_K = \int_K f \, dm_K$. Because f_N is bounded, Fatou's lemma implies

$$\int_K \limsup_{N\to\infty} f_N \, dm_K \geq \limsup_{N\to\infty} \int_K f_N \, dm_K = \int_K f \, dm_K = m_K(D).$$

Therefore, the set

$$\begin{aligned} S :&= \left\{z \in K : \limsup_{N\to\infty} f_N(z) \geq m_K(D)\right\} \\ &= \left\{z \in K : \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} 1_D(z - \tau(c_n)) \geq m_K(D)\right\} \end{aligned}$$

has positive measure; in particular, this set is non-empty.

Let z be a point in S and $A := \{n \in \mathbb{N} : z - \tau(n) \in D\}$. Then,

$$|A \cap \{c_1, \ldots, c_N\}| = \sum_{n=1}^{N} 1_D(z - \tau(c_n)).$$

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Therefore,

$$\overline{d}_E(A) := \limsup_{N\to\infty} \frac{|A \cap \{c_1, \ldots, c_N\}|}{N}$$
$$= \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} 1_D(z - \tau(c_n)) \geq m_K(D). \qquad \square$$

Equipped with these intermediate results, we may prove Proposition 4.1.

Proof of Theorem 4.1. Let $E \subset \mathbb{N}$ be such that $m_{b\mathbb{Z}}(\overline{\tau(E)}) = 0$ and let $\epsilon > 0$. By applying Lemma 4.2 for $K = b\mathbb{Z}$, there exists a compact set $F \subset b\mathbb{Z}$ such that $m_{b\mathbb{Z}}(F) > 1 - \epsilon$ and $\overline{\tau(E)} + F$ has an empty interior. Lemma 4.3 implies that for all $z \in b\mathbb{Z}$,

$$E_z := \{n \in \mathbb{Z} : z + \tau(n) \in \overline{\tau(E)} + F\}$$

is not piecewise syndetic.

In view of Lemma 4.4, we can choose $z \in b\mathbb{Z}$ so that $A := \{n \in \mathbb{N} : z - \tau(n) \in F\} \cap E$ satisfies

$$\overline{d}_E(A) \geq m_{b\mathbb{Z}}(F) > 1 - \epsilon.$$

Fixing this z, for any $e \in E$ and $a \in A$, we have

$$z + \tau(e - a) = \tau(e) + z - \tau(a) \in \overline{\tau(E)} + F.$$

Therefore, $E - A \subset E_z$ and so $E - A$ is not piecewise syndetic.

Under the additional assumption that E is good for the pointwise ergodic theorem, then the upper relative density $\overline{d}_E$ in Lemma 4.4 can be replaced by the relative density d_E. We may then conclude that $d_E(A) > 1 - \varepsilon$ in the construction of A above. $\square$

Remark 3. In the proof of Proposition 4.1, all we need is a compact abelian group K and a group homomorphism $\tau : \mathbb{Z} \to K$ such that $\tau(\mathbb{Z})$ is dense in K, and $b\mathbb{Z}$ is not the only choice for K. See [27] and [25, §3] for a construction of all such groups K. However, as observed by Dressler and Pigno [12, Theorem 1], $m_K(\overline{\tau(E)})$ is minimized when $K = b\mathbb{Z}$ (where m_K is the probability Haar measure on K). Therefore, the choice $K = b\mathbb{Z}$ is optimal in the statement of Theorem 4.1.

5. *The converse: E-intersectivity implies intersectivity*

The goal of this section is to prove Theorem D. It follows from a more general theorem below regarding sets of multiple recurrence.

Definition 5.1. Let E be an infinite subset of $\mathbb{N}$. For $k \in \mathbb{N}$, a set $S \subset \mathbb{N}^k$ is called a *k-intersective set for E* if for any $A \subset E$ such that $\overline{d}_E(A) > 0$, there exists $(n_1, \ldots, n_k) \in S$ such that

$$A \cap (A - n_1) \cap \cdots \cap (A - n_k) \neq \varnothing.$$

For example, Szemerédi's theorem [48] says that for any k, the set $\{(n, 2n, \ldots, kn) : n \in \mathbb{N}\} \subset \mathbb{N}^k$ is k-intersective for $\mathbb{N}$. The polynomial Szemerédi theorem [4] says

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

that if $P_1, \ldots, P_k \in \mathbb{Z}[x]$ are polynomials without constant term, then the set $\{(P_1(n), \ldots, P_k(n)) : n \in \mathbb{N}\}$ is k-intersective for $\mathbb{N}$.

Theorem D corresponds to the case $k = 1$ of the next proposition.

PROPOSITION 5.2. *Let $E \subset \mathbb{N}$ be an infinite set and $k \in \mathbb{N}$. Every k-intersective set for E is a k-intersective set for $\mathbb{N}$.*

Proof. Suppose $S \subset \mathbb{N}^k$ is not a k-intersective set for $\mathbb{N}$. We will prove S is not a k-intersective set for E.

Since S is not a k-intersective set for $\mathbb{N}$, there exists $B \subset \mathbb{N}$ such that $\overline{d}(B) > 0$ and

$$B \cap (B - n_1) \cap \cdots \cap (B - n_k) = \varnothing \quad \text{for all } (n_1, \ldots, n_k) \in S. \tag{12}$$

By Furstenberg's correspondence principle [15, Theorem 1.1], there exist a measure-preserving system $(X, \mathcal{B}, \mu, T)$ and a set $A \subset X$ with $\mu(A) = \overline{d}(B)$ such that

$$\begin{aligned}&\mu(A \cap T^{-h_1} A \cap \cdots \cap T^{-h_k} A)\\ &\quad \le \overline{d}(B \cap (B - h_1) \cap \cdots \cap (B - h_k)) \quad \text{for all } (h_1, \ldots, h_k) \in \mathbb{N}^k.\end{aligned}$$

Therefore, equation (12) implies

$$\mu(A \cap T^{-n_1} A \cap \cdots \cap T^{-n_k} A) = 0 \quad \text{for all } (n_1, \ldots, n_k) \in S.$$

Define $A' = A \setminus \bigcup_{(n_1,\ldots,n_k)\in S}(A \cap T^{-n_1} A \cap \cdots \cap T^{-n_k} A)$. Then, we have $\mu(A') = \mu(A) > 0$ and

$$A' \cap T^{-n_1} A' \cap \cdots \cap T^{-n_k} A' = \varnothing \quad \text{for all } (n_1, \ldots, n_k) \in S.$$

Therefore, by replacing A with A', we can assume

$$A \cap T^{-n_1} A \cap \cdots \cap T^{-n_k} A = \varnothing \quad \text{for all } (n_1, \ldots, n_k) \in S.$$

It follows that for every $x \in X$, the set $A_x := \{n \in \mathbb{N} : T^n x \in A\}$ satisfies

$$A_x \cap (A_x - n_1) \cap \cdots \cap (A_x - n_k) = \varnothing \quad \text{for all } (n_1, \ldots, n_k) \in S.$$

Enumerate $E = \{c_1, c_2, \ldots\}$ in increasing order. Let $f = 1_A \in L^\infty(X)$. For $N \in \mathbb{N}$, define the function

$$f_N(x) = \frac{1}{N} \sum_{n=1}^{N} f(T^{c_n} x).$$

Since μ is T-invariant, $\int_X f_N \, d\mu = \int_X f \, d\mu$. By Fatou's lemma,

$$\int_X \limsup_{N\to\infty} f_N \, d\mu \ge \limsup_{N\to\infty} \int_X f_N \, d\mu = \int_X f \, d\mu = \mu(A).$$

Therefore, the set

$$R := \left\{x \in X : \limsup_{N\to\infty} f_N(x) \ge \mu(A)\right\}$$

has positive measure; in particular, R is non-empty.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Let $x \in R$. Then,

$$|A_x \cap \{c_1, \dots, c_N\}| = \sum_{n=1}^{N} 1_A(T^{c_n}x).$$

Therefore,

$$\overline{d}_E(A_x) = \limsup_{N\to\infty} \frac{|A_x \cap \{c_1, \dots, c_N\}|}{N} = \limsup_{N\to\infty} f_N(x) \geq \mu(A) > 0.$$

In other words, $B := A_x \cap E$ is a subset of E of positive upper density. However,

$$B \cap (B - n_1) \cap \cdots \cap (B - n_k) \subset A_x \cap (A_x - n_1) \cap \cdots \cap (A_x - n_k) = \varnothing$$

for all $(n_1, \dots, n_k) \in S$. In other words, S is not a k-intersective set for E. □

Remark 4. We remark that the chromatic analogues of Theorems 5.2 and D are obvious: because for every $E \subset \mathbb{N}$, a partition of $\mathbb{N}$ automatically induces a partition of E, a chromatically E-intersective set is chromatically intersective.

6. *Chromatic intersectivity versus density intersectivity*

6.1. *For an arbitrary ambient set.*

Theorem E is a corollary of the next proposition, which in turn was implicitly proved in [24, Theorem 1.2]. Since the concept of E-intersectivity is not mentioned in [24], we prove it explicitly here.

PROPOSITION 6.1. *For every infinite set $E \subset \mathbb{N}$ and $\delta \in (0, 1/2)$, there exists a set $S \subset \mathbb{N}$ which is chromatically E-intersective and a set $A \subset \mathbb{N}$ of upper density at least δ such that $S \cap (A - A) = \varnothing$.*

Remark 5. As mentioned in the introduction, the strength of Proposition 6.1 and Theorem E is in their generality. The price for this generality is the lack of understanding of the chromatically E-intersective set S found in these results. This is in contrast with the special case $E = \mathbb{P}$, where we can take S to be thickly syndetic. In this remark, we discuss another difference between the general case and the case $E = \mathbb{P}$.

For every $\delta \in (0, 1/2)$, Proposition 6.1 gives a set S which is chromatically E-intersective and a subset $A \subset \mathbb{N}$ such that $\overline{d}(A) > \delta$ and $S \cap (A - A) = \varnothing$ (in particular, S is not intersective). To show that S is not E-intersective, we apply the proof of Theorem 5.2, which produces a set $A_x \subset E$ such that $\overline{d}_E(A_x) \geq \overline{d}(A) > \delta$ and $S \cap (A_x - A_x) = \varnothing$. Thus, we only know that the *upper* relative density of A_x is bounded below by $\delta < 1/2$. However, when $E = \mathbb{P}$, we can produce a subset having relative density 1 whose difference set is disjoint from a certain chromatically prime intersective set. Indeed, as a consequence of Theorem B, there are a thick set S and a subset $A \subset \mathbb{P}$ such that $d_{\mathbb{P}}(A) = 1$ and $S \cap (A - A) = \varnothing$. Furthermore, Theorem F, which we will prove shortly, says that this set S (more generally, every thick set) is chromatically prime intersective.

Proof of Proposition 6.1. Let $E \subset \mathbb{N}$ be infinite and let $\delta < \frac{1}{2}$. Theorem 1.2 of [24] proves that there is a set $S \subset E - E$ which is chromatically intersective, and such that $(A - A) \cap S = \varnothing$ for some $A \subset \mathbb{N}$ with $\overline{d}(A) > \delta$. We will explain how a modification

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

of the proof of [24, Theorem 1.2] shows that the S constructed there is chromatically E-intersective.

First, note that the set S is constructed as a union of finite subsets provided by [24, Lemma 4.1]. These are the sets denoted $\tilde{H}(\alpha, 2k+1, \varepsilon) \cap (E-E)$; we will abbreviate them here as S_k. Using [24, Lemma 5.1], it is proved that the Cayley graph determined by S_k has chromatic number at least k. Lemma B.1 has the same hypothesis as [24, Lemma 5.1], but permits the stronger conclusion that the graph with vertex set E and edge set $\{\{a,b\} : a,b \in E, a-b \in S_k \cup (-S_k)\}$ has chromatic number at least k. Since $S = \bigcup_{k=1}^{\infty} S_k$, this implies that the graph with vertex set E and edge set $\{\{a,b\} : a,b \in E, a-b \in S \cup (-S)\}$ has infinite chromatic number. In other words, S is chromatically E-intersective. □

6.2. *For $\mathbb{P}$ as the ambient set.* Theorem B says that there is a thick set R which is not $\mathbb{P}$-intersective. However, we may wonder whether it is at least chromatically $\mathbb{P}$-intersective. Theorem F answers this question positively and we prove it here. In fact, we prove two different enhancements of Theorem F: the first one is a quantitative version of Theorem F, whereas the second one is qualitative and applies to a vast collection of sets, beyond the set of all primes.

6.2.1. *A quantitative enhancement of Theorem F.* A set $A \subset \mathbb{N}$ is a Δ_r^**-set* if $A \cap (S-S) \neq \varnothing$ for every subset S of $\mathbb{Z}$ with $|S| = r$.

The $\mathbb{N}$*-syndeticity index* of a set $A \subset \mathbb{N}$ is the smallest cardinality of a set $S \subset \mathbb{Z}$ such that $A + S \supset \mathbb{N}$. If A is a Δ_r^*-set, then A has $\mathbb{N}$-syndeticity index at most $r-1$. To see this, choose $t_1 \in \mathbb{N}$ arbitrary and choose $t_2, t_3, \ldots \in \mathbb{N}$ recursively such that $t_k > t_{k-1}$ and $t_k \notin \bigcup_{i=1}^{k-1}(t_i + A)$ for any $k \geq 2$. Since $A \cap \{t_j - t_i : 1 \leq i < j \leq k\} = \varnothing$ for all k, this process must stop at some $k \leq r-1$. Then, $\mathbb{N} \setminus \bigcup_{i=1}^{k}(t_i + A)$ is finite and let m be the greatest integer in this set. It follows that $\mathbb{N} \subset \bigcup_{i=1}^{k}(t_i - m + A)$.

A set $H = \{h_1, \ldots, h_k\}$ of integers is said to be *admissible* if for every prime p, the elements of H do not occupy all the residues modulo p. The Hardy–Littlewood conjecture says that if H is admissible, then there are infinitely many $n \in \mathbb{N}$ such that all elements of $n + H$ are primes. The Maynard–Tao theorem [38] says that for any $r > 0$, if H is any admissible set with $|H| \gg r^2 e^{4r}$, there exists infinitely many $n \in \mathbb{N}$ such that $|(n+H) \cap \mathbb{P}| \geq r$.

Pintz [41, Theorem 2] proved that $\mathbb{P} - \mathbb{P}$ has finite $\mathbb{N}$-syndeticity index; however, his proof does not give a bound on the $\mathbb{N}$-syndeticity index. Huang and Wu [28] proved that $\mathbb{P} - \mathbb{P}$ is a Δ_{721}^*-set. The next proposition extends Huang and Wu's result to any coloring of $\mathbb{P}$.

PROPOSITION 6.2. *For each $r \geq 1$, there is a number $m = m(r) \ll r^3 e^{4r}$ such that the following holds. For any partition $\mathbb{P} = \bigcup_{i=1}^{r} P_i$, the set $\bigcup_{i=1}^{r}(P_i - P_i) \cap \mathbb{N}$ is Δ_m^* and so has $\mathbb{N}$-syndeticity index $\ll r^3 e^{4r}$. Consequently, $\bigcup_{i=1}^{r}(P_i - P_i)$ is syndetic.*

Proposition 6.2 directly implies Theorem F.

Proof. To begin with, we recall the following observation of Huang and Wu [28].

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

CLAIM. *If $A \subset \mathbb{Z}$, $|A| = m \geq k \prod_{p\in\mathbb{P},\, p\leq k}(1-1/p)^{-1}$ (i.e. $k \ll m/\log m$), then A contains an admissible set B of cardinality k.*

Proof of the claim. By sieving out one residue modulo p, for each $p \leq k$, we have a set $B \subset A$ of cardinality $\geq m \prod_{p\in\mathbb{P},\, p\leq k}(1-1/p) \geq k$ which has the property that for each $p \in \mathbb{P}$, $p \leq k$, B misses at least one residue modulo p. By removing additional elements from B, we may assume $|B| = k$. Clearly, for each $p \in \mathbb{P}$, $p > k$, B does not occupy all the residues modulo p. Hence, B is admissible and so the claim is proved. □

Returning to Proposition 6.2, let A be any set of cardinality $m \gg r^3 e^{4r}$, then A contains an admissible set B of cardinality $k \gg r^2 e^{4r}$. By the Maynard–Tao theorem, there are infinitely many translates $x + B$ that contain at least $r + 1$ primes. Two of these primes must have the same color, so $(B - B) \cap \bigcup_{i=1}^{r}(P_i - P_i) \neq \{0\}$ and therefore, $(A - A) \cap \bigcup_{i=1}^{r}(P_i - P_i) \neq \{0\}$. Thus, $\bigcup_{i=1}^{r}(P_i - P_i)$ is Δ_m^* and so has $\mathbb{N}$-syndeticity index $\ll r^3 e^{4r}$. Since $\bigcup_{i=1}^{r}(P_i - P_i)$ is symmetric, it is syndetic. □

6.2.2. *A generalization of Theorem F to other sets.* The next theorem gives a criterion on $E \subset \mathbb{N}$ for every thick set to be chromatically E-intersective. Its proof is partially inspired by Pintz [41]. Its statement involves a generalization of the notion of admissible sets.

PROPOSITION 6.3. *Let E be a set of positive integers. Suppose that there exists a family $\mathcal{F}$ of finite sets of positive integers, called generalized admissible sets, satisfying the following two properties.*

(1) *For every $k \in \mathbb{N}$, there exists $\ell \in \mathbb{N}$ such that for all $F \in \mathcal{F}$ of cardinality at least ℓ, the set $|\{n \in \mathbb{N} : |(n + F) \cap E| \geq k\}|$ is infinite.*

(2) *For every $\ell \in \mathbb{N}$, there exists $C \in \mathbb{N}$ such that for every family $I_1, \ldots, I_\ell$ of intervals of length at least C, there exists $F = \{f_1, \ldots, f_\ell\} \in \mathcal{F}$ such that $f_j \in I_j$ for all $j \in [\ell]$.*

Then every thick set is chromatically E-intersective.

Proof. Consider a partition $E = \bigcup_{i=1}^{r} E_i$. Applying condition (1) with $k = r + 1$, letting $\ell \in \mathbb{N}$ be given by this condition, we infer that for every generalized admissible set $F \in \mathcal{F}$ of cardinality ℓ, the set X of $n \in \mathbb{N}$ such that $|(n + F) \cap E| \geq r + 1$ is infinite. Then, by pigeonholing, for every $n \in X$, there is $i \in [r]$ such that $|(n + F) \cap E_i| \geq 2$. Pigeonholing again, we find an $i \in [r]$ such that the set Y of integers n for which $|(n + F) \cap E_i| \geq 2$ is infinite.

Let $T \subset \mathbb{N}$ be a thick set. We need to show that $T \cap \bigcup_{i\in r}(E_i - E_i) \neq \varnothing$. By definition of T, there exists a sequence $(N_c)_{c\in\mathbb{N}}$ of integers such that $T \supseteq \bigcup_c [N_c, N_c + c]$. Upon extracting a sequence of c, we may in fact suppose that $T \supseteq \bigcup_{k\geq 1} I_k$, where $I_k = [M_k, M_k + C_k]$ and $M_k > C_k > 4M_{k-1}$; also $C_1 > 0$ may be chosen arbitrarily large. Let ℓ be given by condition (1) for $k = r + 1$. Consider the intervals $I_j' = [M_j + C_j/2, M_j + C_j]$ for $j \in [\ell]$. By condition (2), there exists a generalized admissible set $F = \{f_1, \ldots, f_\ell\} \in \mathcal{F}$, where $f_j \in I_j'$ for each $j \in [\ell]$, if C_1 is large enough. Thus,

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

there exists $i \in [r]$ such that $|(n+F) \cap E_i| \geq 2$ for some $n \in \mathbb{N}$. Therefore, $f_j - f_m \subset E_i - E_i$ for some $1 \leq m < j \leq \ell$.

By assumption, $0 \leq f_m \leq 2M_m \leq 2M_{j-1}$, so $f_j - f_m \in [M_j + C_j/2 - 2M_{j-1}, M_j + C_j] \subset I_j$. We infer $I_j \cap (E_i - E_i) \neq \varnothing$, and hence $T \cap (E_i - E_i) \neq \varnothing$. □

Many sets E satisfy the hypothesis of Proposition 6.3. Here are some examples.

(1) The set of primes, with $\mathcal{F}$ being the family of all admissible sets in the usual sense. Indeed, property (1) is satisfied by Maynard's theorem [38]. Property (2) is satisfied with $C = \prod_{p \leq \ell} p$ for instance, because whenever $F = \{f_1, \ldots, f_\ell\}$ is such that $(f_i, C) = 1$, we have $|F \bmod p| < p$ for every $p \leq \ell$, and $|F \bmod p| \leq \ell < p$ for every $p > \ell$. Therefore, Theorem 6.3 implies Theorem F.

More generally, still by Maynard's theorem, the set of all primes of the form $an + b$, for any given coprime integers a and b satisfies the hypothesis of Theorem 6.3. Since the set E of all sums of two squares contain the set of all primes of the form $4n + 1$, Theorem 6.3 also applies to E.

(2) Subsets of the primes which retain some of the equidistribution enjoyed by the primes, in the form of a Siegel–Walfisz and a Bombieri–Vinogradov theorem, also satisfy the hypothesis of Theorem 6.3, as shown by Benatar [1] and Maynard [39], with $\mathcal{F}$ being again the family of all admissible sets in the usual sense. These authors respectively mention sets of the form $\mathbb{P} \cap g^{-1}((0, d))$, where $g \in \mathbb{R}[x]$ satisfies some diophantine conditions and d is a positive real, and Chebotarev sets, i.e. primes with a prescribed value of the Artin symbol respective to a Galois number field extension. See also [49] for the latter example.

(3) Almost all sets of integers, by [6, Lemma 5]. Here, $\mathcal{F}$ is the family of all finite subsets of $\mathbb{N}$. 'Almost all' refers to the probability measure on (non-cofinite) sets of integers induced by the Lebesgue measure on [0, 1] through the bijection provided by the binary expansion of real numbers.

6.3. *Ambient sets whose gaps go to infinity.* Here, we prove Theorem G, demonstrating the necessity of the gaps between consecutive primes not tending to infinity in Theorem F.

Proof of Theorem G. Let $R = \bigcup_{n \geq 2} I_n$ where $I_n = [f(n), f(n) + n]$, where $f(n)$ is a sufficiently quickly increasing sequence of elements of E. In particular, the intervals I_n may be assumed to be pairwise disjoint and so R is thick.

CLAIM. *Given $a \in E$, there exists at most one integer $b < a$ in E such that $a - b \in R$.*

Proof of the claim. Consider $a \in E$. Assume there exists such an integer $b \in E$, and let $n \in \mathbb{N}$ such that $a - b \in I_n$; n is unique given $a - b$. By hypothesis, there exists a function $g : E \to \mathbb{N}$ increasing to infinity such that $a - c > g(a)$ for any $c \in E$ with $c < a$, in particular, for $c = b$. Thus, $a - b \in [g(a), a - 1]$. If f grows sufficiently quickly in terms of g, there is at most one $n \in \mathbb{N}$ such that $[g(a), a - 1] \cap I_n \neq \varnothing$.

So there is a unique n such that there exists $c < a, c \in E$ such that $a - c \in I_n$. In particular, $a > f(n)$. Since $f(n) \in E$, we infer $a \geq f(n) + g(f(n))$. Since $a - b \leq f(n) + n$, this implies in turn that $b \geq f(g(n)) - n$.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

However, then for any integer $c \in E$, $c \neq b$, we have $|c-b| \geq g(b) \geq g(f(g(n))-n) > n$ if f grows sufficiently quickly in terms of g. Thus, $a-c \notin I_n$. This means that $a-c \notin R$. Whence the uniqueness of $b \in E$ such that $a-b \in R$, which completes the proof of the claim. □

Now the claim above implies that there exists a two-coloring of $c : \mathbb{N} \to \{1, 2\}$ such that

$$\text{for all } \{a, b\} \subset E,\ a-b \in R \Rightarrow c(a) \neq c(b). \tag{13}$$

This is a standard deduction in graph theory (the chromatic number is at most one plus the degeneracy), but we provide it briefly here. We construct $c(e_n)$ inductively, where $e_1 < e_2 < \cdots$ is the increasing sequence of the elements of E. Suppose $c(e_1), \ldots, c(e_n)$ have been constructed and satisfy

$$\text{for all } k, \ell \in [n],\ e_k - e_\ell \in R \Rightarrow c(e_k) \neq c(e_\ell).$$

Then, we color e_{n+1}. Since there is at most one $k < n+1$ such that $e_{n+1} - e_k \in R$, it suffices to take $c(e_{n+1}) \neq c(e_k)$ if such a k exists, and $c(e_{n+1}) \in \{1, 2\}$ arbitrary otherwise.

Now the coloring c induces a bipartition $E = E_1 \cup E_2$ defined by $E_i = \{e \in E : c(e) = i\}$ for $i \in [2]$. In view of equation (13), we have $R \cap (E_i - E_i) = \varnothing$ for any $i \in [2]$, as desired. □

7. *Open questions*

We present some questions suggested by our study. The first four questions involve the set of primes and the last two are about arbitrary sets of zero Banach density.

In Theorem F, we prove that for any finite partition $\mathbb{P} = \bigcup_{i=1}^{k} E_i$, the union $\bigcup_{i-1}^{k}(E_i - E_i)$ is syndetic. It follows that $E_i - E_i$ is piecewise syndetic for some $i \in [k]$. It is not clear whether we can upgrade from piecewise syndeticity to syndeticity.

Question 7.1. For any partition $\mathbb{P} = \bigcup_{i=1}^{k} E_i$, does there exist $i \in [k]$ such that $E_i - E_i$ is syndetic?

Note that the density analogue of Question 7.1 is false, as shown in Theorems B and C.

A natural approach to answer Question 7.1 is to use Theorem B, in which we found a partition $\mathbb{P} = A \cup B$ where $A - A$ is not syndetic and $\overline{d}_{\mathbb{P}}(B) = 0$. Since B is a sparse subset of $\mathbb{P}$, it would be plausible to expect that $B - B$ is not syndetic, and thus gives a negative answer to Question 7.1. However, assuming the Hardy–Littlewood conjecture (see §6.2), we can prove that $B - B$ is syndetic.

Indeed, Theorem B is a special case of Proposition 3.1 when $E = \mathbb{P}$, $T = \mathbb{N}$ and $f(x) = x^2$. First, observe that $B = \mathbb{P} \setminus A \supset \mathbb{P} \setminus C =: \mathbb{P}'$, where C is the set defined in the proof of Proposition 3.1. In turn, $\mathbb{P}' \supset \{p \in \mathbb{P} : p \geq \sqrt{g(2)}, (p + [g(2)-2, g(2)]) \cap \mathbb{P} \neq \varnothing\}$. Certainly, $[g(2)-2, g(2)]$ contains an even number m. So $B \supset \{p \in \mathbb{P} : p \geq g(2), p + m \in \mathbb{P}\} =: \mathbb{P}_m$. Let $k \in \mathbb{N}$ and $H = \{0, m, 6k, m+6k\}$. Since m is even, H is admissible. By the Hardy–Littlewood conjecture, there are infinitely many n such that $n + H \subset \mathbb{P}$. Hence, there are infinitely many $n \in \mathbb{P}_m$ such that $n + 6k \in \mathbb{P}_m$. We conclude that $6 \cdot \mathbb{Z} \subset \mathbb{P}_m - \mathbb{P}_m$ and therefore $\mathbb{P}_m - \mathbb{P}_m$ is syndetic.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

The next question is about the chromatic analog of Theorems B and C. More precisely, it follows from these theorems that that there are intersective sets which are not prime intersective. Therefore, we ask the following question.

Question 7.2. Must every chromatically intersective set be chromatically prime intersective?

For the density counterpart of Question 7.2, we produce thick sets (and so an intersective set) which are not prime intersective. However, the same idea will not work for Question 7.2: it has been shown in Theorem F that every thick set is chromatically prime intersective. We also remark that the converse of Question 7.2 is true; it is easy to see that every chromatically prime intersective set is chromatically intersective.

In the next question, we upgrade chromatic intersectivity to density intersectivity.

Question 7.3. Must every intersective set be chromatically prime intersective?

Questions 7.2 and 7.3 are related to the next conjecture. It is widely believed that $\mathbb{P}-\mathbb{P} \supset 2\cdot\mathbb{Z}$. If this conjecture is true, then for every intersective set R, $R \cap (\mathbb{P}-\mathbb{P}) \neq \varnothing$. The following conjecture seeks to prove the second clause unconditionally.

Conjecture 7.4. For every intersective set R, we have $R \cap (\mathbb{P}-\mathbb{P}) \neq \varnothing$.

This conjecture is equivalent to the statement that $\mathbb{P}-\mathbb{P}$ contains a set of the form $E-E$ where $d^*(E) > 0$. Pintz [**41**] shows that $\mathbb{P}-\mathbb{P}$ is syndetic, proving the special case of the conjecture where R is thick. However, as proved in Theorem B, the analog of the conjecture where $\mathbb{P}$ is replaced by a set A with $d_{\mathbb{P}}(A) = 1$ is false. Among the previous three questions/conjectures, Conjecture 7.4 is the weakest in the following sense: a positive answer to Question 7.2 implies a positive answer to Question 7.3, which in turn implies Conjecture 7.4.

In Proposition 3.2, we show that one cannot replace $\mathbb{P}$ in Theorem B with an arbitrary set E having $d^*(E) = 0$. More precisely, there exists $E \subset \mathbb{N}$ such that $d^*(E) = 0$ and if $d_E(A) = 1$, then $A-A = \mathbb{Z}$ (in particular, syndetic). Currently, we do not know if the same is true for Theorem C. Likewise, the following slightly stronger statement is open.

Question 7.5. Does there exist $E \subset \mathbb{N}$ such that $d^*(E) = 0$ and if $\overline{d}_E(A) > 0$, then is $A-A$ syndetic?

Nevertheless, Propositions 3.3 and 3.4 tend to suggest that the answer is no, perhaps even with the condition $\overline{d}_E(A) > 1-\epsilon$ instead of $\overline{d}_E(A) > 0$.

Our last question aims to generalize the fact that there is an intersective set which is not prime intersective to an arbitrary set of zero Banach density.

Question 7.6. Suppose $d^*(E) = 0$. Does there exist an intersective set which is not E-intersective?

Note that the converse of Question 7.6 is false according to Theorem D.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Acknowledgments. We thank the anonymous referee for feedback and suggestions. The first author is supported by FWF Grant I-5554. The third author is supported by an AMS-Simons Travel Grant. The fourth author is supported by NSF Grant DMS-2246921 and a Travel Support for Mathematicians gift from the Simons Foundation.

A. *Appendix. Subsets of $\mathbb{Z}$ whose closures in $b\mathbb{Z}$ have measure zero*

In this appendix, we exhibit some subsets of $\mathbb{Z}$ whose closures in $b\mathbb{Z}$ have measure 0. Recall from §4 that $b\mathbb{Z}$ is the Bohr compactification of $\mathbb{Z}$ with normalized Haar measure $m_{b\mathbb{Z}}$, and that $\tau : \mathbb{Z} \to b\mathbb{Z}$ is a one-to-one homomorphism such that $\tau(\mathbb{Z})$ is topologically dense in $b\mathbb{Z}$. We now identify $\mathbb{Z}$ with its image $\tau(\mathbb{Z})$ in $b\mathbb{Z}$, and we identify a subset $A \subset \mathbb{Z}$ with its image $\tau(A) \subset b\mathbb{Z}$. With this identification, we let $\overline{A}$ denote the closure of A in $b\mathbb{Z}$. To estimate $m_{b\mathbb{Z}}(\overline{A})$, we follow the approach of Dressler and Pigno [12], which relies on the following observations.

LEMMA A.1.

(1) *For all $c \in \mathbb{N}$, $d \in \mathbb{Z}$, we have $m_{b\mathbb{Z}}(\overline{c \cdot \mathbb{Z} + d}) = 1/c$.*
(2) *For any sets $A, E \subset \mathbb{Z}$ with E finite, we have $m_{b\mathbb{Z}}(\overline{A \cup E}) = m_{b\mathbb{Z}}(\overline{A})$.*
(3) *Suppose $c_1, \ldots, c_k \in \mathbb{N}$ are pairwise coprime and $A \subset \mathbb{Z}$ is such that, except for a finite number of exceptions, every $a \in A$ misses c_i' residues$(\mathrm{mod}\ c_i)$ for each $1 \le i \le k$. Then, $m_{b\mathbb{Z}}(\overline{A}) \le \prod_{i=1}^{k}(1 - c_i'/c_i)$.*

Proof. The first statement simply follows from the fact that $\overline{c \cdot \mathbb{Z}}$ is a closed subgroup of index c in $b\mathbb{Z}$, and $m_{b\mathbb{Z}}$ is translation invariant. The second statement holds because $\overline{A \cup E} = \overline{A} \cup \overline{E} = \overline{A} \cup E$, and $m_{b\mathbb{Z}}(E) = 0$. The third statement follows from the first two and the Chinese remainder theorem. □

Let p_i be the ith prime. Then, except for p_i, all primes miss one residue (mod p_i). Hence, Lemma A.1 implies that for any k, $m_{b\mathbb{Z}}(\overline{\mathbb{P}}) \le \prod_{i=1}^{k}(1 - 1/p_i)$. Letting k go to infinity, we conclude that $m_{b\mathbb{Z}}(\overline{\mathbb{P}}) = 0$. All of our examples are obtained in this way, by taking $\{c_i\}$ to be an appropriate sequence of moduli. More precisely, they will be dense subsets of $\mathbb{P}$ and $\{p^2 : p \in \mathbb{P}\}$. The fact that these subsets are dense follows from various instances of Chebotarev's density theorem. We refer the reader to [35] for the statement and an account of Chebotarev's density theorem.

Our first example is the set of values of a polynomial P. It was proved by Kunen and Rudin [31, Theorem 5.6] in the case deg $P = 2$ or 3. It was also pointed out to them by David Boyd (see [31, footnote 1]) that the results hold whenever deg $P \ge 2$, though no proof was given. We will now supply a proof.

PROPOSITION A.2. *Let $P \in \mathbb{Z}[x]$ have degree ≥ 2 and $A = \{P(n) : n \in \mathbb{Z}\}$. Then, $m_{b\mathbb{Z}}(\overline{A}) = 0$.*

Proof. Suppose deg $P = n > 1$. The polynomial $f(x, y) = P(x) + y$ is clearly irreducible in $\mathbb{Z}[x, y]$. By the Hilbert irreducibility theorem (see e.g. [18, Theorem 4]), there exists infinitely many $r \in \mathbb{Z}$ such that $f(x, r) = P(x) + r$ is irreducible in $\mathbb{Z}[x]$. Since $m_{b\mathbb{Z}}$ is translation invariant, we may assume that P is irreducible in $\mathbb{Z}[x]$.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Let K be a splitting field of P and X be the set of all roots of P in K. Let $G = \mathrm{Gal}(K/\mathbb{Q})$ be the Galois group of P. Then G is a subgroup of S_n that acts transitively on X.

CLAIM. *G has an element without fixed points.*

Proof of the claim. By Burnside's lemma, the number of orbits in X is $1/|G| \sum_{g\in G} |X^g|$, where X^g is the set of all elements of X fixed by g. Since the action is transitive, we have $\sum_{g\in G} |X^g| = |G|$. Since $|X^e| = n > 1$, there must be some $g \in G$ such that $|X^g| = 0$, and the claim is proved. □

We now recall Frobenius's theorem (see [35, p. 11]), which is a precursor of Chebotarev's density theorem. Let p be a prime not dividing the discriminant of P. Then in $\mathbb{F}_p[x]$, P factors as a product of distinct irreducible polynomials of degrees $n_1, \ldots, n_k$, where $n_1 + \cdots + n_k = n$. Frobenius's theorem says that the density of primes p with given decomposition pattern $(n_1, \ldots, n_k)$ exists, and is equal to $m/|G|$, where m is the number of permutations $\sigma \in G$ with cycle pattern $(n_1, \ldots, n_k)$.

Let Q be the set of all primes p such that $p \nmid P(n)$ for all $n \in \mathbb{Z}$. Applying Frobenius's theorem with $n_1 = 1$, we see that $d_{\mathbb{P}}(Q) = m/|G|$, where m is the number of permutations $\sigma \in G$ without a fixed point. Hence, $d_{\mathbb{P}}(Q) > 0$.

Let $\{c_i\}_{i=1}^{\infty}$ be all the elements of Q. Applying Lemma A.1 with $c_i' = 1$, we have that $m_{b\mathbb{Z}}(\overline{A}) \le \prod_{i=1}^{k}(1 - 1/c_i)$ for all k. Letting k go to infinity, we have $m_{b\mathbb{Z}}(\overline{A}) = 0$, as desired. □

Our second example generalizes Dressler and Pigno's example of sums of two squares.

PROPOSITION A.3. *Let $a, b, c \in \mathbb{Z}$ such that $D = b^2 - 4ac$ is not a perfect square, and let $A = \{ax^2 + bxy + cy^2 : x, y \in \mathbb{Z}\}$. Then, $m_{b\mathbb{Z}}(\overline{A}) = 0$.*

Proof. According to [32, Lemma 2.8], we have $m_{b\mathbb{Z}}(\overline{A}) \le 4|a| m_{b\mathbb{Z}}(4a \cdot \overline{A}) = 4|a| m_{b\mathbb{Z}}(\overline{4a \cdot A})$. Since

$$4a(ax^2 + bxy + cy^2) = (2ax + by)^2 - Dy^2,$$

it suffices to show that $m_{b\mathbb{Z}}(\overline{A'}) = 0$, where $A' = \{z^2 - Dt^2 : z, t \in \mathbb{Z}\}$. Let p be a prime number such that D is not a quadratic residue modulo p. If $p|(z^2 - Dt^2)$, then we must have $p|z$ and $p|t$, so $p^2|(z^2 - Dt^2)$. Therefore, elements of A' cannot be congruent to $p, 2p, \ldots, (p-1)p \pmod{p^2}$.

Let $Q = \{q_i\}_{i=1}^{\infty}$ be all the primes for which D is not a quadratic residue. We claim that $\underline{d}_{\mathbb{P}}(Q) > 0$. Indeed, since D is not a perfect square, there exists a prime p such that $D = p^k m$, where k is an odd positive integer and $(p, m) = 1$. Consider two cases.

Case 1: $p = 2$ (and m is odd). Let $Q' = \{q \in \mathbb{P} : q \equiv 5 \pmod 8, q \equiv 1 \pmod m\}$ and let $q \in Q'$. Then, 2 is not a quadratic residue mod q, while q is a quadratic residue of every prime divisor of m. By the law of quadratic reciprocity and the fact that $q \equiv 1 \pmod 4$, every prime divisor of m is a quadratic residue$\pmod q$. Hence, m is a quadratic residue $\pmod q$ (this remains true if $m < 0$, since -1 is a quadratic residue$\pmod q$). It follows that D is not a quadratic residue mod q.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

Case 2: $p > 2$. Let q' be a quadratic non-residue mod p. Let $Q' = \{q \in \mathbb{P} : q \equiv q' \pmod{p}, q \equiv 1 \pmod{8m}\}$ and let $q \in Q'$. Since $q \equiv 1 \pmod 8$, 2 is a quadratic residue $(\text{mod } q)$. By the law of quadratic reciprocity, every odd prime divisor of m is a quadratic residue$(\text{mod } q)$ while p is not a quadratic residue$(\text{mod } q)$. Also, -1 is a quadratic residue $(\text{mod } q)$. It again implies that D is not a quadratic residue mod q.

Both of the sets Q' defined above satisfy $Q' \subset Q$ and by Dirichlet's theorem, $d_{\mathbb{P}}(Q') > 0$. Therefore, we always have $\underline{d}_{\mathbb{P}}(Q) > 0$; our claim is proved.

Applying Lemma A.1 with $c_i = q_i^2$ and $c_i' = q_i - 1$, we have that $m_{b\mathbb{Z}}(\overline{A'}) \le \prod_{i=1}^{k}(1 - (q_i - 1)/q_i^2)$ for all k. Letting k go to infinity, we have $m_{b\mathbb{Z}}(\overline{A'}) = 0$, as desired. □

Our third example generalizes the second one.

PROPOSITION A.4. *Let K be an algebraic number field of degree $n > 1$. Let $\mathcal{O}_K$ be the ring of integers of K and $\{\omega_1, \dots, \omega_n\}$ be an integral basis of $\mathcal{O}_K$. Let $F(x_1, \dots, x_n) = N_{F/\mathbb{Q}}(x_1\omega_1 + \cdots + x_n\omega_n)$ and $A = \{F(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{Z}\}$. Then, $m_{b\mathbb{Z}}(\overline{A}) = 0$.*

Proof. Similarly to the proof of Proposition A.3, there is a set Q of primes with $d_{\mathbb{P}}(Q) > 0$, such that whenever $q \in Q$, $a \in A$, and $q|a$, we have $q^2|a$. Such primes q can be characterized by the residual degrees of prime ideals $\mathfrak{p} \subset \mathcal{O}_K$ lying above q (i.e. $\mathfrak{p} \cap \mathbb{Z} = q\mathbb{Z}$). This was done in detail by Glasscock [**19**, Main Theorem (I)], so we will just sketch the idea.

Recall that the norm $N(I)$ of an ideal $I \subset \mathcal{O}_K$ is the index $[\mathcal{O}_K : I]$, and for $x \in \mathcal{O}_K$, $N(x\mathcal{O}_K) = |N_{K/\mathbb{Q}}(x)|$. Any ideal $I \subset \mathcal{O}_K$ has a unique factorization as

$$I = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_k^{e_k},$$

where $\mathfrak{p}_1, \dots, \mathfrak{p}_k$ are prime ideals in $\mathcal{O}_K$.

In particular, when q is a rational prime and $I = q\mathcal{O}_K$, then $\mathfrak{p}_1, \dots, \mathfrak{p}_k$ are all the prime ideals of $\mathcal{O}_K$ lying above q. For each $i = 1, \dots, k$, $\mathcal{O}_K/\mathfrak{p}_i$ is a finite field extension of $\mathbb{Z}_q$; its dimension f_i is called the residual degree of $\mathfrak{p}_i$. In particular, $N(\mathfrak{p}_i) = q^{f_i}$.

CLAIM. *Let q be a prime and with the property that all prime ideals of $\mathcal{O}_K$ lying above q have residual degrees > 1. Suppose $a \in A$ and $q|a$. Then, $q^2|a$.*

Proof of the claim. Suppose $a = N_{K/\mathbb{Q}}(x)$ for some $x \in \mathcal{O}_K$. Then, factoring $x\mathcal{O}_K$ as a product of prime ideals and taking norms, we have

$$|a| = N(x\mathcal{O}_K) = \prod_{i=1}^{k} N(\mathfrak{p}_i)^{e_i}$$

for some prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_k \in \mathcal{O}_K$. Since $q|a$, there exists i such that $q \mid N(\mathfrak{p}_i)$ and therefore $\mathfrak{p}_i$ lies above q. Since the residual degree of $\mathfrak{p}_i$ is greater than 1, we have $q^2 \mid N(\mathfrak{p}_i) \mid a$, as desired. □

To finish the proof of Proposition A.4, one has to show that the set Q of primes satisfying the claim has positive density in $\mathbb{P}$. This follows from an application of the Chebotarev

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

density theorem, and the fact that a finite group cannot be covered by the conjugates of a proper subgroup. We refer the reader to [19, Lemma 4.1] for details. □

B. *Appendix. Proof of Proposition 6.1*

Here, we prove the modification of [24, Lemma 5.1] required for our proof of Proposition 6.1. We first recall some standard terminology for graphs.

A graph $\mathcal{G}$ is a set V whose elements are called *vertices*, together with a set E of unordered pairs of elements of V, called *edges*. A *k-coloring* of $\mathcal{G}$ is a function $f : V \to \{1, \dots, k\}$. We say f is *proper* if $f(v_1) \neq f(v_2)$ for every edge $(v_1, v_2) \in E$. The *chromatic number* of $\mathcal{G}$ is the smallest $k \in \mathbb{N}$ such that there is a proper k-coloring of $\mathcal{G}$. If there is no such $k \in \mathbb{N}$, we say the chromatic number of $\mathcal{G}$ is infinite.

Given an abelian group G and subsets $V, S \subset G$, the *Cayley graph based on* V, S, denoted $\mathrm{Cay}(V, S)$, is the graph whose vertex set is V, with two vertices x, y joined by an edge if $x - y \in S$ or $y - x \in S$. It follows from the definitions that S is chromatically E-intersective if and only if the chromatic number of $\mathrm{Cay}(E, S)$ is infinite. If the vertex set is the ambient group, i.e. $V = G$, we abbreviate $\mathrm{Cay}(V, S) = \mathrm{Cay}(S)$.

The next lemma is essentially [24, Lemma 5.1], modified to conclude that the copy of the graph $\mathcal{G}$ found in $\mathrm{Cay}(\rho^{-1}(U) \cap (E - E))$ has vertex set E rather than vertex set $\mathbb{Z}$.

LEMMA B.1. *Let G be a discrete abelian group, K be a Hausdorff abelian topological group K, and $\rho : G \to K$ be a homomorphism. Assume $E \subset G$ is such that $\rho(E)$ is dense in K. If $U \subset K$ is open, then every finite subgraph contained in* $\mathrm{Cay}(U)$ *has an isomorphic copy in* $\mathrm{Cay}(E, \rho^{-1}(U) \cap (E - E))$.

Consequently, if $\mathrm{Cay}(U)$ *has a finite subgraph with chromatic number k, then*

$$\mathrm{Cay}(E, \rho^{-1}(U) \cap (E - E))$$

has chromatic number $\geq k$.

Proof. To prove the first statement of the lemma, it suffices to prove that if V is a finite subset of K, then there exists $\{g_v : v \in V\} \subset E$ such that for each $v, v' \in V$, we have

$$v - v' \in U \Rightarrow g_v - g'_v \in \rho^{-1}(U). \tag{B.1}$$

So let V be a finite subset of K. Let $S := (V - V) \cap U$, and let W be a neighborhood of 0 in K so that $S + W \subset U$ (one may take $W = \bigcap_{s \in S}(U - s)$, since $S \subset U$ is finite). Choose a neighborhood W' of 0 so that $W' - W' \subset W$. For each $v \in V$, choose $g_v \in E$ so that $\rho(g_v) \in v + W'$; this is possible since $\rho(E)$ is dense in K. Also, since K is Hausdorff, we can ensure that $g_v \neq g_{v'}$ if $v \neq v'$ by choosing W' small enough that the neighborhoods $\{v + W' : v \in V\}$ are mutually disjoint. We now prove that the implication in equation (B.1) holds with these g_v. Assuming $v - v' \in U$, we have

$$\rho(g_v) - \rho(g'_v) \in v + W' - (v' + W') = (v - v') + (W' - W') \subset v - v' + W \subset S + W \subset U,$$

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

so $g_v - g'_v \in \rho^{-1}(U)$. This proves equation (B.1). Since the chromatic number is invariant under isomorphism of graphs, the second assertion of the lemma follows immediately from the first. □

References

[1] J. Benatar. The existence of small prime gaps in subsets of the integers. *Int. J. Number Theory* **11**(3) (2015), 801–833.

[2] V. Bergelson. Sets of recurrence of $\mathbf{Z}^m$-actions and properties of sets of differences in Zm. *J. Lond. Math. Soc. (2)* **31**(2) (1985), 295–304.

[3] V. Bergelson, B. Host, R. McCutcheon and F. Parreau. Aspects of uniformity in recurrence. *Colloq. Math.* **84/85**(part 2) (2000), 549–576; dedicated to the memory of Anzelm Iwanik.

[4] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.* **9**(3) (1996), 725–753.

[5] A. Bertrand-Mathis. Ensembles intersectifs et récurrence de Poincaré. *Israel J. Math.* **55**(2) (1986), 184–198.

[6] P.-Y. Bienvenu. Metric decomposability theorems on sets of integers. *Bull. Lond. Math. Soc.* **55** (2023), 2653–2659.

[7] P.-Y. Bienvenu, X. Shao and J. Teräväinen. A transference principle for systems of linear equations, and applications to almost twin primes. *Algebra Number Theory* **17**(2) (2023), 497–539.

[8] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publ. Math. Inst. Hautes Études Sci.* **69** (1989), 5–45; with an appendix by the author, H. Furstenberg, Y. Katznelson and D. Ornstein.

[9] T. Brown. An interesting combinatorial method in the theory of locally finite semigroups. *Pacific J. Math.* **36**(2) (1971), 285–289.

[10] T. Browning and S. Prendiville. A transference approach to a Roth-type theorem in the squares. *Int. Math. Res. Not. IMRN* **7** (2017), 2219–2248.

[11] J. Chen. On the representation of a large even integer as the sum of a prime and the product of at most two primes. II. *Sci. Sinica* **21**(4) (1978), 421–430.

[12] R. Dressler and L. Pigno. The Haar measure of certain sets in the Bohr group. *Colloq. Math.* **41**(2) (1979), 297–301.

[13] M. Drmota, M. Kauers and L. Spiegelhofer. On a conjecture of Cusick concerning the sum of digits of n and $n + t$. *SIAM J. Discrete Math.* **30**(2) (2016), 621–649.

[14] N. Frantzikinakis, B. Host and B. Kra. The polynomial multidimensional Szemerédi theorem along shifted primes. *Israel J. Math.* **194**(1) (2013), 331–348.

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

[16] H. Furstenberg. Poincaré recurrence and number theory. *Bull. Amer. Math. Soc. (N.S.)* **5**(3) (1981), 211–234.

[17] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, Princeton, NJ, 1981.

[18] W. Gasarch, M. Villarino and K. Regan. Hilbert's proof of his irreducibility theorem. *Amer. Math. Monthly* **125**(6) (2018), 513–530.

[19] D. Glasscock. Norm forms represent few integers but relatively many primes. *Preprint*, 2019, arXiv:1705.00531.

[20] B. Green. Roth's theorem in the primes. *Ann. of Math. (2)* **161**(3) (2005), 1609–1636.

[21] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)* **167**(2) (2008), 481–547.

[22] B. Green and T. Tao. Linear equations in primes. *Ann. of Math. (2)* **171**(3) (2010), 1753–1850.

[23] J. T. Griesmer. Sumsets of dense sets and sparse sets. *Israel J. Math.* **190** (2012), 229–252.

[24] J. T. Griesmer. Separating topological recurrence from measurable recurrence: exposition and extension of Kříž's example. *Australas. J. Combin.*, to appear, arXiv:2108.01642.

[25] J. T. Griesmer, A. Le and T. H. Lê. Bohr sets in sumsets II: countable abelian groups. *Forum Math. Sigma* **11** (2023), Paper no. e57.

[26] H. Halberstam and H. Richert. *Sieve Methods (London Mathematical Society Monographs, 4)*. Academic Press [Harcourt Brace Jovanovich, Publishers], London, 1974.

[27] P. Halmos and H. Samelson. On monothetic groups. *Proc. Natl. Acad. Sci. USA* **28** (1942), 254–258.

[28] W. Huang and X. Wu. On the set of the difference of primes. *Proc. Amer. Math. Soc.* **145**(9) (2017), 3787–3793.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press

[29] T. Kamae and M. Mendès France. Van der Corput's difference theorem. *Israel J. Math.* **31**(3–4) (1978), 335–342.
[30] I. Kříž. Large independent sets in shift-invariant graphs: solution of Bergelson's problem. *Graphs Combin.* **3**(2) (1987), 145–158.
[31] K. Kunen and W. Rudin. Lacunarity and the Bohr topology. *Math. Proc. Cambridge Philos. Soc.* **126**(1) (1999), 117–137.
[32] A. Le and T. H. Lê. Bohr sets in sumsets I: compact groups. *Preprint*, 2021, arXiv:2112.11997. *Discrete Anal.* to appear.
[33] T. H. Lê. Intersective polynomials and the primes. *J. Number Theory* **130**(8) (2010), 1705–1717.
[34] T. H. Lê. Problems and results on intersective sets. *Combinatorial and Additive Number Theory—CANT 2011 and 2012 (Springer Proceedings in Mathematics and Statistics, 101)*. Ed. M. Nathanson. Springer, New York, 2014, pp. 115–128.
[35] H. W. Lenstra Jr and P. Stevenhagen. Chebotarev and his density theorem. *Math. Intelligencer* **18**(2) (1996), 26–37.
[36] H. Li and H. Pan. Difference sets and polynomials of prime variables. *Acta Arith.* **138**(1) (2009), 25–52.
[37] L. Matthiesen. Linear correlations amongst numbers represented by positive definite binary quadratic forms. *Acta Arith.* **154**(3) (2012), 235–306.
[38] J. Maynard. Small gaps between primes. *Ann. of Math. (2)* **181**(1) (2015), 383–413.
[39] J. Maynard. Dense clusters of primes in subsets. *Compos. Math.* **152**(7) (2016), 1517–1554.
[40] R. Nair. On polynomials in primes and J. Bourgain's circle method approach to ergodic theorems. II. *Studia Math.* **105**(3) (1993), 207–233.
[41] J. Pintz. Polignac numbers, conjectures of Erdős on gaps between primes, arithmetic progressions in primes, and the bounded gap conjecture. *From Arithmetic to Zeta-Functions*. Ed. J. Sander, J. Steuding and R. Steuding. Springer, Cham, 2016, pp. 367–384.
[42] A. Rice. Sárközy's theorem for P-intersective polynomials. *Acta Arith.* **157**(1) (2013), 69–89.
[43] A. Rice. A maximal extension of the best-known bounds for the Furstenberg–Sárközy theorem. *Acta Arith.* **187**(1) (2019), 1–41.
[44] K. Roth. On certain sets of integers. *J. Lond. Math. Soc. (2)* **28** (1953), 245–252.
[45] W. Rudin. *Fourier Analysis on Groups*. Dover Publications, Garden City, NY, 2017.
[46] A. Sárközy. On difference sets of sequences of integers. I. *Acta Math. Acad. Sci. Hungar.* **31**(1–2) (1978), 125–149.
[47] A. Sárközy. On difference sets of sequences of integers. III. *Acta Math. Acad. Sci. Hungar.* **31**(3–4) (1978), 355–386.
[48] E. Szemerédi. On the sets of integers containing no k elements in arithmetic progressions. *Acta Arith.* **27** (1975), 299–345.
[49] J. Thorner. Bounded gaps between primes in Chebotarev sets. *Res. Math. Sci.* **1** (2014), Article no. 4, 16pp.
[50] M. Wierdl. Pointwise ergodic theorem along the prime numbers. *Israel J. Math.* **64**(3) (1988), 315–336.
[51] Y. Zhang. Bounded gaps between primes. *Ann. of Math. (2)* **179**(3) (2014), 1121–1174.

https://doi.org/10.1017/etds.2024.73 Published online by Cambridge University Press