Problems on infinite sumset configurations in the integers and beyond

Bryna Kra   Joel Moreira   Florian K. Richter
Donald Robertson

January 10, 2024

Abstract

In contrast to finite arithmetic configurations, relatively little is known about which infinite patterns can be found in every set of natural numbers with positive density. Building on recent advances showing infinite sumsets can be found, we explore numerous open problems and obstructions to finding other infinite configurations in every set of natural numbers with positive density.

Contents

1   Introduction

A central topic in arithmetic combinatorics is the study of additive structures in the integers, such as arithmetic progressions and sumsets. The foundations for the area were laid by early results in Ramsey Theory. Van der Waerden [59] showed that in any finite coloring of the integers, there are arbitrarily long monochromatic arithmetic progressions. Thereafter Erdős and Turán [18] conjectured and Szemerédi [57] proved a density analog of van der Waerden's theorem: every set of natural numbers with positive density contains arbitrarily long arithmetic progressions. Many other finite arithmetic structures have been studied from a combinatorial perspective, including solutions to systems of linear Diophantine equations [53, 56], polynomial progressions [4], and sum-product configurations [49], to name a few. The techniques developed to tackle these problems are robust enough to play a role in many other contexts, leading to advances on multi-dimensional patterns [25], progressions in the primes [31, 32], and patterns in amenable groups [1]. There have been many and varied developments besides those mentioned here; for more we recommend as a starting point the surveys and monographs [3, 9, 20, 24, 29, 43].

This survey is about the arithmetic combinatorics of *infinite* configurations, a relatively young branch of arithmetic combinatorics that complements and contrasts both classical and ongoing work on finite configurations. What infinite configurations should one expect to be present in every set of natural numbers with positive upper density? It follows from Ramsey's theorem [54] that, no matter how the elements of $\mathbb{N}$ are colored using finitely many colors, there is an infinite set $B \subset \mathbb{N}$ such that

$$\{b_1 + b_2 : b_1, b_2 \in B, \ b_1 \neq b_2\}$$

is monochromatic. Erdős made several conjectures that serve as density analogs of this result. For example, amongst them was the statement that every positive density set of natural numbers contains the sum

$$B + C := \{b + c : b \in B, \ c \in C\}$$

of two infinite sets $B, C \subset \mathbb{N}$. Motivated by recent developments [2, 8, 12, 41, 46, 47, 50, 51] on Erdős's conjectures, we explore various new directions of research pertaining to infinite sumset phenomena in the integers and beyond. We supply examples that provide natural obstructions to infinite sumset configurations and collect open problems of interest. Our focus is predominantly on questions and conjectures within the area of infinitary density combinatorics; for a survey of problems in infinitary partition regularity see [39].

Notions of density

Throughout we let $\mathbb{N} = \{1, 2, 3, \ldots\}$ denote the positive integers and for $k \in \mathbb{N}$, let $k\mathbb{N} = \{kn : n \in \mathbb{N}\}$ denote the set of all positive multiples of k. There are various notions of density we can consider on $\mathbb{N}$, including upper and lower natural density, and upper Banach density. We recall here those that we use most frequently.

If A is a set of natural numbers, define its *upper Banach density* $\mathsf{d}^*(A)$ to be

$$\mathsf{d}^*(A) := \limsup_{N-M \to \infty} \frac{\left|A \cap \{M+1, M+2, \ldots, N\}\right|}{N - M}. \tag{1}$$

The *upper density* $\overline{\mathsf{d}}(A)$ and *lower density* $\underline{\mathsf{d}}(A)$ of $A \subset \mathbb{N}$ are defined respectively as

$$\overline{\mathsf{d}}(A) = \limsup_{N \to \infty} \frac{|A \cap \{1, \ldots, N\}|}{N} \quad \text{and} \quad \underline{\mathsf{d}}(A) = \liminf_{N \to \infty} \frac{|A \cap \{1, \ldots, N\}|}{N}.$$

If $\overline{\mathsf{d}}(A) = \underline{\mathsf{d}}(A)$, then we call this number the *natural density* of A and denote it by $\mathsf{d}(A)$. We formulate most of our conjectures and statements in terms of upper Banach density, the weakest measurement of largeness among these notions, but for most questions and conjectures one can ask about the analog under a stronger density assumption.

In some settings, we use density with respect to a *Følner sequence* in $\mathbb{N}$, meaning a sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ of finite subsets of $\mathbb{N}$ satisfying

$$\lim_{N \to \infty} \frac{|(\Phi_N + t) \cap \Phi_N|}{|\Phi_N|} = 1$$

for all $t \in \mathbb{N}$, where $\Phi_N + t$ denotes the shift of the set Φ_N by t.

Analogous to upper and lower density, the *upper* and *lower density along* Φ of a set $A \subset \mathbb{N}$ are defined as

$$\overline{\mathsf{d}}_\Phi(A) = \limsup_{N \to \infty} \frac{|A \cap \Phi_N|}{|\Phi_N|} \quad \text{and} \quad \underline{\mathsf{d}}_\Phi(A) = \liminf_{N \to \infty} \frac{|A \cap \Phi_N|}{|\Phi_N|}.$$

When the quantities above are equal, we denote the common value by $\mathsf{d}_\Phi(A)$ and call it the *density* of A along Φ.

As needed, we make use of the natural analogs of Følner sequences and densities along Følner sequences in the setting of $\mathbb{N}^d$ for $d \geq 2$ See Section 5.4 for the precise definition of Følner sequences in even more general settings.

Acknowledgements

2 Infinite sumsets in sets of integers with positive density

2.1 Searching for a density version of Hindman's theorem

We begin by exploring what types of infinite sumset configurations can be found in subsets of the integers with positive upper Banach density. Results of this type can be viewed as density analogs of one of the central results in additive combinatorics: Hindman's theorem. Given a set $B \subset \mathbb{N}$, its *finite sums set* is defined by

$$\mathsf{FS}(B) = \left\{ \sum_{n \in F} n : F \subset B \text{ with } 0 < |F| < \infty \right\} \tag{2}$$

and consists of all natural numbers obtained by adding together finitely many distinct elements in B. When B is infinite we say that $\mathsf{FS}(B)$ is an *IP-set*[1].

---

[1]Depending on the source, the acronym IP stands for *infinite parallelepiped* [27] or *idempotent* [26], where the latter relates to the Stone-Čech compactification of the integers.

Hindman's Theorem ([35])**.** *If the natural numbers are partitioned into finitely many cells then at least one of these cells contains an IP-set.*

It follows from Ramsey's theorem [54] that for any finite coloring of $\mathbb{N}$ and any $k \in \mathbb{N}$ there exists an infinite set $B \subset \mathbb{N}$ such that

$$\left\{ \sum_{n \in F} n : F \subset B_k \text{ with } |F| = k \right\}$$

is monochromatic. However, before Hindman's theorem it was not even known that infinite monochromatic configurations of the form $B \cup \{b_1 + b_2 : b_1, b_2 \in B, b_1 \neq b_2\}$ exist for any fininte coloring of $\mathbb{N}$ (see [39, Question 12] and the surrounding discussion).

In a first attempt to formulate a density version of Hindman's theorem, Erdős asked whether every set $A \subset \mathbb{N}$ with positive upper density has a shift

$$A - t := \{n \in \mathbb{N} : t + n \in A\}$$

for some integer $t \geq 0$ that contains an IP-set. Note that the shift is necessary because the set of odd numbers has density $1/2$, yet it does not contain a triple of the form $\{x, y, x+y\}$. However, Straus provided an example (not published by Straus, but published by others, for example in [37, Theorem 11.6]), showing that Erdős's question has a negative answer.

Example 2.1 (Straus example)**.** Let (p_t) be a sufficiently fast growing sequence of primes and let

$$A := \mathbb{N} \setminus \left(\bigcup_{t \in \mathbb{N}} (p_t \mathbb{N} + t) \right)$$

Then $\mathsf{d}(A) \geq 1 - \sum_{t=1}^{\infty} \frac{1}{p_t} > 0$ but any shift $A - t$ of A contains only finitely many multiples of p_t and hence cannot contain an IP-set.

Undeterred by this example and in light of the ongoing advances surrounding Szemerédi's and Hindman's theorems at the time, Erdős [14, Page 305] wrote:

> I have tried to formulate a conjecture which would be in the same relation to Hindman's theorem as Szemerédi's theorem is to van der Waerden's. I have not been very successful so far. Perhaps the following result holds. Let $a_1 < a_2 < \ldots$ be a sequence of integers with positive upper density. Then there is an integer t and an infinite subsequence $a_{i_1} < a_{i_2} < \ldots$ so that all the sums $a_{i_r} + a_{i_s} + t$ are again a's.

Erdős's proposed result, together with the weaker conjecture from [15, Pages 57–58] and [16, Page 105], were proved by the authors.

Theorem 2.2 ([47, Theorem 1.2, part (i)])**.** *Every set $A \subset \mathbb{N}$ of positive upper Banach density contains a sumset*

$$t + \left\{ b_1 + b_2 : b_1, b_2 \in B, \ b_1 \neq b_2 \right\}$$

for some shift $t \geq 0$ and some infinite set $B \subset A$.

We note that there are examples showing that one cannot remove the restriction $b_1 \neq b_2$. One such example was communicated to us by Steven Leth.

Example 2.3. For $c \in (1, 2)$ consider the set

$$A = \mathbb{N} \cap \bigcup_{n \in \mathbb{N}} [4^n, c \cdot 4^n), \tag{3}$$

which has $\mathsf{d}^*(A) = 1$ and $\underline{\mathsf{d}}(A) > 0$. The set A does not contain any set of the form $\{b, 2b\}$ for $b \in \mathbb{N}$. When b is much larger than b' one cannot find $\{b + b', 2b\}$ in A either. One can use this observation to show that A does not contain an unrestricted sumset $B + B$ for an infinite set B. In fact, every shift of A has the same property, and so A cannot contain any set of the form $B + B + t$ with $B \subset \mathbb{N}$ infinite and $t \geq 0$ integer. The proofs of these statements are deferred to the more general setting of Example 3.6.

In an attempt to meet Erdős's aspirations for a density version of Hindman's theorem, we propose several conjectures that extend Theorem 2.2 in different directions without encroaching on the example of Straus. One of the most natural conjectures along these lines is the following.

Conjecture 2.4 ([47, Conjecture 1.5]). *Fix $k \in \mathbb{N}$. Every set of positive upper Banach density $A \subset \mathbb{N}$ has a shift $A - t$ which contains the set*

$$\left\{ \sum_{n \in F} n : F \subset B \text{ with } 0 < |F| \leq k \right\}$$

for some integer $t \geq 0$ and some infinite set $B \subset \mathbb{N}$.

We stress that one cannot refine Conjecture 2.4 by choosing t independent of k because the example of Straus does not contain such configurations.

The case $k = 2$ of Conjecture 2.4, which corresponds to a slight variation of Theorem 2.2, is proved in [47, Theorem 1.2, part (ii)]. For $k \geq 3$ even the following special case of Conjecture 2.4 remains open.

Conjecture 2.5. *Fix $k \in \mathbb{N}$. Every set of positive upper Banach density $A \subset N$ has a shift $A - t$ which contains a sumset*

$$B^{\oplus k} := \left\{ \sum_{n \in F} n : F \subset B \text{ with } |F| = k \right\} \tag{4}$$

for some integer $t \geq 0$ and some infinite set $B \subset \mathbb{N}$.

Throughout the rest of the paper, we continue using the notation $B^{\oplus k}$ introduced in Conjecture 2.5, and for the case $k = 2$, where there is no ambiguity, we use the more explicit notation $B \oplus B$ instead of $B^{\oplus 2}$.

The odd integers make it clear that some shift t is needed in Conjecture 2.5. However, it is possible to restrict the value of the shift to the set $\{0, \ldots, k-1\}$ by analysing which congruence class modulo k is preferred by the set A. In particular, if the intersection $A \cap k\mathbb{Z}$ has positive density, then we can take $t = 0$. We summarize this in the next proposition.

Proposition 2.6 (cf. [37, Theorem 11.8]). *The following are equivalent.*

1. *Conjecture 2.5.*

2. *For every $k \in \mathbb{N}$ and every $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$ there exist an infinite set $B \subset \mathbb{N}$ and integer $t \in \{0, \ldots, k-1\}$ with $B^{\oplus k} + t \subset A$.*

3. *For every $k \in \mathbb{N}$ and every $A \subset \mathbb{N}$ with $\mathsf{d}^*(A \cap k\mathbb{N}) > 0$ there exists an infinite set $B \subset \mathbb{N}$ with $B^{\oplus k} \subset A$.*

Proof. Clearly the second statement implies the first.

To see how the third statement implies the second, let $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$ and find $t \in \{0, \ldots, k-1\}$ such that $A' := A \cap (k\mathbb{N} + t)$ has $\mathsf{d}^*(A') > 0$. Then $A' - t \subset k\mathbb{N}$ and by the third statement there exists an infinite set $B \subset \mathbb{N}$ with $B^{\oplus k} \subset A' - t$.

To show that the first statement implies the third, let $A \subset k\mathbb{N}$ with $\mathsf{d}^*(A) > 0$ and, using the first statement, find $t \geq 0$ and an infinite set $B \subset \mathbb{N}$ such that $B^{\oplus k} + t \subset A \subset k\mathbb{N}$. Now B must contain k elements which are all congruent modulo k, and therefore their sum is a multiple of k. It follows that $t = k\tilde{t}$ for some $\tilde{t} \in \mathbb{Z}$. Letting $\tilde{B} = B + \tilde{t}$ we conclude that $\tilde{B}^{\oplus k} \subset A$. $\qquad\square$

We continue the discussion on the role played by t in Conjecture 2.4 and related contexts in Section 3.5.

2.2   Sums of distinct infinite sets

One can weaken Conjecture 2.5 by replacing (4) with the sum

$$B_1 + \cdots + B_k := \big\{ b_1 + \cdots + b_k : b_1 \in B_1, \ldots, b_k \in B_k \big\}$$

of k infinite sets $B_1, \ldots, B_k \subset \mathbb{N}$. It is then unnecessary to require a shift t as it can be absorbed into one of the infinite sets. The $k = 2$ case of this weakening was established in 2019.

Theorem 2.7 ([50])**.** *If $A \subset \mathbb{N}$ has positive upper Banach density, there exist infinite sets $B, C \subset \mathbb{N}$ such that $B + C \subset A$.*

In [41], Host gave an ergodic theoretic proof of Theorem 2.7, finding the proper ergodic analog for some of the ideas in [50]. The techniques used in both [50] and [41] seem to be unable to handle higher order sumsets. In [46] a different ergodic proof of Theorem 2.7 was discovered, building on the ideas of [41], which works for higher-order sumsets.

Theorem 2.8 ([46, Theorem 1.1])**.** *For every set $A \subset \mathbb{N}$ of positive upper Banach density and integer $k \in \mathbb{N}$, there exist infinite sets $B_1, \ldots, B_k \subset \mathbb{N}$ such that $B_1 + \cdots + B_k \subset A$.*

The conclusion of Theorem 2.8 follows from Conjecture 2.5 by splitting the set B into k disjoint infinite subsets. This motivates several variations of the conjectures and questions posed in Section 2.1, starting with the following.

Conjecture 2.9. *Given $A \subset \mathbb{N}$ with positive upper Banach density and $k \in \mathbb{N}$ there are infinite sets $B_1, \ldots, B_k \subset \mathbb{N}$ such that $B_1 + \cdots + B_i \subset A$ for all $i = 1, \ldots, k$.*

We note that Conjecture 2.9 follows from Conjecture 2.4: if B and t satisfy the conclusion of the latter, then partitioning B into k disjoint infinite subsets $B = B_1' \cup \cdots \cup B_k'$ and letting $B_1 = B_1' + t$ and $B_i = B_i'$ for all $i > 1$ we obtain the conclusion of Conjecture 2.9.

One may be tempted to ask whether A contains all sums

$$B_{i(1)} + \cdots + B_{i(r)}$$

for all $1 \leq i(1) < \cdots < i(r) \leq k$ for some r with $2 \leq r \leq k$. However, there are local obstructions to this and in particular the set $A = 2\mathbb{N} - 1$ does not contain such a configuration.

Another possibility is to strengthen Conjecture 2.9 by essentially making k infinite.

Conjecture 2.10. *Given $A \subset \mathbb{N}$ with positive upper Banach density, there is an infinite sequence of infinite sets $B_1, B_2, \ldots \subset \mathbb{N}$ such that $B_1 + \cdots + B_i \subset A$ for every $i \in \mathbb{N}$.*

We end this section with two related but more ambitious questions. While the Straus Example precludes the possibility that a single shift of an arbitrary set A of positive Banach density contains an IP-set, it may be possible to have a single infinite set B and different shifts depending on the number of summands.

Question 2.11. *Given $A \subset \mathbb{N}$ with positive upper Banach density, does there exist an infinite set $B \subset \mathbb{N}$ such that for every $k \in \mathbb{N}$ there exists a shift $t_k \geq 0$ with*

$$\left\{ \sum_{n \in F} n : F \subset B \text{ with } 0 < |F| \leq k \right\} \subset A - t_k.$$

It is not clear if a positive answer is to be expected, but we were unable to find a counterexample. Perhaps a slightly weaker version where one is allowed to disregard finitely many elements of B as the number of summands increases has higher chances of a positive answer.

Question 2.12. *Given $A \subset \mathbb{N}$ with positive upper Banach density, does there exist an infinite set $B \subset \mathbb{N}$ such that for all $k \in \mathbb{N}$ there is a co-finite set $B' \subset B$ and a shift $t_k \geq 0$ with*

$$\left\{ \sum_{n \in F} n : F \subset B' \text{ with } 0 < |F| \leq k \right\} \subset A - t_k.$$

We note that a positive answer to Question 2.11 implies all the conjectures listed in this section. See Fig. 1 for a schematic.
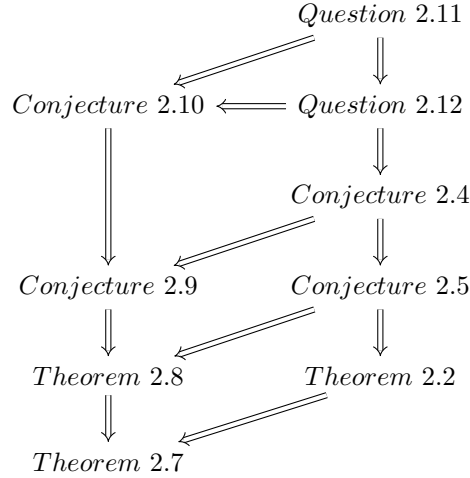
Figure 1: Summary of the relations among the questions (assuming positive answers), conjectures, and theorems in Section 2.

3 Refinements of sumsets in positive density

3.1 Constraining the summands

Theorem 2.7 states that any set $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$ contains a sumset $B + C$ for two infinite sets $B, C \subset \mathbb{N}$. It is natural to ask whether one can impose certain restrictions on the sets B and C. We first show as a straightforward consequence of Theorem 2.7 that, up to a translate, one can draw B and C from any subsets of the integers with bounded gaps. Subsets of $\mathbb{N}$ with bounded gaps are called *syndetic*. Equivalently, a set $S \subset \mathbb{N}$ is syndetic if there exists $h \in \mathbb{N}$ such that $(S - 1) \cup (S - 2) \cup \cdots \cup (S - h) = \mathbb{N}$.

Proposition 3.1. *Fix $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$. For any syndetic subsets P and Q of $\mathbb{N}$ there is $P' \subset P$ infinite and $Q' \subset Q$ infinite and $t \geq 0$ with $P' + Q' + t \subset A$.*

Proof. Let $B, C \subset \mathbb{N}$ be infinite sets with $B + C \subset A$. Since P and Q are syndetic, we can find shifts $p, q \in \mathbb{N}$ such that $(P + p) \cap B$ and $(Q + q) \cap C$ are both infinite. Take

$$P' = ((P + p) \cap B) - p = P \cap (B - p)$$
$$Q' = ((Q + q) \cap C) - q = Q \cap (C - q)$$

and $t = p + q$. We have that $P' + Q' \subset B + C - t \subset A - t$. $\qquad\square$

The shift t in Proposition 3.1 is needed, as taking both P and Q to be the even numbers and A to be the odd numbers shows. Next we consider the situation where only C is constrained; in this setup a shift is no longer needed as it can be absorbed into B. As the following proposition shows, in order to be a repository for C a set need not be syndetic across all of $\mathbb{N}$ but it must be "visible locally along a Følner sequence that sees A".

Proposition 3.2. *Let $W \subset \mathbb{N}$ and let Φ be a Følner sequence for which $\mathsf{d}_\Phi(W)$ exists. The following are equivalent.*

1. *For every $A \subset \mathbb{N}$ with $\overline{\mathsf{d}}_\Phi(A) > 0$, there exist infinite sets $B \subset \mathbb{N}$, $C \subset W$ such that $B + C \subset A$.*

2. *$\mathsf{d}_\Phi(W) > 0$.*

Proof. The implication $(2) \Rightarrow (1)$ follows from a careful analysis of the proof in [50] (or the proofs in [41] and in [46]).

The proof of $(1) \Rightarrow (2)$ is by construction. If $\mathsf{d}_\Phi(W) = 0$ we can find a slowly growing sequence $f \colon \mathbb{N} \to \mathbb{N}$ such that $\lim f(n) = \infty$ and such that the set

$$\tilde{W} := \bigcup_{w \in W} [w - f(w), w + f(w)]$$

still has $\mathsf{d}_\Phi(\tilde{W}) = 0$. Letting $A = \mathbb{N} \setminus \tilde{W}$ we see that $\mathsf{d}_\Phi(A) = 1$. Whenever A contains a sumset $B + C$ and $C \subset W$, if $b \in B$ and $w \in W$ are such that $f(w) > b$, then we cannot have $w \in C$ (as $w + b \notin A$). It follows that C must be finite, finishing the proof. $\qquad\square$

Among other things, Proposition 3.2 implies that not every set $A \subset \mathbb{N}$ with $\mathsf{d}(A) > 0$ contains a sumset $B + C \subset A$ with $B, C \subset \mathbb{N}$ infinite and C consisting entirely of squares or of primes. (However, see Section 3.7 for positive results pertaining to sumsets with restrictions to squares, primes, or other sparse subsets of $\mathbb{N}$.)

While Proposition 3.2 involves local restrictions with respect to a fixed Følner sequence, the following two propositions consider more global assumptions on A and W.

Definition 3.3. *Recall that a set $T \subset \mathbb{N}$ is thick if it contains arbitrarily long blocks of consecutive integers. The properties of being thick and being syndetic are dual in the following sense: a set is thick if and only if it has nonempty intersection with every syndetic set, and vice versa. A set $P \subset \mathbb{N}$ is piecewise syndetic if there is a syndetic set $S \subset \mathbb{N}$ and a thick set $T \subset \mathbb{N}$ such that $P = S \cap T$.*

Proposition 3.4. *Let $W \subset \mathbb{N}$. The following are equivalent.*

1. *For every thick set $A \subset \mathbb{N}$ there exist infinite sets $B \subset \mathbb{N}$, $C \subset W$ such that $B + C \subset A$.*

2. *For every piecewise syndetic set $A \subset \mathbb{N}$ there exist infinite sets $B \subset \mathbb{N}$, $C \subset W$ such that $B + C \subset A$.*

3. *W is syndetic.*

Proof. The implication $(3) \Rightarrow (2)$ follows from Proposition 3.1 and the implication $(2) \Rightarrow$ (1) is trivial. For the final direction $(1) \Rightarrow (3)$, assume that W is not syndetic and so for every $n \in \mathbb{N}$, there exist $a_n, b_n \in \mathbb{N}$ with $|b_n - a_n| \geq 2n$ and such that $W \cap [a_n, b_n] = \varnothing$. Then the set $A = \bigcup_{n \in \mathbb{N}} [a_n + n, b_n]$ is a thick set with the property that any shift of it $A - t$ has finite intersection with W. Therefore one can not find infinite sets $B \subset \mathbb{N}$, $C \subset W$ such that $B + C \subset A$. $\qquad\square$

Proposition 3.5. *Let $W \subset \mathbb{N}$. The following are equivalent.*

1. *For every syndetic set $A \subset \mathbb{N}$ there exist infinite sets $B \subset \mathbb{N}$, $C \subset W$ such that $B + C \subset A$.*

2. *W is infinite.*

Proof. The implication $(1) \Rightarrow (2)$ is obvious. The converse implication uses Ramsey's theorem. Assuming W is infinite and A is syndetic, find $r \in \mathbb{N}$ such that $A \cup (A - 1) \cup \cdots \cup (A - r) = \mathbb{N}$. Then color each pair $\{w_1, w_2\}$ of elements of W with the color $i \in \{0, \ldots, r\}$ such that $w_1 + w_2 + i \in A$. Using Ramsey's theorem one can extract an infinite subset $W' \subset W$ such that every pair $\{w_1, w_2\} \subset W'$ has the same color i. Next split W' into two disjoint infinite sets $W' = C \cup B'$ and let $B = B' + i$. It follows that $C \subset W$ and $B + C \subset A$. $\qquad\square$

3.2 Combinatorial obstructions to $B + B + t$

Recall Theorem 2.2 which guarantees that every set of positive upper Banach density contains a sumset $\{b_1 + b_2 : b_1, b_2 \in B, b_1 \neq b_2\}$ for some infinite set $B \subset \mathbb{N}$. Example 2.3 shows that the restriction $b_1 \neq b_2$ is necessary. More generally, examples in which one cannot find a configuration of the form $\{b, 2b\}$, or perturbations thereof such as $\{b + b', 2b\}$, provide obstructions to many reasonable questions one might pose about sumsets in positive density sets. Nonetheless, beyond a particular density threshold, no such obstructions arise and the following example closes in on these limitations.

Example 3.6. Consider the set

$$A := \left\{ m \in \mathbb{N} : \left\{ \theta \log_2(m) \right\} \in U \right\}$$

for any $\theta > 0$ and any set $U \subset \mathbb{T}$ for which U and $U + \theta$ (calculated mod 1) are separated by a positive distance. Taking $\theta = 1/2$ and $U = \left[0, \frac{1}{2} \log_2(c) \right)$, we recover the set in (3) of Example 2.3.

We claim that A does not contain $B + B + t$ for any infinite set $B \subset \mathbb{N}$ and shift $t \geq 0$. Indeed, for any infinite set $B \subset \mathbb{N}$ and $t \geq 0$, we can take $b' \in B$ arbitrary and then choose $b \in B$ sufficiently large so that the difference

$$\left| \log_2 \left(\frac{2b + t}{b + b' + t} \right) - 1 \right|$$

is less than the distance $d(U, U + \theta)/\theta$. Multiplying by θ and reducing modulo one, it follows that one can not have both $b + b' + t$ and $2b + t$ in A.

Taking $U = [0, 1/2 - \varepsilon]$ for small $\varepsilon > 0$ and $\theta = 1/2$, we have that

$$\overline{\mathsf{d}}(A) = \frac{2}{3} \cdot (2 - 4^{\varepsilon})$$

As $\varepsilon \to 0$ we obtain an example A with upper density arbitrarily close to $2/3$. On the other hand, taking $U = [0, 1/2 - \varepsilon]$ and setting $\theta = 1/2 + n$ for some large $n \in \mathbb{N}$ (this

yields the same set A as setting $\theta = 1/2$ and $U = T^{-1}[0, 1/2 - \varepsilon]$ where $T \colon \mathbb{T} \to \mathbb{T}$ is multiplication by $2n + 1$) one can check that

$$\underline{\mathsf{d}}(A) = \frac{2^{(1/2 - \varepsilon)/\theta} - 1}{2^{1/\theta} - 1}.$$

As $\varepsilon \to 0$ and $\theta \to \infty$ it follows that $\underline{\mathsf{d}}(A)$ comes arbitrarily close to $1/2$.

This leads us to formulate a conjecture on the threshold for the density of obstructions.

Conjecture 3.7. *Let $A \subset \mathbb{N}$.*

1. *If $\overline{\mathsf{d}}(A) > 2/3$, there is an infinite set $B \subset \mathbb{N}$ such that $A \supset B + B$.*

2. *If $\underline{\mathsf{d}}(A) > 1/2$, there is an infinite set $B \subset \mathbb{N}$ such that $A \supset B + B$.*

More generally, it would be interesting to know what properties of $A \subset \mathbb{N}$ guarantee that it contains $B + B + B$ for some infinite set $B \subset \mathbb{N}$. In particular, does any density assumption on A suffice? See also Section 3.4 below for related questions.

We finish with a brief mention of partition regularity of $B + B$ with $B \subset \mathbb{N}$ infinite. Hindman [36, Section 2] produced a three-coloring of the integers that does not admit a monochromatic set of the form $B + B$ for any infinite set $B \subset \mathbb{N}$, but the following question of J. C. Owings is still open.

Question 3.8 ([45, Problem E2494])**.** *Is it true that, no matter how $\mathbb{N}$ is partitioned into two sets, one of the sets must contain $B + B$ for some infinite set $B \subset \mathbb{N}$?*

3.3 Ordered sums

In this section we consider generalizations of $B \oplus B$ of the form

$$\big\{ f(b_1) + g(b_2) : b_1, b_2 \in B, \ b_1 < b_2 \big\}, \tag{5}$$

where f and g are polynomials over $\mathbb{N}$. Note that when $f(x) = g(x) = x$, the expression in (5) is precisely the sumset $B \oplus B$. Although we do not formalize the terminology, we think of expressions like (5) as an "ordered sum" of $f(B)$ with $g(B)$. It is immaterial whether one interprets the ordered sum of $f(B)$ with $g(B)$ as

$$\big\{ f(b_1) + g(b_2) : b_1, b_2 \in B, \ b_1 < b_2 \big\} \text{ or } \big\{ r + s : r \in f(B), \ s \in g(B), \ r < s \big\}$$

because one can obtain either as a subset of the other by refining B. One could also consider ordered sumsets of unrelated sets B and C; some results and open questions on such sumsets are collected in Section 3.6.

In the following example we show that the restriction on the ordering $b_1 < b_2$ cannot in general be replaced with $b_1 \neq b_2$, even when f and g are both linear.

Example 3.9. Fix real numbers $\lambda_1 > \lambda_2 > 0$. We build, using the set in Example 3.6 as a model, a finite coloring of $\mathbb{N}$ such that for any infinite set $B \subset \mathbb{N}$ and any shift $t \geq 0$, the set

$$\big\{ \lfloor \lambda_1 b_1 + \lambda_2 b_2 \rfloor : b_1, b_2 \in B, b_1 \neq b_2 \big\} + t$$

is not monochromatic. Let

$$\chi(n) = \lfloor \log_\rho(n) \rfloor \bmod 3$$

where $\rho = \sqrt{\lambda_1/\lambda_2} > 1$. Fix $t \in \mathbb{N}$ and $B \subset \mathbb{N}$ infinite and let $b_1 \in B$ be arbitrary. Let $b \in B$ be sufficiently large depending on λ_1, λ_2, t and b_1. Let $m_1 = \lfloor \lambda_1 b_1 + \lambda_2 b \rfloor + t$ and $m_2 = \lfloor \lambda_1 b + \lambda_2 b_1 \rfloor + t$. Note that $\rho^2 m_1 \approx m_2$ as $b \to \infty$, so

$$\rho m_1 < m_2 < \rho^3 m_1$$

10

if we choose b large enough. Taking $\log_\rho$ on each term we conclude that

$$\log_\rho(m_2) \in \Big(\log_\rho(m_1) + 1, \log_\rho(m_1) + 3\Big)$$

and after taking floors and reducing modulo 3 we get $\chi(m_1) \neq \chi(m_2)$.

Note that Ramsey's theorem implies (5) is partition regular. This motivates the question whether instances of (5) can be found (up to a shift) in every set with positive density. Our first conjecture deals with the special case when f and g are linear polynomials.

Conjecture 3.10. *If $A \subset \mathbb{N}$ has positive upper Banach density and $\ell, m \in \mathbb{N}$ then there exist an infinite set $B = \{b_1 < b_2 < \cdots\}$ and a shift $t \geq 0$ such that*

$$\{\ell b_i + m b_j : i < j\} \tag{6}$$

is contained in $A - t$.

If $\ell = m$ then Conjecture 3.10 follows from Theorem 2.2.

The following question asks whether it is possible to extend Conjecture 3.10 in a way that would also contain Szemerédi's theorem as a special case (see Section 3.6 for another way to combine Szemerédis's theorem with sumsets).

Question 3.11. *Fix $k \in \mathbb{N}$. If $A \subset \mathbb{N}$ has positive upper Banach density, then does there exists an infinite set $B = \{b_1 < b_2 < \cdots\}$ and a shift $t \geq 0$ such that the set*

$$\bigcup_{\ell=0}^{k} \{\ell b_i + b_j : i < j\} \tag{7}$$

is contained in $A - t$?

If in Question 3.11 one asks for the set B to be arbitrarily large but finite, then a positive answer follows immediately from Szemerédi's theorem, by taking $b_i = ib$ and $t = a$ for some long arithmetic progression $a, a + b, \ldots, a + Lb$ contained in A. It also follows from Rado's theorem [53] that whenever $\mathbb{N}$ is partitioned into finitely many pieces, at least one of the pieces contains (7) for an arbitrarily long but finite sequence $b_1 < b_2 < \cdots < b_N$ without the need for a shift. For B infinite, even the partition variant of Question 3.11 is not known to us; in fact we do not even know whether in every finite partition of $\mathbb{N}$ there exists a cell containing a configuration $\{b_j, 2b_i + b_j : i < j\}$.

Surprisingly, Question 3.11 has a negative answer when (7) is replaced by

$$\bigcup_{\ell=0}^{k} \{b_i + \ell b_j : i < j\} \tag{8}$$

because otherwise we would find $\{b' + b, b' + 2b\}$ within $A - t$ where b is much larger than b', but these are the sort of configurations that Example 2.3 precludes. The same reasoning shows that there are finite colorings of $\mathbb{N}$ that do not admit monochromatic sets of the form (8).

This leads us to inquire which ordered sumset configurations are actually possible.

Question 3.12. *Let $k \in \mathbb{N}$ and $(a_0, d_0), \ldots, (a_k, d_k) \in \mathbb{N}^2$. What are necessary and/or sufficient conditions on $(a_0, d_0), \ldots, (a_k, d_k)$ such that for any set $A \subset \mathbb{N}$ with positive upper Banach density there exist an infinite set $B = \{b_1 < b_2 < \cdots\}$ and a shift $t \geq 0$ with*

$$\bigcup_{\ell=0}^{k} \{a_\ell b_i + d_\ell b_j : i < j\}$$

contained in $A - t$?

Again, the order should play a role, as we expect that (7) appears, up to a shift, in every positive density set, when we know that (8) does not. The reason is that when one considers ordered sumsets, the first and second summands seem to follow different rules. We expect that the rules for restricting the first summand are similar to those for restricting the gaps in sets of positive density. The source of these rules can be understood dynamically as refinements of Poincaré recurrence; more on this perspective is discussed in Section 3.7 and Theorem 3.32 in particular. On the other hand, the second summand can sometimes be drawn from the given set of positive density, but there do not seem to be any arithmetic restrictions that we can place on it.

The Furstenberg-Sárközy theorem [22, 55] states that every set of positive density contains a square difference. This motivates the following question.

Question 3.13. *If $A \subset \mathbb{N}$ has positive upper Banach density, must there exists an infinite set $B = \{b_1 < b_2 < \cdots\}$ and $t \geq 0$ such that*

$$\{b_i^2 + b_j : i < j\} \tag{9}$$

is contained in $A - t$?

Since the configuration (9) is a special case of (5), an application of Ramsey's theorem shows that it is partition regular. By the same argument, the pattern $\{b_i + b_j^2 : i < j\}$ is also partition regular. However, the next example shows that in Question 3.13 one cannot replace (9) with $\{b_i + b_j^2 : i < j\}$, lending further credence to the heuristic that we can not place arithmetic restrictions on the second summand.

Example 3.14. Consider the set

$$A = \mathbb{N} \setminus \bigcup_{k \in \mathbb{N}} \left(k^2 - \log(k), k^2 + \log(k) \right)$$

which has full density, yet no shift of it contains $\{b_i^2 + c_j : i > j\}$ where $B = \{b_1 < b_2 < \cdots\}$ and $C = \{c_1 < c_2 < \cdots\}$ are infinite. Picking $b_i = c_i$ shows that the order in Question 3.13 cannot be reversed, and picking $c_i = b_i^2$ shows that there exists a full-density set with the property that no shift of it contains the ordered sumset of an infinite set of perfect squares.

We note that the only property of the squares that was used to produce this counterexample is its sparsity. We could replace the set of squares with any set of zero density (such as the prime numbers) and obtain analogous counterexamples.

Question 3.13 is inspired by the Furstenberg-Sárközy theorem, but a positive answer to the question does not imply it. Our next question asks for a common generalization of both.

Question 3.15. *If $A \subset \mathbb{N}$ has positive upper Banach density, must there exists an infinite set $B = \{b_1 < b_2 < \cdots\}$ and $t \geq 0$ such that*

$$\{b_j, b_i^2 + b_j : i < j\} \tag{10}$$

is contained in $A - t$?

We conclude this section with the following generalization of Questions 3.13 and 3.15 inspired by the polynomial Szemerédi theorem of Bergelson and Leibman [4].

Question 3.16. *Let $p_1, \ldots, p_k \in \mathbb{Z}[x]$ satisfy $p_\ell(0) = 0$ for all $\ell = 1, \ldots, k$ and fix $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$. Can one find an infinite set $B = \{b_1 < b_2 < \cdots\}$ and $t \geq 0$ such that*

$$\left\{ p_1(b_i) + b_j, \, p_2(b_i) + b_j, \, \ldots, \, p_\ell(b_i) + b_j : i < j \right\}$$

is contained in $A - t$?

Remark 3.17. Though Questions 3.15 and 3.16 are both about sets of positive density, the corresponding partition questions are also interesting and open.

We finish with a variant of Question 3.13 where instead of squares one considers a product set.

Question 3.18. *If $A \subset \mathbb{N}$ has positive upper Banach density, must there exists an infinite set $B = \{b_1 < b_2 < \cdots\}$ and $t \geq 0$ such that*

$$\{b_i b_j + b_k : i < j < k\}$$

is contained in $A - t$?

Similar to the way Question 3.13 is generalized in Questions 3.15 and 3.16, one can inquire about analogous generalizations of Question 3.18.

3.4 Sumsets of the form $B + mB$

Example 3.9 provides for each $m \in \mathbb{N}$ a 3-coloring of $\mathbb{N}$ without an infinite sumset of the form $B + mB$ with B infinite. The coloring in the example depends heavily on the choice of m, raising the question whether this dependence is unavoidable.

Question 3.19. *Is it true that for any finite coloring of $\mathbb{N}$, there is an infinite set $B \subset \mathbb{N}$ and some $m \in \mathbb{N}$ such that $B + mB$ is monochromatic?*

It may be that the set of problematic m is actually small in a multiplicative sense; we make this precise in the next question.

Question 3.20. *Is it true that for any finite coloring of $\mathbb{N}$, there exists $q \in \mathbb{N}$ such that for all $m \in \mathbb{N}$ with $\gcd(m, q) = 1$ there is an infinite set $B \subset \mathbb{N}$ such that $B + mB$ is monochromatic?*

One could also ask for a density version of Question 3.19, but as the following example shows, some care is needed. Note that every thick (Definition 3.3) set $T \subset \mathbb{N}$ has $\mathsf{d}^*(T) = 1$ but need not have $\overline{\mathsf{d}}(T) = 1$.

Example 3.21. There exists a thick set $A \subset \mathbb{N}$ with $\overline{\mathsf{d}}(A) > 0$ such that whenever

$$A - t \supset \{b_1 + mb_2 : b_1, b_2 \in B, b_1 \neq b_2\} \tag{11}$$

for some $m \geq 1$, $t \geq 0$, and set $B \subset \mathbb{N}$, then B must be finite. To construct such a set, let $(x_n)_{n \in \mathbb{N}}$ be a quickly increasing sequence (so that $x_{n+1} > 4x_n^2$) and let

$$A = \bigcup_{n \in \mathbb{N}} [x_n, \tfrac{3}{2}x_n]. \tag{12}$$

We observe that $\overline{\mathsf{d}}(A) = 1/3$. Suppose $t, m \in \mathbb{N}$ and $B \subset \mathbb{N}$ satisfy (11). We show that B can not be infinite. Indeed, take $b_0 \in B$ and suppose that there exists $b \in B$ arbitrarily large, depending on b_0. Since $b_0 + mb + t \in A$ and $mb_0 + b + t \in A$, there exist $n, n' \in \mathbb{N}$ such that $b_0 + mb + t \in [x_n, \tfrac{3}{2}x_n]$ and $mb_0 + b + t \in [x_{n'}, \tfrac{3}{2}x_{n'}]$. Since $b > b_0$ and $m \geq 1$, we have $n \geq n'$. If $n = n'$, then

$$b_0 + mb + t \leq \tfrac{3}{2}(mb_0 + b + t) \iff b \leq \frac{3m - 2}{2m - 3}b_0 + \frac{t}{2m - 3},$$

but for large enough b this is not possible. If $n > n'$ then

$$b_0 + mb + t \geq x_n > 4x_{n'}^2 > 2x_{n'}(mb_0 + b + t) > (mb_0 + b + t)^2 > b^2,$$

whence $b(b - m) < b_0 + t$, and again this is not possible for sufficiently large b.

Just as the idea behind the construction in Example 2.3 is to avoid pairs $\{b, 2b\}$, the idea behind Example 3.21 is to avoid pairs $\{b, mb\}$ for all $m \in \mathbb{N}$ with $m \geq 2$; the set constructed in Example 3.21 has the property that it only contains finitely many pairs $\{b, mb\}$ for any given $m \geq 2$. Sets with positive upper density can exhibit even more extreme behavior. A well-known, but complicated, example of Besicovitch [6] produces sets with positive upper density containing no pair $\{b, mb\}$ for $b, m \in \mathbb{N}$ and $m \geq 2$. However, Davenport and Erdős [10, 11] showed that if one replaces upper density with the notion of upper logarithmic density, which is more robust under dilation, then this behavior is avoided.

Definition 3.22 (Logarithmic density). *Given a set $A \subset \mathbb{N}$, we define its lower and upper logarithmic density, respectively, by the formulas*

$$\underline{\mathsf{d}}_{\log}(A) := \liminf_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{\mathbf{1}_A(n)}{n} \quad and \quad \overline{\mathsf{d}}_{\log}(A) := \limsup_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{\mathbf{1}_A(n)}{n}$$

and when $\underline{\mathsf{d}}_{\log}(A) = \overline{\mathsf{d}}_{\log}(A)$, we say that the logarithmic density of A exists and denote the common value simply by $\mathsf{d}_{\log}(A)$.

For any set $A \subset \mathbb{N}$, one has $\underline{\mathsf{d}}(A) \leq \underline{\mathsf{d}}_{\log}(A) \leq \overline{\mathsf{d}}_{\log}(A) \leq \overline{\mathsf{d}}(A)$. It can be verified directly that the set produced in Example 3.21 has zero upper logarithmic density, and so we ask the following.

Question 3.23. *Does every set $A \subset \mathbb{N}$ with $\overline{\mathsf{d}}_{\log}(A) > 0$ contain a sumset $B + mB + t$ for some $m \geq 2$, $t \geq 0$ and some infinite set $B \subset \mathbb{N}$?*

All our examples thus far of sets with positive density that avoid some unrestricted sumset are built at a logarithmic scale. We would like to know if this is the only obstruction. One could attempt to make this precise by asking whether any set $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$ and $\{\log(n) : n \in A\}$ uniformly distributed modulo 1 contains an unrestricted sumset $B + B + t$. Unfortunately, as written it is not well posed since the sequence $\log(n)$ itself is not uniformly distributed modulo 1. This inconvenience can be avoided by using logarithmic averages, and we formalize this in the following question.

Question 3.24. *If $A \subset \mathbb{N}$ has $\mathsf{d}_{\log}(A) > 0$ and for any interval $(a, b) \subset [0, 1)$ and $\theta > 0$ one has*

$$\mathsf{d}_{\log}\left(\left\{n \in A : \{\theta \log n\} \in (a, b)\right\}\right) = \mathsf{d}_{\log}(A) \cdot (b - a)$$

then does A contain a sumset $B + B + t$ for some $t \geq 0$ and infinite set $B \subset \mathbb{N}$?

3.5 Uniformity norms and shifts

We have seen in Proposition 2.6 that the shift t in Conjecture 2.5 can be fully understood in terms of congruence obstructions. One may wonder whether the shift in Conjecture 2.4 is subject to the same restrictions. However, the next example shows that is not the case.

Example 3.25. For every $m \in \mathbb{N}$, the set $A_m := 2m\mathbb{N} - 1$ has $\mathsf{d}^*(A_m) > 0$, but $A_m - t$ does *not* contain $B \cup (B \oplus B)$ for any $t \in \{0, \dots, m - 1\}$. Indeed, suppose $\{b, b'\} \subset A_m - t$. Then $b, b' \equiv -1 - t \bmod 2m$ and so $b + b' \equiv -2 - 2t \bmod 2m$. Since $-1 - t \equiv -2 - 2t \bmod 2m$ only holds when $t \equiv -1 \bmod 2m$, it follows that $b + b' \notin A_m - t$.

This example tells us that congruence obstructions are not the only obstacle. Nevertheless, we believe that if similar, but higher order, restrictions on A are imposed, then the need for a shift in Conjecture 2.4 is avoided. We formalize this using uniformity norms.

Given a Følner sequence $(\Phi_N)_{N \in \mathbb{N}}$ in $\mathbb{N}$, we say that a bounded function $f \colon \mathbb{N} \to \mathbb{R}$ *admits correlations along* Φ if the limit

$$\lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} f(n + h_1) \cdots f(n + h_k)$$

exists for all $h_1, \ldots, h_k \in \mathbb{N} \cup \{0\}$. Note that for any $f \colon \mathbb{N} \to \mathbb{R}$ and any Følner sequence Φ, there exists a subsequence Ψ of Φ along which f admits correlations. Under this assumption, the *local uniformity seminorms of f along* Φ are defined inductively by

$$\|f\|_{\mathsf{U}^0(\Phi)} = \lim_{N \to \infty} \frac{1}{|\Phi_N|} \sum_{n \in \Phi_N} f(n)$$

$$\|f\|_{\mathsf{U}^{k+1}(\Phi)}^{2^{k+1}} = \lim_{H \to \infty} \frac{1}{H} \sum_{h=1}^{H} \|\Delta_h f\|_{\mathsf{U}^k(\Phi)}^{2^k},$$

where

$$(\Delta_h f)(n) = f(n)f(n + h)$$

for all $n \in \mathbb{N}$ and all $h \in \mathbb{N} \cup \{0\}$. These types of uniformity seminorms were introduced in [42], where the existence of the limits is deduced from analogous ergodic results.

A set $A \subset \mathbb{N}$ is $\mathsf{U}^k(\Phi)$-*uniform with respect to the Følner sequence* Φ if $\mathbf{1}_A$ admits correlations along Φ, $\mathsf{d}_\Phi(A) > 0$, and $\|\mathbf{1}_A - \mathsf{d}_\Phi(A)\|_{\mathsf{U}^k(\Phi)} = 0$. (The restriction of this definition to sets with $\mathsf{d}_\Phi(A) > 0$ is just to avoid trivially uniform sets such as the empty set.) We believe that, for uniform sets, one can take $t = 0$ in Conjecture 2.4.

Conjecture 3.26. *Let $A \subset \mathbb{N}$ and $k \geq 2$. If there exists a Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ with respect to which A is $\mathsf{U}^k(\Phi)$-uniform then for all $\ell_1, \ldots, \ell_k \in \mathbb{N}$ there exists an infinite set $B \subset \mathbb{N}$ such that*

$$B^{\oplus \ell_1}, \ B^{\oplus \ell_2}, \ \ldots, \ B^{\oplus \ell_k} \subset A.$$

One could also formulate a stronger version of this conjecture, with the local uniformity seminorms replaced by the Gowers seminorms (see for example [31] for definitions).

We remark that Conjecture 3.26 is false for $k = 1$. This can be seen by taking $\ell_1 = 2$ and $A = 2\mathbb{N} - 1$, which is a $\mathsf{U}^1(\Phi)$-uniform set for every Følner sequence Φ. Conjecture 3.26 is much simpler when $\ell_i = i$ for all $1 \leq i \leq k$, and we include the proof for $k = 2$. The proof of Proposition 3.27, with appropriate modifications, also works for higher values of k, but for simplicity we restrict attention to the case $k = 2$. Without the constraint $\ell_i = i$, we do not have a proof even for $k = 2$ and we do not expect that the proof of Proposition 3.27 can be modified to prove, for example, that when A is $\mathsf{U}^2(\Phi)$-uniform with respect to a Følner sequence Φ it contains $B \oplus B \oplus B$ without introducing new tools.

Proposition 3.27. *If $A \subset \mathbb{N}$ is $\mathsf{U}^2(\Phi)$-uniform with respect to some Følner sequence Φ then there exists an infinite $B \subset A$ with $B \oplus B \subset A$.*

Proof. Since $\|\mathbf{1}_A - \mathsf{d}_\Phi(A)\|_{\mathsf{U}^2(\Phi)} = 0$ one can show (for completeness, the proof is given in Lemma A.1 in the appendix) that the set A is weakly mixing in the sense that

$$\mathsf{d}^*\left(\left\{n : \left|\mathsf{d}_\Phi\big((A - n) \cap C\big) - \mathsf{d}_\Phi(A)\mathsf{d}_\Phi(C)\right| > \varepsilon\right\}\right) = 0$$

for any $\varepsilon > 0$ and any set $C \subset \mathbb{N}$ for which all the implicit limits exist. Now choose $b_1 \in A$ such that $\mathsf{d}_\Phi\big((A - b_1) \cap A\big) > 0$. (Note that with respect to Φ a full-density set of $b_1 \in A$ satisfies this.) Put $C_1 := A \cap (A - b_1)$.

Recursively, for each $n = 2, 3, \dots$ choose $b_n \in C_{n-1}$ such that

$$\mathsf{d}_\Phi\big(C_{n-1} \cap (A - b_n)\big) > 0$$

and set $C_n := C_{n-1} \cap (A - b_n)$. It is clear that $B := \{b_n : n \in \mathbb{N}\} \subset A$ and that $B \oplus B \subset A$. $\qquad\square$

The following example gives a set A which is k-step uniform but not $(k + 1)$-step uniform and has the property that it contains $B \cup \cdots \cup B^{\oplus k}$ but not $B \cup \cdots \cup B^{\oplus k} \cup B^{\oplus k+1}$. In this sense, the k-step seminorms can be said to be "characteristic" for k-fold sumsets but not for $(k + 1)$-fold sumsets. Our example is inspired by [21], where sets of k-recurrence that are not $(k+1)$-recurrence are constructed in a similar way (see Section 3.7 for definitions and related results).

Example 3.28. For $k \in \mathbb{N}$, $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and $\varepsilon > 0$ sufficiently small, the set

$$A = \left\{ n \in \mathbb{N} : \{n^k \alpha\} \in \left[\tfrac{1}{2}, \, \tfrac{1}{2} + \varepsilon \right] \right\}$$

satisfies $\|\mathbf{1}_A - \varepsilon\|_{\mathsf{U}^k(\Phi)} = 0$ and $\|\mathbf{1}_A - \varepsilon\|_{\mathsf{U}^{k+1}(\Phi)} > 0$ for all Følner-sequences Φ. There is an infinite set B such that $B \cup \cdots \cup B^{\oplus k} \subset A$ but there exists no infinite set B with $B \cup \cdots \cup B^{\oplus k+1} \subset A$. We prove this claim for the case $k = 2$ and leave the general case to the interested reader. The fact that

$$A = \left\{ n \in \mathbb{N} : \{n^2 \alpha\} \in \left[\tfrac{1}{2}, \, \tfrac{1}{2} + \varepsilon \right] \right\}$$

contains $B \cup B^{\oplus 2}$ follows from Proposition 3.27. In fact, since at each stage of the inductive construction in the proof of Proposition 3.27 there is a positive density set from which the members of B can be drawn, there are in a sense many sets $B \subset \mathbb{N}$ for which $B \cup B^{\oplus 2} \subset A$.

We are left with showing that A does not contain $B \cup B^{\oplus 2} \cup B^{\oplus 3}$. Let $B \subset \mathbb{N}$ be infinite. In view of Ramsey's theorem, replacing B by an infinite subset of itself if needed, we can assume that for every distinct $b, b' \in B$, $\{bb'\alpha\} \approx_\varepsilon \gamma$ for some fixed $\gamma \in [0, 1)$. By further refining B we can assume that $\{b^2\alpha\} \approx_\varepsilon \beta$ for every $b \in B$ and some $\beta \in [0, 1)$.

If $B \subset A$ it follows that $\beta \approx_\varepsilon 1/2$. If $B \oplus B \subset A$, expanding the square $(b + b')^2 \alpha$ it follows that $2(\beta + \gamma) \approx_\varepsilon 1/2$. If $B^{\oplus 3} \subset A$, then expanding the square $(b_1 + b_2 + b_3)^2 \alpha$ it follows that $3\beta + 6\gamma \approx_\varepsilon 1/2$. Since it is impossible that β, $2\beta + 2\gamma$, and $3\beta + 6\gamma$ are simultaneously close to $1/2$ modulo 1, it follows that A cannot contain $B \cup B^{\oplus 2} \cup B^{\oplus 3}$.

3.6 The density Ramsey property

The sumset results and Szemerédi's theorem both show the existence of patterns in a set of positive upper Banach density. Seeking a natural unification of these results, the next definition is used to formulate such a statement (see Question 3.11 for a different way to combine Szemerédis's theorem with sumsets).

Definition 3.29. *We say that a collection $\mathcal{D} \subset \{F \subset \mathbb{N} : 0 < |F| < \infty\}$ of finite sets has the density Ramsey property if for any set $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$, there exists some $X \in \mathcal{D}$ such that $X \subset A$.*

By a compactness argument, the set $\mathcal{D}$ has the density Ramsey property if for all $\delta > 0$, there exists $N(\mathcal{D}, \delta) \in \mathbb{N}$ such that for any interval $J \subset \mathbb{N}$ of length at least $N(\mathcal{D}, \delta)$ and any set $E \subset J$ with $|E| \geq \delta|J|$, there is $X \in \mathcal{D}$ with $X \subset E$.

In this language, Szemerédi's theorem is equivalent to the assertion that for every $k \in \mathbb{N}$, the set of all k-term arithmetic progressions has the density Ramsey property.

Theorem 3.30. *If $\mathcal{D} \subset \{F \subset \mathbb{N} \colon 0 < |F| < \infty\}$ has the density Ramsey property, then for any $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$, there exist pairwise disjoint sets $B_1, B_2, \ldots \in \mathcal{D}$ and pairwise disjoint sets $C_1, C_2, \ldots \in \mathcal{D}$ such that $B_i + C_j \subset A$ for all $i, j \in \mathbb{N}$.*

Taking $\mathcal{D}$ to be the collection of singletons, we obtain Theorem 2.7 and taking $\mathcal{D}$ to be k-term arithmetic progressions, we obtain a generalization of Szemerédi's Theorem. The analogous statement also holds for higher order sumsets, meaning for any $k \in \mathbb{N}$, there exist k sequences of disjoint sets $(B_{i,j})_{j \in \mathbb{N}} \in \mathcal{D}$ for $i = 1, \ldots, k$ such that the k-fold sumsets all lie in A. We could also start with two collections $\mathcal{D}_1$ and $\mathcal{D}_2$ with the density Ramsey property, and select the sets B_i from $\mathcal{D}_1$ and C_i from $\mathcal{D}_2$. However, to minimize notation, we restrict ourselves to the two-fold sumset and single collection $\mathcal{D}$.

Proof. Let $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$. By combining the results in [50, Section 2], we can find $\varepsilon > 0$, a Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, and a set $L \subset \mathbb{N}$ such that for every nonempty finite set $F \subset L$, the set

$$\bigcap_{x \in F} (A - x) \cap \left\{n \in \mathbb{N} \colon \mathsf{d}_\Phi\big((A - n) \cap L\big) > \varepsilon\right\}$$

has positive density with respect to Φ. Let $F_1 \subset F_2 \subset \cdots \subset L$ be an exhaustion of L by finite sets. For $i \in \mathbb{N}$, define

$$Q_i = \bigcap_{x \in F_i} (A - x) \cap \left\{n \in \mathbb{N} \colon \mathsf{d}_\Phi\big((A - n) \cap L\big) > \varepsilon\right\}. \tag{13}$$

Pick $N(\mathcal{D}, \delta) \in \mathbb{N}$ such that for any interval $J \subset \mathbb{N}$ of length at least $N(\mathcal{D}, \delta)$ and any set $E \subset J$ with $|E| \geq \delta|J|$, there exists $X \in \mathcal{D}$ with $X \subset E$. Set $\delta_i = \mathsf{d}_\Phi(Q_i)/2$ and for each $i \in \mathbb{N}$, let $J_i \subset \mathbb{N}$ denote an interval whose length is at least $N(\mathcal{D}, \varepsilon\delta_i/2)$ and satisfies

$$|J_i \cap Q_i| \geq \delta_i |J_i|.$$

Without loss of generality, we can further assume that for each $i \in \mathbb{N}$, we have $\max J_i < \min J_{i+1}$. Then by (13), it follows that

1. $F_i + (Q_i \cap J_i) \subset A$ for all $i \in \mathbb{N}$.

2. If $n_1 < n_2 < \cdots$ is an enumeration of $\bigcup_{i \in \mathbb{N}}(J_i \cap Q_i)$, then

$$\mathsf{d}_\Phi\big((A - n_k) \cap L\big) > \varepsilon.$$

Applying Bergelson's Intersectivity Lemma [2, Theorem 2.1], there exists $K \subset \mathbb{N}$ with $\mathsf{d}(K) = \varepsilon$ such that for all nonempty finite sets $H \subset K$, we have

$$\mathsf{d}^*\left(\bigcap_{k \in H}\big((A - n_k) \cap L\big)\right) > 0. \tag{14}$$

Define

$$R_i = \{n_k \colon k \in K\} \cap J_i \cap Q_i$$

and let $I \subset \mathbb{N}$ be an infinite set such that for all $i \in I$, we have

$$|R_i| \geq \frac{\varepsilon\delta_i}{2}|J_i|$$

which exists because $\mathsf{d}(K) = \varepsilon$ and $|J_i \cap Q_i| \geq \delta_i|J_i|$. Since the length of J_i exceeds $N(\mathcal{D}, \varepsilon\delta_i/2)$, we can find $X_i \in \mathcal{D}$ with $X_i \subset R_i$. Then by property 1 and (14), we conclude that

17

3. $F_i + X_i \subset A$ for all $i \in I$.

4. For each $i \in I$, we have

$$\mathsf{d}^*\left(\bigcap_{x \in X_1 \cup \dots \cup X_i} \big((A - x) \cap L\big)\right) > 0.$$

We are now ready to inductively construct the sets $B_1, B_2, \dots$ and $C_1, C_2, \dots$.

First choose $B_1 \in \mathcal{D}$ such that $B_1 \subset L$. Choose $i_1 \in I$ sufficiently large such that $B_1 \subset F_{i_1}$, and define $C_1 = X_{i_1}$. By Property (3), we have that

$$B_1 + C_1 \subset A.$$

Assume we have defined the sets $B_1, \dots, B_n$ and $C_1, \dots, C_n$ such that $B_1, \dots, B_n \subset L$ and $C_1 = X_{i_1}, \dots, C_n = X_{i_n}$ for some $i_1 < i_2 < \dots < i_n \in I$. Then by Property (4), there exists $B_{n+1} \in \mathcal{D}$ such that

$$B_{n+1} \subset \bigcap_{x \in X_{i_1} \cup \dots \cup X_{i_n}} \big((A - x) \cap L\big).$$

It follows that $B_{n+1} \subset L$ and for all $1 \leq k \leq n$, we have

$$B_{n+1} + C_k \subset A.$$

Choose $i_{n+1} \in I$ that is sufficiently large such that

$$B_1 \cup \dots \cup B_n \cup B_{n+1} \subset F_{i_{n+1}}$$

and set $C_{n+1} = X_{i_{n+1}}$. Then it follows from Property (3) that for all $1 \leq k \leq n + 1$, we have

$$B_k + C_{n+1} \subset A,$$

completing the proof. $\qquad\square$

Whilst Theorem 3.30 provides a common generalization of Theorem 2.7 and Szemerédi's theorem, it does not include Theorem 2.2 as a special case. This leads to the following question.

Question 3.31. *If $\mathcal{D} \subset \{F \subset \mathbb{N} \colon 0 < |F| < \infty\}$ has the density Ramsey property, then is it true that for any $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$, there exist $t \in \mathbb{N}$ and disjoint $B_1, B_2, \dots \in \mathcal{D}$ such that $B_i + B_j \subset A$ for all $i \neq j \in \mathbb{N}$?*

Similarly, one can formulate the analog for the k-fold sum of the sets.

3.7 Sumsets and recurrence

The quadruple $(X, \mathcal{B}, \mu, T)$ is a *measure preserving system* if $(X, \mathcal{B}, \mu)$ is a probability space, and $T \colon X \to X$ is a measurable and measure preserving map. As part of his proof of Szemerédi's theorem, Furstenberg [22] introduced a general method, known as the *Correspondence Principle*, for translating a problem about finding configurations in sets of upper Banach density into a question of recurrence in measure preserving systems.

A set $R \subset \mathbb{N}$ is a *set of recurrence* [24] if for any measure preserving system $(X, \mathcal{B}, \mu, T)$ and any $E \in \mathcal{B}$ with $\mu(E) > 0$, there exist infinitely many $n \in R$ with $\mu(E \cap T^{-n} E) > 0$. A set $R \subset \mathbb{N}$ is a *set of strong recurrence* [2] if for any measure preserving system $(X, \mathcal{B}, \mu, T)$ and any $E \in \mathcal{B}$ with $\mu(E) > 0$ one has

$$\limsup_{n \in R} \mu(E \cap T^{-n} E) > 0.$$

Note that any set of strong recurrence is a set of recurrence, but the converse does not hold [19]. There are many known examples of sets of recurrence and strong recurrence (see for example [23]).

Theorem 3.32. *Suppose that the set $R \subset \mathbb{N}$ is a set of strong recurrence. Then for any $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$, there exist infinite sets $B \subset R$ and $C \subset A$ such that*

$$\{b + c : b \in B, \ c \in C, \ b < c\} \subset A.$$

We note that for a set R with Banach density zero, the conclusion can not be strengthened to finding the unrestricted sumset $\{b + c : b \in B, c \in C\}$, and the obstruction to such a configuration is given by Example 3.14 (see also Proposition 3.2).

Proof. Fix $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$. By the Furstenberg correspondence principle (see [24]), there exist a measure preserving system $(X, \mathcal{B}, \mu, T)$ and a set $E \in \mathcal{B}$ such that $\mu(E) = \mathsf{d}^*(A)$ and

$$\mathsf{d}^*((A - n_1) \cap \ldots \cap (A - n_j)) \geq \mu(T^{-n_1} E \cap \ldots \cap T^{-n_j} E) \tag{15}$$

for all $j \in \mathbb{N}$ and all $n_1, \ldots, n_j \in \mathbb{N} \cup \{0\}$. Since R is a set of strong recurrence, the sequence of sets $E_n = E \cap T^{-n} E$ satisfies $\limsup_{n \in R} \mu(E_n) > 0$. Applying Bergelson's intersectivity lemma [2, Corollary 2.4] to the collection $\{E_n : n \in R\}$, there exists an infinite subset $L \subset R$ such that for any finite nonempty set $F \subset L$, the intersection $\bigcap_{n \in F} E_n$ has positive measure. Combining this with (15), we conclude that for any finite nonempty set $F \subset L$, the intersection

$$A \cap \left(\bigcap_{n \in F} (A - n) \right)$$

has positive upper Banach density. In particular, this intersection is infinite and we use it to inductively construct the sets B and C. Namely, let $\ell_1 < \ell_2 < \ldots$ be an enumeration of L and choose $b_1 := \ell_1$ and then choose c_1 to be any element in the intersection $A \cap (A - b_1)$ with $c_1 > b_1$. Let $b_2 \in L$ be any element larger than c_1 and then choose c_2 to be any element in $A \cap (A - b_1) \cap (A - b_2)$ with $c_2 > b_2$. Choose $b_3 \in L$ larger than c_2 and take c_3 to be any element in the intersection $A \cap (A - b_1) \cap (A - b_2) \cap (A - b_3)$ with $c_3 > b_3$. Continuing this procedure, we obtain sets $C \subset A$ and $B \subset R$ with $\{b + c : b \in B, \ c \in C, \ b < c\} \subset A$. □

Since the set of perfect squares is known to be a set of strong recurrence [22], Theorem 3.32 implies the following corollary, lending evidence towards a positive answer to Question 3.13.

Corollary 3.33. *For any $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$, there exist infinite sets $B \subset \mathbb{N}$ and $C \subset A$ such that*
$$\{b^2 + c : b \in B, \ c \in C, \ b < c\} \subset A.$$

If the set $R \subset \mathbb{N}$ in Theorem 3.32 is only assumed to be infinite instead of a set of strong recurrence, then a similar argument (using Furstenberg Correspondence and Fatou's lemma) shows that for any $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$, there exist infinite sets $B \subset R$ and $C \subset \mathbb{N}$ such that
$$\{b + c : b \in B, \ c \in C, \ b < c\} \subset A.$$

The difference in this statement and that of Theorem 3.32 is that C cannot necessarily be taken from the set A with density. On the other hand, it is not hard to see that the notion of recurrence is a necessary condition for the stronger conclusion in the theorem. This motivates us to ask how much we can weaken the dynamical assumption on the set R and obtain the same conclusion.

Question 3.34. *Does the conclusion of Theorem 3.32 hold if we replace the assumption on R being a set of strong recurrence by R being a set of recurrence?*

We can also obtain a higher order extension of Theorem 3.32, involving the analog for k-fold strong recurrence. Given $k \in \mathbb{N}$, a set $R \subset \mathbb{N}$ is a *set of strong k-recurrence* if for any measure preserving system $(X, \mathcal{B}, \mu, T)$ and any $E \in \mathcal{B}$ with $\mu(E) > 0$ one has

$$\limsup_{n \in R} \mu(E \cap T^{-n}E \cap \ldots \cap T^{-kn}) > 0.$$

With slight modifications of the proof of Theorem 3.32, one can obtain the following generalization of Theorem 3.32 involving sets of strong k-recurrence. This gives a weaker version of Question 3.11, yielding a different way of combining sumsets and Szemerédi's Theorem (for which we can ask the higher order analog of Question 3.34).

Theorem 3.35. *Suppose R is a set of strong k-recurrence. Then for any $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$ there exist infinite sets $B \subset R$ and $C \subset \mathbb{N}$ such that*

$$\bigcup_{\ell=0}^{k} \{\ell b + c : b \in B, \ c \in C, \ b < c\} \subset A.$$

4 Sumsets in sets of integers without density

4.1 Sumsets in the primes

The set $\mathbb{P}$ of primes has zero density, and so neither Szemerédi's theorem nor any of the results in Section 2 are applicable. However, in a major breakthrough, Green and Tao [31] proved the analog of Szemerédi's theorem in the primes, showing that $\mathbb{P}$ contains arbitrarily long arithmetic progressions. Their proof utilizes a transference method, adapting techniques developed for the study of configurations in sets of positive density to sets of integers that are sparser but sufficiently pseudo-random. This motivates considering sumset phenomenon in the primes and, more generally, studying which infinite patterns occur in $\mathbb{P}$.

Conjecture 4.1. *The set of primes contains a sumset $B + C$ for some infinite sets $B, C \subset \mathbb{N}$.*

Granville [30] proved Conjecture 4.1 conditionally on the prime tuples conjecture of Dickson, Hardy and Littlewood [13, 34] (see Conjecture B.2 in Appendix B for the statement). The fact that $\mathbb{P}$ contains a sumset $B + C$ where B is infinite and C has two elements is equivalent to Zhang's theorem [60] on bounded gaps in primes. The subsequent work of Maynard [48] and Polymath 8 [52] implies that for any size k, the primes contain a sumset $B + C$ for some infinite set $B \subset \mathbb{N}$ and some set $C \subset \mathbb{N}$ with at least k elements. Recently, Tao and Ziegler [58] adapted Maynard's sieve to show that there are infinite sets $B, C \subset \mathbb{N}$ such that the ordered sumset

$$\{b + c : b \in B, \ c \in C, \ b < c\}$$

is contained in the primes. They further show that, conditionally on the Dickson-Hardy-Littlewood conjecture, $\mathbb{P} - 1$ contains a sumset of the form $\{b_1 + b_2 : b_1, b_2 \in B, \ b_1 \neq b_2\}$ for some infinite set $B \subset \mathbb{N}$.

With the goal of understanding which infinite patterns are contained in the primes, and heeding the examples discussed in Section 2, we first observe that no shift of $\mathbb{P}$ contains an IP-set. Recall the notation FS defined in (2).

Proposition 4.2. *There is no sequence $(x_n)_{n=1}^{\infty}$ and shift $t \in \mathbb{Z}$ such that*

$$\mathbb{P} + t \supseteq \mathsf{FS}(\{x_n\}). \tag{16}$$

Proof. Suppose, for the sake of a contradiction, that $t \in \mathbb{Z}$ and $(x_n)_{n=1}^{\infty}$ satisfy (16). Then $p_n := x_n - t$ is prime for every $n \in \mathbb{N}$. Find a set of indices $F \subset \mathbb{N}$ with size $|F| = p_1$ and $\min F > 1$, such that every prime p_i with $i \in F$ has the same congruence class mod p_1. Then on the one hand

$$q := (p_1 + t) + \sum_{i \in F}(p_i + t) \in \mathbb{P} + t$$

by assumption. On the other hand, $q \equiv t \bmod p_1$, so $q - t$ is both a prime and a multiple of p_1, a contradiction. $\qquad\square$

For comparison, it follows from [32, Example 9] that for any $k \in \mathbb{N}$ the set $\mathbb{P} - 1$ contains $\mathsf{FS}(\{x_1, \ldots, x_k\})$ for some set $x_1, \ldots, x_k \in \mathbb{N}$ (and the same holds for $\mathbb{P} + 1$).

Inspired by Conjecture 2.4, we make the following conjecture.

Conjecture 4.3. *For every $k \in \mathbb{N}$, there exists an infinite set $B \subset \mathbb{N}$ such that*

$$\mathbb{P} - 1 \supset \left\{ \sum_{n \in F} n : F \subset B \text{ with } 0 < |F| \leq k \right\}.$$

The analogous statement holds with $\mathbb{P} - 1$ replaced by $\mathbb{P} + 1$.

Observe that the shift in Conjecture 4.3 is fixed independent of k, and any shift other than ± 1 leads to a false statement. This differs from the case of sets with positive density, where the Straus example shows that the shift must be allowed to depend on k. Conjecture 4.3 can be proved conditionally on the Dickson-Hardy-Littlewood conjecture following similar arguments to those of Granville [30] and of Tao and Ziegler [58]. We present this argument in Appendix B.

In view of the recent result in [58] on ordered sumsets inside the primes, it makes sense to wonder whether the analogous ordered version of Conjecture 4.3 can be proved unconditionally. To be concrete, we explicitly formulate the first two cases which are open.

Conjecture 4.4. *There exist infinite sets $B, C, D \subset \mathbb{N}$ such that*

$$\mathbb{P} \supset \{b + c + d : b \in B, c \in C, d \in D; \ b < c < d\}.$$

Conjecture 4.5. *There exist infinite sets $B \subset \mathbb{N}$ and $C \subset \mathbb{P} - 1$ such that*

$$\mathbb{P} \supset \{b + c : b \in B, c \in C; \ b < c\}.$$

We remark that the analogous statement to Conjecture 4.5 where B is required to be contained in $\mathbb{P} - 1$, instead of C, follows from the results in [58] (together with the fact that there exists an infinite admissible set contained in $\mathbb{P} - 1$).

Whenever $A \subset \mathbb{N}$ contains $B + C$ with $B, C \subset \mathbb{N}$ infinite, the set A must have infinitely many pairs of terms at the same distance. Just as in [58, Remark 1.4] this precludes a version of Conjecture 4.1 that applies to subsets of $\mathbb{P}$ with positive relative density: one can remove a subset of zero relative density from $\mathbb{P}$ such that the resulting set no longer has bounded gaps, and hence no infinite sumset.

We finish this section with a question in a different direction. By Proposition 3.2, we know that not every set of positive density contains a sumset of the form $B \oplus B + t$ where B is an infinite set of primes and $t \in \mathbb{N}$. However, for certain sets of number theoretic origin, such as level sets of the classical Liouville function, this conclusion may hold.

21

Question 4.6. *Let $A \subset \mathbb{N}$ be the set of natural numbers that are divisible by an odd number of primes counting multiplicity. Is there an infinite set $B \subset \mathbb{P}$ such that $B \oplus B \subset A$?*

This is only the easiest case to state of numerous variants of this question. Instead we could consider integers whose number of prime factors are $a \bmod b$ for some $a \geq 0$ and $b \geq 1$, or count the prime factors in any of these classes without multiplicity, or more generally consider the level set of any aperiodic multiplicative function.

4.2   Zero density sets

Here we move away from concrete arithmetic sets like $\mathbb{P}$ and consider sets of zero density more generally. We begin by asking if infinite sumsets occur in sets of sublinear growth.

The analogous question in the case of finite configurations plays an important role in contemporary arithmetic combinatorics and the growth rate on the size of a set A that guarantees the existence of various finite patterns has been extensively studied. For example, for 3-term progressions, there have been recent spectacular breakthroughs on the quantitative bound needed to guarantee the existence of a three term arithmetic progression [7, 44]. For progressions of higher length, the best known bounds are due to Gowers [28], but Erdős conjectured that if the sum of the reciprocals of a set of integers diverges then the set contains arbitrarily long arithmetic progressions. This leads us to consider similar questions for infinite sumsets instead of finite progressions.

A naive attempt is to seek a sublinear growth rate that guarantees the existence of infinite sumsets. In other words, can one find a non-decreasing function $h\colon \mathbb{N} \to \mathbb{N}$ with $h(N) \to \infty$ as $N \to \infty$ such that any set $A \subset \mathbb{N}$ with

$$|A \cap \{1, \ldots, N\}| \geq \frac{N}{h(N)} \tag{17}$$

contains $B + C$ for two infinite sets $B, C \subset \mathbb{N}$. However, the answer to this question is negative, because for any such function h there exists a set $A \subset \mathbb{N}$ whose asymptotic growth rate is larger than $N/h(N)$ but the distance between consecutive elements in A tends to infinity. Considering two distinct elements of C, it follows that the sumset $B + C$ of two infinite sets $B, C \subset \mathbb{N}$ contains infinitely many pairs of numbers at the same distance, so such a set A provides a counterexample. However, we ask the following.

Question 4.7. *Is there a set $S \subset \mathbb{N}$ with upper Banach density 0 such that for any subset $A \subset S$ with positive relative density, meaning that*

$$\liminf_{N \to \infty} \frac{|A \cap \{1, \ldots, N\}|}{|S \cap \{1, \ldots, N\}|} > 0,$$

there exist infinite sets $B, C \subset \mathbb{N}$ with $B + C \subset A$?

Given the behavior of specific examples of sets of zero Banach density, such as the set of primes, or the set of sums of two squares, it seems that Question 4.7 has a negative answer. In fact it is possible that the next question has a positive answer, which would rule out a positive answer to Question 4.7.

Question 4.8. *Is it true that if $S \subset \mathbb{N}$ has zero Banach density, there exists a subset $A \subset S$ with positive relative density such that for all but finitely many $t \in \mathbb{N}$, $A \setminus (A - t)$ has the same relative density (in S) as A.*

To see why a positive answer to Question 4.8 implies a negative answer to Question 4.7, suppose that $Y_t := A \setminus (A - t)$ has the same relative density as A for all but

finitely many $t \in \mathbb{N}$. Choosing N_t sufficiently large, the set

$$Y := A \setminus \bigcup_t \{a \geq N_t : a \in A - t\}$$

also has positive relative density but no gap t "appears" in Y infinitely many times, so Y cannot contain an infinite sumset.

4.3 Quantitative versions

While most of our focus so far has been on qualitative results, we turn now to quantitative analogs. We have noted that there is no growth rate on the size of the set A that guarantees the existence of infinite sumsets (see the discussion in Section 4.2). In [41], Host exhibits an example of a set $A \subset \mathbb{N}$ with $\mathsf{d}^*(A) > 0$ such that whenever $B + C \subset A$, both $\mathsf{d}^*(B) = 0$ and $\mathsf{d}^*(C) = 0$. However, this does not preclude there being constraints on the growth of the summands, and so we ask whether one can impose any growth rate on the sets B and C.

Question 4.9. *Given $\delta > 0$, is there a non-decreasing function $\mathcal{H} \colon \mathbb{N} \to \mathbb{N}$ with $\mathcal{H}(N) \to \infty$ as $N \to \infty$ satisfying the following: for any set $A \subset \mathbb{N}$ with $\underline{\mathsf{d}}(A) \geq \delta$ there exist $B, C \subset \mathbb{N}$ with*
$$\min\{|B \cap \{1, \ldots, N\}|, \, |C \cap \{1, \ldots, N\}|\} \geq \mathcal{H}(N)$$
for all sufficiently large N and such that $B + C \subset A$.

We do not even know if one can impose such conditions in the coloring version of this question, where one seeks a monochromatic sumset $B + C$ for an arbitrary coloring of $\mathbb{N}$ using finitely many (or just two) colors.[2] In a similar direction, a result of Erdős and Galvin [17] states that for any function $\mathcal{H} \colon \mathbb{N} \to \mathbb{N}$ with $\mathcal{H}(N) \to \infty$ as $N \to \infty$ there is a 2-coloring of $\mathbb{N}$ such that whenever the IP-set $\mathsf{FS}(B)$ is monochromatic, one has $|B \cap \{1, \ldots, N\}| \leq \mathcal{H}(N)$.

There is a similar regime that is worth exploring: given $\delta > 0$ and $N \in \mathbb{N}$, let $\phi(\delta, N)$ denote the largest integer such that any set $A \subset \{1, \ldots, N\}$ with $|A| \geq \delta N$ contains $B + C$, where $B, C \subset \{1, \ldots, N\}$ satisfy $\min\{|B|, |C|\} \geq \phi(\delta, N)$. It is not difficult to show that $\phi(\delta, N) \to \infty$ as $N \to \infty$ for any fixed $\delta > 0$. On the other hand, the exact behavior of this function is not clear and we make the following conjecture.

Conjecture 4.10. *For every $\delta > 0$,*

$$\limsup_{N \to \infty} \frac{\phi(\delta, N)}{\log N} < \infty \qquad and \qquad \liminf_{N \to \infty} \frac{\phi(\delta, N)}{\log N} > 0.$$

We give a heuristic explanation for why this is the correct approximation for $\phi(\delta, N)$. Given $\delta > 0$, $N \in \mathbb{N}$ and $A \subset [N]$ with $|A| \geq \delta N$, pick $B \subset [N]$ uniformly at random with $|B| = \varepsilon \log N$, for some small positive ε. Then the expected value of the cardinality of $C := \bigcap_{b \in B}(A - b)$ is $|C| \approx \delta^{\varepsilon \log N} N = N^{\varepsilon \log \delta + 1} \gg \log N$. This shows that $\phi(\delta, N) \gg \log N$. For the converse estimate, one can take A itself to be a random set, and then for any B with size $|B| = \theta \log N$, for some parameter $\theta > 0$, the cardinality of $C := \bigcap_{b \in B}(A - b)$ should be $|C| \approx \delta^{\theta \log N} N \ll \log N$.

Question 4.11. *Does the limit*
$$\lim_{N \to \infty} \frac{\phi(\delta, N)}{\log N}$$
exist, and if so, what is it?

---
[2]We thank Thomas Bloom for discussions surrounding this question.

We can also ask about the behavior of ϕ as $\delta \to 0$ when $N \to \infty$.

Question 4.12. *Is* $\liminf\limits_{N \to \infty} \dfrac{\phi(\delta_N, N)}{\log N} > 0$ *for some sequence* $\delta_N \to 0$ *as* $N \to \infty$? *Can one take* $\delta_N = 1/\log N$?

4.4 Sumsets in random sets

Let $(p_n)_{n \in \mathbb{N}}$ be a sequence taking values in $[0, 1]$ and consider the random set $A \subset \mathbb{N}$ defined such that the event $n \in A$ has probability p_n and these events are independent. A general question is for which sequences $(p_n)_{n \in \mathbb{N}}$ does the set A contain almost surely sumsets.

If the probabilities p_n are not sufficiently small, then A almost surely contains arbitrarily long intervals and hence also infinite sumsets. More precisely, if $(p_n)_{n \in \mathbb{N}}$ is a decreasing sequence and for every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} p_n \cdot n^\varepsilon \to \infty$$

then for every $k, n \in \mathbb{N}$, the probability that the interval $[n+1, n+k]$ is contained in A is $\prod_{i=1}^{k} p_{n+i}$, which is approximately equal to p_n^k. Since for disjoint intervals these events are independent and eventually have probability larger than $1/n$, by the second Borel-Cantelli lemma it follows that A contains infinitely many intervals of length k almost surely. Since this holds for every k, the set A is thick (see Definition 3.3) almost surely.

On the other hand, suppose that

$$\lim_{n \to \infty} p_n \cdot n^\varepsilon = 0$$

for some $\varepsilon > 0$. Then for some $k \in \mathbb{N}$, we have that $p_n \ll n^{-1/k}$. Then for any set $F \subset \mathbb{N}$ with $|F| = k$, the probability that A contains infinitely many shifts of F is 0 and hence almost surely A does not contain the configuration $B + F$ for any infinite set $B \subset \mathbb{N}$.

While random sets are often sampled from a distribution where the events $n \in A$ and $m \in A$ are independent for $n \neq m$, this is of course not a requirement. Random sets drawn using different distributions may still lead to interesting and non-trivial infinite sumsets. As an example, we ask the following question.

Question 4.13. *Fix* $\gamma > 0$ *and, for each* $x \in \mathbb{R}$, *let* $A_x := \{n \in \mathbb{N} : \{n^\gamma x\} \leq 1/\log n\}$. *For which values of* γ *is it true that for almost every* $x \in \mathbb{R}$, *the set* A_x *contains a sumset* $B + C$ *for infinite sets* $B, C \subset \mathbb{N}$?

We note that for $\gamma < 1$, the set A_x is thick for almost every x and hence contains a sumset. On the other hand, when γ is an integer one can show that for every irrational x the set A_x cannot contain an infinite sumset. The deterministic positive density analog of such sets is discussed in Example 3.28.

5 Analogs of sumsets beyond the additive integers

5.1 Product sets

A sequence $(\Phi_N)_{N \in \mathbb{N}}$ of finite subsets of $\mathbb{N}$ is a *multiplicative Følner sequence* if

$$\lim_{N \to \infty} \frac{|\Phi_N \cap t\Phi_N|}{|\Phi_N|} = 1$$

for all $t \in \mathbb{N}$ where $t\Phi_N = \{tn : n \in \Phi_N\}$. It is an immediate consequence of [50, Theorem 1.3] that if $A \subset \mathbb{N}$ has positive upper density with respect to a Følner sequence on $\mathbb{N}$ for multiplication, then A contains a product set

$$BC = \{bc : b \in B, c \in C\}$$

for some infinite sets $B, C \subset \mathbb{N}$. On the other hand, the following multiplicative analog of Theorem 2.2 is open.

Conjecture 5.1. *If $A \subset \mathbb{N}$ has positive density with respect to a Følner sequence for multiplication, then it contains $\{b_1 b_2 t : b_1 \neq b_2\}$ for some infinite set $B \subset \mathbb{N}$ and $t \in \mathbb{N}$.*

Next, we ask if sets with positive density but without local obstructions contain both sums and products of infinite sets. Say that $A \subset \mathbb{N}$ is *aperiodic* if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_A(n) e(\alpha n) = 0 \quad \text{for all } \alpha \in \mathbb{Q} \setminus \mathbb{Z} \tag{18}$$

where $e(\alpha n) = e^{2\pi i \alpha n}$. If the density of A exists then (18) is equivalent to

$$\mathsf{d}(A \cap (a\mathbb{Z} + b)) = \frac{1}{a}\mathsf{d}(A) \quad \text{for all } a \in \mathbb{N} \text{ and all } b \in \mathbb{Z},$$

which we can think of as saying A has no bias towards any infinite arithmetic progression. Many sets of integers with number-theoretic or combinatorial origins, such as the set of all numbers with an even number of prime factors or Beatty sequences $\{\lfloor n\alpha + \gamma \rfloor : n \in \mathbb{N}\}$ for some irrational $\alpha > 0$ and $\gamma \in \mathbb{R}$, have this property, warranting our interest in this class of sets.

Question 5.2. *Fix $A \subset \mathbb{N}$ that has positive density and is aperiodic. Are there infinite sets $B, C \subset \mathbb{N}$ with $B + C \subset A$ and $BC \subset A$?*

Question 5.2 is related to [46, Question 8.4], which we repeat here.

Question 5.3. *Is it true that for every finite partition of $\mathbb{N}$, one of the sets in the partition contains $B + C$ and BC for some infinite sets $B, C \subset \mathbb{N}$?*

If the assumption of aperiodicity in Question 5.2 is dropped, the answer is negative due to local obstructions. For example, the modified question is false when A is the set of odd numbers. The following variant of Question 5.2 circumvents this issue by introducing a shift. It makes use of logarithmic density (see Definition 3.22).

Question 5.4. *Suppose $A \subset \mathbb{N}$ has positive upper logarithmic density. Are there infinite sets $B, C \subset \mathbb{N}$ with $B + C \subset A$ and $BC + 1 \subset A$?*

This question is a generalization of one asked previously by the second author, seeking the result when the sets B and C are singletons. In Question 5.4 one might replace the assumption of positive upper logarithmic density with the stronger assumption of syndeticity as defined in Definition 3.3. The question is open even under that assumption, and in fact we do not know the answer to the following special case.

Question 5.5. *Does every syndetic set $S \subset \mathbb{N}$ contain a product set BC for infinite sets $B, C \subset \mathbb{N}$?*

When dealing with Ramsey theoretic questions involving both sums and products, it is often easier to consider the analogous questions over $\mathbb{Q}$ or, more generally, countably infinite fields. To measure density in $\mathbb{Q}$ we make use of a *double Følner sequence* $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, as introduced in [5, Definition 1.3]; this means that Φ is simultaneously a Følner sequence in group $(\mathbb{Q}, +)$ and in the group $(\mathbb{Q}^*, \times)$. Such sequences lead to a notion of density, d_Φ, in $\mathbb{Q}$ that is invariant under both addition and multiplication.

Question 5.6. *Let $A \subset \mathbb{Q}$ have $\mathsf{d}_\Phi(A) > 0$ for some double Følner sequence Φ. Are there infinite sets $B, C \subset \mathbb{Q}$ such that $B + C \subset A$ and $BC \subset A$?*

We end this section with the following related question of Hindman regarding partitions of $\mathbb{N}$.

Question 5.7 ([38, 3.1 Question (a)])**.** *Does there exist a partition of $\mathbb{N}$ into two sets such that neither of them contains an infinite set $B \subset \mathbb{N}$ together with*

$$\{b_1 + b_2 : b_1, b_2 \in B, \ b_1 \neq b_2\} \cup \{b_1 b_2 : b_1, b_2 \in B, \ b_1 \neq b_2\}?$$

Hindman [38] showed this question has a positive answer if one instead considers partitions of $\mathbb{N}$ into seven sets.

5.2 Cartesian product sets

We turn to Cartesian product configurations that can be found in subsets of $\mathbb{N}^2$ with positive density. The starting point of our inquiry is the infinite Ramsey theorem for 2-sets, which we recall.

Theorem 5.8 (Ramsey's theorem for 2-sets, [54])**.** *For any finite coloring of $\mathbb{N}^2$ there exists an infinite set $B \subset \mathbb{N}$ such that*

$$\{(b_1, b_2) : b_1, b_2 \in B, \ b_1 < b_2\} \tag{19}$$

is monochromatic.

Motivated by Theorems 5.8 and 2.2, we discuss the weakest hypotheses on a set $A \subset \mathbb{N}^2$ that guarantees the existence of an infinite set $B \subset \mathbb{N}$ and a shift $t \in \mathbb{N}^2$ such that

$$\big\{(b_1, b_2) : b_1, b_2 \in B, \ b_1 < b_2\big\} \subset A - t. \tag{20}$$

Given $A \subset \mathbb{N}^2$, we write $A_n = \{m \in \mathbb{N} : (n, m) \in A\}$ for the fiber of A with first coordinate n. A necessary condition for A to satisfy (20) is that there are infinitely many vertical fibers A_n such that any finite subcollection of them has infinite intersection. Using this condition, the following example shows that the naive density version of Theorem 5.8 – that a set $A \subset \mathbb{N}^2$ having positive density must contain a translate of (19) for some infinite set $B \subset \mathbb{N}$ – is false.

Example 5.9. Let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be a Følner sequence in $\mathbb{N}^2$. Then there exists a set $A \subset \mathbb{N}^2$ with $\mathsf{d}_\Phi(A) = 1$ and such that each vertical fiber A_n is finite. In particular, any such A cannot satisfy (20).

For each $j \in \mathbb{N}$, using the Følner property, we choose the smallest N_j such that for any $N > N_j$, the first j columns of Φ_N account for less than $1/j$ of the total cardinality of Φ_N. More precisely, if $N > N_j$, then

$$\sum_{n > j} \big|(\Phi_N)_n\big| \geq |\Phi_N|(1 - 1/j).$$

Let $M_j = \max\{m : (j, m) \in \Phi_N, N \leq N_j\} = \max \bigcup_{N \leq N_j} (\Phi_N)_j$ and take A to be the set described in terms of its columns by $A_j = \{1, \ldots, M_j\}$. Note that if $n \geq j$ and $N \leq N_j$, then $A_j \supset (\Phi_N)_j$. It follows that for $N \in [N_j, N_{j+1}]$,

$$|A \cap \Phi_N| = \sum_n |A_n \cap (\Phi_N)_n| \geq \sum_{n > j} |A_n \cap (\Phi_N)_n| = \sum_{n > j} |(\Phi_N)_n| \geq |\Phi_N|(1 - 1/j),$$

and hence $\mathsf{d}_\Phi(A) = 1$ as claimed.

The following question can be viewed as a density variant of Theorem 5.8.

Question 5.10. *Suppose $A \subset \mathbb{N}^2$ satisfies*

$$\liminf_{n \in \mathbb{N}} \mathsf{d}_\Phi(A_n) > 0 \tag{21}$$

for some Følner sequence Φ on $\mathbb{N}$. Does there exist an infinite set $B \subset \mathbb{N}$ and an element $t \in \mathbb{N}^2$ satisfying (20)*?*

Using the map $(x, y) \mapsto x + y$, one can show that a positive answer to Question 5.10 implies Theorem 2.2. However, a positive answer to Question 5.10 is significantly more powerful: for instance, via the map $(x, y) \mapsto \ell x + my$, it implies Conjecture 3.10.

The following example, constructed by Felipe Hernández Castro, shows the $\liminf$ in (21) can not be replaced with $\limsup$.

Example 5.11. Take $1 < c < c' < c'' < 2$ and define

$$X = \mathbb{N} \cap \bigcup_{n \in \mathbb{N}} [4^n, c \cdot 4^n) \qquad \text{and} \qquad Y = \mathbb{N} \cap \bigcup_{n \in \mathbb{N}} [c'4^n, c'' \cdot 4^n),$$

and take $A = X \times Y$. Note that for every integer $t \geq 0$ the intersection $X \cap (Y - t)$ is finite, implying that no shift of A can contain a set of the form $\{(b_1, b_2) : b_1, b_2 \in B, \ b_1 < b_2\}$ for some infinite $B \subset \mathbb{N}$. Yet, there is $\delta > 0$ such that $\underline{\mathsf{d}}(A_n) > \delta$ holds for a set of $n \in \mathbb{N}$ that has positive lower density.

Our next theorem provides some evidence for a positive answer to Question 5.10, by showing that a weaker condition on the set A implies that there exist infinite sets $B, C \subset \mathbb{N}$ with

$$\{(b, c) : b \in B, \ c \in C, \ b < c\} \subset A.$$

Theorem 5.12. *If $A \subset \mathbb{N}^2$ satisfies $\limsup_{n \in \mathbb{N}} \mathsf{d}_\Phi(A_n) > 0$ for some Følner sequence Φ, then there exist infinite sets $B, C \subset \mathbb{N}$ such that $\{(b, c) : b \in B, \ c \in C, \ b < c\} \subset A$.*

Proof. By assumption on the set A, there are $a > 0$ and $n(1) < n(2) < \ldots \in \mathbb{N}$ such that $\mathsf{d}_\Phi(A_{n(i)}) \geq a$ for all $i \in \mathbb{N}$. It follows from [2, Theorem 2.1] that, after possibly passing to a sub-Følner sequence, there is a subsequence $r(1), r(2), \ldots$ of the sequence $n(1), n(2), \ldots$ with

$$\mathsf{d}_\Phi(A_{r(1)} \cap A_{r(2)} \cap \cdots \cap A_{r(i)}) > 0$$

for all $i \in \mathbb{N}$. Inductively choose $b(i)$ and $c(i)$ with $c(i) \in A_{b(1)} \cap \cdots \cap A_{b(i)}$ and $b(i+1) > c(i) > b(i)$. Taking $B = \{b(i) : i \in \mathbb{N}\}$ and $C = \{c(i) : i \in \mathbb{N}\}$, the proof is complete. $\square$

A tantalizing possibility is to upgrade the conclusion of Theorem 5.12 to $A \supset B \times C$, but a counterexample is the set $A = \{(n, m) : n < m\}$ which can not contain such a Cartesian product with both B and C infinite. An obvious way to rule out this example is to impose that the set A must be symmetric with respect to the diagonal (meaning that $(a_1, a_2) \in A$ if and only if $(a_2, a_1) \in A$). However, the next example shows that even this is not sufficient.

Example 5.13. Let A be the set $\{(r, s) \in \mathbb{N}^2 : \max\{r, s\} \text{ is even}, \ \min\{r, s\} \text{ is odd}\}$, which is a symmetric subset of $\mathbb{N}^2$ whose density equals $1/4$ with respect to any Følner sequence in $\mathbb{N}^2$. Moreover, $d_\Phi(A_n) = 1/2$ for every odd n and every Følner sequence Φ in $\mathbb{N}$. However, there are no infinite sets $B, C \subset \mathbb{N}$ such that $B \times C \subset A$.

5.3 Sumsets in general abelian groups

Many questions on sums of infinite sets make sense in all countable abelian groups, where one can define a notion of upper Banach density. With care one can even work with general abelian semigroups. Although we have done so in specific examples such as $(\mathbb{N}, \times)$, here we restrict our attention to countable abelian groups.

Question 5.14. *Is it true in every countable abelian group $(G, +)$ that every set A of positive upper Banach density contains $B_1 + \cdots + B_k$ for infinite sets $B_1, \ldots, B_k \subset G$?*

The techniques of [46] together with work of Griesmer [33] may suffice to give a positive answer in the case $G = \mathbb{Z}^d$. In other settings one cannot proceed in the same way, because the analogous ergodic structure theory is not currently available.

Conjecture 5.15. *Let $(G, +)$ be a countable abelian group. Every set A of positive upper Banach density contains $B \oplus B + t$ for an infinite set $B \subset G$ and some $t \in G$.*

When $G = \mathbb{Z}$, we can restrict the value of t to lie in $\{0, 1\}$ (see Proposition 2.6). We would like to know if similar restrictions can be placed in other groups.

Question 5.16. *Assuming that Conjecture 5.15 holds, what values of the shift t suffice?*

We suspect that there is a direct relation between coset representatives of the subgroup $\{g + g : g \in G\}$ and the restrictions that can be placed on the shift t. This is already hinted at by Proposition 2.6 in $\mathbb{Z}$. The case where G is the direct sum of infinitely many copies of a finite cyclic group sheds some further light on this issue. When G is the direct sum of infinitely many copies of $\mathbb{Z}/3\mathbb{Z}$ one can always take $t = 0$ because all elements in this group are divisible by 2. In contrast, if G is the direct sum of infinitely many copies of $\mathbb{Z}/2\mathbb{Z}$ then one can consider the set of all elements in G whose Hamming distance to the origin is odd, yielding a set of positive density that does not contain $B \oplus B$ for infinite B, showing that one can not always take $t = 0$ in this case.

5.4 Analogs in general amenable groups

Many questions in combinatorial number theory have analogs in countable amenable groups. Recall that amenability of a countable group $(G, \cdot)$ is characterized by the existence of a *Følner sequence*: a sequence $N \mapsto \Phi_N$ of finite, nonempty subsets of G satisfying
$$\lim_{N \to \infty} \frac{|\Phi_N \cap g\Phi_N|}{|\Phi_N|} = 1 = \lim_{N \to \infty} \frac{|\Phi_N g \cap \Phi_N|}{|\Phi_N|}$$
for all $g \in G$. Given a Følner sequence, one speaks of the upper density of a set $A \subset G$ with respect to a Følner sequence
$$\overline{\mathsf{d}}_\Phi(A) = \limsup_{N \to \infty} \frac{|A \cap \Phi_N|}{|\Phi_N|}$$

and can enquire about those configurations that are guaranteed in any set whose upper density is positive. The main result of [50] holds in this context: every subset of a countable, amenable group with positive upper density contains the product
$$B \cdot C = \{bc : b \in B, c \in C\}$$

of two infinite sets $B, C \subset G$ ([50, Theorem 1.3]).

At this level of generality, it is not clear how to formulate appropriate analogs of $B \oplus B + t$ and $B + C + D$. We collect here some questions refining [50, Theorem 1.3] in various directions.

Question 5.17. *If $A \subset G$ has positive upper-density with respect to a Følner sequence, is there an unbounded sequence b in G and $t \in G$ with*

$$\{b(i)b(j) : i < j\}$$

contained in $t^{-1}A$?

Question 5.18. *If $A \subset G$ has positive upper-density with respect to a Følner sequence, is there an unbounded sequence b in G and $s, t \in G$ with*

$$\{b(i)b(j) : i < j\} \cup \{b(i)b(j) : i > j\}\}$$

contained in $t^{-1}Es^{-1}$?

Question 5.19. *If $A \subset G$ has positive upper density with respect to a Følner sequence, are there infinite sets $B, C, D \subset G$ with $BCD \subset A$?*

By a *left Følner* sequence we mean a sequence $N \mapsto \Phi_N$ of finite, nonempty subsets of G satisfying

$$\lim_{N \to \infty} \frac{|\Phi_N \cap g\Phi_N|}{|\Phi_N|} = 1$$

for all $g \in G$. One analogously defines *right Følner* sequences.

Question 5.20. *What are the answers to the above questions if one only assumes A has positive upper density with respect to a left Følner sequence or a right Følner sequence?*

In a non-abelian group $(G, \cdot)$ there are two versions of finite sumsets: given a sequence $n \mapsto g(n)$ in G one can define

$$\mathsf{FPL}(g) = \{g(i_1) \cdots g(i_r) : i_1 < \cdots < i_r\}$$
$$\mathsf{FPR}(g) = \{g(i_r) \cdots g(i_1) : i_r > \cdots > i_1\}$$

for the left and right finite products sets defined by g. Both types of product set are partition regular by Hindman's theorem. It is immediate that every $\mathsf{FPL}(g)$ contains $\{g(i)g(j) : i < j\}$ for some unbounded sequence $n \mapsto g(n)$ in G. Unlike the special case of abelian groups, in general finite products sets need not contain BC with $B, C \subset G$ infinite. In particular, we do not know how to answer the following question in full generality.

Question 5.21. *Fix a countable group $(G, \cdot)$. Does every finite partition of G have a cell that contains BC for infinite sets $B, C \subset G$?*

We have drawn several parallels in the integers between results about sumsets and results about arithmetic progressions. We conclude this section by drawing one more such parallel, posing a question on infinite patterns in cartesian products of amenable groups based on Austin's celebrated result [1] on corners.

Theorem 5.22 ([1, Corollary after Theorem B])**.** *Let G^d be the direct product of d copies of an amenable group $(G, \cdot)$ and let $N \mapsto \Phi_N$ be a left Følner sequence in G. If $A \subset G^d$ satisfies $\overline{\mathsf{d}}_{\Phi^d}(A) > 0$ then*

$$\overline{\mathsf{d}}_\Phi \Big(\big\{ g \in G : \exists (a_1, a_2 \ldots, a_d) \in G^d \text{ such that } (a_1, a_2 \ldots, a_d),$$

$$(ga_1, a_2 \ldots, a_d), (ga_1, ga_2 \ldots, a_d), \ldots, (ga_1, \ldots, ga_d) \in A \big\} \Big) > 0.$$

This theorem gives, for example, that every set $A \subset G \times G$ with $\bar{\mathsf{d}}_{\Phi \times \Phi}(A) > 0$ contains many sets of the form $\{(a, b), (ga, b), (ga, gb)\}$. The following question is an attempt to ask a sumset analog, which we only state in $G \times G$ for convenience, but which can be generalized to G^d for $d > 2$ in a straightforward manner.

Question 5.23. *Let $(G, \cdot)$ be an amenable group and let $N \mapsto \Phi_N$ be a left Følner sequence in G. If $A \subset G^2$ satisfies*

$$\liminf_{g \to \infty} \liminf_{M \to \infty} \frac{1}{|\Phi_M|} \sum_{h \in \Phi_M} \mathbf{1}_A(g, h) > 0, \tag{22}$$

then there exists $(t, s) \in G^2$ and an unbounded sequence b in G such that

$$\{(t, s), \ (b(i)t, s), \ (b(i)t, b(j)s) : i < j\} \subset A.$$

If $G = \mathbb{N}$, then an affirmative answer to Question 5.23 implies an affirmative answer to Question 5.10.

5.5 Ultrafilters

We conclude with one question regarding ultrafilters on $\mathbb{N}$. An *ultrafilter* is a non-empty collection p of subsets of $\mathbb{N}$ closed under intersections and supersets which satisfies $A \notin \mathsf{p} \iff (\mathbb{N} \setminus A) \in \mathsf{p}$ for every $A \subset \mathbb{N}$. An ultrafilter is called *non-principal* if all sets belonging to the ultrafilter are infinite.

Recall that one defines

$$\mathsf{p} + \mathsf{q} = \{A \subset \mathbb{N} : \{n \in \mathbb{N} : A - n \in \mathsf{q}\} \in \mathsf{p}\}$$

whenever p and q are ultrafilters on $\mathbb{N}$. While associative, this operation on the set of all ultrafilters is not commutative. For exposition on ultrafilters and their role in Ramsey theory see, for instance, [3, Section 3] or [40].

A set $A \subset \mathbb{N}$ contains $B + C$ for infinite sets $B, C \subset \mathbb{N}$ if and only if there are non-principal ultrafilters p and q with A in both $\mathsf{p} + \mathsf{q}$ and $\mathsf{q} + \mathsf{p}$. It therefore follows from Theorem 2.7 that every positive density set belongs to such a pair of ultrafilters.

Question 5.24. *For which sets $A \subset \mathbb{N}$ does one have*

$$A \in \mathsf{p} + \mathsf{q} = \mathsf{q} + \mathsf{p}$$

for some pair p, q of non-principal ultrafilters?

The answer to Question 5.24 is positive when $A \subset \mathbb{N}$ is a piecewise syndetic set (see Definition 3.3). Indeed, in view of [40, Theorem 4.43] there exists $s \in \mathbb{N}$ such that $A - s$ contains an idempotent ultrafilter $\mathsf{p} = \mathsf{p} + \mathsf{p}$ and hence taking $\mathsf{q} = \mathsf{p} + s$ yields $A \in \mathsf{p} + \mathsf{q} = \mathsf{q} + \mathsf{p}$ as desired. We do not know the answer for sets of positive density, which is perhaps the most interesting aspect of the question.

This question also makes sense in arbitrary groups, if one multiplies the ultrafilters in the appropriate way (see [50, Page 642]).

Appendices

A Uniformity norms and weak mixing functions

Fix a triple (X, μ, T), where μ is a T-invariant probability measure on X. A function $f \in \mathsf{L}^2(X, \mu)$ is *weak mixing* if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} |\langle T^n f, g \rangle| = 0$$

for every $g \in \mathsf{L}^\infty(X, \mu)$. (Note that this makes sense even when the system is not ergodic). The function f being weak mixing is equivalent to saying that

$$\{n \in \mathbb{N} : |\langle T^n f, g \rangle| < \varepsilon\}$$

has full density for every $\varepsilon > 0$ and every $g \in \mathsf{L}^\infty(X, \mu)$. We recall here that if f in $\mathsf{L}^\infty(X, \mu)$ satisfies $\|f\|_2 = 0$, then f is weak mixing. To see this, combine the observation

$$\langle T^n f, g \rangle = \int \mathbb{E}(T^n f \cdot \overline{g} \mid \mathcal{I}) \, d\mu$$

with the Cauchy-Schwarz to obtain

$$\left(\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} |\langle T^n f, g \rangle| \right)^2 \leq \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int |\mathbb{E}(T^n f \cdot \overline{g} \mid \mathcal{I})|^2 \, d\mu$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} \int T^n f \cdot \overline{g} \cdot T^{n+m} \overline{f} \cdot T^m g \, d\mu$$

then note that this last expression is zero whenever $\|f\|_2 = 0$ by [43, Theorem 8.13].

We next present the analog of this result for the ℓ^∞ seminorms in Section 3.5. Fix $f : \mathbb{Z} \to \mathbb{R}$ bounded and assume that f admits correlations and satisfies $\|f\|_{\mathsf{U}^2(\Phi)} = 0$. We check that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \lim_{J \to \infty} \frac{1}{|\Phi_J|} \sum_{j \in \Phi_J} f(n+j)g(j) \right| = 0$$

holds for every bounded function $g : \mathbb{Z} \to \mathbb{R}$. First estimate

$$\left| \lim_{J \to \infty} \frac{1}{|\Phi_J|} \sum_{j \in \Phi_J} f(j+n)g(j) \right|^2$$

$$\leq \left| \frac{1}{M} \sum_{m=1}^{M} \lim_{J \to \infty} \frac{1}{|\Phi_J|} \sum_{j \in \Phi_J} f(j+n+m)g(j+m)f(j+n)g(j) \right|$$

using the van der Corput inequality. The limit as $M \to \infty$ of the quantity inside $|\cdot|$ is non-negative, and so we may dispense in the limit with the absolute values. Then taking the Cesàro average in N, we obtain

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left| \lim_{J \to \infty} \frac{1}{|\Phi_J|} \sum_{j \in \Phi_J} f(j+n)g(j) \right|^2$$

$$\leq \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} \lim_{J \to \infty} \frac{1}{|\Phi_J|} \sum_{j \in \Phi_N} f(j+n+m)g(j+m)f(j+n)g(j)$$

and Hölder's inequality shows that this last term is zero whenever $\|f\|_{\mathsf{U}^2(\Phi)}$ equals zero. This allows us to conclude the following.

Lemma A.1. *If $A \subset \mathbb{N}$ admits correlations and $\|\mathbf{1}_A - \delta\|_{\mathsf{U}^2(\Phi)} = 0$, then for every $B \subset \mathbb{N}$ the set*

$$\{n \in \mathbb{N} : |\mathsf{d}(A - n \cap B) - \mathsf{d}(A)\,\mathsf{d}(B)| < \varepsilon\}$$

has full density.

B   Conditional results in the primes

In this section we show that Conjecture 4.3, postulating infinite patterns in the primes, is implied by the well known Dickson-Hardy-Littlewood conjecture. We begin by recalling the notion of admissible set, needed to formulate the latter conjecture.

Definition B.1. *A (finite or infinite) set $H \subset \mathbb{N}$ is called* admissible *if for every prime $p \in \mathbb{P}$ there exists $i = i(p) \in \mathbb{N}$ such that no element of H is congruent to i mod p.*

Conjecture B.2 (Dickson-Hardy-Littlewood conjecture)**.** *If $H \subset \mathbb{N}$ is finite and admissible, then there exist infinitely many $n \in \mathbb{N}$ such that $n + H \subset \mathbb{P}$.*

Theorem B.3. *Conjecture B.2 implies Conjecture 4.3.*

We first need a lemma imposing congruence restrictions in the shift n arising in Conjecture B.2.

Lemma B.4. *Let $H \subset \mathbb{N}$ be a finite admissible set, let q be a squarefree number and let $a \in \mathbb{N}$ such that $(a + h, q) = 1$ for all $h \in H$. Assuming Conjecture B.2, there exist infinitely many $n \in \mathbb{N}$ such that $n + H \subset \mathbb{P}$ and $n \equiv a$ mod q.*

Proof. By shifting H we may assume without loss of generality that $0 \in H$. Let $N = |H| + \sum_{p|q} p$, where the sum runs over all prime divisors of q. Let $\tilde{Q} = \prod_{p \leq N} p$ where the product is taken over all primes smaller than N, use the Chinese Remainder Theorem to find L such that $L \equiv (\tilde{Q}/p)^{-1}$ mod p for every prime $p < N$ and let $Q = \tilde{Q}L$. Note that $Q/p \equiv 1$ mod p for every prime $p < N$.

For each prime divisor p of q, let $H_p = \{x \in \{0, \ldots, p-1\} : x + a \not\equiv 0 \text{ mod } p\}$, and let

$$\tilde{H} = H \cup \bigcup_{p|q} \frac{Q}{p} \cdot H_p.$$

We claim that $\tilde{H}$ is admissible. Assuming the claim for now, notice that if $n + \tilde{H} \subset \mathbb{P}$, then $n + H \subset \mathbb{P}$. Moreover, if $n \not\equiv a$ mod q, then for some prime $p|q$ we have $n \not\equiv a$ mod p, and hence there is some $x \in H_p$ with $x + n \equiv 0$ mod p. Therefore $h := x\frac{Q}{p} \in \tilde{H}$ and satisfies $h + n \equiv 0$ mod p which implies that $h + n$ is not a prime (unless $n < q$). From Conjecture B.2, we conclude that there exist infinitely many $n \in \mathbb{N}$ such that $n + H \subset \mathbb{P}$ and $n \equiv a$ mod q.

It remains to show that $\tilde{H}$ is admissible. Note that

$$|\tilde{H}| \leq |H| + \sum_{p|q} |H_p| \leq |H| + \sum_{p|q} p = N$$

so for every prime $p > N$ we trivially have less than p residue classes mod p represented in $\tilde{H}$. For each $p < N$, if p does not divide q then any $h \in \tilde{H} \setminus H$ is a multiple of p; since $0 \in H$ and H is admissible we conclude that there are less than p residue classes mod p represented in $\tilde{H}$. Finally, if $p|q$, then for any $p' \neq p$, each element of $\frac{Q}{p'} \cdot H_{p'}$ is a multiple of p. Since no element of H is congruent to $-a$ mod p, no element of H_p is congruent to $-a$ mod p and $\frac{Q}{p} \equiv 1$ mod p, we conclude that no element of $\tilde{H}$ is congruent to $-a$ mod p. $\qquad \square$

Proof of Theorem B.3. Given $n, k \in \mathbb{N}$ denote by $[n] := \{1, 2, \ldots, n\}$ and denote by $[n]^{<k} := \{F \subset [n] : |F| < k\}$. We stress that $[n]^{<k}$ includes the empty set.

Our goal (cf. Conjecture 4.3) is to construct, for a given fixed $k \in \mathbb{N}$, a strictly increasing sequence $b : \mathbb{N} \to \mathbb{N}$ such that

$$\forall n \in \mathbb{N}, \ \forall F \in [n]^{<k+1}, \ F \neq \varnothing, \qquad \sum_{i \in F} b(i) \in \mathbb{P} - 1. \tag{23}$$

We construct additionally, for each prime p, some $x_p \in \mathbb{N} \cup \{0\}$ such that for every $n \in \mathbb{N}$ and $p \in \mathbb{P}$,

$$p < k2^n \ \Rightarrow \ b(n) \equiv x_p \bmod p \tag{24}$$

and

$$\forall F \in [n]^{<k}, \ \forall \ell \leq k - |F|, \qquad p \nmid \left(1 + \ell x_p + \sum_{i \in F} b(i)\right) \tag{25}$$

both hold.

Let $b(1) \in \mathbb{P} - 1$ be divisible by each prime $p < 2k$, and set $x_p := 0$ for those primes. That such a $b(1)$ exists follows from Dirichlet's theorem.

Suppose now that $m > 1$ and we have chosen $b(1), \ldots, b(m-1)$ and $(x_p)_{p < k2^{m-1}}$ satisfying (23), (24) and (25) for all $n < m$ and each prime $p < k2^{m-1}$. For each prime $p \in (k2^{m-1}, k2^m)$ choose an arbitrary x_p such that

$$\forall F \in [m-1]^{<k}, \ \forall \ell \leq k, \qquad p \nmid \left(1 + \ell x_p + \sum_{i \in F} b(i)\right). \tag{26}$$

Note that there are at most $k2^{m-1}$ values of $x_p \bmod p$ that do *not* satisfy (26); therefore for each $p > k2^{m-1}$ we can choose a value of x_p that does satisfy (26). Observe that this choice of x_p satisfies (25) for all $n < m$ and $p < k2^m$. Let $q = \prod_{p < k2^m} p$ and use the Chinese remainder theorem to find $a \in \mathbb{N}$ such that $a \equiv x_p \bmod p$ for every $p < k2^m$. Next let

$$H := \left\{1 + \sum_{i \in F} b(i) : F \in [m-1]^{<k}\right\}.$$

Note that $|H| \leq 2^{m-1}$. It follows from (24) that for each $p < 2^{m-1} \leq k2^{m-1}$, $|\{b(n) \bmod p : n \leq m-1\}| \leq 1 + \log_2(p/k) < p$ and hence that H is an admissible set. It follows from (25) that $(a + h, q) = 1$ for every $h \in H$.

In view of Lemma B.4, there exists $b(m) > b(m-1)$ such that $b(m) + H \subset \mathbb{P}$ and $b(m) \equiv a \bmod q$. From $b(m) + H \subset \mathbb{P}$ we deduce that (23) holds with $n = m$. From $b(m) \equiv a \bmod q$ and the fact that $a \equiv x_p \bmod p$ for every prime $p < k2^m$ we deduce that (24) holds with $n = m$. Finally, from (26), we deduce that (25) holds with $n = m$ for every prime $p < k2^m$. $\square$

References

[1] T. Austin, *Non-conventional ergodic averages for several commuting actions of an amenable group*, J. Anal. Math. **130** (2016), 243–274. MR 3574655

[2] V. Bergelson, *Sets of recurrence of $\mathbf{Z}^m$-actions and properties of sets of differences in $\mathbf{Z}^m$*, J. London Math. Soc. (2) **31** (1985), no. 2, 295–304. MR 809951

[3] ———, *Ergodic Ramsey theory—an update*, Ergodic theory of $\mathbf{Z}^d$ actions (Warwick, 1993–1994), London Math. Soc. Lecture Note Ser., vol. 228, Cambridge Univ. Press, Cambridge, 1996, pp. 1–61. MR 1411215

[4] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753. MR 1325795

[5] V. Bergelson and J. Moreira, *Ergodic theorem involving additive and multiplicative groups of a field and $\{x + y, xy\}$ patterns*, Ergodic Theory Dynam. Systems **37** (2017), no. 3, 673–692. MR 3628915

[6] A. S. Besicovitch, *On the density of certain sequences of integers*, Math. Ann. **110** (1935), no. 1, 336–341. MR 1512943

[7] T. Bloom and O. Sisask, *Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions*, 2020.

[8] Y. Choi and M. J. Heath, *Translation-finite sets and weakly compact derivations from $\ell^1(\mathbb{Z}_+)$ to its dual*, Bull. Lond. Math. Soc. **42** (2010), no. 3, 429–440. MR 2651938

[9] D. Conlon, J. Fox, and Y. Zhao, *The Green-Tao theorem: an exposition*, EMS Surv. Math. Sci. **1** (2014), no. 2, 249–282. MR 3285854

[10] H. Davenport and P. Erdös, *On sequences of positive integers*, Acta Arith **2** (1936), 147–151.

[11] ———, *On sequences of positive integers*, J. Indian Math. Soc. (N.S.) **15** (1951), 19–24. MR 43835

[12] M. Di Nasso, I. Goldbring, R. Jin, S. Leth, M. Lupini, and K. Mahlburg, *On a sumset conjecture of Erdős*, Canad. J. Math. **67** (2015), no. 4, 795–809. MR 3361013

[13] E.L. Dickson, *A new extension of Dirichlet's theorem on prime numbers*, Messenger of Mathematics **33** (1904), 155–161.

[14] P. Erdős, *Problems and results in combinatorial number theory*, Journées Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, Bordeaux, 1974), Astérisque, Nos. 24-25, Soc. Math. France, Paris, 1975, pp. 295–310. MR 0374075

[15] ———, *Problems and results on combinatorial number theory. III*, Number theory day (Proc. Conf., Rockefeller Univ., New York, 1976), Lecture Notes in Math., Vol. 626, Springer, Berlin, 1977, pp. 43–72. MR 0472752

[16] ———, *A survey of problems in combinatorial number theory*, Ann. Discrete Math. **6** (1980), 89–115. MR 593525

[17] P. Erdős and F. Galvin, *Some Ramsey-type theorems*, Discrete Math. **87** (1991), no. 3, 261–269. MR 1095471

[18] P. Erdös and P. Turán, *On Some Sequences of Integers*, J. London Math. Soc. **11** (1936), no. 4, 261–264. MR 1574918

[19] A. H. Forrest, *The construction of a set of recurrence which is not a set of strong recurrence*, Israel J. Math. **76** (1991), no. 1-2, 215–228. MR 1177341

[20] N. Frantzikinakis, *Some open problems on multiple ergodic averages*, Bull. Hellenic Math. Soc. **60** (2016), 41–90. MR 3613710

[21] N. Frantzikinakis, E. Lesigne, and M. Wierdl, *Sets of k-recurrence but not $(k + 1)$-recurrence*, Ann. Inst. Fourier (Grenoble) **56** (2006), no. 4, 839–849. MR 2266880

[22] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR 498471

[23] _____, *Poincaré recurrence and number theory*, Bull. Amer. Math. Soc. (N.S.) **5** (1981), no. 3, 211–234. MR 628658

[24] _____, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures. MR 603625

[25] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291 (1979). MR 531279

[26] _____, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. Analyse Math. **45** (1985), 117–168. MR 833409

[27] H. Furstenberg and B. Weiss, *Topological dynamics and combinatorial number theory*, J. Analyse Math. **34** (1978), 61–85 (1979). MR 531271

[28] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588. MR 1844079

[29] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey theory*, second ed., John Wiley & Sons, Inc., New York, 1990. MR 1044995 (90m:05003)

[30] A. Granville, *A note on sums of primes*, Canad. Math. Bull. **33** (1990), no. 4, 452–454. MR 1091350

[31] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547. MR 2415379

[32] _____, *Linear equations in primes*, Ann. of Math. (2) **171** (2010), no. 3, 1753–1850. MR 2680398 (2011j:11177)

[33] J. T. Griesmer, *Ergodic averages, correlation sequences, and sumsets*, ProQuest LLC, Ann Arbor, MI, 2009, Thesis (Ph.D.)–The Ohio State University. MR 2713738

[34] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), no. 1, 1–70. MR 1555183

[35] N. Hindman, *Finite sums from sequences within cells of a partition of N*, J. Combinatorial Theory Ser. A **17** (1974), 1–11. MR 349574

[36] _____, *Partitions and sums of integers with repetition*, J. Combin. Theory Ser. A **27** (1979), no. 1, 19–32. MR 541340

[37] _____, *Ultrafilters and combinatorial number theory*, Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), Lecture Notes in Math., vol. 751, Springer, Berlin, 1979, pp. 119–184. MR 564927

[38] _____, *Partitions and sums and products—two counterexamples*, J. Combin. Theory Ser. A **29** (1980), no. 1, 113–120. MR 577548

[39] N. Hindman, I. Leader, and D. Strauss, *Open problems in partition regularity*, Special issue on Ramsey theory, vol. 12, Cambridge Univ. Press, 2003, pp. 571–583. MR 2037071

[40] Neil Hindman and Dona Strauss, *Algebra in the Stone-Čech compactification*, De Gruyter Textbook, Walter de Gruyter & Co., Berlin, 2012, Theory and applications, Second revised and extended edition [of MR1642231]. MR 2893605

[41] B. Host, *A short proof of a conjecture of Erdős proved by Moreira, Richter and Robertson*, Discrete Anal. (2019), Paper No. 19, 10. MR 4042162

[42] B. Host and B. Kra, *Uniformity seminorms on ℓ^∞ and applications*, J. Anal. Math. **108** (2009), 219–276. MR 2544760

[43] _____, *Nilpotent structures in ergodic theory*, Mathematical Surveys and Monographs, vol. 236, American Mathematical Society, Providence, RI, 2018. MR 3839640

[44] Z. Kelley and R. Meka, *Strong bounds for 3-progressions*, 2023.

[45] D. Knuth, C. D. H. Cooper, J. C. Owings, Jr., M. S. Klamkin, R. D. Whittekin, J. King, P. Hosford, and M. Geraghty, *Problems and Solutions: Elementary Problems: E2492-E2496*, Amer. Math. Monthly **81** (1974), no. 8, 901–902. MR 1544067

[46] B. Kra, J. Moreira, F. K. Richter, and D. Robertson, *Infinite sumsets in sets with positive density*, 2022.

[47] B. Kra, J. Moreira, F. K. Richter Richter, and D. Robertson, *A proof of Erdős's $B + B + t$ conjecture*, 2022.

[48] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), no. 1, 383–413. MR 3272929

[49] J. Moreira, *Monochromatic sums and products in $\mathbb{N}$*, Ann. of Math. (2) **185** (2017), no. 3, 1069–1090. MR 3664819

[50] J. Moreira, F. K. Richter, and D. Robertson, *A proof of a sumset conjecture of Erdős*, Ann. of Math. (2) **189** (2019), no. 2, 605–652. MR 3919363

[51] M. B. Nathanson, *Sumsets contained in infinite sets of integers*, J. Combin. Theory Ser. A **28** (1980), no. 2, 150–155. MR 563552

[52] D. H. J. Polymath, *New equidistribution estimates of Zhang type*, Algebra Number Theory **8** (2014), no. 9, 2067–2199. MR 3294387

[53] R. Rado, *Studien zur Kombinatorik*, Math. Z. **36** (1933), no. 1, 424–470. MR 1545354

[54] F. P. Ramsey, *On a Problem of Formal Logic*, Proc. London Math. Soc. (2) **30** (1929), no. 4, 264–286. MR 1576401

[55] A. Sárkőzy, *On difference sets of sequences of integers. I*, Acta Math. Acad. Sci. Hungar. **31** (1978), no. 1-2, 125–149. MR 466059

[56] I. Schur, *Über die kongruenz $x^m + y^m \equiv z^m (mod\ p)$*, Jahresbericht der Deutschen Math. Verein. **25** (1916), 114–117.

[57] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245. MR 369312

[58] T. Tao and T. Ziegler, *Infinite partial sumsets in the primes*, 2023.

[59] B. L. van der Waerden, *Beweis einer baudetschen vermutung*, Nieuw. Arch. Wisk. **15** (1927), 212–216.

[60] Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174. MR 3171761