LARGE SETS ARE SUMSETS (DRAFT)

BENJAMIN BAILY, JUSTINE DELL, SOPHIA DEVER, ADAM DIONNE, HENRY FLEISCHMANN, FAYE JACKSON, LEO GOLDMAKHER, GAL GROSS, ETHAN PESIKOFF, HUY PHAM, LUKE REIFENBERG, AND VIDYA VENKATESH

ABSTRACT. Let $[n] := \{0, 1, 2, \dots, n\}$. Intuitively, all large subsets of $[n]$ have additive structure, and there exist various ways to formalize this. For instance, Roth showed in 1953 that for large enough n, any subset of $[n]$ containing more than $n(\log n)^{-2+\epsilon}$ elements contains an arithmetic progression of length 3. Similarly, Szemeredi finds thresholds $0 < \delta_k < 1$ and $N_k > 0$ such that for $n \geq N_k$, subsets of $[n]$ containing more than $\delta_k n$ elements contain arithmetic progressions of length k.

We establish a new variation: for all $d \geq 1$, we find constants $0 < \alpha_d < \beta$ such that for any box $B = [n_1] \times \cdots \times [n_d] \subset \mathbb{N}^d$:

(i) any subset of B with more than $|B| - \alpha_d \log |B|$ elements is additively decomposable, and

(ii) there exists a subset of B of size at least $|B| - \beta \log |B|$ that is has no nontrivial additive decomposition. To do this, we introduce a tool for working with explicit indecomposable sets, a condition on subsets of abelian groups which is efficiently checkable and implies their indecomposability.

In the one dimensional case, this answers a 2005 question of Kim and Roush. In higher dimensions, our result represents new additive structure within generalized arithmetic progressions.

1. INTRODUCTION

As usual in arithmetic combinatorics, set

$$[n] := \{0, 1, 2, \dots, n\}.$$

Let $[n]^d = [n] \times [n] \times \cdots [n]$. Large subsets of $[n]$ have a lot of additive structure. For example, Roth's theorem... Szemeredi's theorem... Green's theorem...

Definition 1. Let $[n] \coloneqq \{0, 1, \dots, n\}$ and $(n] \coloneqq \{1, \dots, n\}$.

Definition 2. For sets $A, B \subseteq \mathbb{Z}^{\geq 0}$, we define $A + B$ as the set of sums $\{a + b : a \in A, b \in B\}$. If $A + B = C$, then we say $A + B$ is an additive decomposition of C. Whereas any set can be decomposed as $A = A + \{0\}$, it is interesting to consider only nontrivial additive decompositions, that is $A + B = C$ with $|A|, |B| \geq 2$. A set without such a nontrivial decomposition is called irreducible. Elsewhere in the literature, they may be referred to as "Ostmann irreducible," "prime," or "primitive."

1.1. History.

In [1], K. H. Kim and F. W. Roush show that the problem of determining whether a set of nonnegative integers exhibits a nontrivial additive decomposition is NP-complete. In the same section, they ask 2 questions:

(1) Is a randomly-chosen subset of $[n]$ more likely to reducible or irreducible?
(2) Of all irreducible subsets $\mathcal{P} \subseteq [n]$, what is the maximum value of $|\mathcal{P}|$?

2020 *Mathematics Subject Classification.* 11P70, 11B13.

In [4], Shitov answers Kim's first question: the number of reducible subsets of $[n]$ is $o(2^{n+1})$ and confirm [2], conjecture 10 for $b = 2$. We answer (2) by finding bounds for $n - |\mathcal{P}|$. Defining sumsets in the same way in higher dimensions, we extend these bounds to irreducible subsets of $[n]^d$.

Remark. K. H. Kim, F. W. Roush, and Y. Shitov use the terminology of boolean tropical polynomials. The difference, however, is only cosmetic. The monoid of sets of nonnegative integers under additive composition is isomorphic to the monoid of boolean tropical polynomials under multiplication, via the natural isomorphism.

1.2. **Main Results.**
Our paper concerns a certain set of functions $f_d : \mathbb{Z}^+ \to \mathbb{Z}^+$ which given as follows:

$$f_d(n) = \min_{\mathcal{P} \subset [n]^d \text{ irreducible}} \#([n]^d \setminus \mathcal{P}).$$

Proposition 1.1. *Let $S \subseteq \prod_{i=1}^d [n_i]$ with $k := \#\prod_{i=1}^d [n_i] \setminus S$ and $n_1, n_2, \dots n_d \in \mathbb{Z}^+$. If*

$$dk \log 2 + 2dH_k + dH_{\lceil \binom{k}{2}/d \rceil} + 2d < \sum_{i=1}^{d} H_{n_i}, \tag{1}$$

where H_m denotes the mth harmonic number, then S is reducible.

To prove this, we show that dense sets exhibit additive structure; in particular, $\{0, r\} + B = S$ for some $r \in \prod_{i=1}^d [n_i]$ and some set $B \subset S$. It is a simple corollary that this proposition extends to subsets of shifted boxes as well. This gives an lower bound for the value $\#[n]^d \setminus \mathcal{P}, \mathcal{P} \subset [n]^d$ of $\log_2(n) + o(\log n)$ when $\mathcal{P}$ is irreducible.

Proposition 1.2. *There exists an irreducible set $\mathcal{P}$ such that $\mathcal{P} \cap [n]$ is irreducible for each n and with*

$$\#([n] \setminus \mathcal{P}) \leq \log_\lambda(n) + o(1),$$

where $\lambda \approx 1.325$ is the real root of $x^3 - x - 1$.

This set $\mathcal{P}$ arises as the complement of a sequence which is eventually a linear recurrence sequence. To show that it is irreducible, we define a stronger notion of irreducibility, which we refer to as "local irreducibility." This property is much easier to verify; in fact it can be verified in $\mathcal{O}(|S|^2)$ time, whereas in general, factoring sets is NP-Complete.

Using this one dimensional construction, we are able to construct large irreducible subsets of $[n]^d$. Indeed, we have the following:

Proposition 1.3. *Let $d \geq 1$. Then,*

$$f_d(n) \leq d \log_\lambda n + o(1),$$

where $\lambda \approx 1.325$ is the real root of $x^3 - x - 1$.

Theorem 1.4. *We have*

$$\log_2(n) + o(\log n) \leq f_d(n) \leq d \log_\lambda n + o(1),$$

where $\lambda \approx 1.325$ is the real root of $x^3 - x - 1$.

Proof. This follows from Proposition 1.1 and Proposition 1.2. The asymptotics are outlined in more detail in Corollaries 2.4 and 3.7. □

ACKNOWLEDGEMENTS. Thanks to NSF and Williams College for funding SMALL, to Williams College for funding Adam Dionne and Ben Baily.

2. BOUNDING IRREDUCIBLE SETS

We first prove Proposition 1.1. To do so, we begin with a lemma.

Lemma 2.1. *Let $S, A \subseteq G$ for any abelian group G. Let $0_G \in A \cap S$. Then there exists $B \subseteq G$ such that $A + B = S$ if and only if for each $s \in S$, $s - a \in S$ for some a and $\{s - a\} + A \subseteq S$.*

Proof. Provided A satisfies the condition, set $B = \{s \in S : s + A \subseteq S\}$ and it follows $A + B = S$. If instead B exists, for each $s \in S$, write $s = a + b$ for some $a \in A, b \in B$. Then since $0 \in A, 0 + b = b \in S$ and $\{b\} + A \subseteq B + A = S$ as desired. □

A special case of this which we will use throughout this section is the one where $A = \{0, r\}$ for some $r \in \mathbb{Z}^d$. In this case, our hypothesis requires at least one of $s - r, s + r \in S$ for each $s \in S$. In particular, we have the following corollary.

Corollary 2.2. *Let $S \subset \mathbb{Z}^d$ such that $|S| \geq 3$ and $0 \in S$. Let $A = \{0, r\}$ for $r \in \mathbb{Z}^d \setminus \{0\}$. $S = A + B$ for some $B \subset S$ if and only if for all $s \in S$, $s - r \in S$ or $s + r \in S$.*

This corollary demonstrates that, for $\#\prod_{i=1}^{d}[n_i] \setminus S$ too small, S must be reducible with one of summand sets being a two element set. This observation is the crux of the proof of Proposition 1.1.

Proof of Proposition 1.1. Let $\mathcal{R} = \prod_{i=1}^{d}[n_i]$. Fix $S \subseteq \mathcal{R}$ and let $k = \#\mathcal{R} \setminus S$. We may shift S by $l \in \mathbb{Z}^d_{\geq 0}$ such that $S' = S - l$, $0 \in S'$, and $S' \subseteq \mathcal{R} - l$. Then, Corollary 2.2 applies.

Suppose $\{0, r\}$ is not a summand of S'. Then, there exists $a \in S'$ such that $\{a - r, a + r\} \cap S' = \varnothing$. From Corollary 2.2, there are three ways for this to happen:

(1) $\{a - r, a + r\} \cap S' = \varnothing, \{a - r, a + r\} \subseteq \mathcal{R} - l$.
(2) $\{a - r, a + r\} \cap S' = \varnothing, \#\{a - r, a + r\} \cap \mathcal{R} - l = 1$.
(3) $\{a - r, a + r\} \cap \mathcal{R} - l = \varnothing$.

Our objective is to exhibit a minimal size for k such that $\{0, r\}$ cannot be a summand of S for each $r \in \mathcal{R} \setminus \{0\}$. For $r \in \mathcal{R} \setminus \{0\}$, let $r \in B_i$ if there exists $a \in S$ such that (i) holds. Now, define

$$B_2' = \bigcup_{q \in \mathcal{R} \setminus S} \left\{ v \,:\, v_i \neq 0 \text{ for exactly one } i, \left\lfloor \frac{s_i + 1}{2} \right\rfloor \leq v_i \leq s_i \right\},$$

where $s_i = \min(q_i, n_i - q_i)$.

Similarly, define

$$B_3' = \mathcal{R} \setminus \left\{ v \,:\, v_i \neq 0 \text{ for exactly one } i, 0 \leq v_i \leq \left\lfloor \frac{n_i}{4} \right\rfloor \right\}.$$

We show $B_2 \cup B_3 \subseteq B_2' \cup B_3'$.

First, let $r \in B_2$ and $a \in S'$ such that (2) holds. If $a_i - r_i < -l_i$ for some i, then $a + r \in (\mathcal{R} - l) \setminus S'$. If $r_i \neq 0$ for exactly one i and $0 < r_i \leq \lfloor n_i/4 \rfloor$, then

$$r \in \{v \,:\, v_i \neq 0 \text{ for exactly one } i, \left\lfloor \frac{a_i + l_i + r_i + 1}{2} \right\rfloor \leq v_i \leq a_i + r_i + l_i\},$$

since $r_i > a_i + l_i$, so $r \in B_2'$. Otherwise, $r \in B_3'$.

Next, suppose $r \in B_2$, $a \in S'$ such that (2) holds and $a_i + r_i > n - l_i$ for some i. Then, $a - r \in (\mathcal{R} - l) \setminus S'$. If $r_i \neq 0$ in exactly one i and $0 < r_i \leq \lfloor n_i/4 \rfloor$, then

$$r \in \{v \,:\, v_i \neq 0 \text{ for exactly one } i, \left\lfloor \frac{n_i - (a_i + l_i) + r_i + 1}{2} \right\rfloor \leq v_i \leq n_i - (a_i + l_i) + r_i\},$$

so $r \in B_2'$. Otherwise, $r \in B_3'$.

Finally, if $r \in B_3$, $r \in B_3'$ if r has more than one nonzero index. If it has only one such r_i then, since there exists $a \in S$ such that $a_i + r_i > n_i$ and $a_i - r_i < 0$, $r_i > n_i/2$ and $r \in B_3'$.

Thus, we have $B_2 \cup B_3 \subseteq B_2' \cup B_3'$, as desired.

Now, define a function $\mu : \mathbb{Z}_{\geq 0}^d \to \mathbb{Z}_{\geq 0}$ where

$$\mu(x) = \begin{cases} \frac{1}{a_i} & a_i \neq 0 \text{ for exactly one } i, \\ 0 & \text{otherwise.} \end{cases}$$

For $A \subset \mathbb{Z}_{\geq 0}^d$, let

$$\mu(A) \coloneqq \sum_{a \in A} \mu(a).$$

Note that each pair $q, r \in (\mathcal{R} - l) \setminus S'$ corresponds to at most one member of B_1. In particular, they correspond to $(|q_1 - r_1|/2, \ldots |q_d - r_d|/2)$ if each index is integral and the point between q and r is in S'. Thus, $\#B_1 \leq \binom{k}{2}$ and $\mu(B_1) \leq dH_{\lceil \binom{k}{2}/d \rceil}$.

Moreover, since $H_{2l} - H_\ell \nearrow \ln 2$,

$$\mu(\{v \,:\, v_i \neq 0 \text{ for exactly one } i, \left\lfloor \frac{s_i + 1}{2} \right\rfloor \leq v_i \leq s_i\}) < \ln 2 + \frac{2}{s_i}$$

for a fixed s_i. Since there are dk sets unioned to form B_2' and each s_i can appear in each axis, we have

$$\mu(B_2') < dk \ln 2 + 2dH_k.$$

Finally,

$$\mu(B_3') \leq d \sum_{i=\lceil \frac{n}{4} \rceil}^{n} \frac{1}{i} < 2d.$$

Thus,

$$\mu(B_1 \cup B_2 \cup B_3) \leq \mu(B_1 \cup B_2' \cup B_3') < dk \ln 2 + 2dH_k + dH_{\lceil \binom{k}{2}/d \rceil} + 2d.$$

As $\mu(\mathcal{R}) = \sum_{i=1}^d H_{n_i}$, to prevent all options of r for factoring S with a set of the form $\{0, r\}$, we require $\mu(B_1 \cup B_2 \cup B_3) \geq \sum_{i=1}^d H_{n_i}$.

Therefore, for any k such that

$$\mu(B_1 \cup B_2 \cup B_3) \leq \mu(B_1 \cup B_2' \cup B_3') < dk \ln 2 + 2dH_k + dH_{\lceil \binom{k}{2}/d \rceil} + 2d < \sum_{i=1}^{d} H_{n_i},$$

S is reducible. $\square$

Corollary 2.3. *Let $S \subseteq \prod_{i=1}^{d}[n_i] - l$ with $n_1, n_2, \dots n_d \in \mathbb{Z}^+$ and $l \in \mathbb{Z}^d$. Let $k := \#\prod_{i=1}^{d}[n_i] \setminus S$. Then, S is reducible if*

$$d\left(k\ln 2 + 2H_k + H_{\lceil \binom{k}{2}/d \rceil} + 2\right) < \sum_{i=1}^{d} H_{n_i}. \tag{2}$$

Proof. Suppose k satisfies (2). Note that $S + l \subseteq \prod_{i=1}^{d}[n_i]$. Thus, from Proposition 1.1, $S + l = A + B$ for $A, B \subset \mathbb{Z}^d$, $\#A, \#B \geq 2$. Thus, $S = A - l + B$ and S is reducible. □

Letting $n_1, n_2, \dots n_d = n$, we can translate Proposition 1.1 into an upper bound on the size of the largest irreducible subset of $[n]^d$.

Corollary 2.4. *We have*

$$f_d(n) \geq \log_2(n) + o(\log n).$$

Proof. From Proposition 1.1, we have that, for all n,

$$m := \min_{\mathcal{P} \subseteq [n]^d \text{ irreducible}} \#([n]^d \setminus \mathcal{P})$$

satisfies

$$m\ln(2) + 2H_m + H_{\lceil \binom{m}{2}/d \rceil} + 2 \geq H_n.$$

As $H_l = \ln(l) + o(1)$, we have

$$m\ln(2) + 4\ln(m) + o(1) \geq \ln(n),$$

implying the desired result. □

3. Constructing Large Irreducible Sets

We begin by constructing a large 1-dimensional irreducible set by introducing *local irreducibility.*

Definition 3. Let G be an abelian group and let $S \subseteq G$ with $0 \in S$. Let $(S, \prec)$ be a totally ordered set such that for $a, b \in S \setminus \{0\}$, we have $a \prec a + b$. We say that S is locally irreducible with respect to $\prec$ or $\prec$-locally irreducible if the following hold:

(1) $s + s \notin S$ where s covers 0, and
(2) for all $t \in S$ with $s \prec t$, there exists $u \in S$ such that $u \prec t$ and $u + t \notin S$.

This has a natural monotonicity: if S is locally irreducible, then the set of the first n elements of S with respect to $\prec$ is locally irreducible. Moreover, any finite locally irreducible set can be extended arbitrarily.

Proposition 3.1. *If S is $\prec$-locally irreducible, then S is irreducible.*

To prove this, we first prove a helpful lemma.

Lemma 3.2. *Let G be an abelian group and $S \subset G$ such that $0 \in S$. If S is reducible, $S = A + B$ for $A, B \subseteq S$ with $\#A, \#B \geq 2$.*

Proof. Suppose S is reducible and $A + B = S$ is a nontrivial reduction of S. Then, since S contains 0, there exists $a \in A$ and $b \in B$ such that $a + b = 0$. Without loss of generality, assume $A = A - a$ and $B = B + a$. Thus, both contain 0. This implies $A, B \subset S$, as desired. □

We now proceed to a proof of Proposition 3.1

Proof. Let $A + B = S \neq \{0\}$ and by Lemma 3.2 let $A, B \subseteq S$. Both must contain 0 as $0 \in S$ and the sum of all nonzero rank elements is nonzero rank. Let $s \in S$ be the element covering 0. Since $s \in S$ and $s \prec s + t$ for $t \in S \setminus 0$, we must have $s \in A$ or B. Without loss of generality, let $s \in A$. Since $s + s \notin S$, we must have $s \notin B$. Next, define ρ_b to be the minimal rank of an element of a nonzero element of B or ∞ if $B = \{0\}$. Let b be the corresponding element in B if it exists. For all $c \in S$ with rank less than ρ_b, we have $c = a_c + b_c$. Since $b_c = 0$, we have $a_c = c$ and thus $c \in A$. However, if b exists, then there exists some $d \in S$ with rank less than ρ_b with $b + d \notin S$. Since $d \in A$, it follows $A + B \neq S$. Hence we conclude $B = \{0\}$ and thus S is irreducible. □

In the case of $S \subset \mathbb{Z}^+$, we use the standard total ordering denoted by $<$. Moreover, this is the only total ordering which is additive, hence we will refer to $<$-locally irreducible subsets of $\mathbb{Z}^+$ as just locally irreducible, as no other ordering is possible.

Example 1. Not every irreducible set is locally irreducible. For instance, $[n] \cup \{2n + 1\}$ is irreducible for any n but not locally irreducible for $n \geq 2$.

We will now construct a locally irreducible set $\mathcal{P}$ with $\#([n] \setminus \mathcal{P}) = \mathcal{O}(\log n)$.

Let $(a_i)_{i=1}^{10} = (2, 4, 8, 11, 16, 22, 27, 44, 54, 91)$. For $i \geq 11$, let $a_i = a_{i-2} + a_{i-3}$. Our locally irreducible set $\mathcal{P}$ will be given by $\mathbb{Z}^{\geq 0} \setminus (a_i)_{i=1}^{\infty}$. The following identity, satisfied for $i \geq 12$, will be useful going forward.

$$a_i + a_{i-4} = a_{i-2} + a_{i-3} + a_{i-4} = a_{i-1} + a_{i-2} = a_{i+1}$$

Proposition 3.3. *$\mathcal{P}$ is locally irreducible.*

Proof. It is clear that $0 \in \mathcal{P}$, that 1 is the minimal nonzero element of $\mathcal{P}$, and that $2 \notin \mathcal{P}$. The main obstacle will be to show the second condition, that is, for $1 < p \in \mathcal{P}$, there exists $q < p$ such that $p + q \notin \mathcal{P}$.

It can be verified directly that $\mathcal{P} \cap [a_{18}]$ is locally irreducible using algorithm 2 in the appendix[1]. For any $p \in \mathcal{P}$ such that $p < a_{15}$, let $q \in \mathcal{P}$ with $q < p, q + p \notin \mathcal{P} \cap [a_19]$. Such a q exists as we directly verified above. We have:

$$p + q < 2a_{15} < a_{15} + a_{16} = a_{18}$$

Therefore $p + q = a_i$ for some $1 \leq i \leq 17$, and it follows that $p + q \notin \mathcal{P}$. This will serve as our base case.

Next, let $a_{i-1} < p < a_i$ for $i \geq 16$.

(1) If $a_i - p \neq a_j$ for any j, then set $q = a_i - p$.

$$q = a_i - p < a_i - a_{i-1} < a_{i-5} < p$$

Thus $q < p$, and since $q \neq a_j$, we know $q \in \mathcal{P}$. Thus $p + q = a_i \notin \mathcal{P}$, and $\mathcal{P}$ is locally prime at p.

(2) If $a_i - p = a_k$ for $k = i - 1, i - 2, i - 3, i - 5$ then $p \notin \mathcal{P}$ and we don't need to check this case.

(3) If $a_i - p = a_{i-4}$, then let $q = 2a_{i-4}$. We have $p + q = a_i + a_{i-4} = a_{i+1} \notin \mathcal{P}$. Furthermore, $a_{i-2} = a_{i-5} + a_{i-4} < q < a_{i-4} + a_{i-3} = a_{i-1}$, hence $q \in \mathcal{P}$. Since $q < a_{i-1} < p$, this verifies that $\mathcal{P}$ is locally irreducible at p.

[1] No computer is needed. In fact, the complement of this set is small enough that this can be done, only moderately painfully, using pencil and paper!

(4) If $a_i - p = a_k$ for $k = i-6, i-7$, then let $q = a_{i+1} - a_i + a_k$. Then we have $q = a_{i+1} - a_i + a_k = a_{i-4} + a_k$. Since $a_{i-3} = a_{i-4} + a_{i-8} < q < a_{i-5} + a_{i-4} = a_{i-2}$, we have $p + q = a_{i+1} \notin \mathcal{P}$, $q \in \mathcal{P}$ and $q < p$. Thus $\mathcal{P}$ is locally irreducible at p.

(5) If $a_i - p = a_k$ for $k \leq i - 8$, then let $q = a_{i+2} - a_i + a_k$. We have $q = a_{i-1} + a_k$. Then $a_{i-1} < q$ clearly. Furthermore:

$$q < a_{i-1} + a_{i-8} < a_{i-1} + a_{i-7} + a_{i-8} - a_{i-8} = a_{i-1} + a_{i-5} - a_{i-8} = a_i - a_{i-8} < p$$

Thus, $q < p < a_i$. Since $p + q = a_{i+2} \notin \mathcal{P}$, $q \in \mathcal{P}$, and $q < p$, (2) holds at p.

□

Proposition 3.4. *$\#([n] \setminus \mathcal{P}) \sim \log_\lambda(n)$, where $\lambda \approx 1.325$ is the real root of $x^3 - x - 1$.*

Proof. By [3], Lemma 8, $\frac{a_{i+1}}{a_i} \to \lambda$. It follows that $\log_\lambda \frac{a_{i+1}}{a_i} \to 1$. For $\delta > 0$, let N_δ denote the minimum N such that $|\log_\lambda \frac{a_{n+1}}{a_n} - 1| < \delta$ for all $n \geq N_\delta$. Then, for all $k \geq 0$,

$$\left| \frac{\log_\lambda(a_{N_\delta+M+k})}{\#[a_{N_\delta+M+k}] \setminus \mathcal{P}} - 1 \right| \leq \frac{\log_\lambda a_{N_\delta} + N_\delta}{N_\delta + M + k} + \frac{M}{N_\delta + M + k}\delta$$

Which, with sufficiently small δ and large M, can be made as small as we like. If $a_{N_\delta+M+k} < n < a_{N_\delta+M+k+1}$, then similarly

$$\left| \frac{\log_\lambda(n)}{\#[n] \setminus \mathcal{P}} - 1 \right| \leq \frac{\log_\lambda a_{N_\delta} + N_\delta + 1 + \delta}{N_\delta + M + k} + \frac{M}{N_\delta + M + k}\delta$$

□

This completes the proof of Proposition 1.2. We are now able to extend this construction into higher dimensions to construct a large irreducible set $\mathcal{Q}(n_1, \ldots, n_d) \subset [n_1] \times [n_d] \times \cdots \times [n_d]$. In doing so, we make heavy use of the generality of the definition of $\prec$-local irreducibility. First, we define the sets $\mathcal{P}(n)$ for each n.

Definition 4. For $n \geq a_{18}$, let $\mathcal{P}(n) = (\mathcal{P} \cap [n]) \cup \{n\}$. Otherwise, set $\mathcal{P}(n) = \{0, n\}$.

Lemma 3.5. *$\mathcal{P}(n)$ is locally irreducible.*

Proof. First, if $n < a_{18}$ then clearly $\mathcal{P}(n)$ is locally irreducible. Otherwise, if $n \geq a_{18}$ and $n \in \mathcal{P}$, then $\mathcal{P}(n) = \mathcal{P} \cap [n]$ and $\mathcal{P}(n)$ is locally irreducible. Otherwise, let $n = a_i$ for some $i \geq 18$. Then, let $p \in \mathcal{P}(a_i)$.

Case 1: If $p < a_{i-3}$, then let $q \in \mathcal{P} \cap [n]$ such that $p + q \notin \mathcal{P} \cap [a_i]$, a subset of $\mathcal{P}(n)$ which we know to be locally irreducible. We know $p + q < a_{i-3} + a_{i-2} = a_i$, hence $p + q \notin \mathcal{P}(a_i)$ as the only difference in the two sets is whether they contain the element a_i, and $p + q \neq a_i$.

Case 2: If $a_{i-3} < p < a_{i-2} - a_{i-10}$, then the proof of 3.3 shows that we can pick $q \in \mathcal{P} \cap [p-1] = \mathcal{P}(a_i) \cap [p-1]$ such that $p + q = a_{i-2}$ or a_{i-1}.

Case 3: Else, $a_{i-2} - a_{i-10} \leq p$. Either $q = p-1, p-2 \in \mathcal{P}(a_i)$ as the sequence a_i is sparse. Then:

$$\begin{aligned} p + q &\geq 2a_{i-2} - 2a_{i-10} - 2 \\ &= a_{i-2} + (a_{i-3} + a_{i-7}) - 2a_{i-10} - 2 \\ &= (a_{i-2} + a_{i-3}) + (a_{i-7} - a_{i-10}) - a_{i-10} - 2 \\ &= a_i + a_{i-9} - a_{i-10} - 2 \\ &= a_i + a_{i-14} - 2 \\ &> a_i \end{aligned}$$

Thus there exists $q < p$ such that $p + q \notin \mathcal{P}(a_i)$.

In any case, for $p \in \mathcal{P}(a_i)$, there exists $q < p \in \mathcal{P}(a_i)$ such that $p_q \notin \mathcal{P}(a_i)$. Thus $\mathcal{P}(a_i)$ is locally irreducible. □

Definition 5. We define $\mathcal{Q}_d(n_1, \ldots, n_d)$ by building its complement. For convenience, let $e_1 = (1, 0, 0 \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_d = (0, 0, \ldots, 0, 1)$.

$$\left(\mathcal{Q}_d(n_1, \ldots, n_d)\right)^C = \bigcup_{i=1}^{d} \{pe_i : p \in [n_i] \setminus \mathcal{P}(n_i)\} \cup \{n_1 e_1 + e_i : 2 \leq i \leq d\}$$

What we have just done is defined $\mathcal{Q}(n_1, \ldots, n_d)$ by turning the ith axis from $[n_i]$ into $\mathcal{P}(n_i)$, then removing the neighbors of ne_1 except for $(n-1)e_1$.

Proposition 3.6. *The set $\mathcal{Q}_d(n_1, \ldots, n_d)$ is $\prec$-locally irreducible, where $\prec$ denotes the lexicographical order.*

Proof. We check the conditions directly. For ease of notation, denote $\mathcal{Q}(n_1, \ldots, n_d)$ by $\mathcal{Q}$. That $0 \in \mathcal{Q}$ follows from the fact that $0 \in \mathcal{P}(n)$ for all n. Let r denote the minimal nonzero element of $\mathcal{Q}$. If $n_1 \geq a_{20}$ then $r = e_1$, else $r = n_1 e_1$. In either case, $2r \notin \mathcal{Q}$ as neither 2 nor $2n_1$ is ever in $\mathcal{P}(n_1)$. Next, let $p \in \mathcal{Q}$ and $p \succ r$.

Case 1: $p = ke_1$. Since $\mathcal{P}(n_1)$ is locally irreducible, choose $j < k$ such that $j + k \notin \mathcal{P}(n_1)$, and set $q = je_1 \prec ke_1 = p$.

Case 2: $p = e_i, i \geq 2$. By construction $n_1 e_1 \in \mathcal{Q}$ and $n_1 e_1 \prec p$. Since $p + n_1 e_1 \notin \mathcal{Q}$, local irreducibility holds at p.

Case 3: $p = ke_i, i \geq 2, k > 1$. As in the first case, the local irreducibility of $\mathcal{P}(n_i)$ allows us to construct such a q.

Case 4: $p = k_1 e_{i_1} + k_2 e_{i_2} + \cdots + k_m e_{i_m}, m \geq 2, k_j > 0$, where the components are in increasing lexicographical order. Then $q := n_{i_1} e_{i_1} \prec p$ and $p+q$ is outside of the box $[n_1] \times \cdots \times [n_d]$, thus $p + q \notin \mathcal{Q}$.

We are able to construct a q as desired in all cases, thus we conclude that $\mathcal{P}_d(n)$ is $\prec$-locally irreducible. □

Corollary 3.7.

$$f_d(n) \leq d \log_\lambda(n) + o(1),$$

where $\lambda \approx 1.325$ is the real root of $x^3 - x - 1$.

Proof. In Proposition , we constructed an irreducible subset of $[n]^d$ with $d\log_\lambda(n) + d - 1 + o(1)$ elements in its complement. The number of elements removed from each axis is $\log_\lambda(n) + o(1)$ from Proposition 3.4. For a fixed d, the result then follows. □

4. Future Work

4.1. **Exploring $f_1(n)$.** Naturally, one can work to refine the constants α, β. There is particular interest in refining the bounds on $f_1(n)$. Our computations suggest that both of our bounds can be tightened substantially.

In the future, we will	Have values of $f_1(n)$ here
Probably from like	$n = 1$ to 30 or so

The following conjecture is natural to make based on our available data and also follows from one's intuition. If proven, it would substantially help understand the behavior of $f_1(n)$.

Conjecture 4.1. *The function $n - f_1(n)$ is monotone increasing.*

4.2. **Finding large irreducibles inside arbitrary sets.** We begin by defining a more general version of the functions f_d.

Definition 6. Let S be a finite subset of $\mathbb{N}$. Then we define:

$$g(S) = \max\{|\mathcal{A}| : \mathcal{A} \subset S \text{ and } \mathcal{A} \text{ is irreducible}\}$$

It suffices to ask this question in the case $S \subset \mathbb{N}$ due to the following idea of Frieman:

Proposition 4.2.

References

[1] K. H. Kim and F. W. Roush *Factorization of Polynomials in One Variable over the Tropical Semiring*, arXiv preprint (uploaded April 2005), https://arxiv.org/pdf/math/0501167.pdf

[2] D. Applegate, M. LeBrun, and N. J. A. Sloane *Dismal Arithmetic* Journal of Integer Sequences, 14 (2011), 9. https://cs.uwaterloo.ca/journals/JIS/VOL14/Sloane/carry2.pdf

[3] F. Dubeau, W. Motta, M. Rachidi, and O. Saeki *On weighted r-generalized Fibonacci sequences* Fibbonaci Quarterly, 35 (1997), 102-110. https://www.fq.math.ca/Scanned/35-2/dubeau.pdf

[4] Y. Shitov *How Many Boolean Polynomials are Irreducible?*, International Journal of Algebra and Computation, 24(08):1183-1189. DOI: 10.1142/S0218196714500520

[5] N. Alon, *Large sets in finite fields are sumsets*, Journal of Number Theory, Volume 126, Issue 1, 2007, Pages 110-118, ISSN 0022-314X, https://doi.org/10.1016/j.jnt.2006.11.007.

5. Appendix: Algorithms

Algorithm 1: Checking LI Directly

$S \subset \mathbb{Z}^{\geq 0}$ with $|S| < \infty$ and m the least positive element of S;
if $2m \in S$ **then**
  return False;
for $s \in S^+$ **do**
  let $A_s = \{t \in S^+ : t < s\}$;
  if $A_s + \{s\} \subset S$ **then**
    return False;
return True;

Proposition 5.1. *Algorithm 1 is correct.*

Proof. As Algorithm 1 directly checks if the 2 conditions of local irreducibility hold, it clearly correctly decides the local irreducibility of $\mathcal{P}$. □

Algorithm 2: Checking LI Via S^C

Result: True if S is LI, False otherwise
$S \subset \mathbb{Z}^{\geq 0}$ with $|S| < \infty$ and m the least positive element of S;
let $T = \{t \in S^C : t > m\}$;
if $2m \in S$ **then**
  return False;
let $T_t^- = \{t' \in T, t' < t\}$;
let $s_j = \min\{s \in S, s \geq j\} + \max\{s \in S, s < j\}$;
let $T_t^+ = \{t' \in T, t < t' \leq s_t\}$;
for $t \in T$ **do**
  if $\overline{T_t^+} = \varnothing$ **then**
    return False;
  else
    for $t' \in T_t^-$ **do**
      if $t - t' \notin T$ **then**
        if $T_{t-t'}^+ - \{t - t'\} \subset T$ **then**
          return False;
return True;

Proposition 5.2. *Algorithm 2 is correct.*

Proof. Suppose S is not locally irreducible. If $2m \in S$ where m is the least positive element of S, then Algorithm 2 will correctly return False. Otherwise, let p denote the minimal element such that $p + \{s \in S : s < p\} \subset S$. Let $t = \min T \setminus [p]$ and $t' = \max T \cap [p]$. If $t > s_{t'}$, then $T_{t'}^+$ is empty and Algorithm 2 will return False. Otherwise, $t \leq 2p - 1$. By assumption that (2) fails at p, we know $t - p \in T$. We can therefore write $p = t - t''$ for some $t'' \in T$. If $t''' - p \in S$ for any $t''' \in T_p^+$ then (2) succeeds at p, hence $T_p^+ - p \in T$ and Algorithm 2 will return False.

Suppose instead that Algorithm 2 returns False. If $2m \in S$, then S is not locally irreducible. Otherwise, let t denote the index at which Algorithm 2 returns False. If $T_t^+ = \varnothing$, then since $t < \min(S \setminus [t]) + q \leq s_t$ for all $q < p, q \in S$, it follows (2) fails at $\min(S \setminus [t])$. Otherwise, there exists t' such that $t - t' \in S$ and $T_{t-t'}^+ - \{t - t'\} \subseteq T$. □

Remark. In most cases, Algorithm 1, running in $\mathcal{O}(\#S^2)$ time, is faster. However for very dense sets, Algorithm 2, running in $\mathcal{O}((\#[\max(S)] \setminus S)^3)$ time, is faster.

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA, USA
Email address: bmb2@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, HAVERFORD COLLEGE, HAVERFORD, PA, USA
Email address: jdell@haverford.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA, USA
Email address: abd2@williams.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI, USA
Email address: henryfl@umich.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI, USA
Email address: alephnil@umich.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA, USA
Email address: Leo.Goldmakher@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF TORONTO, TORONTO, ON, CANADA
Email address: ggross.mail@gmail.com

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CT, USA
Email address: ethan.pesikoff@yale.edu

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CA, USA
Email address: huypham@stanford.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NOTRE DAME, NOTRE DAME, IN, USA
Email address: lreifenb@nd.edu