Discrete Mathematics 338 (2015) 1928–1937

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

On the growth of Stanley sequences

David Rolnick *, Praveen S. Venkataramana

Massachusetts Institute of Technology, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 30 September 2014
Received in revised form 2 April 2015
Accepted 3 April 2015
Available online 6 June 2015

Keywords:
Stanley sequence
3-free set
Arithmetic progression
Roth's theorem
Greedy algorithm

ABSTRACT

From an initial list of nonnegative integers, we form a *Stanley sequence* by recursively adding the smallest integer such that the list remains increasing and no three elements form an arithmetic progression. Odlyzko and Stanley conjectured that every Stanley sequence (a_n) satisfies one of two patterns of asymptotic growth, with no intermediate behavior possible. Sequences of Type 1 satisfy $\alpha/2 \leq \liminf_{n\to\infty} a_n/n^{\log_2 3} \leq \limsup_{n\to\infty} a_n/n^{\log_2 3} \leq \alpha$, for some constant α, while those of Type 2 satisfy $a_n = \Theta(n^2/\log n)$. In this paper, we consider the possible values for α in the growth of Type 1 Stanley sequences. Whereas Odlyzko and Stanley considered only those Type 1 sequences for which α equals 1, we show that α can in fact be any rational number that is at least 1 and for which the denominator, in lowest terms, is a power of 3.

1. Introduction

Since the work of Paul Erdős, discrete mathematicians have recognized that the behavior of random objects may be predictable and interesting. In graph theory, for example, Erdős–Rényi random graphs satisfy many properties that are extremely difficult to construct deterministically. Conversely, other properties are not satisfied by random objects, but may appear when a specific structure is imposed. Stanley sequences straddle the line between randomness and determinism, and have largely remained a mystery since their discovery in 1978. While most examples are disorderly, a select few admit beautifully succinct descriptions. Strikingly, the two types appear to follow two very different types of asymptotic growth, with no intermediate behavior possible; however, a proof of this dichotomy has remained elusive. In this paper, we show how the asymptotic growth rate of a "well-structured" Stanley sequence can fall anywhere on a relatively broad spectrum.

A set is *3-free* if no three elements form an arithmetic progression. Odlyzko and Stanley [4] introduced the natural idea of constructing 3-free sets by the greedy algorithm, starting with some finite set of elements. Specifically, let A be a 3-free set of nonnegative integers $\{a_0, a_1, \dots, a_k\}$ satisfying $0 = a_0 < a_1 < \cdots < a_k$. The *Stanley sequence* $S(A)$ is the infinite sequence (a_n) of nonnegative integers defined greedily such that the 3-free property is preserved. That is, for $n > k$, we pick $a_n > a_{n-1}$ to be the smallest integer for which the set $\{a_0, a_1, \dots, a_n\}$ is 3-free. For simplicity we will often denote $S(\{a_0, a_1, \dots, a_k\})$ by $S(a_0, a_1, \dots, a_k)$.

The simplest Stanley sequence is $S(0)$, which begins $0, 1, 3, 4, 9, 10, 12, 13, 27, \dots$. It is easy to show that the nth term of this sequence is the number obtained by writing n in binary and interpreting it in ternary. In particular, the term a_{2^k} equals 3^k. Odlyzko and Stanley [4] found equally explicit expressions, involving ternary digits, for $S(0, 3^n)$ and $S(0, 2\cdot 3^n)$, again finding that the term a_{2^k} equals 3^k for large enough k.

Odlyzko and Stanley observed that some Stanley sequences, such as $S(0)$, have a regular structure and that their asymptotic behavior resembles $a_{2^k} = 3^k$, while all other Stanley sequences are more disorderly and grow at a faster rate. The conjecture is never stated formally in [4]; we phrase it as follows:

* Corresponding author.
E-mail addresses: drolnick@math.mit.edu (D. Rolnick), venkap@mit.edu (P.S. Venkataramana).

http://dx.doi.org/10.1016/j.disc.2015.04.006
0012-365X/© 2015 Elsevier B.V. All rights reserved.

Conjecture 1.1 (*Based on Work by Odlyzko and Stanley*)**.** *Every Stanley sequence (a_n) follows one of two types of asymptotic growth.*

Type 1: $\alpha/2 \le \liminf_{n\to\infty} a_n/n^{\log_2 3} \le \limsup_{n\to\infty} a_n/n^{\log_2 3} \le \alpha$, where α is a constant, or

Type 2: $a_n = \Theta(n^2/\log n)$.

Odlyzko and Stanley [4] observed Type 1 behavior only in the case of α equal to 1, for which the sequences $S(0)$, $S(0, 3^n)$, and $S(0, 2\cdot 3^n)$ are all examples (see Proposition 2.1 and Corollary 2.3). Erdős et al. [1] later found that the sequence $S(0, 1, 4)$ satisfies $a_{2^k} = 3^k + 2^{k-1}$ (for $k \ge 2$) and is of Type 1 with $\alpha = 1$. However, Rolnick [5] demonstrated that many Stanley sequences follow Type 1 growth for other values of α. One example is the sequence $S(0, 1, 7)$, for which we have $a_{2^k} = (10/9)\cdot 3^k$ and $\alpha = 10/9$. Given a Type 1 sequence, we refer to α as a *scaling factor* for the sequence. For all known Type 1 Stanley sequences, the scaling factor is unique.

To date, no Stanley sequence has been proven conclusively to follow Type 2 growth, even though it is believed that almost all Stanley sequences are of this form. Empirical observations by Lindhurst [2] suggest that the sequence $S(0, 4)$ is indeed of Type 2; however it remains possible that the behavior changes suddenly and unexpectedly after a million terms. A probabilistic argument by Odlyzko and Stanley [4] considered a "random" Stanley sequence defined in terms of probability distributions, and showed that such a "sequence" follows Type 2 growth, but does not prove that any actual Stanley sequence is of this form.

In a recent paper, Moy [3] solved a problem posed by Erdős et al. [1], showing that every Stanley sequence (a_n) satisfies $a_n \le n^2/(2+\epsilon)$ for large enough n. Another problem of [1] remains open, that of finding a Stanley sequence (a_n) satisfying $\lim_{n\to\infty}(a_{n+1} - a_n) = \infty$. However, a related question of [1] was resolved by Savchev and Chen [8], who constructed a sequence (a_n) (not a Stanley sequence) satisfying $\lim_{n\to\infty}(a_{n+1} - a_n) = \infty$ and such that (a_n) defines a *maximal* 3-free set, that is, a 3-free set that is not a proper subset of any other 3-free set.

In this paper, we consider which growth rates are possible for Type 1 Stanley sequences. Results by Rolnick [5] imply that scaling factors of Type 1 Stanley sequences may be arbitrarily high. Here we prove a much stronger result, given in Theorem 2.5. Let α be a rational number at least 1 and for which the denominator is a power of 3. Then, there exists a Type 1 Stanley sequence with α as a scaling factor. We also consider the *repeat factor* of certain Type 1 sequences. Informally, the repeat factor is the integer a_n at which the sequence begins to exhibit its asymptotic pattern of behavior; a formal definition is given in the next section. We demonstrate that every sufficiently large integer is the repeat factor of some Type 1 Stanley sequence.

2. Preliminaries

Some preliminary definitions and results are required before we can state our main result, Theorem 2.5. We begin by verifying that the simplest Stanley sequence, $S(0)$, does indeed follow Type 1 growth. We will use this fact to prove that many other Stanley sequences also follow Type 1 growth.

Proposition 2.1. *The sequence $S(0)$ follows Type 1 growth with 1 as its unique scaling factor.*

Proof. Let (s_n) denote the sequence $S(0)$. We will prove a slightly stronger result than Type 1 growth; we claim that, for each n, we have

$$1/2 \le s_n/n^{\log_2 3} \le 1.$$

We begin by writing n in binary: $n = 2^{d_1} + \cdots + 2^{d_k}$, where we have $d_1 > \cdots > d_k > 0$. We have already noted that s_n equals $3^{d_1} + \cdots + 3^{d_k}$. We conclude

$$\frac{s_n}{n^{\log_2 3}} = \frac{3^{d_1} + \cdots + 3^{d_k}}{\left(2^{d_1} + \cdots + 2^{d_k}\right)^{\log_2 3}} =: f(d_1, \ldots, d_k).$$

Observe that we have

$$(2^{d_1} + \cdots + 2^{d_k})^{\log_2 3} \ge 3^{d_1} + \cdots + 3^{d_k},$$

from which we conclude: $s_n/n^{\log_2 3} \le 1$.

Now, we compute:

$$\frac{\partial f}{\partial d_k} = \frac{(\ln 3)\left(3^{d_k}\right) - (\ln 2)(\log_2 3)\left(2^{d_k}\right)\left(3^{d_1} + \cdots + 3^{d_k}\right)\left(2^{d_1} + \cdots + 2^{d_k}\right)^{-1}}{\left(2^{d_1} + \cdots + 2^{d_k}\right)^{\log_2 3}}.$$

Observe that we have $(\ln 2)(\log_2 3) = \ln 3$. Hence, the numerator is negative under the following condition:

$$\frac{3^{d_k}}{2^{d_k}} < \frac{3^{d_1} + \cdots + 3^{d_k}}{2^{d_1} + \cdots + 2^{d_k}}.$$

This condition holds if we have $d_k < d_{k-1}$. We conclude that increasing d_k as far as possible can only decrease the value of $f(d_1, \ldots, d_k)$. Moreover, creating a new value d_{k+1} may be thought of as increasing d_{k+1} from $-\infty$. Therefore, for a given value d_1, the function $f(d_1, \ldots, d_k)$ is minimized by setting the following values:

$$\begin{aligned} d_2 &= d_1 - 1 \\ d_3 &= d_1 - 2 \\ &\vdots \\ d_k &= d_1 - (k-1) \\ k &= d_1 + 1, \end{aligned}$$

where we have used the fact that the d_i must be distinct nonnegative integers. We conclude:

$$f(d_1, \ldots, d_k) \geq \frac{3^{d_1} + \cdots + 9 + 3 + 1}{\left(2^{d_1} + \cdots + 4 + 2 + 1\right)^{\log_2 3}} = \frac{(3^{d_1+1} - 1)/(3-1)}{\left((2^{d_1+1} - 1)/(2-1)\right)^{\log_2 3}} > 1/2.$$

This completes our proof that $S(0)$ follows Type 1 growth.

Finally, we verify that 1 is the unique scaling factor of this sequence. Note that the upper bound $s_n/n^{\log_2 3} \leq 1$ is attained when n equals 2^k, for every value of k. The lower bound $1/2 \leq s_n/n^{\log_2 3}$ is approximated to arbitrary precision for n equal to $2^k - 1$, as k approaches ∞. Hence, 1 is the only possible scaling factor. □

In this paper, we consider the behavior of *independent* Stanley sequences, defined by Rolnick [5]. In Corollary 2.3, we will show that these sequences follow Type 1 growth. The term "independent" is a consequence of the closely related notion of *dependent* Stanley sequences. Every dependent Stanley sequence is associated with a unique independent Stanley sequence, and the structure of the dependent sequence depends on that of the corresponding independent sequence; for a full definition, see [5]. Rolnick proved that every dependent Stanley sequence follows Type 1 growth, and conjectured that every Type 1 Stanley sequence is either independent or dependent.

Definition 1 (*Rolnick [5]*)**.** Let $S(A) = (a_n)$. We say the Stanley sequence (a_n) is *independent* if there exists a constant $\lambda(A)$ such that the following conditions hold for sufficiently large k:

- $a_{2^k+i} = a_{2^k} + a_i$ for every i such that $0 \leq i < 2^k$, and
- $a_{2^k} = 2a_{2^k-1} - \lambda(A) + 1$.

We let $\kappa(A)$ be the minimum integer such that these conditions hold for all k at least $\kappa(A)$.

It is straightforward to verify that the sequences $S(0)$, $S(0, 3^k)$, and $S(0, 2 \cdot 3^k)$ are independent by using the closed-form descriptions presented in [4]. Rolnick identified several broader classes of independent Stanley sequences (see Theorems 1.2 and 1.4 of [5]). As an example, consider:

$$S(0) = 0, 1, 3, 4, 9, 10, 12, 13, 27, 28, 30, 31, 36, 37, 39, 40, 81, \ldots.$$

The first condition of independence states that the first 2^k terms of the sequence are translated to give the next 2^k terms; for instance, the first 4 terms $\{0, 1, 3, 4\}$ of $S(0)$ are translated by 9 to give the next 4 terms $\{9, 10, 12, 13\}$. The second condition of independence states that the translation approximately doubles the value of the sequence between a_{2^k-1} and a_{2^k}. The parameter λ is 0 in this case; indeed λ is nonnegative for every independent Stanley sequence and equals 0 only for the sequence $S(0)$ (see [5] for details).

For every k at least $\kappa(A)$, we conclude from the definition of an independent Stanley sequence:

$$a_{2^{k+1}} = 2a_{2^{k+1}-1} - \lambda(A) + 1 = 2(a_{2^k} + a_{2^k-1}) - \lambda(A) + 1 = 3a_{2^k}.$$

Hence, for every independent Stanley sequence $S(A)$ there exists a positive number $\alpha = \alpha(A)$ such that for sufficiently large k, we have

$$a_{2^k} = \alpha \cdot 3^k.$$

We now show that, in fact, each term of an independent Stanley sequence is approximately α times the corresponding term of $S(0)$.

Proposition 2.2. *Let $S(A) = (a_n)$. Then, (a_n) is independent if and only if for every n, we have*

$$a_n = \alpha s_n + b_n, \tag{1}$$

where α is a constant, (s_n) denotes the Stanley sequence $S(0)$, and (b_n) is a periodic integer sequence with period 2^κ, for some nonnegative integer κ. Furthermore, if (1) holds, then we have $\kappa = \kappa(A)$.

Proof. We first prove that if $S(A)$ is independent then (1) holds.

Let (b_n) be the sequence in which the values $a_0 - \alpha s_0, a_1 - \alpha s_1, \ldots, a_{2^\kappa-1} - \alpha s_{2^\kappa-1}$ repeat periodically. Pick $k \geq \kappa$, and suppose towards induction that we have $a_n = \alpha s_n + b_n$ whenever n is less than 2^k. The base case of $k = \kappa$ holds from the definition of the sequence (b_n).

For each n less than 2^k, we have

$$\begin{aligned} a_{n+2^k} &= a_{2^k} + a_n \\ &= \alpha s_{2^k} + (\alpha s_n + b_n) \\ &= \alpha(s_{2^k} + s_n) + b_n \\ &= \alpha s_{2^k+n} + b_{2^k+n}. \end{aligned}$$

The last step follows since b_n is periodic. Therefore, we have $a_n = \alpha s_n + b_n$ whenever n is less than 2^{k+1}, completing the induction. One last check remains: The period of (b_n) might be some proper divisor of 2^κ. However, it is easily verified that this implies $\kappa(A) < \kappa$ holds; hence the period of (b_n) is indeed 2^κ.

Now, assume that (1) holds, and pick k, i satisfying $k \geq \kappa$ and $0 \leq i < 2^k$. Then, we have

$$\begin{aligned} a_{2^k+i} &= \alpha s_{2^k+i} + b_{2^k+i} \\ &= \alpha s_{2^k} + \alpha s_i + b_{2^k} + b_i \\ &= a_{2^k} + a_i \end{aligned}$$

and

$$\begin{aligned} a_{2^k} &= \alpha s_{2^k} + b_{2^k} \\ &= \alpha(2s_{2^k-1} + 1) + b_{2^\kappa} \\ &= 2(\alpha s_{2^k-1} + b_{2^k-1}) - (2b_{2^\kappa-1} - b_{2^\kappa} - \alpha + 1) + 1 \\ &= 2a_{2^k-1} - (2b_{2^\kappa-1} - b_{2^\kappa} - \alpha + 1) + 1. \end{aligned}$$

Thus, we see that $S(A)$ is independent and satisfies $\lambda(A) = 2b_{2^\kappa-1} - b_{2^\kappa} - \alpha + 1$. Since the period of (b_n) is 2^κ and not a proper divisor, it follows that κ is the minimum k for which the independence conditions hold; hence $\kappa = \kappa(A)$. □

Corollary 2.3. *Every independent Stanley sequence follows Type* 1 *growth. Moreover it has a unique scaling factor α, given by Eq.* (1)*.*

Proof. This follows from Propositions 2.1 and 2.2. □

Given an independent Stanley sequence $S(A)$, we define the *repeat factor* $\rho(A)$ to be the element a_{2^κ}, which equals $\alpha \cdot 3^\kappa$. Thus, we have

- $\rho(\{0\}) = 1$, because $\kappa = 0$ and $\alpha = 1$;
- $\rho(\{0, 3^n\}) = \rho(\{0, 2 \cdot 3^{n-1}\}) = 3^{n+1}$, because $\kappa = n + 1$ and $\alpha = 1$.

Proposition 2.4. *Let $S(A) = (a_n)$. In order for (a_n) to be independent, it is necessary and sufficient that the following condition holds, for some integers κ, ρ satisfying $\kappa \geq 0$ and $\rho > a_{2^\kappa-1}$:*

$$\{a_n\} = \{\rho x + y \mid x \in S(0),\ y \in \{a_0, a_1, \ldots, a_{2^\kappa-1}\}\}.$$

Moreover if κ is the smallest nonnegative integer such that this equality holds, then we have $\kappa = \kappa(A)$ and $\rho = \rho(A)$.

The proof of Proposition 2.4 follows from a straightforward induction argument similar to that of Proposition 2.2.

Just as the scaling factor of an independent Stanley sequence measures how the sequence behaves asymptotically, so the repeat factor measures how fast the sequence converges to its asymptotic behavior. In this paper, we consider the possible values for the scaling factor and repeat factor of an independent Stanley sequence.

We define a *triadic number* to be a rational number for which the denominator, in lowest terms, is a power of 3. We are now ready to state our main result.

Theorem 2.5. (i) *Every triadic number α that is at least* 1 *is a scaling factor.* (ii) *Every sufficiently large integer ρ is a repeat factor.*

3. Proof of Theorem 2.5

To prove the theorem, we develop a construction for independent Stanley sequences that allows us carefully to control the scaling factor and repeat factor. We begin with several lemmas.

We say that an integer x is *covered* by a set S if there is a 3-term arithmetic progression of the form y, z, x (in that order) satisfying $y, z \in S$. Likewise, we say an integer x is *jointly covered* by sets S and T if there is a 3-term arithmetic progression of the form y, z, x satisfying $y \in S$ and $z \in T$. Given a Stanley sequence $S(A)$, let $O(A)$ be the set of nonnegative integers neither in $S(A)$ nor covered by it. By the definition of a Stanley sequence, $O(A)$ must be a finite set. For $O(A)$ nonempty, let $\omega(A)$ be the maximum element of $O(A)$.

Lemma 3.1 (*Rolnick [5]*). *Let $S(A) = (a_n)$. Suppose that there are integers λ and k satisfying $a_{2^k-1} \geq \lambda + \omega(A)$, such that the following conditions hold:*

- $a_{2^k+i} = a_{2^k} + a_i$ *for every i such that $0 \leq i < 2^k$, and*
- $a_{2^k} = 2a_{2^k-1} - \lambda + 1$.

Then, $S(A)$ is an independent Stanley sequence satisfying $\lambda(A) = \lambda$ and $\kappa(A) \leq k$.

Lemma 3.2. *Let $S(A) = (a_n)$, and suppose that nonnegative integers k and ℓ satisfy $\ell > k \geq \kappa(A)$. Then, we have*

$$a_{2^\ell-2^k-1} = a_{2^\ell-1} - a_{2^k} \tag{2}$$

$$a_{2^\ell-2^k} = a_{2^\ell-1} - a_{2^k-1}. \tag{3}$$

Proof. We prove the result by induction on ℓ. The base case is where ℓ equals $k+1$; we apply the definition of independence to conclude:

$$a_{2^\ell-2^k-1} = a_{2^k-1} = a_{2^{k+1}-1} - a_{2^k} = a_{2^\ell-1} - a_{2^k} \quad \text{and}$$
$$a_{2^\ell-2^k} = a_{2^k} = a_{2^{k+1}-1} - a_{2^k-1} = a_{2^\ell-1} - a_{2^k-1}.$$

Now suppose towards induction that the result holds for ℓ. We apply the definition of independence, together with our inductive hypothesis:

$$a_{2^{\ell+1}-2^k-1} = a_{2^\ell} + a_{2^\ell-2^k-1} = a_{2^\ell} + a_{2^\ell-1} - a_{2^k} = a_{2^{\ell+1}-1} - a_{2^k} \quad \text{and}$$
$$a_{2^{\ell+1}-2^k} = a_{2^\ell} + a_{2^\ell-2^k} = a_{2^\ell} + a_{2^\ell-1} - a_{2^k-1} = a_{2^{\ell+1}-1} - a_{2^k-1}.$$

This completes the induction. □

If x is an integer and S is a set, we will use the notation $S + x$ to denote the set $\{y + x \mid y \in S\}$. The next lemma is based on methods used by Rolnick [5].

Lemma 3.3. *Let $S(A)$ be an independent Stanley sequence. For some k at least $\kappa(A)$, let $c = a_{2^k}$ and let $A^* = \{a_0, a_1, \ldots, a_{2^k-1}\}$. Then, the following statements hold for all integers x and y such that $x < y$.*

a. *The set $A^* + x$ covers*

$$[x, c + x) \setminus ((A^* + x) \cup (O(A) + x)) \cup (O(A) + c + x).$$

b. *The sets $A^* + x$ and $A^* + y$ jointly cover the set*

$$[2y - x, c + 2y - x) \setminus (O(A) + 2y - x) \cup (O(A) + c + 2y - x).$$

c. *The set $(A^* + x) \cup (A^* + c + x)$ covers*

$$[x, 3c + x) \setminus ((A^* + x) \cup (A^* + c + x) \cup (O(A) + x)) \cup (O(A) + 3c + x).$$

d. *The sets $(A^* + x) \cup (A^* + c + x)$ and $(A^* + y) \cup (A^* + c + y)$ jointly cover the set*

$$[2y - x, 3c + 2y - x) \setminus (O(A) + 2y - x) \cup (O(A) + 3c + 2y - x).$$

e. *The set $(A^* + x) \cup (A^* + c + x) \cup (A^* + 3c + x) \cup (A^* + 4c + x)$ covers the set*

$$[x, 9c + x) \setminus ((A^* + x) \cup (A^* + c + x) \cup (A^* + 3c + x) \cup (A^* + 4c + x) \cup (O(A) + x)) \cup (O(A) + 9c + x).$$

Proof. We first prove part (a). Observe that the set A^* must cover every integer in $[0, c) \setminus (A^* \cup O(A))$ because these integers are assumed to be covered by $S(A)$. Hence, $A^* + x$ covers $[x, c + x) \setminus ((A^* + x) \cup (O(A) + x))$.

We now must prove that $O(A) + c + x$ is covered by $A^* + x$. This is equivalent to proving that $O(A) + c$ is covered by A^*. Pick some $z \in O(A) + c$. For ℓ a large integer, let $z' = z + a_{2^\ell} - c$, so that we have $z' \in O(A) + a_{2^\ell}$. Because $S(A)$ is an independent sequence, we have $z' \notin S(A)$, so we have $z' = 2a_j - a_i$ for some i and j satisfying $i < j$. There are three cases: (1) $i, j < \ell$, (2) $i < \ell$ and $j \geq \ell$, or (3) $i, j \geq \ell$. Case (3) is impossible, since A does not cover any element of $O(A)$ and therefore $A + a_{2^\ell}$ does not cover any element of $O(A) + a_{2^\ell}$. In Case (2), we have

$$2a_j - a_i \geq 2a_{2^\ell} - a_{2^\ell-1} = a_{2^\ell} + (a_{2^\ell} - a_{2^\ell-1}).$$

Since we picked ℓ large, $a_{2^\ell} - a_{2^\ell-1}$ is larger than all elements of $O(A)$, so we have $2a_j - a_i > z'$. Therefore, Case (1) is the only possibility.

We claim that $0 \leq i < 2^k$ and $2^\ell - 2^k \leq j < 2^\ell$ hold. Suppose to the contrary that we have $i \geq 2^k$. Then we conclude

$$\begin{aligned} 2a_j - a_i &\leq 2a_{2^\ell-1} - a_{2^k} \\ &= a_{2^\ell} + \lambda - 1 - a_{2^k}. \end{aligned}$$

Note that we have $2a_{2^k} > 2a_{2^k-1} = a_{2^k} + \lambda - 1$, thus $a_{2^\ell} + \lambda - 1 - a_{2^k} < a_{2^\ell}$. This contradicts the assumption $2a_j - a_i = z' > a_{2^\ell}$. Likewise, in the case $j < 2^\ell - 2^k$, we apply Lemma 3.2, Eq. (2) to conclude:

$$\begin{aligned} 2a_j - a_i &\leq 2a_{2^\ell-2^k-1} \\ &= 2(a_{2^\ell-1} - a_{2^k}) \\ &= a_{2^\ell} + \lambda - 1 - 2a_{2^k} \\ &< a_{2^\ell}, \end{aligned}$$

which again is a contradiction. Hence, $0 \leq i < 2^k$ and $2^\ell - 2^k \leq j < 2^\ell$ hold true. Let $h = j - 2^\ell + 2^k$, so $a_i, a_h \in A^*$. By Lemma 3.2, Eq. (3), we have

$$\begin{aligned} 2a_h - a_i &= z' - 2a_{2^\ell-2^k} \\ &= z' - 2(a_{2^\ell-1} - a_{2^k-1}) \\ &= z' - a_{2^\ell} + a_{2^k} \\ &= z. \end{aligned}$$

Thus, z is covered by A^*. It follows that $O(A) + c + x$ is covered by $A^* + x$, completing our proof of part (a).

We now prove part (b). Note that if z is covered by A^*, then $z+2y-x$ is jointly covered by A^*+x and A^*+y. Applying part (a), then, we conclude that A^*+x and A^*+y jointly cover the set $[2y-x, c+2y-x)\setminus((A^*\cup O(A))+2y-x)\cup(O(A)+c+2y-x)$. Furthermore, $A^* + x$ and $A^* + y$ jointly cover $A^* + 2y - x$ because, for each $a \in A^*$, the integers $a + x$, $a + y$, $a + 2y - x$ form an arithmetic progression.

Parts (c) and (d) follow from parts (a) and (b), respectively, by setting $k \leftarrow k + 1$. Part (e) follows from part (a) by setting $k \leftarrow k + 2$. □

The following proposition is the driving force behind the proof of Theorem 2.5.

Proposition 3.4. *Let $S(A) = (a_n)$, and suppose that (a_n) is independent. Pick k greater than $\kappa(A)$, and let $A^* = \{a_0, a_1, \ldots, a_{2^k-1}\}$. Pick d an integer satisfying $\omega(A) < d \leq a_{2^k} - \lambda(A)$ and let*

$$A_k^d = A^* \cup (A^* + a_{2^k}) \cup (A^* + 7a_{2^k} - d) \cup (A^* + 8a_{2^k} - d).$$

Then, $S(A_k^d)$ is independent and satisfies

$$\begin{aligned} \rho(A_k^d) &= 10a_{2^k} - d \\ \alpha(A_k^d) &= \frac{10\alpha(A)}{9} - \frac{d}{3^{k+2}}. \end{aligned}$$

Before proving the proposition, we provide a motivating result from [5].

Proposition 3.5 (*Rolnick [5]*). *Let $S(A) = (a_n)$ and $S(B) = (b_n)$. Pick k at least $\kappa(A)$ and let $A^* = \{a_0, a_1, \ldots, a_{2^k-1}\}$. We define:*

$$A \otimes_k B := \{a_{2^k}b + a \mid a \in A^*, b \in B\}.$$

Then, if $S(A)$ and $S(B)$ are independent, $S(A \otimes_k B)$ is independent and admits the following description:

$$S(A \otimes_k B) = \{a_{2^k}b + a \mid a \in A^*, b \in S(B)\}.$$

Remark 3.6. Proposition 2.4 implies that a Stanley sequence $S(A)$ satisfies $S(A \otimes_\kappa \{0\}) = S(A)$ (for some κ) if and only if $S(A)$ is independent.

Remark 3.7. It is readily verified that the scaling factor of $S(A \otimes_k B)$ is simply the product of the scaling factors of $S(A)$ and $S(B)$. Taking $A_0 = \{0\}$ and $B = \{0, 1, 7\}$, so that $\alpha(B) = 10/9$, it follows that iterated products $S(A_n) = S(A_{n-1} \otimes_k B)$ satisfy $\alpha(A_n) = (10/9)^n$. Hence, the scaling factor can be made arbitrarily large. Theorem 2.5 clearly proves a much stronger result.

In the light of Proposition 3.5, Proposition 3.4 may be seen as defining Stanley sequences that are in some sense "intermediate" between

$$\begin{aligned} &S(A \otimes_k \{0, 1, 6, 7\}) = S\left(A^* \cup (A^* + a_{2^k}) \cup (A^* + 6a_{2^k}) \cup (A^* + 7a_{2^k})\right) \\ \text{and} \quad &S(A \otimes_k \{0, 1, 7, 8\}) = S\left(A^* \cup (A^* + a_{2^k}) \cup (A^* + 7a_{2^k}) \cup (A^* + 8a_{2^k})\right). \end{aligned}$$

Proof of Proposition 3.4. Make the following definitions: $\lambda := \lambda(A)$, $\omega := \omega(A)$, $b := a_{2^k-1}$, $c := a_{2^k}$. We also define the following sets:

$$\begin{aligned}
B &:= A^* \cup (A^* + c)\\
C &:= (A^* + 7c - d) \cup (A^* + 8c - d)\\
D &:= (A^* + 10c - d) \cup (A^* + 11c - d)\\
E &:= (A^* + 17c - 2d) \cup (A^* + 18c - 2d)\\
F &:= (A^* + 30c - 3d) \cup (A^* + 31c - 3d)\\
G &:= (A^* + 37c - 4d) \cup (A^* + 38c - 4d)\\
H &:= (A^* + 40c - 4d) \cup (A^* + 41c - 4d)\\
I &:= (A^* + 47c - 5d) \cup (A^* + 48c - 5d)\\
J &:= B \cup C \cup D \cup E \cup F \cup G \cup H \cup I.
\end{aligned}$$

Thus we have $A_k^d = B \cup C$. Our approach is as follows. We prove that (i) J is 3-free and (ii) the set J covers all integers between $\max(C)$, which equals $b+8c-d$, and $\max(I)$, which equals $b+48c-5d$, with the exception of J itself. This implies that J is a prefix of $S(A_k^d)$. We now may apply Lemma 3.1 to prove that $S(A_k^d)$ is independent. We will use (a'_n) to denote the elements of $S(A_k^d)$. In order to apply this lemma, we require the condition $a'_{2^{k+3}-1} \geq \lambda(A_k^d) + \omega(A_k^d)$. (This is the reason we must consider such a large prefix subsequence of $S(A_k^d)$.) We may verify this condition as follows:

$$\begin{aligned}
\lambda(A_k^d) &= 2(b + 8c - d) - (10c - d) + 1\\
&= 2b + 6c - d + 1\\
&< 8c - d\\
\omega(A_k^d) &= \omega(A) + 8c - d\\
&< b + 8c - d.
\end{aligned}$$

Hence,

$$\begin{aligned}
a'_{2^{k+3}-1} &= b + 18c - 2d\\
&> (8c - d) + (b + 8c - d)\\
&> \lambda(A_k^d) + \omega(A_k^d).
\end{aligned}$$

We now show that J is 3-free. Suppose towards contradiction that $x, y, z \in J$ form an arithmetic progression with $x < y < z$. Observe that J reduces modulo $10c - d$ to $B \cup C$:

$$J = (B \cup C) \cup (B \cup C + (10c - d)) \cup (B \cup C + 3(10c - d)) \cup (B \cup C + 4(10c - d)).$$

There is no 3-term arithmetic progression in the set $\{w, w + (10c - d), w + 3(10c - d), w + 4(10c - d)\}$ for any value of w; hence, x and y must be distinct modulo $10c - d$.

Notice that we have $C \cup D \equiv B \cup C \pmod{10c - d}$. Let $x', y' \in C \cup D$ be congruent, respectively, to x, y modulo $10c - d$, and let $z' = 2y' - x'$, so that x', y', z' form a 3-term arithmetic progression (possibly decreasing). Because we have $2y' - x' \equiv z \pmod{10c - d}$, we conclude $z' \in B \cup C \cup D \cup E$. Since x', y' are both at least $7c - d$ and at most $b + 11c - d$, we have

$$-b + 3c - d \leq 2y' - x' \leq 2b + 15c - d.$$

From $d \leq c - \lambda$, we conclude:

$$\begin{aligned}
&z' \geq -b + 3c - d = b - \lambda + 1 + 2c - d \geq b + c + 1,\\
&\text{and} \quad z' \leq 2b + 15c - d = 16c + \lambda - 1 - d \leq 17c - 2d - 1.
\end{aligned}$$

Hence, z' cannot be in B or E, so z' is in $C \cup D$. But $C \cup D$ is 3-free since we have

$$C \cup D = \{a_n + 7c - d \mid 0 \leq n < 2^{k+2}\},$$

and we know that (a_n) is 3-free. Therefore, x', y', z' cannot form an arithmetic progression, a contradiction. We conclude that J is 3-free, as desired.

We now use repeated applications of Lemma 3.3 to prove that the set J covers all elements of $[b+8c-d, b+48c-5d] \setminus J$.

By part (e) of Lemma 3.3, $C \cup D$ covers the set

$$[7c - d, 16c - d) \setminus (C \cup D \cup (O(A) + 7c - d)) \cup (O(A) + 16c - d). \tag{4}$$

By part (d), B and C jointly cover

$$[14c - 2d, 17c - 2d) \setminus (O(A) + 14c - 2d) \cup (O(A) + 17c - 2d). \tag{5}$$

By part (c), E covers

$$[17c - 2d, 20c - 2d) \setminus (E \cup (O(A) + 17c - 2d)) \cup (O(A) + 20c - 2d). \tag{6}$$

By part (d), B and D jointly cover

$$[20c - 2d, 23c - 2d) \setminus (O(A) + 20c - 2d) \cup (O(A) + 23c - 2d). \tag{7}$$

By part (b), $(A^* + 11c - d)$ and $(A^* + 17c - 2d)$ jointly cover

$$[23c - 3d, 24c - 3d) \setminus (O(A) + 23c - 3d) \cup (O(A) + 24c - 3d). \tag{8}$$

By part (d), D and E jointly cover

$$[24c - 3d, 27c - 3d) \setminus (O(A) + 24c - 3d) \cup (O(A) + 27c - 3d). \tag{9}$$

By part (d), C and E jointly cover

$$[27c - 3d, 30c - 3d) \setminus (O(A) + 27c - 3d) \cup (O(A) + 30c - 3d). \tag{10}$$

By part (c), F covers

$$[30c - 3d, 33c - 3d) \setminus (F \cup (O(A) + 30c - 3d)) \cup (O(A) + 33c - 3d). \tag{11}$$

By part (d), E and F jointly cover

$$[33c - 4d, 36c - 4d) \setminus (O(A) + 33c - 4d) \cup (O(A) + 36c - 4d). \tag{12}$$

By part (d), B and E jointly cover

$$[34c - 4d, 37c - 4d) \setminus (O(A) + 34c - 4d) \cup (O(A) + 37c - 4d). \tag{13}$$

By part (e), $G \cup H$ covers

$$[37c - 4d, 46c - 4d) \setminus (G \cup H \cup (O(A) + 37c - 4d)) \cup (O(A) + 46c - 4d). \tag{14}$$

By part (d), F and G jointly cover

$$[44c - 5d, 47c - 5d) \setminus (O(A) + 44c - 5d) \cup (O(A) + 47c - 5d). \tag{15}$$

By part (c), I covers

$$[47c - 5d, 50c - 5d) \setminus (I \cup (O(A) + 47c - 5d)) \cup (O(A) + 50c - 5d). \tag{16}$$

Combining the sets in (4)–(7), we conclude that J covers the set

$$[7c - d, 23c - 2d) \setminus (C \cup D \cup E). \tag{17}$$

Combining the sets in (8)–(11), we conclude that J covers

$$[23c - 3d, 33c - 3d) \setminus (F \cup (O(A) + 23c - 3d)). \tag{18}$$

Combining the sets in (12)–(16), we conclude that J covers

$$[33c - 4d, 50c - 5d) \setminus (G \cup H \cup I \cup (O(A) + 33c - 4d)). \tag{19}$$

The largest element of $O(A) + 23c - 3d$ is $\omega + 23c - 3d$, which is less than $23c - 2d$ because we have $d > \omega$. Likewise, the largest element of $O(A) + 33c - 4d$ is $\omega + 33c - 4d$, which is less than $33c - 3d$. Hence, we can combine the sets in (17)–(19), into

$$[7c - d, 50c - 5d) \setminus (C \cup D \cup E \cup F \cup G \cup H \cup I). \tag{20}$$

In particular, J covers the set $[b + 8c - d, b + 48c - 5d] \setminus J$, which is a subset of (20).

Since J is 3-free and covers all elements of $[b+8c-d, b+48c-5d] \setminus J$, we conclude that J is a prefix of $S(A_k^d)$. Therefore, by Lemma 3.1, the sequence $S(A_k^d)$ is independent. □

As a result of the construction given in Proposition 3.4, we obtain the following result.

Proposition 3.8. *Let $S(A) = (a_n)$, and suppose that (a_n) is independent with scaling factor α and repeat factor ρ.*

a. *Suppose that α' is a triadic number satisfying $\alpha \le \alpha' < 10\alpha/9$. Then, α' is the scaling factor of some independent Stanley sequence.*
b. *Let ϵ be strictly positive. There exists an integer $N_\epsilon(A)$ such that for all k at least $N_\epsilon(A)$, every integer in the interval $[3^k(1+\epsilon)\rho, 3^k(10/9-\epsilon)\rho]$ is the repeat factor of some independent Stanley sequence.*

Proof. Let $\lambda = \lambda(A)$, $\omega = \omega(A)$, and $\kappa = \kappa(A)$.

(i) Clearly α itself is a scaling factor, so suppose we have $\alpha' > \alpha$. From Proposition 3.4, we have

$$\alpha(A_k^d) = \frac{10\alpha(A)}{9} - \frac{d}{3^{k+2}},$$

for k large enough. Let $d = 3^k(10\alpha - 9\alpha')$, so that we have

$$\alpha' = \frac{10\alpha}{9} - \frac{d}{3^{k+2}}.$$

Note that d is an integer for large k, because α' is a triadic number. Since we have $\alpha' < 10\alpha/9$, the condition $d > \omega(A)$ is satisfied for k large enough. Likewise, since we have $\alpha' > \alpha$, we conclude that $d = t\alpha 3^k$ is satisfied for some value t independent of k, such that $t < 1$. Hence, we can make k large enough to satisfy the condition $d \le a_{2^k} - \lambda = \alpha 3^k - \lambda(A)$. We conclude that there exists an independent Stanley sequence for which α' is the scaling factor.

(ii) Proposition 3.4 implies

$$\rho(A_k^d) = 10a_{2^k} - d,$$

for k large enough. Thus, $\rho(A_k^d)$ may take on any integral value ρ' satisfying

$$9a_{2^k} + \lambda = 10a_{2^k} - (a_{2^k} - \lambda) \le \rho' < 10a_{2^k} - \omega.$$

Pick k_ϵ large enough so that for every k at least k_ϵ, we have

$$9a_{2^k} + \lambda = 3^{k-\kappa+2} \cdot \rho(A) + \lambda < 3^{k-\kappa+2}(1+\epsilon)\rho$$

$$\text{and} \quad 10a_{2^k} - \omega = 3^{k-\kappa+2} \cdot \frac{10}{9}\rho(A) - \omega > 3^{k-\kappa+2}\left(\frac{10}{9} - \epsilon\right)\rho,$$

where we used the equality $\rho(A) = a_{2^\kappa} = a_{2^k}/3^{k-\kappa}$. Let $N_\epsilon(A) = k_\epsilon - \kappa + 2$. Then, for each k at least $N_\epsilon(A)$, every integer in the interval $[3^k(1+\epsilon)\rho, 3^k(10/9-\epsilon)\rho]$ is the repeat factor of some independent Stanley sequence. □

We now are able to prove Theorem 2.5.

Proof of Theorem. (i) We apply Proposition 3.8(a) to the sequence $S(A)$ with $A = \{0\}$, which satisfies $\alpha(A) = 1$. Hence, every triadic number $\alpha' \in [1, 10/9)$ is a valid scaling factor. Applying Proposition 3.8 again, we see that every triadic number $\alpha'' \in [1, 100/81)$ is a scaling factor. Continuing in this way, we conclude that every triadic number in $[1, (10/9)^n)$ is a scaling factor, for any value of n. Since $(10/9)^n$ can be made arbitrarily large, we conclude that every triadic number α that is at least 1 is a valid scaling factor.

(ii) Pick some small ϵ greater than 0. We apply Proposition 3.8(b) to the sequence $S(A_1)$ with $A_1 = \{0\}$, which satisfies $\rho(A_1) = 1$. For every k_1 satisfying $k_1 \ge N_\epsilon(A_1) = N_1$, each integer x in the interval $[3^{k_1}(1+\epsilon), 3^{k_1}(10/9-\epsilon)]$ is the repeat factor of some independent sequence $S(A_x)$.

We next apply Proposition 3.8(b) to each sequence $S(A_x)$, for each $x \in [3^{N_1}(1+\epsilon), 3^{N_1}(10/9-\epsilon)]$. For every k_x satisfying $k_x \ge N_\epsilon(A_x)$, each integer in the interval $[3^{k_x}(1+\epsilon)x, 3^{k_x}(10/9-\epsilon)x]$ is a repeat factor. These intervals overlap as x varies over the integers in $[3^{N_1}(1+\epsilon), 3^{N_1}(10/9-\epsilon)]$. Hence, for every k satisfying $k \ge N_1 + \max_x N_\epsilon(A_x) = N_2$, each integer y in the interval $[3^k(1+\epsilon)^2, 3^k(10/9-\epsilon)^2]$ is the repeat factor of some independent sequence $S(A_y)$.

We may now apply Proposition 3.8(b) to each sequence $S(A_y)$ such that $y \in [3^{N_2}(1+\epsilon)^2, 3^{N_2}(10/9-\epsilon)^2]$. Continuing in this manner, we conclude that, for each n, there exists N_n such that the following property holds: For k at least N_n, every integer in the interval $[3^k(1+\epsilon)^n, 3^k(10/9-\epsilon)^n]$ is the repeat factor of some independent sequence. For ϵ small, we can pick n satisfying $(10/9-\epsilon)^n > 3\cdot(1+\epsilon)^n$. Then, we have $3^k(10/9-\epsilon)^n > 3^{k+1}(1+\epsilon)^n$, so the intervals $[3^k(1+\epsilon)^n, 3^k(10/9-\epsilon)^n]$ overlap for k at least N_n. Every sufficiently large integer is contained in one of these intervals and hence must be the repeat factor of an independent Stanley sequence. □

4. Open problems

There remain many unanswered questions related to the growth of Stanley sequences. Our proof leaves open the question of which integers are not the repeat factor of any independent sequence. Rolnick additionally posed the problem of identifying which values of $\lambda(A)$ are attainable.

Conjecture 4.1 (*Rolnick [5]*). *Let* λ *be any integer other than* 1, 3, 5, 9, 11, 15. *Then, there exists an independent Stanley sequence* $S(A)$ *that satisfies* $\lambda(A) = \lambda$.

Dependent Stanley sequences, which are described in [5], follow Type 1 growth like independent sequences. However, while independent sequences satisfy $a_{2^k} = \alpha \cdot 3^k$, dependent sequences satisfy $a_{2^k - \sigma} = \alpha \cdot 3^k + \beta \cdot 2^k$, where β and σ are constants. Rolnick conjectured that β is nonnegative; further investigation is called for.

It appears very hard to show that every Stanley sequence follows either Type 1 or Type 2 growth. Weaker results have been proven for general 3-free sets that do not use the Stanley sequence property. A classic result by Roth [6] implies that no 3-free sequence (a_n) can have linear density. Sanders [7] recently improved this to $a_n = \Omega(n \log^{1-o(1)} n)$. Erdős et al. posed the problem of showing that every Stanley sequence (a_n) satisfies $a_n = \Omega(n^{1+\epsilon})$ for some positive ϵ. This remains open.

Acknowledgments

This research was conducted at the University of Minnesota Duluth REU, funded by NSF Grant 1358659 and NSA Grant H98230-13-1-0273. The authors would like to thank Joseph Gallian for his advice and Adam Hesterberg, Noah Arbesfeld, Daniel Kriz, and the other visitors and participants in the Duluth REU for their helpful discussions and valuable suggestions.

References

[1] P. Erdős, V. Lev, G. Rauzy, C. Sándor, A. Sárközy, Greedy algorithm, arithmetic progressions, subset sums and divisibility, Discrete Math. 200 (1999) 119–135.
[2] S. Lindhurst, An investigation of several interesting sets of numbers generated by the greedy algorithm (Senior thesis), Princeton University, 1990.
[3] R.A. Moy, On the growth of the counting function of Stanley sequences, Discrete Math. 311 (2011) 560–562.
[4] A.M. Odlyzko, R.P. Stanley, Some Curious Sequences Constructed with the Greedy Algorithm, Bell Laboratories internal memorandum, 1978.
[5] D. Rolnick, On the classification of Stanley sequences. arXiv:1408.1940.
[6] K.F. Roth, On certain sets of integers, J. London Math. Soc. 28 (1953) 104–109.
[7] T. Sanders, On Roth's theorem on progressions, Ann. of Math. 174 (2011) 619–636.
[8] S. Savchev, F. Chen, A note on maximal progression-free sets, Discrete Math. 306 (2006) 2131–2133.