On generalized Stanley sequences

Sándor Z. Kiss [*], Csaba Sándor [†], Quan-Hui Yang [‡]

Abstract

Let $\mathbb{N}$ denote the set of all nonnegative integers. Let $k \geq 3$ be an integer and $A_0 = \{a_1, \ldots, a_t\}$ ($a_1 < \ldots < a_t$) be a nonnegative set which does not contain an arithmetic progression of length k. We denote $A = \{a_1, a_2, \ldots\}$ defined by the following greedy algorithm: if $l \geq t$ and $a_1, \ldots, a_l$ have already been defined, then a_{l+1} is the smallest integer $a > a_l$ such that $\{a_1, \ldots, a_l\} \cup \{a\}$ also does not contain a k-term arithmetic progression. This sequence A is called the Stanley sequence of order k generated by A_0. In this paper, we prove some results about various generalizations of the Stanley sequence.

2010 Mathematics Subject Classification: Primary 11B75.

Keywords and phrases: Stanley sequences, arithmetic progression, probabilistic method

1 Introduction

Let $\mathbb{N}$ denote the set of all nonnegative integers. For a finite set $A_0 \subset \mathbb{N}$, $A_0 = \{a_1, \ldots, a_t\}$ ($a_1 < \ldots < a_t$) which does not contain an arithmetic progression of length k, we denote

$A = \{a_1, a_2, \dots\}$ the sequence defined by the following greedy algorithm: if $l \geq t$ and $a_1, \dots, a_l$ have already been defined, then a_{l+1} is the smallest integer $a > a_l$ such that $\{a_1, \dots, a_l\} \cup \{a\}$ does not contain an arithmetic progression of length k. This sequence is called Stanley sequence of order k generated by A_0.

Remark 1. *If $k = 3$, then A is a Stanley sequence of order 3 if and only if $n \in A \Leftrightarrow n \neq 2b - a$, where a, $b < n$ and $a, b \in A$.*

To investigate the density of sets without arithmetic progressions is one of the most popular topic in additive combinatorics. In 1953, Roth [9] proved that every subset of the set of integers with positive upper density contains an arithmetic progression of length three. On the other hand, Behrend [2] constructed a dense set without any arithmetic progression of length three. The name Stanley sequences established by Erdős et al. [4] and the definition originates with Odlyzko and Stanley from 1978. In their joint paper, they [8] constructed sets without arithmetic progression of length three by using the greedy algorithm. In this paper we generalize the concept of Stanley sequences in two directions. First, we will define the AP_k - covering sequences. In the first three theorems, we study the density of these sequences. In the other direction, we extend the definition of Stanley sequence according to Remark 1. In the last theorem we give a fully description of the structure of such sets when $A_0 = \{a_0\}$. Now we give the notations and definitions we are working with.

Let $A(n)$ be the number of elements of A up to n i.e.,

$$A(n) = \sum_{\substack{a \in A \\ a \leq n}} 1.$$

We denote $f = O(g)$ by $f \ll g$. Gerver and Ramsey [5] proved that if A is a Stanley sequence of order 3, then

$$\liminf_{n \to \infty} \frac{A(n)}{\sqrt{n}} \geq \sqrt{2}.$$

A few years later, Moy [7] rediscovered this inequality. Recently Chen and Dai [3] proved that if A is a Stanley sequence of order 3, then

$$\limsup_{n \to \infty} \frac{A(n)}{\sqrt{n}} \geq 1.77.$$

We say a sequence $A \subseteq \mathbb{N}$ is an AP_k - *covering sequence* if there exists an integer n_0 such that if $n > n_0$, then there exist $a_1 \in A, \dots, a_{k-1} \in A$, $a_1 < \cdots < a_{k-1} < n$ such that

$a_1, \ldots, a_{k-1}, n$ form a k-term arithmetic progression. Clearly, if A is a Stanley sequence of order k, then A is also an AP_k - covering sequence.

Using Gerver and Ramsey' idea, we can give a lower bound for $A(n)$ if A is an AP_k-covering sequence. Obviously

$$n - n_0 \leq |\{(a_m, b_m) : n_0 < m \leq n, a_m, b_m < m, a_m, b_m \in A,$$
$$a_m, b_m, m \text{ form an arithmetic progression of length three}\}| \leq \binom{A(n)}{2}.$$

Hence we have $A(n) \geq \sqrt{2n - 2n_0 + 0.25} + 0.5$, which implies

$$\liminf_{n \to \infty} \frac{A(n)}{\sqrt{n}} \geq \sqrt{2}.$$

Similarly, using Chen and Dai's proof, we may verify that if A is an AP_3 - covering sequence, then

$$\limsup_{n \to \infty} \frac{A(n)}{\sqrt{n}} \geq 1.77.$$

We omit the details.

In this paper we prove the following theorems.

Theorem 1. *There exists an AP_3 - covering sequence A such that*

$$\liminf_{n \to \infty} \frac{A(n)}{\sqrt{n}} \leq 2.$$

Theorem 2. *There exists an AP_3 - covering sequence A such that*

$$\limsup_{n \to \infty} \frac{A(n)}{\sqrt{n}} \leq 34.$$

Theorem 3. *There exists an AP_k - covering sequence A such that*

$$A(n) \ll_k (\log n)^{1/(k-1)} n^{\frac{k-2}{k-1}}.$$

We pose the following conjecture.

Conjecture 1. *(i) For any integer $k \geq 3$, there exists an AP_k - covering sequence A such that*

$$\limsup_{n \to \infty} \frac{A(n)}{n^{\frac{k-2}{k-1}}} < \infty.$$

(ii) For any AP_k - covering sequence A, we have

$$\liminf_{n \to \infty} \frac{A(n)}{n^{\frac{k-2}{k-1}}} > c_k,$$

where $c_k > 0$ is a constant and $k \geq 3$.

Finally we change the number 2 in Remark 1 to any integer k and obtain the following result.

Theorem 4. *Let $a_0 \geq 3$ and $k \geq 4$ be fixed. Let $A = \{a_0, \ldots\}$ be defined by the following greedy algorithm: for any integer $n > a_0$, $n \in A$ if and only if $n \neq kb - a$, where $a, b < n$ and $a, b \in A$. Then we have*

$$A = \bigcup_{n=0}^{\infty} [a_n, b_n],$$

where $b_0 = \lfloor \frac{ka_0}{2} \rfloor$, $a_l = kb_{l-1} - a_0 + 1$ and $b_l = \lfloor \frac{ka_l}{2} \rfloor$ for all integers $l \geq 1$.

Remark 2. *If one of the conditions $a_0 \geq 3$ and $k \geq 4$ does not hold, then some sequences generated by $\{a_0\}$ seems to be chaotic, without nice structure.*

2 Proof of Theorem 1

We define the sequence n_k recursively. Let $n_1 = 1$ and $n_{k+1} = 2^{2n_k+2}$ for $k = 1, 2, \ldots$. Define sets

$$A_k = \{n_k + 1, n_k + 2, \ldots, n_k + 2^{n_k+1}\} \cup \{3 \cdot 2^{n_k}, 4 \cdot 2^{n_k}, \ldots, (2^{n_k+1} + 2) \cdot 2^{n_k}\}$$

and $A = \cup_{k=1}^{\infty} A_k$.

Now we prove that for any integer n, there exist $a, b \in A$ with $a < b < n$ such that a, b, n form an arithmetic progression of length three.

Take an integer k such that $n_k + 3 \leq n < n_{k+1} + 3$. It is enough to prove that there exist $a, b \in A_k$ with $a < b < n$ such that a, b, n form an arithmetic progression of length three.

Case 1. $n_k + 3 \leq n \leq n_k + 2^{n_k+1}$. In this case, we take $a = n - 2$, $b = n - 1$. Then $a, b \in A$ and a, b, n form an arithmetic progression of length three.

Case 2. $n_k + 2^{n_k+1} + 1 \leq n < n_{k+1} + 3$. It follows that $n \leq 2^{2n_k+2} + 2$. Let

$$c = 2^{n_k} \cdot \left\lceil \frac{n}{2^{n_k+1}} \right\rceil,$$

where $\lceil x \rceil$ denotes the least integer not less than x. Then

$$\frac{n}{2} = 2^{n_k} \cdot \frac{n}{2^{n_k+1}} \leq c < 2^{n_k} \cdot \left(\frac{n}{2^{n_k+1}} + 1 \right) = \frac{n}{2} + 2^{n_k}.$$

Let $d = 2c - n$. Then $0 \leq d < 2^{n_k+1}$.

Subcase 2.1. $d > n_k$. It follows that $d \in A_k$. Noting that $2 \le \lceil \frac{n}{2^{n_k+1}} \rceil \le 2^{n_k+1} + 1$, we have

$$c = 2^{n_k} \cdot \left\lceil \frac{n}{2^{n_k+1}} \right\rceil \ge 2^{n_k+1} > d$$

and $c \in A_k$. Take $a = d$, $b = c$. Obviously, $a, b \in A_k$, $a < b < n$ and a, b, n form an arithmetic progression of length three.

Subcase 2.2. $d \le n_k$. Let $a = d + 2^{n_k+1}$, $b = c + 2^{n_k}$. Then $2b = a + n$. By $d = 2c - n \le n_k$ and $n_k + 1 + 2^{n_k+1} \le n$, we have

$$b = c + 2^{n_k} = \frac{d+n}{2} + 2^{n_k} \le \frac{n_k + n}{2} + 2^{n_k} < \frac{n}{2} + \frac{n}{2} = n,$$

$$a = 2b - n < 2b - b = b.$$

Noting that $2^{n_k+1} \le a \le 2^{n_k+1} + n_k$, b is a multiple of 2^{n_k} and $3 \cdot 2^{n_k} \le b \le (2^{n_k+1}+2) \cdot 2^{n_k}$, we have $a, b \in A_k$. Hence, there exist $a, b \in A_k$ with $a < b < n$ such that a, b, n form an arithmetic progression of length three.

Noting that $\min A_{k+1} = n_{k+1} + 1 > 2^{2n_k+2}$, we have

$$A(2^{2n_k+2}) \le n_k + |A_k| = n_k + 2 \cdot 2^{n_k+1}.$$

Thus

$$\liminf_{n \to \infty} \frac{A(n)}{\sqrt{n}} \le \liminf_{k \to \infty} \frac{A(2^{2n_k+2})}{2^{n_k+1}} \le \liminf_{k \to \infty} \frac{n_k + 2 \cdot 2^{n_k+1}}{2^{n_k+1}} = 2.$$

This completes the proof of Theorem 1.

3   Proof of Theorem 2

Let

$$B_k = \left\{ \sum_{i=0}^{k-1} \varepsilon_i 4^i : \varepsilon_i \in \{1, 2\} \right\}.$$

We will prove that the set

$$A = \bigcup_{k=1}^{\infty} \left(\bigcup_{i=0}^{8} (i \cdot 4^{k-1} + B_k) \right)$$

satisfies the conditions of Theorem 2. Clearly $B_k \subseteq A$ for any integer k.

We first prove that for any positive integers k and n with $3 \cdot 4^{k-1} \le n < 4^k$, there exist integers $a, b \in B_k$ such that $a < b < n$ and a, b, n form an arithmetic progression of length three. Write

$$n = \sum_{i=0}^{k-1} \mu_i 4^i,$$

where $\mu_i \in \{0, 1, 2, 3\}$. Take

$$a = \sum_{i=0}^{k-1} \varepsilon_i^{(1)} 4^i, \quad b = \sum_{i=0}^{k-1} \varepsilon_i^{(2)} 4^i,$$

where $\varepsilon_i^{(1)} = 2$, $\varepsilon_i^{(2)} = 1$ if $\mu_i = 0$; $\varepsilon_i^{(1)} = 1$, $\varepsilon_i^{(2)} = 1$ if $\mu_i = 1$; $\varepsilon_i^{(1)} = 2$, $\varepsilon_i^{(2)} = 2$ if $\mu_i = 2$; $\varepsilon_i^{(1)} = 1$, $\varepsilon_i^{(2)} = 2$ if $\mu_i = 3$. Since $3 \cdot 4^{k-1} \le n < 4^k$, it follows that $\mu_{k-1} = 3$, and so $\varepsilon_{k-1}^{(1)} = 1$, $\varepsilon_{k-1}^{(2)} = 2$. Hence $a < b < 3 \cdot 4^{k-1} \le n$ and $a, b \in B_k$. It is easy to see that $2b = a + n$, and so a, b, n form an arithmetic progression of length three.

Next we will prove that for any integer n, there exist $a, b \in A$ such that $a < b < n$ and a, b, n form an arithmetic progression.

Write

$$A_k = \bigcup_{i=0}^{8} \left(i \cdot 4^{k-1} + B_k \right), \quad k = 1, 2, \ldots.$$

For any integer n, let $3 \cdot 4^{t-1} \le n < 3 \cdot 4^t$ and let

$$n' = n - \left(\left\lfloor \frac{n}{4^{t-1}} \right\rfloor - 3 \right) \cdot 4^{t-1}.$$

Then we obtain $3 \cdot 4^{t-1} \le n' < 4^t$. By arguments above, it follows that there exist integers a', $b' \in B_t$ such that $a' < b' < n'$ form an arithmetic progression of length three. Now let

$$a = a' + \left(\left\lfloor \frac{n}{4^{t-1}} \right\rfloor - 3 \right) \cdot 4^{t-1},$$

$$b = b' + \left(\left\lfloor \frac{n}{4^{t-1}} \right\rfloor - 3 \right) \cdot 4^{t-1}.$$

Noting that $3 \cdot 4^{t-1} \le n < 3 \cdot 4^t$, we have $0 \le \left\lfloor \frac{n}{4^{t-1}} \right\rfloor - 3 \le 8$. Hence $a, b \in A_t$ and

$$b = b' + \left(\left\lfloor \frac{n}{4^{t-1}} \right\rfloor - 3 \right) \cdot 4^{t-1} < 3 \cdot 4^{t-1} + \left(\frac{n}{4^{t-1}} - 3 \right) \cdot 4^{t-1} = n.$$

By $2b = a + n$, we have $a < b < n$. Therefore, $a, b \in A$, $a < b < n$ and a, b, n form an arithmetic progression of length three.

In the next step we give an upper estimation of $A(n)$. It is clear that

$$A = \left(\bigcup_{k=1}^{\infty} B_k \right) \cup \left(\bigcup_{k=1}^{\infty} \left(\bigcup_{i=1}^{8} (i \cdot 4^{k-1} + B_k) \right) \right).$$

Write

$$B_1 = \bigcup_{k=1}^{\infty} B_k, \quad B_2 = \bigcup_{k=1}^{\infty} \left(\bigcup_{i=1}^{8} (i \cdot 4^{k-1} + B_k) \right).$$

Then $A = B_1 \cup B_2$. If

$$\left(\bigcup_{i=1}^{8} \left(i \cdot 4^{k-1} + B_k \right) \right) \cap [1, n] \neq \emptyset,$$

then we have $4^{k-1} \leq n$, and so $k \leq \log_4 n + 1$. It follows that

$$B_2(n) \leq \left| \bigcup_{k \leq \log_4 n+1} \left(\bigcup_{i=1}^{8} (i \cdot 4^{k-1} + B_k) \right) \right| \leq \sum_{k \leq \log_4 n+1} 8|B_k| = 8 \sum_{k \leq \log_4 n+1} 2^k < 32\sqrt{n}.$$

Let $4^{s-1} \leq n < 4^s$. Then

$$B_1(n) \leq B_1(4^s - 1) = 2^s \leq 2^{\log_4 n+1} = 2\sqrt{n}.$$

Hence, we obtain

$$A(n) \leq B_1(n) + B_2(n) < 34\sqrt{n}.$$

This completes the proof of Theorem 2.

4 Proof of Theorem 3

The proof of Theorem 3 is based on the probabilistic method due to Erdős and Rényi. There is an excellent summary of the probabilistic method in the books [1] and [6]. Let $P(E)$ denote the probability of an event E. Define the random set A by

$$P(n \in A) = \min \left\{ 1, c \left(\frac{\log n}{n} \right)^{\frac{1}{k-1}} \right\},$$

where c is a positive constant. Let

$$\frac{n}{2k} \leq u \leq \frac{n}{2(k-1)}$$

be fixed. Let

$$Y_{n,u} = \{n - iu : 1 \leq u \leq k - 1\}.$$

We prove that if $u \neq v$, then $Y_{n,u} \cap Y_{n,v} = \emptyset$. Otherwise, if $n - iu = n - jv$, then $iu = jv$, where $i \neq j$. We can assume that $i > j$, thus

$$\frac{k-1}{k-2} \leq \frac{i}{j} = \frac{u}{v} \leq \frac{k}{k-1},$$

which is impossible. Let $X_{n,u}$ denotes the event $Y_{n,u} \subset A$. For every $n \geq n_0$ we have

$$P(\nexists l : n - l, n - 2l, \ldots, n - (k-1)l \in A)$$

$$\leq P\left(\bigcap_{\frac{n}{2k} \leq u \leq \frac{n}{2(k-1)}} \bar{X}_{n,u} \right) = \prod_{\frac{n}{2k} \leq u \leq \frac{n}{2(k-1)}} \left(1 - \prod_{i=1}^{k-1} c \cdot \left(\frac{\log(n - iu)}{n - iu} \right)^{1/(k-1)} \right)$$

$$\leq \prod_{\frac{n}{2k} \leq u \leq \frac{n}{2(k-1)}} \left(1 - \prod_{i=1}^{k-1} c \cdot \left(\frac{\log n}{n} \right)^{1/(k-1)} \right) = \prod_{\frac{n}{2k} \leq u \leq \frac{n}{2(k-1)}} \left(1 - c^{k-1} \frac{\log n}{n} \right)$$

$$\leq \prod_{\frac{n}{2k} \leq u \leq \frac{n}{2(k-1)}} \exp\left(-c^{k-1} \frac{\log n}{n} \right) \leq \exp\left(-c^{k-1} \frac{\log n}{n} \cdot \frac{n}{2k(k-1)} \right)$$

$$\leq \exp(-2 \log n) = \frac{1}{n^2}.$$

if c is large enough. We will apply the following important lemma.

Lemma 1. *[6, Borel-Cantelli, See p.135] Let E_1, E_2, ... be a sequence of events in a probability space. If*

$$\sum_{j=1}^{+\infty} P(E_j) < +\infty,$$

then with probability 1, at most a finite number of the events E_j can occur.

It follows from Lemma 1 that with probability 1, there are only finitely many n such that there does not exist l such that $n - l, n - 2l, \ldots, n - (k-1)l \in A$. It is easy to see from the method of the proofs of Lemmas 10 and 11 in [6], pp. 144 - 145 that with probability 1, $A(n) \ll_k (\log n)^{1/(k-1)} \cdot n^{\frac{k-2}{k-1}}$. Thus, with probability 1, there exist AP_k - covering sets A with $A(n) \ll_k (\log n)^{1/(k-1)} \cdot n^{\frac{k-2}{k-1}}$.

5 Proof of Theorem 4

Let $I_l = [a_l, b_l]$ and $J_l = [b_l + 1, a_{l+1} - 1]$. First we prove that for any $n \in I_l$, a, $b \in A$ and a, $b < n$, we have $n \neq ka - b$.

Suppose that $n \in I_l$ and $n = ka - b$, where a, $b \in A$ and a, $b < n$. Then if $a \in I_j$ for some $j \leq l - 1$, we have

$$kb_{l-1} - a_0 + 1 = a_l \leq n = ka - b \leq kb_{l-1} - a_0,$$

a contradiction. If $a \in I_l$, and $b < n$ then

$$\left\lfloor \frac{ka_l}{2} \right\rfloor = b_l \geq n = ka - b \geq ka_l - (n - 1) \geq ka_l - \left\lfloor \frac{ka_l}{2} \right\rfloor + 1 \geq \left\lfloor \frac{ka_l}{2} \right\rfloor + 1,$$

which is also a contradiction.

Hence, for any $n \in I_l$, $a, b \in A$ and $a, b < n$, we have $n \neq ka - b$.

In the next step, we prove that for any integers $l \geq 0$ and $n \in J_l$, there exist $a, b \in A$ with $a, b < n$ such that $n = ka - b$.

Suppose that $n \in J_l$. For $h = 0, 1, \ldots, b_l - a_l$, we define

$$J_l^{(h)} = k(a_l + h) - I_l = \{k(a_l + h) - i : i \in I_l\}.$$

It is easy to see that the smallest element of $J_l^{(h+1)}$ is

$$\min J_l^{(h+1)} = k(a_l + h + 1) - \left\lfloor \frac{ka_l}{2} \right\rfloor$$

and the largest element of $J_l^{(h)}$ is

$$\max J_l^{(h)} = k(a_l + h) - a_l.$$

Since $k \geq 4$ and $a_0 \geq 3$, it follows that for any h with $0 \leq h \leq b_l - a_l - 1$, we have

$$\min J_l^{(h+1)} - \max J_l^{(h)} = k(a_l + h + 1) - \left\lfloor \frac{ka_l}{2} \right\rfloor - (k(a_l + h) - a_l) = k + a_l - \left\lfloor \frac{ka_l}{2} \right\rfloor \leq 1.$$

It follows that

$$[b_l + 1, kb_l - a_l] \subseteq \bigcup_{h=0}^{b_l - a_l} J_l^{(h)}.$$

Hence, for any integer $n \in [b_l + 1, kb_l - a_l]$, there exist integers h with $0 \leq h \leq b_l - a_l$ and $i \in I_l$ such that $n = k(a_l + h) - i$. Clearly $i \leq b_l < n$ and $a_l + h \leq a_l + (b_l - a_l) = b_l < n$. Thus we have $i \in A$, $a_l + h \in A$.

It remains to show that for any $kb_l - a_l + 1 \leq n \leq kb_l - a_0$ there exist $a, b \in A$, a, $b < n$ such that $n = ka - b$. If $l = 0$, then $kb_l - a_l = kb_0 - a_0 = a_1 - 1 = a_{l+1} - 1$, and so

$$[b_l + 1, kb_l - a_l] = [b_l + 1, a_{l+1} - 1].$$

Now we suppose that $l \geq 1$.

Let $K_l = \{ka - b : a \in I_l, b \in I_0\}$. Since $l \geq 1$, it follows that $a > b$. By $k \geq 4$, we have $a < ka - b$ and $b < ka - b$. By $k \geq 4$ and $a_0 \geq 3$, we have

$$|I_0| = b_0 - a_0 + 1 = \left\lfloor \frac{ka_0}{2} \right\rfloor - a_0 + 1 \geq k.$$

It follows that $K_l = [ka_l - b_0, kb_l - a_0]$. By $b_l = \lfloor \frac{ka_l}{2} \rfloor$ and $b_0 \geq k \geq 4$, we have

$$kb_l - a_l = k \left\lfloor \frac{ka_l}{2} \right\rfloor - a_l > k \left(\frac{ka_l}{2} - 1 \right) - a_l = \frac{k^2}{2} a_l - k - a_l > ka_l - b_0.$$

Hence

$$[b_l + 1, kb_l - a_l] \cup [ka_l - b_0, kb_l - a_0] = [b_l + 1, kb_l - a_o] = [b_l + 1, a_{l+1} - 1].$$

Therefore, if $n \in J_l$, then there exist $a, b \in A$ and $a, b < n$ such that $n = ka - b$.

This completes the proof of Theorem 4.

6 Acknowledgement

References

[1] N. Alon, J. H. Spencer, *The Probabilistic Method*, 4th ed., Wiley, 2016.

[2] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. of Sci. USA , **32** (1946), 331-332.

[3] Y-G. Chen, L-X. Dai, *On the counting function of Stanley sequences*, Publ. Math. Debrecen, **82** (2013), 91-95.

[4] P. Erdős, V. Lev, G. Rauzy, C. Sándor, A. Sárközy, *Greedy algorithm, arithmetic progressions, subset sums and divisibility*, Discrete Math., **200** (1999), 119-135.

[5] J. Gerver, L. T. Ramsey, *Sets of integers with no long arithmetic progressions generated by the greedy algorithm*, Mathematics of Computation, **33** (1979), 1353-1359.

[6] H. Halberstam, K. F. Roth, *Sequences*, Springer - Verlag, New York, 1983.

[7] R. A. Moy, *On the growth of the counting function of Stanley sequences*, Discrete Math. **311** (2011), 560-562.

[8] A. M. Odlyzko, R. P. Stanley, *Some curious sequences constructed with the greedy algorithm*, Bell Laboratories internal memorandum, 1978.

[9] K. F. Roth, *On certain sets of integers*, J. London Math. Soc., **28** (1953), 104-109.