ON A SYMMETRICITY PROPERTY CONNECTED TO THE EUCLIDEAN ALGORITHM

Srikanth Cherukupally
Department of Computer Science and Automation, Indian Institute of Science, Bangalore, Karnataka, India
cherukupal@iisc.ac.in, sricheru1214@gmail.com

Received: 7/7/17, Revised: 1/9/20, Accepted: 4/25/20, Published: 5/8/20

Abstract

We study an object: a sequence (collection) of arithmetic progressions with the property that the j^{th} terms of the i^{th} and $(i+1)^{th}$ progressions are the multiplicative inverses of each other, modulo the $(j+1)^{th}$ term of the i^{th} progression. In the study we address some combinatorial and algorithmic issues on a mirror symmetry (called the symmetricity property) satisfied by leading terms of progressions of such an object. The issues are in connection with the number of divisors k of integers of the form $x^2 - y^2$, with k falling in specific intervals. Our study explores a new perspective on the quotient sequence of the standard Euclidean algorithm on relatively-prime input pairs. Some open issues are left concerning the symmetricity property.

1. Introduction

Let $A(a, d)$ denote an arithmetic progression of integers with leading term a and common difference d.

The starting point of our study is the fact that, for a given progression $A(a, d)$, there exists a progression $A(a', d')$ such that the terms of the two progressions satisfy the property:

$$(a + jd)(a' + jd') \equiv 1 \pmod{a + (j+1)d},\ j \geq 0.$$

In other words, the property states that the j^{th} terms of $A(a, d)$ and $A(a', d')$ are the multiplicative inverses of each other modulo the $(j+1)^{th}$ term of $A(a, d)$. We refer to this invertible property as *Property* $\mathcal{P}$. Note that this property holds only when a and d are co-prime.

We prove the following main result, which shows the uniqueness of existence of $A(a', d')$ when $d' \leq d$.

Theorem 1. *For a given progression $A(a, d)$ with a and d being co-prime, there exists only one progression $A(a', d')$ with $1 \leq d' \leq d$ such that, for any $i \geq 0$,*

$$(a + id)(a' + id') \equiv 1 \pmod{a + (i + 1)d}.$$

The above result allows us to build inductively, from a given progression $A(a_0, d_0)$, a unique sequence of progressions $A(a_0, d_0)$, $A(a_1, d_1)$, $A(a_2, d_2)$, ..., where the terms of any two consecutive progressions of the sequence satisfy Property $\mathcal{P}$, i.e.,

$$(a_i + jd_i)(a_{i+1} + jd_{i+1}) \equiv 1 \pmod{a_i + (j + 1)d_i},\ i, j \geq 0,$$

and the common differences of the progressions satisfy $d_i \geq d_{i+1} \geq 1$. The entire sequence is constructed uniquely from the base progression $A(a_0, d_0)$ for co-prime integers a_0, d_0. Hence, we denote the sequence by $\mathfrak{S}(a_0, d_0)$.

As an illustration, the sequence $\mathfrak{S}(11, 25)$ is as follows.

$$\begin{array}{lllll} 11, & 36, & 61, & 86, & \ldots \\ 23, & 39, & 55, & 71, & \ldots \\ 17, & 24, & 31, & 38, & \ldots \\ 17, & 22, & 27, & 32, & \ldots \\ 13, & 16, & 19, & 22, & \ldots \\ 5, & 6, & 7, & 8, & \ldots \\ 5, & 6, & 7, & 8, & \ldots \end{array}$$

The common differences are in decreasing order and any two consecutive common differences are co-prime. So, the sequence $\mathfrak{S}(a_0, d_0)$ eventually has a progression with common difference 1. After that point the same progression repeats. This shows that the sequence has only finitely many distinct arithmetic progressions. The above example illustrates this fact.

The object $\mathfrak{S}(a_0, d_0)$ displays some interesting properties, which are connected to the quotient sequence arising in the standard Euclidean algorithm of finding the greatest common divisor of two given numbers. Essentially, alternating quotients of the quotient sequence of the Euclidean algorithm give the sizes of certain sub-collections of the sequence $\mathfrak{S}(a_0, d_0)$. We prove this connection. There has been extensive study of the Euclidean algorithm. The average behavior of the quotient sequence and the average number of iterations of the algorithm are studied in [3, 4, 5]. Our connection is, to the best of our knowledge, the first time it has been shown in the literature that the quotients of the Euclidean algorithm, on a given input co-prime pair, correspond to the cardinalities (or sizes) of certain sub-collections belonging to the defined object $\mathfrak{S}(a_0, d_0)$ for some pair (a_0, d_0). Further, a mirror symmetry exhibited by leading terms of progressions of $\mathfrak{S}(a_0, d_0)$ is subject to certain conditions on the quotient sequence of the Euclidean algorithm. The property is also related to the number of divisors k of integers of the form $x^2 - y^2$, with k falling in specific intervals. We prove these connections.

In the literature, collections of progressions have been considered in the context of covering integers. A collection of progressions is called a *regular covering system* if every integer belongs to at least one progression in the collection. Results on covering systems can be found in [7]. In this paper, our object is also a collection of progressions, but we have a different motivation. Our collection is ordered, thus we call it a sequence rather than a collection.

The following notation is used throughout the paper. Let $A(a, d)$ denote an arithmetic progression (AP) of integers with leading term a, and common difference d. The natural logarithm of x is denoted by $\log x$. The logarithm of x to base a is denoted by $\log_a x$. The number of divisors of an integer x is denoted by $\sigma_0(x)$. A function $f(x)$ is said to be $O(g(x))$ if $f(x) \leq cg(x)$ for some constant $c > 0$. An algorithm is said to be a *polynomial-time algorithm* if its running time is $O(n^k)$ for some constant $k > 0$, where n is the binary length of input given to the algorithm. For example, the standard Euclidean algorithm is called a polynomial-time algorithm, since its running time on input pair (a, b) is $O(\log^2 c)$ where $c = \max\{a, b\}$ (see [2]).

1.1. Our Results

We consider certain combinatorial and algorithmic aspects of the defined sequence $\mathfrak{S}(a_0, d_0)$. Before describing our main results on these aspects, we make the following remark.

Remark. The sequence $\mathfrak{S}(a_0, d_0)$ exists for any pair of co-prime integers a_0, d_0. In the present work, we maintain the condition: $1 \leq a_0 < d_0$, i.e., the common difference of the first progression of the sequence is greater than its leading term. This condition is enforced to quantify some sets which we define in studying the following aspects of $\mathfrak{S}(a_0, d_0)$.

Let $N(a_0, d_0)$ denote the number of distinct progressions of $\mathfrak{S}(a_0, d_0)$. The inductive procedure of computing leading terms and common differences is inefficient when the number of progressions is exponential in $\log d_0$. We establish a connection between $\mathfrak{S}(a_0, d_0)$ and the quotient sequence of the Euclidean algorithm. This result yields a polynomial-time (quick) algorithm for computing $N(a_0, d_0)$. The algorithm is a minor modification of the Euclidean algorithm. Using results [3, 4, 5] on the average case analysis of the Euclidean algorithm, the average value of $N(a_0, d_0)$ is proved to be $O(\log d_0)$ in Section 3.3.

Let d_0 be a fixed integer. For some $1 \leq a_0 < d_0$, the leading terms of a few initial progressions of $\mathfrak{S}(a_0, d_0)$ follow a *symmetricity* property in the following sense: there exists an integer $k < d_0$ such that the leading terms of the first $k + 1$ progressions of $\mathfrak{S}(a_0, d_0)$ follow the property that $a_i = a_{k-i}$, $0 \leq i \leq k$. In other words, the leading terms of the first $k+1$ progressions exhibit a mirror image symmetry about

the index $\lfloor \frac{k}{2} \rfloor$. We call such an integer a_0 a *symmetric number* for d_0.

The following example illustrates the fact that 17 is a symmetric number for 23. The leading terms of six initial progressions of $\mathfrak{S}(17, 23)$ follow symmetricity:

$$\begin{array}{llllll} 17, & 40, & 63, & 86, & 109 & \ldots \\ 33, & 52, & 71, & 90, & 109 & \ldots \\ 41, & 56, & 71, & 86, & 101 & \ldots \\ 41, & 52, & 63, & 74, & 85 & \ldots \\ 33, & 40, & 47, & 54, & 61 & \ldots \\ 17, & 20, & 23, & 26, & 29 & \ldots \\ 13, & 15, & 17, & 19, & 21 & \ldots \\ 7, & 8, & 9, & 10, & 11 & \ldots \end{array}$$

It can be verified that $\mathfrak{S}(1, d_0)$, for $d_0 > 1$, consists of two progressions whose leading terms are $1, 1$, respectively. So, 1 is a trivial symmetric number for any $d_0 > 1$. With these observations, it is natural to ask:

- What is the total number of symmetric numbers for a given d_0?

We show that for every divisor x of $d_0^2 - 1$ with $x < d_0$, there is a corresponding symmetric number for d_0. Precisely, for every divisor x of $d_0^2 - 1$ with $x < d_0$, $a_0 = d_0 - (x^{-1} \pmod{d_0})$ is a symmetric number for d_0. For example, 4 is a divisor of $23^2 - 1$ and, corresponding to it, $a_0 = 23 - (4^{-1} \pmod{23}) = 17$ is a symmetric number for 23. This result proves that the number of symmetric numbers for d_0 is equal to half the number of divisors of $d_0^2 - 1$ (Corollary 3).

In $\mathfrak{S}(17, 23)$, the leading terms of the first six progressions follow the symmetricity property. Notice that the property is also followed by the j^{th} terms of the same progressions for $1 \leq j \leq 5$. As the value of j increases, the number of j^{th} terms following symmetricity decreases. We prove this symmetricity property extended to higher terms of the progressions.

Having observed the symmetricity property of leading terms of some consecutive initial progressions of $\mathfrak{S}(a_0, d_0)$, we ask whether the leading terms of some later progressions of $\mathfrak{S}(a_0, d_0)$ also follow the symmetricity property. A general question is: do there exist integers $\alpha, \beta > 0$ such that the leading terms a_α, $a_{\alpha+1}$, ..., a_β satisfy: $a_{\alpha+i} = a_{\beta-i}$ for $0 \leq i \leq \beta - \alpha$? The answer is yes. For example, the leading terms of the i^{th} progressions[1] of $\mathfrak{S}(17, 37)$, $2 \leq i \leq 5$, satisfy the symmetricity property:

[1]Note that the indexing of progressions starts from 0.

17,	54,	91,	128,	...
35,	59,	83	107,	...
27,	38,	49,	60,	...
31,	40,	49,	58,	...
31,	38,	45,	52,	...
27,	32,	37,	42,	...
19,	22,	25,	28,	...
7,	8,	9,	10,	...
7,	8,	9,	10,	...

We prove that there are only certain sub-collections of consecutive progressions of $\mathfrak{S}(a_0, d_0)$ whose leading terms follow the symmetricity property. A *grouping* of $\mathfrak{S}(a_0, d_0)$ is a maximal sub-collection of consecutive progressions with respect to a common difference property (see Section 2.2 for the formal definition of grouping and illustrations). We prove that symmetricity happens only within groupings. Further, we establish a connection between the number of groupings of $\mathfrak{S}(a_0, d_0)$ with symmetricity and the quotient sequence of the Euclidean algorithm. The connection yields a polynomial time algorithm for computing the number of groupings of $\mathfrak{S}(a_0, d_0)$ with symmetricity. This (quick) algorithm enables us to conduct experiments concerning symmetricity and obtain empirical data on the size of

$$\mathcal{T}(d_0, k) = \{a_0 : \mathfrak{S}(a_0, d_0) \text{ has } k \text{ groupings with symmetricity}\},$$

for $k \geq 1$. One important observation from the data is that the number of groupings of $\mathfrak{S}(a_0, d_0)$ with symmetricity is at most 2 for any pair (a_0, d_0) with $1 \leq a_0 < d_0 \leq 10^6$. In other words, the observation is that $|\mathcal{T}(d_0, k)| = 0$ for $k \geq 3$. Our other observations are reported in Section 5. We are able to provide only partial answers toward proving the sizes. The observations are left as open issues.

The rest of the paper is organized as follows. In Section 2, we prove the main properties of $\mathfrak{S}(a_0, d_0)$, which will be used for proving the main results of the paper. In Section 3, we study the size of $\mathfrak{S}(a_0, d_0)$ and establish a connection between $\mathfrak{S}(a_0, d_0)$ and the Euclidean algorithm. In Section 4, we derive a necessary and sufficient condition for the symmetricity property and present combinatorial results on the property. In Section 5, we give some open combinatorial issues on the symmetricity property based on our empirical data.

2. Properties of $\mathfrak{S}(a_0, d_0)$

We first establish the existence and uniqueness of the sequence $\mathfrak{S}(a_0, d_0)$ by proving Theorem 1.

2.1. Proof of Theorem 1

The congruence property (Property $\mathcal{P}$) can be rewritten as

$$(-d)(a' + id') \equiv 1 \pmod{a + (i+1)d}.$$

From this, we obtain

$$z_i = \frac{(a' + id')d + 1}{a + (i+1)d} \text{ is an integer.}$$

So, for any $i \geq 1$,

$$z_{i+1} - z_i = \frac{Cd}{(a + id)(a + (i+1)d)} \text{ is an integer}$$

where $C = a'd + dd' - ad' - 1$. Since a, d, $a + id$ and $a + (i+1)d$ are pair-wise co-prime, we have C divisible by $a + id$ for any $i \geq 1$. This implies that $C = 0$. The expression for C can be rewritten as follows:

$$a' = d' + \frac{ad' - 1}{d}. \tag{1}$$

In the above formula, a' is an integer only when $\frac{ad'-1}{d}$ is an integer. For $d > 1$, $\frac{ad'-1}{d}$ is an integer only when d' is the multiplicative inverse of $a \pmod{d}$. There exists only one value for d' which is less than d. When $d = 1$, we have $d' = 1$ and $a' = a$. So, the progression $A(a', d')$ is unique if $1 \leq d' \leq d$. □

We reiterate that the above result allows us to construct inductively, from a given progression $A(a_0, d_0)$, a unique sequence of arithmetic progressions $\mathfrak{S}(a_0, d_0)$. The formal definition of the sequence is as follows.

Definition 1. $\mathfrak{S}(a_0, d_0)$ is the sequence of distinct arithmetic progressions

$$\langle A(a_0, d_0), A(a_1, d_1), A(a_2, d_2), \ldots \rangle,$$

where the terms of any two consecutive progressions of the sequence satisfy Property $\mathcal{P}$, i.e.,

$$(a_i + jd_i)(a_{i+1} + jd_{i+1}) \equiv 1 \pmod{a_i + (j+1)d_i},\ i, j \geq 0$$

and the common differences of the progressions satisfy the condition $d_i \geq d_{i+1} \geq 1$.

By Equation (1), leading terms a_i and common differences d_i satisfy the property:

$$\begin{aligned} d_{i+1} &\equiv a_i^{-1} \pmod{d_i}, \\ a_{i+1} &= d_{i+1} + \frac{a_i d_{i+1} - 1}{d_i}. \end{aligned} \tag{2}$$

We note that the definition of $\mathfrak{S}(a_0, d_0)$ captures the sequence to contain a progression with common difference 1 as its last progression. As remarked in the beginning of Section 1.1, we assume the condition that $1 \leq a_0 < d_0$ throughout the paper.

2.2. Groupings of $\mathfrak{S}(a_0, d_0)$

A sub-collection $\mathcal{G}$ of consecutive progressions of $\mathfrak{S}(a_0, d_0)$ is called a *grouping* if it satisfies the following two properties.

- The difference between the common differences of any two consecutive progressions in $\mathcal{G}$ is the same.
- $\mathcal{G}$ is maximal.

We refer to the difference between the consecutive common differences as the *second common difference* corresponding to $\mathcal{G}$. The size of $\mathcal{G}$ is the number of progressions in it, and is denoted by $|\mathcal{G}|$. Note that any two consecutive groupings share an arithmetic progression.

Example 1. Consider the sequence $\mathfrak{S}(11, 25)$. Let us pay attention to the common differences of the progressions: $(25, 16, 7, 5, 3, 1)$. With respect to the above definition of grouping, the common differences can be partitioned into two sets. So, the sequence $\mathfrak{S}(11, 25)$ has two groupings:

$$\begin{aligned} \mathcal{G}_1 &= \langle (11, 25), A(23, 16), A(17, 7)\rangle, \\ \mathcal{G}_2 &= \langle A(17, 7), A(17, 5), A(13, 3), A(5, 1)\rangle. \end{aligned}$$

The second common difference corresponding to $\mathcal{G}_1$ is 9. The second common difference corresponding to $\mathcal{G}_2$ is 2. The groupings share the progression $A(17, 7)$. The sizes of $\mathcal{G}_1$ and $\mathcal{G}_2$ are 3 and 4, respectively.

Example 2. The sequence $\mathfrak{S}(17, 37)$ has two groupings:

$$\begin{aligned} \mathcal{G}_1 &= \langle A(17, 37), A(35, 24), A(27, 11)\rangle, \\ \mathcal{G}_2 &= \langle A(27, 11), A(31, 9), A(31, 7), A(27, 5), A(19, 3), A(7, 1)\rangle. \end{aligned}$$

The second common difference corresponding to $\mathcal{G}_1$ is 13. The second common difference corresponding to $\mathcal{G}_2$ is 2. The size of $\mathcal{G}_1$ is 3. The size of $\mathcal{G}_2$ is 6. The progression $A(27, 11)$ belongs to the two groupings.

2.3. Properties of Terms Within a Grouping

Let $\mathcal{G}$ be a grouping of $\mathfrak{S}(a_0, d_0)$ consisting of progressions

$$A(a_\alpha, d_\alpha), A(a_{\alpha+1}, d_{\alpha+1}), \ldots, A(a_\beta, d_\beta),$$

for some $0 \leq \alpha < \beta$. Let $\triangle$ be the second common difference corresponding to $\mathcal{G}$. By the definition of grouping, $\triangle = d_r - d_{r+1}$, $\alpha \leq r \leq \beta - 1$.

We derive a defining equation satisfied by leading terms of the progressions in $\mathcal{G}$. The equation is derived though the following two lemmas.

Lemma 1. $\frac{a_{i+1}\triangle+1}{d_{i+1}} = \frac{a_i\triangle+1}{d_i} + \triangle$, $\alpha \le i \le \beta - 1$.

Proof. From Equation (2), $a_i d_{i+1} + d_i d_{i+1} - a_{i+1} d_i - 1 = 0$, $\alpha \le i < \beta$. Thus,

$$\begin{aligned} \frac{a_{i+1}\triangle+1}{d_{i+1}} - \frac{a_i\triangle+1}{d_i} &= \triangle\left(\frac{a_{i+1}d_i - a_i d_{i+1} + 1}{d_i d_{i+1}}\right) \\ &= \triangle. \end{aligned}$$

□

Lemma 2. *For $\alpha \le i \le \beta - 1$, $a_{i+1} - a_i = (\alpha + \beta - 1 - 2i)\triangle + d_\beta - z_\alpha$, where*

$$z_\alpha = \frac{a_\alpha\triangle+1}{d_\alpha}.$$

Proof. By introducing the terms $a_{i+1}d_{i+1} + (-a_{i+1}d_{i+1})$ and $d_{i+1}^2 + (-d_{i+1}^2)$ in the equation $a_i d_{i+1} + d_i d_{i+1} - a_{i+1} d_i - 1 = 0$, we obtain

$$\begin{aligned} a_{i+1} - a_i &= -\frac{a_{i+1}\triangle+1}{d_{i+1}} + \triangle + d_{i+1} \\ &= -\frac{a_{i+1}\triangle+1}{d_{i+1}} + \triangle + (d_\beta + (\beta - i - 1)\triangle) \\ &= -(z_\alpha + (i + 1 - \alpha)\triangle) + d_\beta + (\beta - i)\triangle \\ &= (\alpha + \beta - 1 - 2i)\triangle + d_\beta - z_\alpha. \end{aligned}$$

□

Corollary 1. $a_i = a_\alpha + \triangle(\beta - i)(i - \alpha) + (i - \alpha)(d_\beta - z_\alpha)$, $\alpha \le i \le \beta$.

Proof. From Lemma 2,

$$\begin{aligned} a_i - a_\alpha &= \sum_{j=\alpha}^{i-1}(a_{j+1} - a_j) \\ &= \sum_{j=\alpha}^{i-1}\big(\triangle(\alpha + \beta - 1 - 2j) + d_\beta - z_\alpha\big) \\ &= \triangle(\beta - i)(i - \alpha) + (i - \alpha)(d_\beta - z_\alpha). \end{aligned}$$

□

The above result shows that leading terms of the progressions (in $\mathcal{G}$) are evaluations of a polynomial $f(x) = ax^2 + bx + c$ at integer values, where a, b, c are some constants specific to $\mathcal{G}$. Since $a < 0$, $f(x)$ defines an inverted parabola. This indicates that there is a chance of some leading terms appearing more than once. This behavior of leading terms essentially gives rise to a mirror image symmetry, which

has been observed for the example sequences $\mathfrak{S}(11, 25)$, $\mathfrak{S}(17, 23)$ and $\mathfrak{S}(17, 37)$ in Section 1.1.

For the same example sequences, one can verify that the ratio of the leading term of a progression to its common difference gives the position (or index) of that progression in the sequence. This ratio property is inferred from the fact that the same holds within any grouping of $\mathfrak{S}(a_0, d_0)$. The following lemma proves the property for grouping $\mathcal{G}$.

Lemma 3. *For any* $\alpha \leq i \leq \beta$, $\left\lfloor \frac{a_i}{d_i} \right\rfloor = \left\lfloor \frac{a_\alpha}{d_\alpha} \right\rfloor + i - \alpha$.

Proof. From Lemma 1, we have $\frac{a_i \triangle + 1}{d_i} = \frac{a_\alpha \triangle + 1}{d_\alpha} + (i - \alpha)\triangle$. □

The ratio property is formally stated as follows.

Property 1. For a given sequence $\mathfrak{S}(a_0, d_0)$ consisting of progressions

$$\langle A(a_0, d_0), A(a_1, d_1), \ldots, A(a_l, d_l)\rangle,$$

we have $\left\lfloor \frac{a_j}{d_j} \right\rfloor = j$, for $0 \leq j \leq l$.

3. Size of $\mathfrak{S}(a_0, d_0)$

In this section, we consider the following three issues.

1. Computing $N(a_0, d_0)$, the total number of progressions of $\mathfrak{S}(a_0, d_0)$.
2. Establishing a connection between $\mathfrak{S}(a_0, d_0)$ and the Euclidean algorithm.
3. Estimating the average value of $N(a_0, d_0)$.

3.1. Computing $N(a_0, d_0)$

The quantity $N(a_0, d_0)$ can be computed by inductively generating the leading terms and common differences of the progressions of $\mathfrak{S}(a_0, d_0)$ using Equation (2). This inductive procedure is inefficient as it involves one inverse computation and division operations in each step (see Equation (2)). A much faster method would be to evaluate the formula for $N(a_0, d_0)$ given below.

Let $\mathcal{G}_1$, $\mathcal{G}_2$, …, $\mathcal{G}_m$ be the groupings of $\mathfrak{S}(a_0, d_0)$ with corresponding second common differences $\triangle_1$, $\triangle_2$, …, $\triangle_m$. Let j_r denote the index of the first progression of $\mathcal{G}_r$ in $\mathfrak{S}(a_0, d_0)$, $1 \leq r \leq m$. Let d_{j_r} be the common difference of the first progression of $\mathcal{G}_r$. Then, we have

$$N(a_0, d_0) = 1 - m + \sum_{r=1}^{m} \left\lceil \frac{d_{j_r}}{\triangle_r} \right\rceil.$$

The above identity holds true, since the size of $\mathcal{G}_r$ is $\left\lceil \frac{d_{j_r}}{\triangle_r} \right\rceil$ and any two consecutive groupings share an arithmetic progression.

The main idea of evaluating the formula for $N(a_0, d_0)$ is to compute the sizes of groupings of $\mathfrak{S}(a_0, d_0)$. The size of a grouping can be calculated without actually computing all leading terms of its progressions. This quick calculation is possible due to a relation between second common differences of successive groupings, which we observe below.

Let $d_0, d_1, \ldots, d_l$ be the common differences of the progressions of $\mathfrak{S}(a_0, d_0)$, and let $a_0, a_1, \ldots, a_l$ be the corresponding leading terms of the progressions. Here, $l = N(a_0, d_0) - 1$, which we want to compute. A *jump* is the difference between consecutive common differences. We count the number of common differences that are involved in the same jump[2]. The same approach is followed to get the formula for $N(a_0, d_0)$. But, here the common differences that are involved in two jumps are accounted for in the count corresponding to the smaller jump.

To illustrate the above rule, consider the common differences d_α, $d_{\alpha+1}$, ..., $d_{\beta-1}$, d_β, $d_{\beta+1}$, ..., d_γ. Let $\triangle = d_i - d_{i+1}$, $\alpha \leq i \leq \beta - 1$, and let $\triangle' = d_i - d_{i+1}$, $\beta \leq i \leq \gamma - 1$. Then, the count corresponding to $\triangle$ is $\lfloor d_\alpha / \triangle \rfloor$ and the count corresponding to $\triangle'$ is $\lfloor d_\beta / \triangle' \rfloor$. The common difference d_γ will be accounted for in the count corresponding to the jump that comes after $\triangle'$. The following lemma shows that $\lfloor d_\beta / \triangle' \rfloor$ can be expressed in terms of d_α and $\triangle$.

Lemma 4. *$d_\beta \equiv d_\alpha \pmod{\triangle}$, and $\triangle' \equiv \triangle \pmod{d_\beta}$.*

Proof. It is easy to verify that $d_\beta \equiv d_\alpha \pmod{\triangle}$. We prove the second congruence relation. From the proven properties of terms in Section 2.3, it is known that $\frac{a_\beta \triangle + 1}{d_\beta}$ is an integer. This implies that $\triangle \equiv -a_\beta^{-1} \pmod{d_\beta}$. By Equation (2), $d_{\beta+1} \equiv a_\beta^{-1} \pmod{d_\beta}$. Thus, $\triangle \equiv d_\beta - d_{\beta+1} \equiv \triangle' \pmod{d_\beta}$. □

The proven relation (above) results in a fast algorithm (given as Algorithm 1) for computing $N(a_0, d_0)$ and yields a connection between $\mathfrak{S}(a_0, d_0)$ and the Euclidean algorithm (given as Algorithm 2). The two congruences in the result reflect the two modular reduction steps (Steps 8 and 9) of Algorithm 1. Note that the algorithm replaces the ceiling function by the floor function in the formula for $N(a_0, d_0)$ and thereby removes the additive term $1 - m$.

Notation. The following notation is used in the descriptions of Algorithm 1 and Algorithm 2.

- $u \leftarrow v$ means the value of v is assigned to u.
- $x \leftarrow y \pmod{z}$ means the unique remainder resulting from the division of y by z is assigned to x.

[2]Note that the common differences that are involved in the same jump belong to the same grouping. So, the count of such common differences gives the size of the grouping

For example, if $y = 12$ and $z = 5$, then the value of x will be 2.

Algorithm 1 : Algorithm for computing $N(a_0, d_0)$

Input : a_0, d_0 with $d_0 > a_0 \geq 1$
Output : $N(a_0, d_0)$, the number of progressions of $\mathfrak{S}(a_0, d_0)$

1: $b \leftarrow a_0^{-1} \pmod{d_0}$
2: $\triangle \leftarrow d_0 - b$
3: $d \leftarrow d_0$
4: $n \leftarrow 0$
5: **while** $d > 1$ **do**
6: $\quad n \leftarrow n + \lfloor \frac{d}{\triangle} \rfloor$
7: $\quad$ If $\triangle \leq 1$, break the loop
8: $\quad d \leftarrow d \pmod{\triangle}$
9: $\quad \triangle \leftarrow \triangle \pmod{d}$
10: **end while**
11: $n \leftarrow n + 1$
12: Return n

3.2. Connection With the Euclidean Algorithm

In addition to computing $N(a_0, d_0)$, Algorithm 1 computes the greatest common divisor (gcd) of two numbers. The final value of d gives the gcd of d_0 and $\triangle$, where $\triangle$ is the quantity computed in Step 2 of the algorithm. Since our d_0 and $\triangle$ are co-prime, the gcd will be 1. Instead of setting $\triangle$ to $d_0 - b$ in Step 2, if $\triangle$ is chosen to be a random value $< d_0$, then the algorithm computes $gcd(d_0, \triangle)$. Algorithm 2 is the standard Euclidean algorithm for computing the gcd of two numbers.

Algorithm 2 : Euclidean algorithm

Input : $\triangle$, d with $d > \triangle \geq 1$
Output : $gcd(\triangle, d)$

1: **while** $\triangle \geq 1$ **do**
2: $\quad temp \leftarrow \triangle$
3: $\quad \triangle \leftarrow d \pmod{\triangle}$
4: $\quad d \leftarrow temp$
5: **end while**
6: Return d

By observing the steps of the two algorithms, we find that Algorithm 1 is just a minor modification of the Euclidean algorithm. Essentially, we prove the following connection.

Lemma 5. *The modular reduction steps of two consecutive iterations of Algorithm 2 are wrapped in a single iteration of Algorithm 1.*

Proof. Suppose $(d, \triangle)$ assume the values (r_0, r_1), (r_1, r_2), (r_2, r_3), $\ldots$, (r_x, r_{x+1}) in successive iterations of Algorithm 2 respectively staring from the first iteration. This means that there exist integers q_i, $1 \leq i \leq x$, such that $r_{i-1} = q_i r_i + r_{i+1}$. Here, q_i are quotients in successive Euclidean divisions[3].

Suppose $(d, \triangle) = (r_0, r_1)$ in the first iteration of Algorithm 1. Since there are two modular reduction steps in each iteration, it can be verified that $(d, \triangle) = (r_2, r_3)$ in the second iteration. In general, in the i^{th} iteration, $(d, \triangle) = (r_{2i-2}, r_{2i-1})$. This proves the claim. □

Illustration. Let us look at how Algorithm 1 computes the total number of progressions of $\mathfrak{S}(11, 25)$. Suppose the input to Algorithm 1 is $(a_0, d_0) = (11, 25)$. Step 2 of the algorithm computes $\triangle = d_0 - b = d_0 - a_0^{-1} \pmod{d_0} = 9$. Notice that 9 is the second common difference of the first grouping of $\mathfrak{S}(11, 25)$. Suppose the input to the Euclidean algorithm (Algorithm 2) is $(\triangle, d) = (9, 25)$. Then during the while loop execution of the algorithm, the values of $(\triangle, d)$ change as shown in the following table. Notice that the alternating pairs $(9, 25)$ and $(2, 7)$ correspond to the

Iteration number of while loop (Algorithm 2)	1	2	3	4
$(\triangle, d)$	(9,25)	(7,9)	(2,7)	(1,2)
Quotient ($\lfloor \frac{d}{\triangle} \rfloor$)	2	1	3	1

Table 1: Algorithm 2 on input pair $(9, 25)$

defining parameters of the first and the second grouping of $\mathfrak{S}(11, 25)$, respectively.

- 9 is the second common difference of $\mathcal{G}_1$, 25 is the common difference of the first progression of $\mathcal{G}_1$ and thus $\lceil \frac{25}{9} \rceil$ gives the size of $\mathcal{G}_1$.
- 2 is the second common difference of $\mathcal{G}_2$, 7 is the common difference of the first progression of $\mathcal{G}_2$ and thus $\lceil \frac{7}{2} \rceil$ gives the size of $\mathcal{G}_2$.

The quotient sequence of the Euclidean algorithm is $(2, 1, 3, 1)$, of which the alternating quotients, $(2, 3)$, when incremented by one, correspond to the sizes of groupings of $\mathfrak{S}(11, 25)$.

With the above illustration, we now formally state a property of the quotient sequence of the Euclidean algorithm on co-prime input pairs.

Property 2. Suppose the quotient sequence of the Euclidean algorithm on co-prime input pair $(d_0, \triangle)$ with $d_0 > \triangle$ is $Q = (q_1, q_2, q_3, \ldots, q_{l-1}, q_l)$. Then, the quotient

[3]The Euclidean division of an integer y by an integer x ($\neq 0$) gives two integers q, r such that $y = qx + r$, where $0 \leq r < |x|$. The integer r is called the remainder. The integer q is called the quotient of the division operation.

sequence of Algorithm 1, on input pair (a_0, d_0) with $a_0 \equiv d_0 - \triangle^{-1} \pmod{d_0}$, will be

$$Q' = (q_1, q_3, q_5, \ldots).$$

Further,

- The size of the i^{th} grouping of $\mathfrak{S}(a_0, d_0)$ is $q_{2i-1} + B_i$
- $N(a_0, d_0) = B + \sum_{1 \le i \le \lceil l/2 \rceil} q_{2i-1}$

Here, $B_i = 1$, for $1 \le i \le \lceil l/2 \rceil - 1$, and

$$B = B_{\lceil l/2 \rceil} = \begin{cases} 0 & \text{if } \triangle_{\lceil l/2 \rceil} = 1 \\ 1 & \text{if } \triangle_{\lceil l/2 \rceil} > 1 \end{cases}$$

$\triangle_{\lceil l/2 \rceil}$ is the second common difference corresponding to the last grouping of $\mathfrak{S}(a_0, d_0)$.

The average case analysis of the Euclidean algorithm is studied in [1, 3, 4, 5]. These results show that the average number of iterations of the Euclidean algorithm on co-prime input pair $(\triangle, d_0)$, with $\triangle < d_0$, is about

$$\frac{12 \log 2}{\pi^2} \log d_0 + C + O({d_0}^{-1/6} + \epsilon),$$

where C is the *Porter's constant* [5], whose value is approximately 1.4670. Thus, the average number of iterations of the Euclidean algorithm is about $0.842 \log d_0$. Hence, the average number of groupings of $\mathfrak{S}(a_0, d_0)$ is about $0.421 \log d_0$.

For $1 \le a_0 < d_0$, the maximum value for the number of groupings of $\mathfrak{S}(a_0, d_0)$ is attained when

$$\begin{aligned} d_0 &= F_i, \\ a_0 &\equiv -F_{i-1}^{-1} \pmod{d_0}. \end{aligned} \tag{3}$$

Here, F_i, F_{i-1} are the i^{th} and $(i-1)^{th}$ Fibonacci numbers, respectively. The number of iterations of the Euclidean algorithm on input pair $(F_i,\ F_{i-1})$ will be $i-2$. So, the number of iterations of Algorithm 1 on the input pair $(a_0,\ d_0)$ defined in Equation (3) will be $\lfloor \frac{i-2}{2} \rfloor$. The i^{th} Fibonacci number is $F_i = \frac{(\phi^i - (-1)^i \phi^{-i})}{\sqrt{5}}$, where $\phi = \frac{\sqrt{5}+1}{2}$. The number F_i is approximately $\frac{\phi^i}{\sqrt{5}}$ for large i. Hence, we have the following property.

Property 3. The maximum number of groupings of $\mathfrak{S}(a_0, d_0)$ is less than $\frac{\log_\phi \sqrt{5} d_0 - 2}{2}$.

3.3. Expected Value of $N(a_0, d_0)$

For a co-prime pair (a_0, d_0) with $1 \le a_0 < d_0$, we have $2 \le N(a_0, d_0) \le d_0$. The smallest value of $N(a_0, d_0)$ occurs at $a_0 = 1$, and the greatest value occurs at $a_0 = d_0 - 1$. Here, we estimate the average value of $N(a_0, d_0)$.

By Property 3.2, $N(a_0, d_0)$ is the sum of alternating quotients in the quotient sequence of the Euclidean algorithm on input pair $(d_0, \triangle)$, where $\triangle \equiv -a_0^{-1} \pmod{d_0}$. We work with the sum of all quotients (of the Euclidean algorithm) in estimating the average value of $N(a_0, d_0)$, with the assumption that a_0 is uniformly randomly chosen from $[1, d_0 - 1]$.

Suppose the number of iterations of the Euclidean algorithm on co-prime pair $(d_0, \triangle)$ is l. Suppose the divisor in the i^{th} iteration of the algorithm reduces by k_i bits. In other words, the quotient q_i in the i^{th} Euclidean division contains k_i binary bits. By its size, $2^{k_i - 1} \leq q_i \leq 2^{k_i} - 1$. Then, we have $\sum_{i=1}^{l} k_i = \lfloor \log_2 d_0 \rfloor$ and

$$\frac{1}{2} \sum_{i=1}^{l} 2^{k_i} \leq \sum_{i=1}^{l} q_i \leq \sum_{i=1}^{l} (2^{k_i} - 1).$$

It is known that the expected value of l is $0.842 \log d_0 \approx 0.583 \log_2 d_0$. The expected value of l indicates that in each iteration of the Euclidean algorithm the numbers are likely to decrease by at most two bits most of the time. So the quotients $q_i \in \{1, 2, 3, 4\}$ occur more often. The behavior of the quotients is precisely stated in the following result from [2].

Result (Theorem 1.3.4 in [2]). *The probability $P(q)$ that a Euclidean quotient is equal to q is $P(q) = \log_2(u/(u-1))$, where $u = (q+1)^2$.*

From the above result, the quotients 1, 2, 3 and 4 occur with corresponding probabilities about 0.415, 0.169, 0.093 and 0.0589. Further, $\sum_{1 \leq q \leq 50} P(q) = 0.97$. This result suggests that the expected value of the sum of all quotients is $O(\log d_0)$, and so is the average value of $N(a_0, d_0)$ for uniformly randomly chosen (a_0, d_0).

However, there exist a_0 for which $N(a_0, d_0)$ is much bigger than $\log d_0$. For example, for (a_0, d_0), the size of the first grouping is $\frac{d_0 - r}{\triangle}$, where $\triangle = -a_0^{-1} \pmod{d_0}$ is the second common difference of the first grouping and $r \equiv d_0 \pmod{\triangle}$. For small $\triangle$, $\frac{d_0 - r}{\triangle}$ is $O(d_0)$. However, the number of of such a_0 is a small fraction of $\phi(d_0)$, where $\phi(.)$ is Euler's totient function.

4. Symmetricity Property

In this section, we address the main question: when will leading terms of consecutive progressions of $\mathfrak{S}(a_0, d_0)$ exhibit the symmetricity property? In answering the question, we first prove that symmetricity happens only within groupings. We then derive a necessary and sufficient condition for the symmetricity property to happen in a given grouping. Using this condition, we prove some combinatorial results on the property.

Lemma 6. *Symmetricity happens only within a grouping.*

Proof. Suppose leading terms of progressions belonging to two successive groupings $\mathcal{G}$ and $\mathcal{G}'$ follow symmetricity. Since consecutive groupings share a progression, the number of leading terms involved in symmetricity should be at least 3. Let d be the common difference of the progression shared by both $\mathcal{G}$ and $\mathcal{G}'$. Let L be the leading term of the progression with common difference $d+\triangle$. Then, by Property 2.3,

$$\left\lfloor \frac{L}{d+\triangle} \right\rfloor + 2 = \left\lfloor \frac{L}{d-\triangle'} \right\rfloor, \tag{4}$$

where $\triangle$ and $\triangle'$ are the second common differences of $\mathcal{G}$ and $\mathcal{G}'$, respectively. Since $\triangle' < d < \triangle$, we have $\left\lfloor \frac{L}{d+\triangle} \right\rfloor < \frac{L}{2d}$ and $\frac{L}{d} < \left\lfloor \frac{L}{d-\triangle'} \right\rfloor$, which together contradict Equation (4). □

Corollary 2. *The maximum number of occurrences of symmetricity is at most* $\frac{\log_\phi \sqrt{5}d_0 - 2}{2}$.

Proof. Since symmetricity happens within a grouping, the result follows from Property 3.2. □

Through the following two main results (Theorem 2 and Theorem 3), we derive two different conditions which should meet simultaneously for symmetricity to occur in a grouping. The conditions seldom meet together. This shows that the bound given in Corollary 2 is trivial and unrefined. But the proven conditions do not provide any hint about a refined upper-bound.

Theorem 2. *Let $\mathcal{G}$ be a grouping of $\mathfrak{S}(a_0, d_0)$ that consists of progressions*

$$A(a_\alpha, d_\alpha), A(a_{\alpha+1}, d_{\alpha+1}), \ldots, A(a_\beta, d_\beta),$$

for some $0 \le \alpha < \beta$. Let $\triangle$ be the second common difference of $\mathcal{G}$. If $\triangle$ divides $d_\alpha^2 - 1$, then the number of leading terms of the progressions in $\mathcal{G}$ that follow symmetricity is $|\mathcal{G}| - \alpha$.

Proof. By Corollary 1, leading terms of the progressions in $\mathcal{G}$ satisfy the equation:

$$a_i = a_\alpha + \triangle(\beta - i)(i - \alpha) + (i - \alpha)(d_\beta - z_\alpha),\ \alpha \le i \le \beta.$$

Here,

$$z_\alpha = \frac{a_\alpha \triangle + 1}{d_\alpha} = \left\lfloor \frac{a_\alpha}{d_\alpha} \right\rfloor \triangle + \left(d_\beta^{-1} \pmod{\triangle}\right).$$

Since $\triangle$ divides $d_\alpha^2 - 1$, $d_\beta^2 \equiv 1 \pmod{\triangle}$ and thus d_β is a self-invertible element $\pmod{\triangle}$. So $z_\alpha = \lfloor a_\alpha / d_\alpha \rfloor + d_\beta$. By substituting the value of z_α in the leading

term equation, we obtain

$$a_i = a_\alpha + \triangle(i-\alpha)\big(\beta - i - \left\lfloor \frac{a_\alpha}{d_\alpha} \right\rfloor\big).$$

The size of $\mathcal{G}$ is $\beta - \alpha + 1 = \lceil \frac{d_\alpha}{\triangle} \rceil$. By substituting the value of β in the above equation, we obtain

$$a_i = a_\alpha + \triangle(i-\alpha)\big(\gamma - i\big), \tag{5}$$

where $\gamma = \alpha + \left\lceil \frac{d_\alpha}{\triangle} \right\rceil - \left\lfloor \frac{a_\alpha}{d_\alpha} \right\rfloor - 1$. It can be verified that $a_{\alpha+j} = a_{\gamma-j}$, $0 \le j \le \left\lceil \frac{d_\alpha}{\triangle} \right\rceil - \left\lfloor \frac{a_\alpha}{d_\alpha} \right\rfloor - 1$. Thus, the number of leading terms satisfying symmetricity is equal to $\left\lceil \frac{d_\alpha}{\triangle} \right\rceil - \left\lfloor \frac{a_\alpha}{d_\alpha} \right\rfloor$. We have $|\mathcal{G}| = \left\lceil \frac{d_\alpha}{\triangle} \right\rceil$ and by Property 2.3, $\left\lfloor \frac{a_\alpha}{d_\alpha} \right\rfloor = \alpha$. □

From the above result we observe the following for a grouping $\mathcal{G}$ with symmetricity.

- If $\mathcal{G}$ is the first grouping, the number of leading terms (in $\mathcal{G}$) satisfying the symmetricity property will be equal to $|\mathcal{G}|$, since the leading term a_α is less than the common difference d_α.
- If $\mathcal{G}$ is not the first grouping, then the leading term a_α is greater than the common difference d_α. Thus, the number of leading terms satisfying the symmetricity property will be less than $|\mathcal{G}|$.

For example, all leading terms in the first grouping of $\mathfrak{S}(17, 23)$ satisfy the symmetricity property. On the other hand, only two leading terms in the second grouping of $\mathfrak{S}(11, 25)$ satisfy symmetricity; only four leading terms in the second grouping of $\mathfrak{S}(11, 37)$ satisfy symmetricity.

Theorem 3. *Let $\mathcal{G}$ be the grouping as defined in Theorem 2. Let $\mathcal{G}_1$, $\mathcal{G}_2$, ..., $\mathcal{G}_m$ be the groupings of $\mathfrak{S}(a_0, d_0)$ that precede $\mathcal{G}$. The leading terms of the progressions in $\mathcal{G}$ follow symmetricity if and only if $\triangle$ divides $d_\alpha^2 - 1$, and*

$$|\mathcal{G}| \ge \sum_{i=1}^{m} |\mathcal{G}_i| - m + 2. \tag{6}$$

Proof. By definition, $\mathcal{G}$ has symmetricity only when at least one leading term in it appears twice, which is equivalent to the condition $|\mathcal{G}| - \alpha \ge 2$. Since any two consecutive groupings share a progression, we have $|\mathcal{G}_1| + |\mathcal{G}_2| + \ldots + |\mathcal{G}_m| = \alpha + m$. This completes the proof. □

In the rest of the paper, we refer to the condition that $\triangle$ divides $d_\alpha^2 - 1$ as the *divisibility condition*, and the condition given in Equation (6) as the *sum condition*.

The sum condition on the size of groupings can be converted into an equivalent condition on the quotient sequence of Algorithm 1. We know that each iteration of Algorithm 1 corresponds to a grouping. Suppose q is the quotient arising in the iteration corresponding to grouping $\mathcal{G}$, and q_i is the quotient in the iteration corresponding to grouping $\mathcal{G}_i$, $1 \leq i \leq m$. Then, we have $q = |\mathcal{G}| - c$ and $q_i = |\mathcal{G}_i| - 1$. Thus, the sum condition can be expressed as

$$q \geq 2 - c + \sum_{i=1}^{m} q_i. \tag{7}$$

By Property 3.2, $c = 0$ if $\mathcal{G}$ is the last grouping of $\mathfrak{S}(a_0, d_0)$, and $c = 1$ otherwise.

By Theorem 3, the number of groupings of $\mathfrak{S}(a_0, d_0)$ with symmetricity is equal to the number of quotients $q = \left\lfloor \frac{d}{\triangle} \right\rfloor$ arising in Algorithm 1 (on input (a_0, d_0)), where $\triangle$ divides $d^2 - 1$ and q is greater than the sum of all quotients that arise before q in the algorithm. Thus one can determine the number of groupings of $\mathfrak{S}(a_0, d_0)$ with symmetricity just by verifying the two conditions on the quotients in each iteration of Algorithm 1. The running time of the algorithm remains $O(\log^2 d_0)$. We refer to the algorithm for computing the number of groupings with symmetricity as Algorithm 3. We have run the algorithm and obtained empirical data on the number of groupings with symmetricity. Our observations from the empirical data are reported in Section 5.

4.1. Symmetricity of Higher Terms

For the example sequence $\mathfrak{S}(17, 23)$ we have observed that the symmetricity property is exhibited by some higher terms of progressions also. The same can be observed for $\mathfrak{S}(11, 37)$. The number of higher terms (of the progressions in $\mathcal{G}$) satisfying symmetricity is dependent on the width of symmetricity of leading terms, i.e., $|\mathcal{G}| - \alpha$ (the quantity established in Theorem 2). The following result proves this fact. The proof of the result is just an extension of the proof of Theorem 2.

Lemma 7. *Suppose $\mathcal{G}$ is a grouping as defined in Theorem 2. If $\triangle$ divides $d_\alpha^2 - 1$, then the number of j^{th} ($j \geq 1$) terms of the progressions in $\mathcal{G}$ that follow symmetricity is $|\mathcal{G}| - \alpha - j + 1$.*

Proof. If $\triangle$ divides $d_\alpha^2 - 1$, then we obtain Equation (5). Using the equation, it can be verified that, for $0 \leq i \leq \gamma - \alpha$ and $0 \leq j \leq \gamma - \alpha - i$,

$$a_{\alpha+i} + j d_{\alpha+i} = a_{\gamma-i-j} + j d_{\gamma-i-j}.$$

For a fixed value of j, the number of possible values for i for which the above equality holds is $\gamma - \alpha - j + 2 = |\mathcal{G}| - \alpha - j + 1$. □

4.2. Number of Symmetric Numbers for d_0

Recall the definition that a_0 is a symmetric number for d_0 if leading terms of the initial progressions of $\mathfrak{S}(a_0, d_0)$ satisfy the symmetricity property. By Lemma 6, symmetricity occurs within a grouping of $\mathfrak{S}(a_0, d_0)$. Thus the symmetric numbers for d_0 are connected to the first grouping of $\mathfrak{S}(a_0, d_0)$. As a corollary to Theorem 2, we have following result.

Corollary 3. *The number of symmetric numbers for d_0 is $\frac{\sigma_0(d_0^2-1)}{2}$. Here, $\sigma_0(x)$ denotes the number of divisors of integer x.*

Proof. For the first grouping, the sum condition (i.e., Equation (7)) is already met. Thus, by Theorem 3, a_0 is a symmetric number for d_0 if and only if the second common difference of the first grouping $\triangle$ divides $d_0^2 - 1$. For each divisor $\triangle(< d_0)$ of $d_0^2 - 1$, there is a corresponding symmetric number $a_0 \equiv -\triangle^{-1} \pmod{d_0}$. The result follows. □

For example, leading terms of the progressions in the first grouping of $\mathfrak{S}(17, 23)$ display symmetricity. So, 17 is a symmetric number for 23. This fact can be verified from the above result as the second common difference of the first grouping of the sequence $\triangle = 23 - 17^{-1} \pmod{23} = 4$ is a divisor of $23^2 - 1$. In general, symmetric numbers of 23 can be computed from $A = \{1, 2, 3, 4, 6, 8, 11, 12, 16, 22\}$, the set of all divisors $\triangle$ of $23^2 - 1$ with $\triangle < 23$. For each $\triangle \in A$, $a_0 = 23 - \triangle^{-1} \pmod{23}$ is a symmetric number for 23. The set $B = \{1, 2, 11, 15, 17, 19, 20, 21, 22\}$ comprises of all symmetric numbers for 23.

4.3. Symmetricity in Later Groupings

We have established a result in Corollary 3 on the number of a_0 $(< d_0)$ for which the first grouping of $\mathfrak{S}(a_0, d_0)$ has symmetricity. It is of interest to ask: what is the count of a_0 for which later groupings of $\mathfrak{S}(a_0, d_0)$ have symmetricity? Being motivated by the question, we define the following set for a given d_0:

$$\mathcal{S}_i(d_0) = \{a_0 : i^{th} \text{ grouping of } \mathfrak{S}(a_0, d_0) \text{ has symmetricity}\}.$$

We have proved that $|\mathcal{S}_1(d_0)| = \frac{1}{2}\sigma_0(d_0^2 - 1)$. Other than the result on $\mathcal{S}_1(d_0)$, we lack a proper understanding of $\mathcal{S}_i(d_0)$ for $i > 1$. Lemma 8 provides only a partial answer on the size of $\mathcal{S}_2(d_0)$.

In what follows, $\tau(n; z) = \#\{x \in (0, \frac{n}{2z+1}] : x | n^2 - z^2, z | n - x\}$. In other words, $\tau(n; z)$ is the number of divisors x of $n^2 - z^2$, with x falling in the interval $\left(0, \frac{n}{2z+1}\right]$ and such that z divides $n - x$.

Lemma 8. *The number of a_0, with $a_0^{-1} \pmod{d_0} < \frac{d_0}{2}$, such that the second grouping of $\mathfrak{S}(a_0, d_0)$ has symmetricity, is equal to $\sum_{z=1}^{\lfloor\sqrt{d_0}\rfloor-1} \tau(d_0; z)$.*

Proof. Let $\mathcal{G}_1, \mathcal{G}_2$ be the first and the second grouping of $\mathfrak{S}(a_0, d_0)$ with corresponding second common differences $\triangle_1$ and $\triangle_2$. If $\triangle_1 \equiv d_0 - a_0^{-1} \pmod{d_0}$ is greater than $\frac{d_0}{2}$, then $d_1 = d_0 - \triangle_1$ is the common difference of first progression of $\mathcal{G}_2$. For $\triangle_1 \in \left(\frac{(z-1)d_0}{z}, \frac{zd_0}{z+1}\right)$, with $z \geq 2$, it can be verified that $\triangle_2 = d_0 - zd_1$. It is known that symmetricity occurs in $\mathcal{G}_2$ when (i) $\triangle_2 | d_1^2 - 1$ (divisibility condition) and (ii) $d_1/\triangle_2 \geq 2$ (sum condition). The divisibility condition implies that

$$z^2 \equiv (zd_1)^2 \equiv d_0^2 \pmod{\triangle_2}.$$

Thus, $\triangle_2$ should be a divisor of $d_0^2 - z^2$, where $d_1 = \frac{d_0 - \triangle_2}{z}$ is an integer. By the sum condition, $\frac{d_0 - \triangle_2}{z\triangle_2} \geq 2$, which is equivalent to the condition $\triangle_2 \leq \frac{d_0}{2z+1}$. □

For a_0, with $a_0^{-1} \pmod{d_0} > \frac{d_0}{2}$, the size of the first grouping of $\mathfrak{S}(a_0, d_0)$ is at least 2 (bigger for some a_0) and thus the probability that the sum condition is met diminishes. With this observation, the number of a_0, with $a_0^{-1} \pmod{d_0} > \frac{d_0}{2}$, such that the second grouping of $\mathfrak{S}(a_0, d_0)$ has symmetricity, will be less than the sum $\sum_{z=1}^{\lfloor\sqrt{d_0}\rfloor-1} \tau(d_0; z)$. We thus have

$$|\mathcal{S}_2(d_0)| < 2 \sum_{z=1}^{\lfloor\sqrt{d_0}\rfloor-1} \tau(d_0; z).$$

By the average behavior of the divisor function, the expected value of the sum will be about $\sqrt{d_0} \log d_0$. The actual value will be much less as the numbers in $\tau(d_0; z)$ need to obey two other conditions (by the definition of τ). Further, the sum attains the maximum value when it gets highly composite numbers. From [6], for a highly composite number t, $\sigma_0(t)$ is about $2^{\frac{\log t}{\log\log t} + O\left(\frac{\log t}{(\log\log t)^2}\right)}$. However, the number of highly composite numbers at most x is less than $(\log x)^c$ for some constant c. The conclusion that we draw from the discussion is that even with presence of highly composite numbers the value of the sum will be much less than $\sqrt{d_0} \log d_0$. Hence, $|\mathcal{S}_2(d_0)|$ will be much less than $2\sqrt{d_0} \log d_0$.

5. Open Combinatorial Issues

In studying the symmetricity property in general, we have investigated the size of the following set:

$$\mathcal{T}(d_0, k) = \{a_0 : \mathfrak{S}(a_0, d_0) \text{ has } k \text{ groupings with symmetricity}\},$$

for different values of k.

Using Algorithm 3 as a subroutine, we have computed the size of $\mathcal{T}(d_0, k)$ for each $d_0 \leq 10^5$ and also for some random $d_0 > 10^5$. Our empirical observations are summarized as follows:

$$|\mathcal{T}(d_0, 1)| = c_{d_0} \times \frac{\phi(d_0)}{\log d_0} \tag{8}$$

$$|\mathcal{T}(d_0, 2)| = c'_{d_0} \times \frac{\phi(d_0)}{\log^2 d_0}. \tag{9}$$

In the above, $\phi(d_0)$ is the count of numbers co-prime to d_0. From the data, it is observed that both c_{d_0}, c'_{d_0} decrease with increasing d_0.

The case $k \geq 3$ requires a special mention, as the size of $\mathcal{T}(d_0, k)$ drops to zero. It is expected that the size of $\mathcal{T}(d_0, k)$ decreases as k increases. But the sudden fall from a large value of about $\phi(d_0)/\log^2(d_0)$ (when $k = 2$) to 0 (when $k = 3$) is surprising. Even for randomly chosen $d_0 \in (10^6, 10^{50})$, we did not encounter any value of a_0 for which the number of groupings of $\mathfrak{S}(a_0, d_0)$ with symmetricity is ≥ 3. This observation motivates us to put forth the following conjecture.

Conjecture 1. The number of groupings of $\mathfrak{S}(a_0, d_0)$ with symmetricity is ≤ 2.

The conjecture predicts that $\sum_{k=0}^{2} |\mathcal{T}(d_0, k)| = \phi(d_0)$, for any $d_0 \geq 2$. By Property 3.2, the conjecture can be stated in terms of the quotient sequence of the Euclidean algorithm on co-prime input pairs.

5.1. Remarks on $|\mathcal{T}(d_0, k)|$, for $k = 1, 2$

We firstly note that $a_0 \in \mathcal{S}_i(d_0) \Rightarrow a_0 \in \mathcal{T}(d_0, k)$ for some k. We have proved that $|\mathcal{S}_1(d_0)| = \frac{1}{2}\sigma_0(d_0^2 - 1)$. As the expected value of $\sigma_0(x)$ is $\log x$, the contribution of $\mathcal{S}_1(d_0)$ toward $\mathcal{T}(d_0, k)$ is very small compared to the observed magnitudes of $|\mathcal{T}(d_0, k)|$ for $k = 1, 2$. From the discussion after Lemma 8, it is known that $|\mathcal{S}_2(d_0)|$ will be much less than $2\sqrt{d_0}\log(d_0)$. The conclusion from these observations is that the combined contribution of both $\mathcal{S}_1(d_0)$ and $\mathcal{S}_2(d_0)$ toward $|\mathcal{T}(d_0, k)|$ is very small. This clearly indicates that the combined contribution from $\mathcal{S}_i(d_0)$, $i \geq 3$, will be much more. Proving the sizes of the sets is beyond our present understanding.

5.2. Remarks on Conjecture 1

For a given quotient sequence $Q = (q_1, q_2, q_3, \ldots, q_l)$, there exist infinitely many co-prime pairs (a_0, d_0) such that the sequence $\mathfrak{S}(a_0, d_0)$ has two properties: (i) the number of groupings of the sequence is l and (ii) the size of the i^{th} grouping is q_l. The proof of the statement involves reversing the steps of Algorithm 1. Thus, if the conjecture is false, there exist infinitely many counter-examples.

Acknowledgements. I am very thankful to the anonymous referee for the comments that helped me to improve the presentation and quality of the paper. I am grateful to my doctoral advisor Prof. Veni Madhavan for his invaluable inputs on the research problem. I am very thankful to Sam Hartburn, a freelance maths editor, for proofreading the paper.

References

[1] E. Bach and Jeffrey O. Shallit, *Algorithmic Number Theory*, Vol. 1, Foundations of Computing Series, MIT Press, Cambridge, MA, 1996.

[2] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, Berlin Heidelberg, 1996.

[3] D.E. Knuth, *The Art of Computer Programming: Semi Numerical Algorithms*, Second edition, Vol. 2, Pearson Education India, 1998.

[4] G. H. Norton, On the asymptotic analysis of the Euclidean algorithm, *J. Symbolic Comput.* **10** (1990), 53–58.

[5] J.W. Porter, On a theorem of Heilbronn, *Mathematika* **22** (1975), 20–28.

[6] Srinivasa Ramanujan, Highly composite numbers, *Proc. Lond. Math. Soc.* **2 XIV** (1915), 347–409.

[7] Zhi-Wei Sun, Problems and results on covering systems (a survey article), available at *http://maths.nju.edu.cn/ zwsun/Cover.pdf*.