SETS OF CARDINALITY 6 ARE NOT SUM-DOMINANT

Hùng Việt Chu
Department of Mathematics, University of Illinois at Urbana-Champaign, Illinois
hungchu2@illinois.edu

Received: 6/23/19, Accepted: 2/6/20, Published: 2/14/20

Abstract

Given a finite set $A \subseteq \mathbb{N}$, define the sum set

$$A + A = \{a_i + a_j \mid a_i, a_j \in A\}$$

and the difference set

$$A - A = \{a_i - a_j \mid a_i, a_j \in A\}.$$

The set A is said to be sum-dominant if $|A + A| > |A - A|$. Hegarty used a nontrivial algorithm to find that 8 is the smallest cardinality of a sum-dominant set. Since then, Nathanson has asked for a human-understandable proof of the result. However, due to the complexity of the interactions among numbers, it is still questionable whether such a proof can be written down in full without computers' help. In this paper, we present a computer-free proof that a sum-dominant set must have at least 7 elements. We also answer the question raised by the author of the current paper et al about the smallest sum-dominant set of primes, in terms of its largest element. Using computers, we find that the smallest sum-dominant set of primes has 73 as its maximum, smaller than the value found before.

1. Introduction

1.1. Background

Given a finite set $A \subseteq \mathbb{N}$, define $A + A = \{a_i + a_j \,|\, a_i, a_j \in A\}$ and $A - A = \{a_i - a_j \,|\, a_i, a_j \in A\}$. The set A is said to be

- *sum-dominant*, if $|A + A| > |A - A|$;
- *balanced*, if $|A + A| = |A - A|$; and
- *difference-dominant*, if $|A + A| < |A - A|$.

Because addition is commutative while subtraction is not, sum-dominant sets are very rare. However, it was first proved by Martin and O'Bryant [10] that as $n \to \infty$, the proportion of sum-dominant subsets of $\{0, 1, 2, \dots, n-1\}$ is bounded below by a positive constant (about $2 \cdot 10^{-7}$), which was latter improved by Zhao [25] to about $4 \cdot 10^{-4}$. The last few years have seen an explosion of papers examining the properties of sum-dominant sets: see [6, 9, 16, 19, 20, 21] for history and overview, [7, 11, 12, 17, 23] for explicit constructions, [3, 10, 25] for positive lower bounds for the percentage of sum-dominant sets, [8, 14] for generalized sum-dominant sets, and [1, 2, 4, 13, 24] for extensions to other settings.

In response to Nathanson's question of the smallest sum-dominant set [16], Hegarty [7] used a clever algorithm to find that a sum-dominant set must have at least 8 elements. (The computer program was reported to run for about 15 hours.) However, a human-understandable proof of the result has not been produced because of the complexity lurking behind the interactions of numbers in addition and subtraction. Nathanson [15, 18] asked for a human-understandable proof of the smallest cardinality of a sum-dominant set. Hegarty, through personal communication, also said that it would be nice to have such a proof written down in full. This paper proves that a set of cardinality 6 is not sum-dominant without the use of computers. In combination with [5, Theorem 1], we have a computer-free proof that a sum-dominant set must have at least 7 elements.

1.2. Notation

We introduce some notation.

- Let A and B be sets. We write $A \to B$ to mean the introduction of elements in A to B. For example, $\{2\} \to \{4, 9, 12\}$ means that we introduce the number 2 into the set $\{4, 9, 12\}$.

- We write $a_n + \cdots + a_m$ for some $n \leq m$ to mean the sum $a_n + a_{n+1} + \cdots + a_{m-1} + a_m$.

- We use a different notation to write a set, which was first introduced by Spohn [22]. Given a set $S = \{m_1, m_2, \dots, m_n\}$, we arrange its elements in increasing order and find the differences between two consecutive numbers to form a sequence. Suppose that $m_1 < m_2 < \cdots < m_n$, then our sequence is $m_2 - m_1, m_3 - m_2, m_4 - m_3, \dots, m_n - m_{n-1}$, and we represent $S = (m_1 \mid m_2 - m_1, m_3 - m_2, m_4 - m_3, \dots, m_n - m_{n-1}) = (m_1 | a_1, \dots, a_{n-1})$, where $a_i = m_{i+1} - m_i$. Finally, any difference in $S - S$ must be equal to at least a sum $a_i + \cdots + a_j$ for some $1 \leq i \leq j \leq n-1$. Take $S = \{3, 2, 15, 10, 9\}$, for example. We arrange the elements in increasing order to have 2, 3, 9, 10, 15, form a sequence by looking at the difference between two consecutive numbers: 1, 6, 1, 5, and write $S = (2\,|\,1, 6, 1, 5)$. All information about a set is preserved in this notation.

1.3. Main Results

Theorem 1. *A set of cardinality* 6 *is not sum-dominant.*

Remark 1. Combined with [5, Theorem 1], Theorem 1 says that a sum-dominant set must have at least 7 elements. This is one step closer to the result of Hegarty; that is, a sum-dominant set must have at least 8 elements.

The interactions of 6 numbers to form the sum set and the difference set are so complicated that we need a clever division of the problem into cases and reduce the complexity considerably. We believe that to prove this theorem, case analysis is inevitable. Therefore, the question is whether the proof can be written down in full without being too overwhelming. Our main technique is to argue for a lower bound for the number of pairs of equal positive differences from $A - A$, which confines set A to certain structures. The lower bound in turn gives an upper bound for the number of distinct positive differences given by numbers in A.

For simplicity of notation, we denote our set $A = (0\,|\,a_1, a_2, a_3, a_4, a_5)$ for $a_i \in \mathbb{N}_{\geq 1}$. In proving that A is not sum-dominant, we split our proof into two sections considering whether $a_1 = a_2$ or $a_1 \neq a_2$. In particular, Section 2 provides tools to eliminate or simplify cases in our proof as well as restricts A to certain structures; Section 3 and Section 4 consider the two cases $a_1 = a_2$ and $a_1 \neq a_2$, respectively; Section 5 proves one of our lemmas; Section 6 investigates sum-dominant sets of primes; finally, Section 7 mentions some open problems for future research.

Our next result is to find the smallest sum-dominant set of primes, in terms of its largest element. The Green-Tao theorem states that the primes contain arbitrarily long arithmetic progressions. Chu et al. [4] used this theorem to prove that there are infinitely many sum-dominant sets of primes. However, sum-dominant sets of primes are expected to appear much earlier before we see a long arithmetic progression. For example, the authors found $\{19, 79, 109, 139, 229, 349, 379, 439\}$ as a sum-dominant set of primes [4]. The following theorem answers their question about the smallest sum-dominant set of primes, in terms of its largest element; equivalently, about how early in the prime sequence, we see a sum-dominant set.

Theorem 2. *The smallest sum-dominant set of primes, in terms of its largest element, is* $\{3, 5, 7, 13, 17, 19, 23, 43, 47, 53, 59, 61, 67, 71, 73\}$. *This set is also unique in the sense that there is no other sum-dominant set with* 73 *as its largest element.*

Lastly, we also have an observation about the minimum number of elements added to an arithmetic progression to have a sum-dominant set.

Remark 2. Let c be the smallest number of elements to be added to an arithmetic progression to form a sum-dominant set. Then $3 \leq c \leq 4$. This is due to two previous works. The author of the current paper proved that adding two arbitrary numbers into an arithmetic progression does not give a sum-dominant set [5]. So, $3 \leq c$. It is also known that $A^* = \{0,2\} \cup \{3,7,11,\ldots,4k-1\} \cup \{4k,4k+2\}$ is sum-dominant [17]. Another example is the set $\{0,1,3\} \cup \{7,8,\ldots,17\} \cup \{24\}$. Hence, $c \leq 4$.

2. Important Results

In this section, we provide all necessary tools that help reduce the complexity of the problem considerably. We use the definition of a symmetric set given in [17]: a set A is symmetric if there exists a number a such that $a - A = A$. If so, we say that the set A is symmetric about a. The following proposition was proved in [17].

Proposition 1. *A symmetric set is balanced.*

Proof. Let A be a symmetric set about a. We have $|A+A| = |A+(a-A)| = |a+(A-A)| = |A-A|$. Hence, A is balanced. $\square$

Though symmetric sets are not sum-dominant, adding a few numbers into these sets (in a clever way) can produce sum-dominant sets. Both [7] and [17] provide examples of such a technique. Note that a set of numbers from an arithmetic progression is symmetric about the sum of the maximum and the minimum of the arithmetic progression. For example, the set $E = \{3,5,7,9,11\}$ is symmetric about 14. Next, we prove a very useful lemma that establishes an upper bound for the number of distinct positive differences in $A - A$.

Lemma 1. *Let A be a sum-dominant set with $|A| = 6$. If there exist m_1, m_2, and $m_3 \in A$ such that $m_2 - m_1 = m_3 - m_2$, then A has at most 7 distinct positive differences.*

Proof. Let x be the number of pairs of equal positive differences given by the interaction of numbers in A when we take $A - A$. For example, in our set E above, $11 - 7 = 9 - 5$. So, $(11,7)$ and $(9,5)$ form a pair of equal positive differences. We need the following two inequalities:

$$|A+A| \leq |A|(|A|+1)/2, \quad (1)$$

$$|A-A| \leq |A|(|A|-1)+1. \quad (2)$$

These inequalities are not hard to prove and were used in [7] and [5]. Inequality (1) gives $|A+A| \leq 21$, while Inequality (2) gives $|A-A| \leq 31$. The equality in (1) is achieved if the sum of any two numbers is distinct, and the equality in (2) is achieved if the difference between any two different numbers is distinct. Because we have x pairs of equal positive differences, we have $|A-A| = 31 - 2x$ (taking into account equal negative differences). We find a lower bound for $|A+A|$ by using [5, Observation 13]. Because $m_2 - m_1 = m_3 - m_2$, we have the pair of equal positive differences: (m_2, m_1) and (m_3, m_2). According to [5, Observation 13], this pair does not give another new pair of equal positive differences. So, the existence of this pair reduce the maximum number of differences in $A - A$ by exactly 2 while reduces the maximum number of sums in $A+A$ by exactly 1. The remaining $x-1$ pairs reduces the maximum number of differences by $2(x-1)$ and reduces the maximum number of sums by at least $(x-1)/2$. Therefore, $|A+A| \le 21-1-(x-1)/2 = 20-(x-1)/2$. Because A is sum-dominant,

$$20 - (x-1)/2 \ \ge\ |A+A| \ >\ |A-A| \ =\ 31 - 2x.$$

We have $x \ge 8$. Hence, $|A-A| \le 31 - 2\cdot 8 = 15$. Because $0 \in A - A$, the number of distinct positive differences is at most $(15-1)/2 = 7$, as desired. □

Remark 3. In Spohn's notation, if we write $A = (0|a_1, a_2, \ldots, a_5)$, then the existence of m_1, m_2, and m_3 as above is equivalent to the existence of i, j, and k such that $a_i + \cdots + a_j = a_{j+1} + \cdots + a_k$. Equivalently, we have an arithmetic progression of length 3.

Lemma 2. *Let A be a sum-dominant set with $|A| = 6$. Then $A - A$ has at most 8 distinct positive differences.*

Proof. Let x be the number of pairs of equal positive differences given by the interaction of numbers in A when we take $A - A$. From the proof of Lemma 1, we know that $|A+A| \le 21$ and $|A-A| \le 31$. By [5, Observation 13], we know that $|A-A| = 31 - 2x$, while $|A+A| \le 21 - x/2$. Because

$$21 - x/2 \ \ge\ |A+A| \ >\ |A-A| \ =\ 31 - 2x,$$

we have $x \ge 7$. Hence, $|A-A| \le 31 - 2\cdot 7 = 17$. Because $0 \in A - A$, the number of distinct positive differences is at most $(17-1)/2 = 8$, as desired. □

Remark 4. If we have numbers that form an arithmetic progression of length 3, the upper bound for the number of distinct positive differences is reduced by 1 (from 8 to 7). This is a big advantage in reducing the number of cases as we will utilize this fact latter.

The following proposition will also be used intensively.

Proposition 2. *Let $|A| = 6$ and A contains an arithmetic progression of length 4, then A is not sum-dominant.*

The proof follows immediately from [5, Theorem 2]. Finally, we present 15 sets that are not sum-dominant. Most of our cases are reduced to one of these forms.

Lemma 3. *Let d, a, and b be positive real numbers. The following sets are not sum-dominant:* $S_1 = (0\,|\,d, d, 2d, a, b)$, *with* $a + b = d$; $S_2 = (0\,|\,d, d, 2d, d, a)$; $S_3 = (0\,|\,d, d, 2d, a, d)$; $S_4 = (0\,|\,2d, d, d, a, 2d)$; $S_5 = (0\,|\,a, b, b, a, a)$; $S_6 = (0\,|\,a+b, a, a, b, a+b)$; $S_7 = (0\,|\,a+b, a, a, b, a)$; $S_8 = (0\,|\,a, 2a, a, a, b)$; $S_9 = (0\,|\,a+b, a, a+b, a, b)$; $S_{10} = (0\,|\,a+b, 2a+b, a+b, a, b)$; $S_{11} = (0\,|\,a, b, a, a+b, a)$; $S_{12} = (0\,|\,a, b, a+b, a, a)$; $S_{13} = (0\,|\,2a+b, a, a, b, a)$; $S_{14} = (0\,|\,a+b, a, a, b, 2a)$; $S_{15} = (0\,|\,a, a+b, a, b, a)$.

Note that we use 0 as the minimum element, but the minimum can be any number since sum-dominance is preserved under affine transformations. Because the proof is tedious and is not the main focus of this paper, we move the proof to Section 5.

3. When $a_1 = a_2$

Because $a_1 = a_2$, Lemma 1 says that $|A - A| \leq 7$. If $a_1 = a_3$, then $a_1 = a_2 = a_3$, and we have an arithmetic progression of length 4. By Proposition 2, we do not have a sum-dominant set. We consider $a_1 \neq a_3$.

Case 1: $a_3 = a_1 + a_2$. Our distinct positive differences include

$$a_1 \;<\; a_1 + a_2 \;<\; a_2 + a_3 \;<\; a_1 + a_2 + a_3 \;<\; a_1 + \cdots + a_4 \;<\; a_1 + \cdots + a_5.$$

We are allowed to have at most one more positive difference. Let $a_1 = a_2 = d$. It follows that $a_3 = 2d$. Consider two cases.

Subcase 1.1: $a_4 = d$. We have S_2, which is not sum-dominant.

Subcase 1.2: $a_4 \neq d$. Then the difference $a_2 + a_3 + a_4 = 3d + a_4$ is another positive difference, meaning that all other differences must be equal to one of the following 7 positive differences:

$$d \;<\; 2d \;<\; 3d \;<\; 4d \;<\; 4d + a_4 \;<\; 4d + a_4 + a_5,\; 3d + a_4.$$

Indeed, $3d + a_4$ is a new difference because $3d < 3d + a_4 < 4d + a_4$ but $3d + a_4 \neq 4d$. Consider the difference $a_2 + a_3 + a_4 + a_5 = 3d + a_4 + a_5$. Either $3d + a_4 + a_5 = 4d$ or $3d + a_4 + a_5 = 4d + a_4$. The former gives $a_4 + a_5 = d$, while the latter gives $a_5 = d$. None of these produces a sum-dominant set because neither S_1 nor S_3 is sum-dominant.

Case 2: $a_3 \neq a_1 + a_2$. Because $a_2 + a_3 > a_1$ and $a_1 \neq a_3$, we have the following list of 7 distinct differences:

$$\begin{aligned}&a_1 \;<\; a_1 + a_2 \;<\; a_1 + a_2 + a_3 \;<\; a_1 + \cdots + a_4 \;<\; a_1 + \cdots + a_5,\\ &a_3 \;<\; a_2 + a_3.\end{aligned}$$

Consider the difference $a_2 + a_3 + a_4$. Either $a_2 + a_3 + a_4 = a_1 + a_2 + a_3$ or $a_2 + a_3 + a_4 = a_1 + a_2$. The former gives $a_1 = a_4$, while the latter gives $a_1 = a_3 + a_4$.

Case 2.1: $a_1 = a_4$. Then $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$, so $a_1 = a_5$. Because $a_1 = a_2 = a_4 = a_5$, we have a symmetric set, which is not sum-dominant.

Case 2.2: $a_1 = a_3 + a_4$. Because $a_1 = a_2 = a_3 + a_4$, we have an arithmetic progression of length 4. By Proposition 2, we do not have a sum-dominant set.

4. When $a_1 \neq a_2$

The following are distinct positive differences:

$$\begin{aligned} &a_1 \ < \ a_1 + a_2 \ < \ a_1 + a_2 + a_3 \ < \ a_1 + \cdots + a_4 \ < \ a_1 + \cdots + a_5, \\ &a_2. \end{aligned} \tag{3}$$

`Case 1.` $a_2 + a_3 = a_1$. Because $a_1 = a_2 + a_3$, we know, by Lemma 1, that the number of positive differences is at most 7. Hence, we are allowed to have at most one more positive difference. We consider $a_2 + a_3 + a_4$.

`Subcase 1.1.` $a_2 + a_3 + a_4 = a_1 + a_2 + a_3$. So, $a_1 = a_4$. Because $a_1 = a_2 + a_3 = a_4$, we have an arithmetic progression of length 4. By Proposition 2, we do not have a sum-dominant set.

`Subcase 1.2.` $a_2 + a_3 + a_4 = a_1 + a_2$. So, $a_1 = a_3 + a_4$. Since $a_1 = a_2 + a_3$, we have $a_2 = a_4$.

1. Subcase 1.2.1: $a_2 = a_3$. Since $a_2 = a_3 = a_4$, we have an arithmetic progression of length 4. By Proposition 2, we do not have a sum-dominant set.

2. Subcase 1.2.2: $a_2 \neq a_3$. Our 7 distinct positive differences are

   $$\begin{aligned} &a_1 \ < \ a_1 + a_2 \ < \ a_1 + a_2 + a_3 \ < \ a_1 + \cdots + a_4 \ < \ a_1 + \cdots + a_5, \\ &a_2,\ a_3. \end{aligned}$$

   (Note that $a_3 < a_1$ since $a_1 = a_3 + a_4$.) Because we cannot have a new difference besides these 7 differences, either $a_2 + \cdots + a_5 = a_1 + a_2 + a_3$ or $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$ because $a_2 + \cdots + a_5 = 2a_2 + a_3 + a_5 > a_1 + a_2$.

   - If the former, we have $a_1 = a_4 + a_5$. Because $a_1 = a_2 + a_3 = a_4 + a_5$, we have an arithmetic progression of length 4 and thus, do not have a sum-dominant set.
   - If the latter, we have $a_1 = a_5$. Because $a_2 = a_4$, we have a symmetric set, which is not sum-dominant.

`Subcase 1.3.` $a_2 + a_3 + a_4$ is not equal to any difference in our List (3). By adding $a_2 + a_3 + a_4$ to our list, we have 7 distinct positive differences and this new list is exhaustive. Consider the difference a_3. It must be that $a_3 = a_2$. Consider $a_2 + \cdots + a_5$.

1. Subcase 1.3.1: $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$. Equivalently, $a_1 = a_5$. Let $a_2 = a_3 = d$. It follows that $a_5 = a_1 = a_2 + a_3 = 2d$. Our set is of the form $(0|2d, d, d, a_4, 2d)$, which is S_4, not a sum-dominant set.

2. Subcase 1.3.2: $a_2 + \cdots + a_5 = a_1 + a_2 + a_3$. Equivalently, $a_1 = a_4 + a_5$. The fact that $a_1 = a_2 + a_3 = a_4 + a_5$ gives us an arithmetic progression of length 4. Hence, our set is not sum-dominant.

3. Subcase 1.3.3: $a_2 + \cdots + a_5 = a_1 + a_2$. Equivalently, $a_1 = a_3 + a_4 + a_5$. So, $a_4 + a_5 = a_1 - a_3 = a_2$. The fact that $a_2 = a_3 = a_4 + a_5$ gives us an arithmetic progression of length 4. Hence, our set is not sum-dominant.

`Case 2.` $a_2 + a_3 \neq a_1$.

`Subcase 2.1.` $a_2 + a_3 = a_1 + a_2$. Equivalently, $a_1 = a_3$. There are two possibilities for $a_2 + a_3 + a_4$ because $a_1 + \cdots + a_4 > a_2 + a_3 + a_4 > a_1 + a_2$.

1. `Subcase 2.1.1.` $a_2 + a_3 + a_4 = a_1 + a_2 + a_3$. Equivalently, $a_1 = a_4$. Because $a_1 = a_3$, we know that $a_3 = a_4$. We thus have at most 7 distinct positive differences. Consider the difference $a_2 + \cdots + a_5$. We know that either $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$ or $a_2 + \cdots + a_5$ is a new difference. If the former, we have $a_1 = a_5$, which implies that $a_3 = a_4 = a_5$, giving us an arithmetic progression of length 4. Hence, we do not have a sum-dominant set. We consider the case where $a_2 + \cdots + a_5$ is a new difference. All of the positive differences are

$$\begin{aligned} &a_1 \ < \ a_1 + a_2 \ < \ a_1 + a_2 + a_3 \ < \ a_1 + \cdots + a_4 \ < \ a_1 + \cdots + a_5, \\ &a_2 \ < \ a_2 + \cdots + a_5. \end{aligned}$$

   Consider $a_3 + a_4$. The only two possible values for $a_3 + a_4$ are $a_1 + a_2$ and a_2. If the former, we have $a_1 = a_2 = a_3 = a_4$, which does not give a sum-dominant set by [5, Lemma 8]. If the latter, we have $a_3 + a_4 = a_2$, which gives us S_8, not a sum-dominant set.

2. `Subcase 2.1.2.` $a_2 + a_3 + a_4$ is a new difference. Our set of positive differences contains

$$\begin{aligned} &a_1 \ < \ a_1 + a_2 \ < \ a_1 + a_2 + a_3 \ < \ a_1 + \cdots + a_4 \ < \ a_1 + \cdots + a_5, \\ &a_2 \ < \ a_2 + a_3 + a_4. \end{aligned}$$

   We consider three possibilities for $a_2 + \cdots + a_5$.

- $a_2+\cdots+a_5 = a_1+a_2+a_3$. Equivalently, $a_1 = a_4+a_5$. Because $a_1 = a_3$, we have $a_3 = a_4 + a_5$. Hence, the above list of differences is exhaustive. Consider $a_3 + a_4$. Either $a_3 + a_4 = a_1 + a_2$ or $a_3 + a_4 = a_2$. Neither of these is a sum-dominant set because neither S_9 nor S_{10} is sum-dominant.
- $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$. Equivalently, $a_1 = a_5$. Consider $a_3 + a_4$. There are four possibilities.

  If $a_3 + a_4 = a_2$, we arrive at S_{15}.

  If $a_3 + a_4 = a_1 + a_2$, we have $a_2 = a_4$ and thus, a symmetric set, which is not sum-dominant.

  If $a_3 + a_4 = a_1 + a_2 + a_3$, then $a_2 + a_3 = a_4$ because $a_1 = a_3$. We arrive at S_{11}.

  If $a_3 + a_4$ is a new difference, then we have exactly 8 distinct differences by Lemma 2. Consider $a_3 + a_4 + a_5$. To have 8 distinct differences, the only possibility is that $a_3 + a_4 + a_5 = a_1 + a_2 + a_3$; equivalently, $a_1 + a_2 = a_4 + a_5$. Because $a_1 = a_5$, it follows that $a_2 = a_4$. So, we have a symmetric set, which is not sum-dominant.
- $a_2 + \cdots + a_5$ is a new difference. Consider $a_3 + a_4$. Note that $a_3 + a_4 \notin \{a_2, a_1 + a_2\}$ because we have 8 distinct positive differences. Indeed, the only possibility is that $a_3+a_4 = a_1+a_2+a_3$. So, $a_1+a_2 = a_4$. However, because $a_1 = a_3$, we have $a_2 + a_3 = a_4$, which contradicts that $A - A$ has 8 positive differences.

`Subcase 2.2.` $a_2 + a_3 \neq a_1 + a_2$. Equivalently, $a_1 \neq a_3$. The following are distinct positive differences:

$$\begin{aligned} &a_1 \ < \ a_1 + a_2 \ < \ a_1 + a_2 + a_3 \ < \ a_1 + \cdots + a_4 \ < \ a_1 + \cdots + a_5, \\ &a_2 \ < \ a_2 + a_3. \end{aligned}$$

1. `Subcase 2.2.1.` $a_3 = a_1+a_2$. The above list contains all positive differences. It must be that $a_2 + a_3 + a_4 = a_1 + a_2 + a_3$. So, $a_1 = a_4$. We also have $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$. Hence, $a_1 = a_5$. We arrive at S_{12}, which is not sum-dominant.
2. `Subcase 2.2.2.` $a_3 = a_2$. The above list contains all positive differences. There are three possibilities for $a_2 + a_3 + a_4$.

   - $a_2 + a_3 + a_4 = a_1 + a_2 + a_3$. Equivalently, $a_1 = a_4$. It follows that $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$. So, $a_1 = a_5$. We arrive at S_5.
   - $a_2 + a_3 + a_4 = a_1 + a_2$. So, $a_1 = a_3 + a_4$. The difference $a_2 + \cdots + a_5$ is either equal to $a_1 + a_2 + a_3$ or $a_1 + \cdots + a_4$. If the former, we obtain $a_1 = a_4 + a_5$ and arrive at S_7. If the latter, we obtain $a_1 = a_5$ and arrive at S_6.

- $a_2 + a_3 + a_4 = a_1$. There are three possibilities for $a_2 + \cdots + a_5$.

  If $a_2 + \cdots + a_5 = a_1 + a_2$, then we have $a_1 = a_3 + a_4 + a_5$. We arrive at S_{13}.

  If $a_2 + \cdots + a_5 = a_1 + a_2 + a_3$, then we have $a_1 = a_4 + a_5$. We arrive at S_{14}.

  If $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$, then $a_1 = a_5$. So, we have $a_1 = a_2 + a_3 + a_4 = a_5$ and thus, have an arithmetic progression of length 4. Our set is not sum-dominant.

3. `Subcase 2.2.3.` a_3 is a new difference. The exhaustive list of positive differences is

$$
\begin{aligned}
&a_1,\ a_1 + a_2,\ a_1 + a_2 + a_3,\ a_1 + \cdots + a_4,\ a_1 + \cdots + a_5,\\
&a_2,\ a_2 + a_3,\\
&a_3.
\end{aligned}
$$

There are three possibilities for $a_2 + a_3 + a_4$. We analyze each possibility.

- $a_2+a_3+a_4 = a_1$. Then we can have at most 7 positive differences, which is a contradiction.

- $a_2 + a_3 + a_4 = a_1 + a_2$. So, $a_1 = a_3 + a_4$. There are two possibilities for $a_2 + \cdots + a_5$.

  If $a_2 + \cdots + a_5 = a_1 + \cdots + a_4$, then $a_1 = a_5$, implying that $a_5 = a_3 + a_4$. Then we have at most 7 positive differences, a contradiction.

  If $a_2 + \cdots + a_5 = a_1 + a_2 + a_3$, then $a_1 = a_4 + a_5$. Consider a_4. There are three possibilities for a_4. If $a_4 = a_2$, then because $a_1 = a_3 + a_4$, we have $a_1 = a_2 + a_3$. So, we have at most 7 positive differences, a contradiction. If $a_4 = a_3$ or $a_4 = a_2 + a_3$, we again have at most 7 positive differences, a contradiction.

- $a_2 + a_3 + a_4 = a_1 + a_2 + a_3$. Equivalently, $a_1 = a_4$. It follows that $a_2 + a_3 + a_4 + a_5 = a_1 + a_2 + a_3 + a_4$. Equivalently, $a_1 = a_5$. So, $a_4 = a_5$, implying that we have at most 7 distinct differences, a contradiction.

5. Proof of Lemma 3

We first prove that S_1 is not sum-dominant. Note that $(0 \,|\, d, d, 2d)$ represents the set $K = \{0, d, 2d, 4d\}$ and $|K - K| - |K + K| = 1$. In particular,

$$
\begin{aligned}
K + K &= \{0, d, 2d, 3d, 4d, 5d, 6d, 8d\},\\
K - K &= \{0, \pm d, \pm 2d, \pm 3d, \pm 4d\}.
\end{aligned}
$$

With $\{4d+a\} \to K$, we have at most 5 new sums. However, the set of new positive differences is $\{a, a+2d, a+3d, a+4d\}$. (These are new differences because $0 < a < d$.) Denote $\{4d+a\} \cup K = K_1$. Then $|K_1 - K_1| - |K_1 + K_1| \geq (|K - K| - |K+K|) + (2 \cdot 4 - 5) = 4$. Finally, $\{4d+a+b\} \to K_1$ gives at least one new positive difference, which is $4d+a+b$ itself and gives at most 6 new sums. Hence, $|S_1 - S_1| - |S_1 + S_1| \geq (|K_1 - K_1| - |K_1 + K_1|) + (2-6) \geq 4 + 2 - 6 = 0$. Hence, S_1 is not sum-dominant.

Next, we prove that S_2 is not sum-dominant. If $a = d$ or $a = 2d$, it is an easy check that S_2 is not sum-dominant. Because the set $(0|d, d, 2d, d)$ is not sum-dominant, it suffices to show that adding $5d + a$ to the set gives at least as many differences as sums. We proceed by considering two cases.

- `Case 1.` $a < d$ or $d < a < 2d$. The set of new positive differences is $\{a, a+d, a+3d, a+4d, a+5d\}$, while there are at most 6 new sums. We are done.

- `Case 2.` $a > 2d$. Then $a+3d, a+4d, a+5d$ are new differences because they are all greater than $5d$. Hence, the number of new differences is at least 6, while there are at most 6 new sums. We are done.

We have shown that S_2 is not sum-dominant.

We prove that S_3 is not sum-dominant. It is easily checked that if $a = d$ or $a = 2d$, we do not have a sum-dominant set. We proceed by considering three cases.

- `Case 1.` $a < d$. The proof follows exactly the proof that S_1 is not sum-dominant.

- `Case 2.` $d < a < 2d$. We have $\{a+4d, a+5d\} \to K$ gives at most 11 new sums. Because $|K - K| - |K + K| = 1$, it suffices to show that there are at least 5 new positive differences. Indeed, new differences include $a+d$, $a+2d$, $a+3d$, $a+4d$, and $a+5d$. We are done.

- `Case 3.` $a > 2d$. We have $\{a+4d, a+5d\} \to K$ gives at most 11 new sums. Because $|K - K| - |K + K| = 1$, it suffices to show that there are at least 5 new positive differences. New differences include $a+2d$, $a+3d$, $a+4d$, and $a+5d$ because each of these is greater than $4d$. If $a + d \neq 4d$, we have a new difference and we are done. If $a + d = 4d$, then $a = 3d$. It can be checked that S_3 is not sum-dominant.

Therefore, S_3 is not sum-dominant.

Let $K_4 = (0|2d, d, d) = \{0, 2d, 3d, 4d\}$. It is easy to check that $|K_4 - K_4| - |K_4 + K_4| = 3$. If $a = d$ or $a = 2d$, it is also easily checked that S_4 is not sum-dominant. Because $\{4d+a, 6d+a\} \to K_4$ gives at most 11 new sums. It suffices to show that the number of new differences is at least 8. We consider two following cases.

- `Case 1.` $a < d$ or $d < a < 2d$. The set of new positive differences includes a, $a+d$, $a+2d$, $a+3d$, and $a+4d$. We are done.

- `Case 2.` $a > 2d$. The set of new positive differences includes $a+2d$, $a+3d$, $a+4d$, and $a+6d$. We are done.

Hence, S_4 is not sum-dominant.

We prove that S_5 is not sum-dominant. Denote $K_5 = S_5\backslash\{3a+2b\}$. The set of all possible differences in $K_5 - K_5$ is $D_5 = \{a, a+b, a+2b, 2a+2b, b, 2b\}$. It is an easy check that if either $2a \in D_5$ or $2a+b \in D_5$, then we do not have a sum-dominant set. To illustrate, we give an example. Suppose that $2a = a+2b$. Equivalently, $a = 2b$. We have $S_5 = (0\,|\,2b, b, b, 2b, 2b) = \{0, 2b, 3b, 4b, 6b, 8b\}$. Because $\{0, 2, 3, 4, 6, 8\}$ is not sum-dominant, S_5 is not sum-dominant. Suppose that $\{2a, 2a+b\} \cap D_5 = \emptyset$. Adding $3a+2b$ to K_5 gives us three new positive differences $\{2a, 2a+b, 3a+2b\}$. Because the number of new sums is at most 6, we know that S_5 is not sum-dominant.

We prove S_6 is not sum-dominant. Denote $K_6 = S_6\backslash\{4a+3b\}$. The set of all possible differences in $K_6 - K_6$ is $D_6 = \{a+b, 2a+b, 3a+b, 3a+2b, a, 2a, b\}$. It is an easy check that if $\{a+2b, 2a+2b\} \cap D_6 \neq \emptyset$, we do not have a sum-dominant set. Suppose that $\{a+2b, 2a+2b\} \cap D_6 = \emptyset$. Adding $4a+3b$ to K_6 gives us three new positive differences $\{a+2b, 2a+2b, 4a+3b\}$ while at most 6 new sums. Hence, S_6 is not sum-dominant.

We prove S_7 is not sum-dominant. Denote $K_7 = S_7\backslash\{4a+2b\}$. Adding $4a+2b$ to K_7 gives us at most two possible new sums $\{7a+3b, 7a+4b\}$. Because $4a+2b$ is a new difference, S_7 is not sum-dominant.

We prove that S_8 is not sum-dominant. Denote $K_8 = S_8\backslash\{4a+b\}$. We have $K_8 - K_8 = D_8 = \{0, a, 2a, 3a, 4a, 5a\}$. If $\{b, a+b, 2a+b\} \cap D_8 \neq \emptyset$, S_8 is not sum-dominant. If otherwise, $4a+b \rightarrow K_8$ gives at least 3 new positive differences while at most 6 new sums. Hence, S_8 is not sum-dominant.

The proof that S_i for $9 \leq i \leq 15$ are not sum-dominant is similar to the proof that S_5 is not sum-dominant. So, we omit the proof.

6. The Smallest Sum-dominant Set of Primes

The Green-Tao theorem guarantees that there are infinitely many sum-dominant sets of primes; that is, sum-dominant sets can be constructed using long arithmetic progressions of primes. However, sum-dominant sets are expected to appear much earlier in the prime sequence. Chu et al. constructed the set $P = \{19, 79, 109, 139, 229, 349, 379, 439\}$ using the Hardy-Littlewood k-tuple conjecture [4]. We summarize the idea of the construction below.

An m-tuple $(b_1, b_2, \ldots, b_m)$ is said to be admissible if for all integers $k \geq 2$, $\{b_1, b_2, \ldots, b_m\}$ does not cover all values modulo k. Clearly, we only need to check for all values of k between 2 and m. An integer n matches the tuple if $b_1 + n, b_2 + n, \ldots, b_m + n$ are all primes. The Hardy-Littlewood conjecture implies that every admissible m-tuple is matched by infinitely many integers.

We apply this construction to A_8 and A_{11} [7] to find new sum-dominant sets that appear earlier in the prime sequence. In particular,

$$12A_8 = \{0, 24, 48, 96, 108, 120, 180, 204, 228\}$$

is an admissible 9-tuple. A quick search shows that

$$A'_8 = 103 + 12A_8 = \{103, 127, 151, 199, 211, 223, 283, 307, 331\}$$

is a set of primes. Because sum-dominance is preserved under affine transformation, $103 + 12A_8$ is also sum-dominant. Similarly,

$$A'_{11} = 23 + 6A_{11} = \{23, 47, 59, 71, 89, 107, 137, 149, 173\}$$

is sum-dominant. Both A'_8 and A'_{11} are smaller than the previous set P in terms of the largest element.

We can do better with computers' help. We run an algorithm to find all sum-dominant subsets of $\{3, 5, \ldots, 109\}$ (all primes from 3 to 109). We find 2725 sets with

$$\{3, 5, 7, 13, 17, 19, 23, 43, 47, 53, 59, 61, 67, 71, 73\}$$

being the uniquely smallest. We exclude 2 from our original set of primes because if a set S of primes containing 2 is sum-dominant, then $S \backslash \{2\}$ is also sum-dominant. (The reason is that adding 2 to a set of odd primes gives at least 7 more differences than sums.) This reduces our running time by a half. Because all of our 2725 sets have their sum sets larger than their difference sets by at most 4, adding 2 to any of these sets does not give a sum-dominant set.

7. Future Work

We end with two questions for future research.

- Is there a human-understandable proof that a set of cardinality 7 is not sum-dominant? Let A be a set of cardinality 7. Then $|A + A| \leq 7 \cdot 8/2 = 28$, while $|A - A| \leq 7 \cdot 6 + 1 = 43$. Using the same argument in the proof of Lemma 2, we know that $A - A$ has at least 11 pairs of equal positive differences. Hence, $A - A$ has at most 21 distinct differences; equivalently, $A - A$ has at most 10 distinct positive differences. This bound is not good enough and requires us to consider a lot more cases than when we have only 8 distinct positive differences. Hence, it is unknown whether a human-understandable proof can be written down in full.

- What is the minimum number of elements to be added to an arithmetic progression to form a sum-dominant set?

Acknowledgement The author would like to thank the anonymous referee and the editor for a careful reading of this article.

References

[1] M. Asada, S. Manski, S. J. Miller, and H. Suh, Fringe pairs in generalized MSTD sets, *Int. J. Number Theory* **13** (2017), 2653-2675.

[2] H. V. Chu, N. Luntzlara, S. J. Miller, and L. Shao, Infinite families of partitions into MSTD subsets, *Integers* **19** (2019), #A60.

[3] H. V. Chu, N. Luntzlara, S. J. Miller, and L. Shao, Generalizations of a curious family of MSTD sets hidden by interior blocks, *Integers*, to appear.

[4] H. V. Chu, N. McNew, S. J. Miller, V. Xu, and S. Zhang, When sets can and cannot have sum-dominant subsets, *J. Integer Seq* **18** (2018).

[5] H. V. Chu, When sets are not sum-dominant, *J. Integer Seq* **22** (2019).

[6] G. A. Freiman and V. P. Pigarev, *Number Theoretic Studies in the Markov Spectrum and in the Structural Theory of Set Addition* (Russian), Kalinin. Gos. Univ., Moscow, 1973.

[7] P. V. Hegarty, Some explicit constructions of sets with more sums than differences, *Acta Arith.* **130** (2007), 61-77.

[8] G. Iyer, O. Lazarev, S. J. Miller, and L. Zhang, Generalized more sums than differences sets, *J. Number Theory* **132** (2012), 1054-1073.

[9] J. Marica, On a conjecture of Conway, *Canad. Math. Bull.* **12** (1969), 233-234.

[10] G. Martin and K. O'Bryant, Many sets have more sums than differences, in *Additive Combinatorics*, CRM Proc. Lecture Notes, Vol. 43, Amer. Math. Soc., 2007, pp. 287-305.

[11] S. J. Miller, B. Orosz, and D. Scheinerman, Explicit constructions of infinite families of MSTD sets, *J. Number Theory* **130** (2010), 1221-1233.

[12] S. J. Miller and D. Scheinerman, Explicit constructions of infinite families of MSTD sets, in *Additive Number Theory*, Springer, 2010, pp. 229-248.

[13] S. J. Miller and K. Vissuet, Most subsets are balanced in finite groups, in M. B. Nathanson, ed., *Combinatorial and Additive Number Theory — CANT 2011 and 2012*, Springer Proceedings in Mathematics & Statistics, Vol. 101, 2014, pp. 147-157.

[14] S. J. Miller, S. Pegado, and L. Robinson, Explicit constructions of large families of generalized more sums than differences sets, *Integers* **12** (2012), #A30.

[15] M. B. Nathanson, Problems in additive number theory, V: affinely inequivalent MSTD sets, *North-West. Eur. J. Math* **3** (2017), 123-141.

[16] M. B. Nathanson, Problems in additive number theory. I, in *Additive Combinatorics*, CRM Proc. Lecture Notes, Vol. 43, *Amer. Math. Soc.*, 2007, pp. 263-270.

[17] M. B. Nathanson, Sets with more sums than differences, *Integers* **7(1)** (2007), #A5.

[18] M. B. Nathanson, MSTD sets and Freiman isomorphism, *Funct. Approx. Comment. Math* **58** (2018), 187-205.

[19] I. Z. Ruzsa, On the cardinality of $A+A$ and $A-A$, in *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, *Coll. Math. Soc. J. Bolyai.*, Vol. 18, 1978, pp. 933-938.

[20] I. Z. Ruzsa, Sets of sums and differences, in *Seminar on Number Theory, Paris 1982-83*, Progr. Math., Vol. 51, Birkhäuser, 1984, pp. 267-273.

[21] I. Z. Ruzsa, On the number of sums and differences, *Acta Math. Sci. Hungar.* **59** (1992), 439-447.

[22] W. G. Spohn, On Conway's conjecture for integer sets, *Canad. Math. Bull.* **14** (1971), 461-462.

[23] Y. Zhao, Constructing MSTD sets using bidirectional ballot sequences, *J. Number Theory* **130** (2010), 1212-1220.

[24] Y. Zhao, Counting MSTD sets in finite abelian groups, *J. Number Theory* **130** (2010), 2308-2322.

[25] Y. Zhao, Sets characterized by missing sums and differences, *J. Number Theory* **131** (2011), 2107-2134.