Generalized More-Sum-Than Difference Sets

Geoffrey Iyer¹, Liyang Zhang², Oleg Lazarev³


Advisor Steven J Miller (sjml@williams.edu)

Maine-Quebec Number Theory Conference
October 2, 2011
Motivation

We often care about the sum/difference of a set $A \subseteq \mathbb{Z}$. 
Motivation

We often care about the sum/difference of a set $A \subseteq \mathbb{Z}$.

- Goldbach’s Conjecture: $\{\text{Evens}\} \setminus \{2 \subseteq P + P$
- Fermat’s Last Theorem: If $A_n$ is the set of positive $n^{\text{th}}$ powers, then $A_n + A_n \cap A_n = \emptyset$ for all $n \geq 3$
Motivation

We often care about the sum/difference of a set $A \subseteq \mathbb{Z}$.

- Goldbach’s Conjecture: $\{\text{Evens}\} \setminus \{2 \subseteq P + P$
- Fermat’s Last Theorem: If $A_n$ is the set of positive $n^{th}$ powers, then $A_n + A_n \cap A_n = \emptyset$ for all $n \geq 3$

Natural question: What are the sizes of the sum/difference sets?
Definitions

A finite set of integers, \(|A|\) its size. Form

- Sumset: \(A + A = \{a_i + a_j : a_i, a_j \in A\}\).
- Difference set: \(A - A = \{a_i - a_j : a_i, a_j \in A\}\).
A finite set of integers, $|A|$ its size. Form

- **Sumset:** $A + A = \{a_i + a_j : a_i, a_j \in A\}$.
- **Difference set:** $A - A = \{a_i - a_j : a_i, a_j \in A\}$.

**Definition**

- **Difference dominated:** $|A - A| > |A + A|$  
- **Balanced:** $|A - A| = |A + A|$  
- **Sum dominated (or MSTD):** $|A + A| > |A - A|$. 
What could cause a set to be sum-dominated? Difference-dominated?

- \( x + x = 2x, \) but \( x - x = 0. \)
- \( x + y = y + x, \) but \( x - y \neq y - x. \)
History

What could cause a set to be sum-dominated?
Difference-dominated?

- $x + x = 2x$, but $x - x = 0$.
- $x + y = y + x$, but $x - y \neq y - x$.

Nathanson, *Problems in Additive Number Theory*. "With the right way of counting the vast majority of sets satisfy $|A - A| > |A + A|$."
Theorem (Martin-O’Bryant): If each set $A \subseteq [0, n - 1]$ is equally likely, then a positive percentage of sets are sum-dominant in the limit. More precisely:

$$\lim_{n \to \infty} \frac{\# \{ A \subseteq [0, n - 1]; A \text{ is sum-dominant} \}}{2^n} \approx 0.00045.$$
How is it possible for a positive percent of sets to be sum-dominant?
How is it possible for a positive percent of sets to be sum-dominant?

**Martin-O’Bryant:** We have the expected values

- $|A + A| \sim 2n - 11,$
- $|A - A| \sim 2n - 7.$
Say $A \subseteq [0, n - 1]$, $x \in A + A$ if we can find $a_1, a_2 \in A$ such that $a_1 + a_2 = x$.
Say $A \subseteq [0, n - 1]$, $x \in A + A$ if we can find $a_1, a_2 \in A$ such that $a_1 + a_2 = x$.

If $x$ is near $n$ there are many possibilities for $a_1, a_2$. 
Say $A \subseteq [0, n - 1]$, $x \in A + A$ if we can find $a_1, a_2 \in A$ such that $a_1 + a_2 = x$.

If $x$ is near $n$ there are many possibilities for $a_1, a_2$.

With high probability, the middle will be full.
Say $A \subseteq [0, n - 1]$, $x \in A + A$ if we can find $a_1, a_2 \in A$ such that $a_1 + a_2 = x$.

If $x$ is near $n$ there are many possibilities for $a_1, a_2$.

With high probability, the middle will be full.

The trick is to control the fringes.
As adding sets and not multiplying, set

\[ kA = \underbrace{A + \cdots + A}_{k \text{ times}}. \]

\[ [a, b] = \{a, a + 1, \ldots, b\}. \]
Questions

- Can we find a set $A$ such that $|kA + kA| > |kA - kA|$?

- Can we find a set $A$ such that $|A + A| > |A - A|$ and $|2A + 2A| > |2A - 2A|$?

- Can we find a set $A$ such that $|kA + kA| > |kA - kA|$ for all $k$?
Questions

- Can we find a set $A$ such that $|kA + kA| > |kA - kA|$? Yes.

- Can we find a set $A$ such that $|A + A| > |A - A|$ and $|2A + 2A| > |2A - 2A|$? Yes.

- Can we find a set $A$ such that $|kA + kA| > |kA - kA|$ for all $k$? No. (No such set exists.)
Initial Observations

Question: Can we find \( A \) with \( |kA + kA| > |kA - kA| \)?
Initial Observations

Question: Can we find $A$ with $|kA + kA| > |kA - kA|$?

- One set is enough to show a positive percentage.
Initial Observations

Question: Can we find $A$ with $|kA + kA| > |kA - kA|$?

- One set is enough to show a positive percentage.
- How do we find one set?
Initial Observations

Question: Can we find $A$ with $|kA + kA| > |kA - kA|$?

- One set is enough to show a positive percentage.
- How do we find one set?
- Computer simulations? We couldn’t find a set for $k = 2$; the probability of finding some of these sets is less than $10^{-8}$. 
Initial Observations

Question: Can we find $A$ with $|kA + kA| > |kA - kA|$?

- One set is enough to show a positive percentage.
- How do we find one set?
- Computer simulations? We couldn’t find a set for $k = 2$; the probability of finding some of these sets is less than $10^{-8}$.

If $A$ is symmetric ($A = c - A$ for some $c$) then

$$|A + A| = |A + (c - A)| = |A - A|.$$
$$|2A + 2A| > |2A - 2A|$$

Example:  $$|2A + 2A| > |2A - 2A|$$
Example: $|2A + 2A| > |2A - 2A|$

$$A = \{0, 1, 3, 4, 5, 9\} \cup [33, 56] \cup \{79, 83, 84, 85, 87, 88, 89\}$$
|2A + 2A| > |2A − 2A|

Example: |2A + 2A| > |2A − 2A|

\[ A = \{0, 1, 3, 4, 5, 9\} \cup [33, 56] \cup \{79, 83, 84, 85, 87, 88, 89\} \]
\[ |2A + 2A| > |2A - 2A| \]

\[ A + A = [0, 9] \cup \{10, 12, 13, 14, 18\} \cup [33, 145] \]
\[ \cup \{158, 162, 163, 164, 166, 167\} \cup [168, 178] \]
$|2A + 2A| > |2A - 2A|$
\[ |2A + 2A| > |2A - 2A| \]
\[2A + 2A > |2A - 2A|\]
\[ 2A + 2A > 2A - 2A \]

\[ A + A + A + A \]

\[ \begin{align*}
0 & \quad 9 & \quad 18 & \quad 33 & \quad 56 & \quad 79 & \quad 89 \\
0 & \quad 9 & \quad 18 & \quad 33 & \quad 145 & \quad 158 & \quad 168 & \quad 178 \\
0 & \quad 9 & \quad 18 & \quad 27 & \quad 33 & \quad 234 & \quad 247 & \quad 257 & \quad 267 \\
0 & \quad 9 & \quad 18 & \quad 27 & \quad 33 & \quad 326 & \quad 336 & \quad 346 & \quad 356
\end{align*} \]
\[ 2A + 2A > 2A - 2A \]

\[ A + A \]
\[ |2A + 2A| > |2A - 2A| \]
$|2A + 2A| > |2A - 2A|$
\[2A + 2A > |2A - 2A|\]

\[A + A - A\]

\[\begin{align*}
0 & \quad 9 & \quad 18 & \quad 33 & \quad 56 & \quad 79 & \quad 89 \\
\text{ Background } & \quad \text{ Generalized MSTD } & \quad \text{ Generations } & \quad \text{ Limiting behavior of } kA & \quad \text{ Conclusion }
\end{align*}\]
\[|2A + 2A| > |2A - 2A|\]

Say that \(L\) is the left fringe of \(A\), \(R\) the right fringe.
Say that $L$ is the left fringe of $A$, $R$ the right fringe.

The left fringe of $A + A$ is $L + L$, the right fringe is $R + R$. 
$|2A + 2A| > |2A - 2A|$

Say that $L$ is the left fringe of $A$, $R$ the right fringe.

The left fringe of $A + A$ is $L + L$, the right fringe is $R + R$.

The left fringe of $A - A$ is $L - R$, the right fringe is $R - L$. 
\[ |2A + 2A| > |2A - 2A| \]

\[ A + A - A - A \]
After dealing with some technical obstructions, we can generalize:
After dealing with some technical obstructions, we can generalize:

For all nontrivial choices of $s_1, d_1, s_2, d_2$, $\exists A \subseteq \mathbb{Z}$ such that $|s_1A - d_1A| > |s_2A - d_2A|$.
After dealing with some technical obstructions, we can generalize:

For all nontrivial choices of $s_1, d_1, s_2, d_2$, $\exists A \subseteq \mathbb{Z}$ such that $|s_1 A - d_1 A| > |s_2 A - d_2 A|$.

Example: We can have $|A + A + A + A| > |A + A + A - A|$:  

$$A = \{0, 1, 3, 4, 5, 9, 33, 34, 35, 50, 54, 55, 56, 58, 59, 60\}$$
Question: Does a set $A$ exist such that $|A + A| > |A - A|$ and $|A + A + A + A| > |A + A - A - A|$?
Question: Does a set $A$ exist such that $|A + A| > |A - A|$ and $|A + A + A| > |A + A - A - A|$?

Say $A$ is $k$-generational if $A, 2A, \ldots, kA$ all sum-dominant.
$k$-Generational Sets

Question: Does a set $A$ exist such that $|A + A| > |A - A|$ and $|A + A + A + A| > |A + A - A - A|$?
Question: Does a set $A$ exist such that $|A + A| > |A - A|$ and $|A + A + A + A| > |A + A - A - A|$?

Yes!

$$A = \{0, 1, 3, 4, 7, 26, 27, 29, 30, 33, 37, 38, 40, 41, 42, 43, 46, 49, 50, 52, 53, 54, 72, 75, 76, 78, 79, 80\}$$

In fact, we can find a $k$-generational set for all $k$. 
Idea of proof: We can find \( A_j \) such that
\[
|jA_j + jA_j| > |jA_j - jA_j| \quad \text{for a specific} \quad 1 \leq j \leq k.
\]
Idea of proof: We can find $A_j$ such that
$|jA_j + jA_j| > |jA_j - jA_j|$ for a specific $1 \leq j \leq k$.

Combine the $A_j$ using the method of base expansion.
Base Expansion: For sets $A_1, A_2$ and $m \in \mathbb{N}$ sufficiently large (relative to $A_1, A_2$) the set

$$A = m \cdot A_1 + A_2$$

behaves like the direct product $A_1 \times A_2 \subseteq \mathbb{Z} \times \mathbb{Z}$. 
Base Expansion: For sets $A_1, A_2$ and $m \in \mathbb{N}$ sufficiently large (relative to $A_1, A_2$) the set

$$A = m \cdot A_1 + A_2$$

behaves like the direct product $A_1 \times A_2 \subseteq \mathbb{Z} \times \mathbb{Z}$.

(here multiplication is the usual scalar multiplication)
Base Expansion: For sets $A_1, A_2$ and $m \in \mathbb{N}$ sufficiently large (relative to $A_1, A_2$) the set

$$A = m \cdot A_1 + A_2$$

behaves like the direct product $A_1 \times A_2 \subseteq \mathbb{Z} \times \mathbb{Z}$.

(here multiplication is the usual scalar multiplication)

In particular:

$$|xA - yA| = |xA_1 - yA_1| \cdot |xA_2 - yA_2|$$

whenever $x + y$ is small relative to $m$. 


Generalization

For nontrivial $x_j, y_j, w_j, z_j \ (2 \leq j \leq k)$, we can find an $A$ such that $|x_jA - y_jA| > |w_jA - z_jA|$ for all $j$. 
For nontrivial \( x_j, y_j, w_j, z_j \) (\( 2 \leq j \leq k \)), we can find an \( A \) such that \( |x_jA − y_jA| > |w_jA − z_jA| \) for all \( j \).

Example: We can find an \( A \) such that

\[
|A + A| > |A − A| \\
|A + A − A| > |A + A + A| \\
|5A − 2A| > |A − 6A| \\
\vdots \\
|1870A − 141A| > |1817A − 194A|
\]
Limiting behavior of \( kA \)

Question: Can we find \( A \) with \( |kA + kA| > |kA - kA| \) for all \( k \)?

No. No such set exists. It turns out that all sets have a sort of limiting behavior.
Question: Can we find $A$ with $|kA + kA| > |kA - kA|$ for all $k$?

No. No such set exists.
Question: Can we find $A$ with $|kA + kA| > |kA - kA|$ for all $k$?

No. No such set exists.

It turns out that all sets have a sort of limiting behavior.
Example: \( A = \{0, 3, 5, 6, 8, 9, 10, 11, 12, 15, 16, 20\} \)

![Diagram of stabilizing fringes with example set A]
Stabilizing Fringes

Example: \( A = \{0, 3, 5, 6, 8, 9, 10, 11, 12, 15, 16, 20\} \)

\[ \begin{align*}
0 & \quad 8 & \quad 12 & \quad 20 \\
\text{(20–8)} & \\
\end{align*} \]

**Figure:** \( A \)

\[ \begin{align*}
0 & \quad 8 & \quad 12 & \quad 32 & \quad 40 \\
\text{(40–8)} & \\
\end{align*} \]

**Figure:** \( A + A \)
Nathanson: For any set $A$, $kA$ becomes stabilized before $k$ reaches $\max(A)^2 \cdot |A|$.

We improve this bound to $\max(A)$. 
Theorem

For any set $A$, $kA$ will become difference-dominated or balanced before $k$ reaches $2 \cdot \max(A)$.

Proof Idea:

- The middle will quickly become full, and the remaining fringes are finite.
- $kA \subseteq kA - kA$. Any sum can eventually be written as a difference.

Because the form stabilizes, this means $kA - kA \supseteq kA + kA$ when $k$ large.
Other Results

**Arbitrary Differences**: We can create sets $A$ where

$$|kA + kA| - |kA - kA| = m.$$ 

More generally, $|s_1 A - d_1 A| - |s_2 A - d_2 A| = m$.

**Simultaneous Comparison**: We can create a set $A$ where


More generally, any order and number of (nontrivial) comparisons.
Thanks

Thanks to:
- Williams College,
- NSF Grant DMS0850577,
- Conference Coordinators.