Sum and Difference Sets in Semidirect Products of Groups

Andrew Keisling and Matthew Phang

SMALL REU 2022 at Williams College
Probability and Number Theory Group
Joint work with Ruben Ascoli, Justin Cheigh,
Guilherme Zeus Dantas e Moura, Ryan Jeong,
Astrid Lilly, Steven J. Miller, and Prakod Ngamlamai

August 13, 2022

Definition

Definition

Given a set of integers A, we define the sumset and difference set of A as follows:

$$A + A = \{a_1 + a_2 : a_1, a_2 \in A\},$$

$$A - A = \{a_1 - a_2 : a_1, a_2 \in A\}.$$

We want to compare the sizes of these two sets:

- $|A+A| > |A-A|$: A has more sums than differences (MSTD).
- $|A+A| = |A-A|$: A is sum-difference balanced.
- $|A+A| < |A-A|$: A has more differences than sums (MDTS).

Why do we care?

Problems in additive number theory can be written in terms of sumsets and difference sets:

Goldbach's conjecture says $\{4, 6, 8, 10, \ldots\} \subseteq P + P$.

Fermat's last theorem says that $(A_n + A_n) \cap A_n = \emptyset$, where A_n is the set of positive n^{th} powers for $n \geq 3$.

Expectation

We expect that most sets of integers are MDTS rather than MSTD.

Addition is commutative, subtraction is not.

Theorem (Martin-O'Bryant, 2006)

Let P be any arithmetic progression with length n. On average, the difference set of a subset of P has 4 more elements than its sumset:

$$\frac{1}{2^n}\sum_{S\subseteq P}|A-A| \sim 2n-7,$$

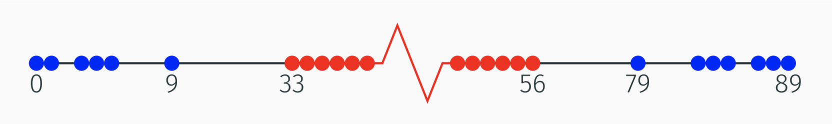
$$\frac{1}{2^n}\sum_{S\subseteq P}|A+A| \sim 2n-11.$$

MSTD sets of integers

Theorem (Martin-O'Bryant, 2006)

For $n \geq 15$, the number of sum-dominant subsets of $\{0, 1, 2, \ldots, n-1\}$ is at least $(2 \cdot 10^{-7})2^n$.

MSTD subsets can be constructed by carefully controlling the "fringes" (elements close to 0 or $n-1$).



Example

Example

Let $A = \{0, 2, 3, 4, 7, 11, 12, 14\}$.

$$A + A = \{0, 1, ..., 28\} \setminus \{1, 20, 27\}, \ |A + A| = 26,$$

$$A - A = \{-14, -13, ..., 14\} \setminus \{-13, -6, 6, 13\}, \ |A - A| = 25.$$

Finite Groups

Miller and Vissuet considered the analogous problem with a finite group G (not necessarily abelian) in place of $\mathbb{Z}$.

Definition

Given $A \subseteq G$, we define the sumset and difference set of A as follows:

$$A + A = \{a_1 a_2 : a_1, a_2 \in A\},$$

$$A - A = \{a_1 a_2^{-1} : a_1, a_2 \in A\}.$$

The lack of fringes or commutativity significantly affect the methods and results in these cases.

Theorem (Miller-Vissuet 2014)

Let G_n be a family of finite groups such that $|G_n| \to \infty$. If $A_n \subseteq G_n$ is chosen uniformly at random, then

$$\mathbb{P}(A_n + A_n = A_n - A_n = G_n) \to 1 \text{ as } n \to \infty.$$

Dihedral groups

More MSTD than MDTS

Conjecture (Miller-Visuet, 2014)

For all $n \geq 3$, D_{2n} has more MSTD subsets than MDTS subsets.

Intuition comes from splitting $A \subseteq D_{2n}$ into R (rotation elements) and F (flip elements):

Set	Rotations in set	Flips in set
A	R	F
$A+A$	$R+R, F+F$	$R+F, -R+F$
$A-A$	$R-R, F+F$	$R+F$

$R+R$ and $-R+F$ contribute to $A+A$ and not $A-A$.

$R-R$ contributes to $A-A$ and not $A+A$.

Partitioning by size

SMALL 2020 made progress toward proving the conjecture by partitioning the subsets of D_{2n} by size.

Notation: Let $\mathcal{S}_m$ denote the set of subsets of D_{2n} of size m.

Lemma (Kim et al. 2020)

$\mathcal{S}_2$ has strictly more MSTD subsets than MDTS subsets.

We further extended this piecemeal approach:

Lemma (SMALL 2022)

$\mathcal{S}_3$ has strictly more MSTD subsets than MDTS subsets.

Large subsets

SMALL 2020 also showed that sufficiently large subsets must be sum-difference balanced:

Lemma (Kim et al. 2020)

Given $A \subseteq D_{2n}$, if $|A| > n$, then $A + A = A - A = D_{2n}$.

It remains to show that $\mathcal{S}_m$ does not have more MDTS sets than MSTD sets for $4 \leq m \leq n$.

Composition of A

We further partitioned $\mathcal{S}_m$ by the number of rotation elements versus flip elements.

Recall that we write each A as $R \cup F$, where R is the rotations and F is the flips.

Lemma (SMALL 2022)

If $|R| > \frac{n}{2}$ or $|F| > \frac{n}{2}$, then A cannot be MDTS.

Proof: We have $|A - A| > |A + A|$ only if $R - R$ contributes more than $R + R$ and $-R + F$.

But if $|R| > \frac{n}{2}$ or $|F| > \frac{n}{2}$, then $R + R$ contributes all of the possible rotations in D_{2n}.

Counting collisions

Results

For large n, we extended to certain values in $4 \leq m \leq n$ by probabilistic methods.

Theorem (SMALL 2022)

For all $n \gg 0$, more of the subsets in $\mathcal{S}_m$ are MSTD than MDTS for $6 \leq m \leq c \cdot \sqrt{n}$ where c is a global constant.

We proved this for $c = 0.1134$.

Even more can be said if we further restrict m:

Theorem (SMALL 2022)

For any $\epsilon > 0$, there exist m_ϵ and c_ϵ such that for all $n \gg 0$, if $m_\epsilon \leq m \leq c_\epsilon\sqrt{n}$, the proportion of MSTD sets in $\mathcal{S}_m$ is at least $1 - \epsilon$.

More MSTD for minimal overlaps

The proof relies on limiting the number of overlapping sums in $A + A$.

Let $|A| = m$, $|F| = k$, and $|R| = m - k$. Assuming minimal overlaps, and not counting $F + F$:

Type	A+A	A-A
Rotations	$\binom{m-k}{2} + (m-k)$	$2\binom{m-k}{2}$
Flips	$2(m-k)k$	$(m-k)k$

This implies that a set is MSTD for $m > k > \frac{m}{3} - 1$.

Additional sum overlaps will subtract from the left hand side of this inequality, so m must be sufficiently small.

Collisions

Definition

Let $A \in \mathcal{S}_m$, and let $i = (a, b, c, d) \in A^4$. We call the event that $ab = cd$ (or equivalently, $d = c^{-1}ab$) a *collision*.

For our purposes, we will disregard three types of collisions:

$$(a, b, a, b),$$
$$(a, b, b, a) : a, b \in R,$$
$$(a, b, c, d) : a, b, c, d \in F.$$

These *degenerate collisions* have already been accounted for in the previous analysis.

Expectation for collisions

For the random variable (A, i) with $A \in \mathcal{S}_m$, $i \in A^4$, let $X_{A,i}$ be the indicator that i is a collision.

Lemma (SMALL 2022)

Define the function $f(a, b, c)$ on D_{2n}^3 as

$$f_m(a,b,c) = \begin{cases} \frac{m-3}{2n-3}, & a, b, c, c^{-1}ab \text{ all distinct,} \\ \frac{m-2}{2n-2}, & (a=b), c, c^{-1}ab \text{ distinct, or} \\ & a, (b=c), c^{-1}ab \text{ distinct,} \\ 1, & \text{otherwise.} \end{cases}$$

Then,

$$\mathbb{E}[X_{A,i}] = \frac{1}{m^4\binom{2n}{m}} \sum_{A' \subseteq D_{2n} : |A'| = m} \sum_{(a,b,c) \in A'^3} f_m(a,b,c).$$

Finishing the proof

Let X_A denote the number of non-degenerate collisions of A.

Apply linearity of expectation and discard degenerate collisions to obtain an upper bound for $\mathbb{E}(X_A)$ using $\mathbb{E}(X_{A,i})$.

Assume that all non-degenerate collisions only decrease $|A+A|$ and not $|A-A|$.

Apply Markov's inequality to obtain both theorems.

Generalizations

Generalized dihedral groups

Recall that for an abelian group G, the *generalized dihedral group* of G is

$$\mathrm{Dih}(G) = \mathbb{Z}/2 \ltimes G$$

with the non-identity element of $\mathbb{Z}/2$ acting on G by inversion.

Conjecture (GenDihMMSTDTMDTS)

$\mathrm{Dih}(G)$ has more MSTD subsets than MDTS subsets for all finite abelian groups G that contain an element of order at least 3.

Our main theorems and methods for D_{2n} should translate to $\mathrm{Dih}(G)$ as half the group elements are still "flips."

We also seek to use bijections $\mathrm{Dih}(G) \to D_{2|G|}$ to directly apply results from D_{2n}.

Infinite dihedral groups

We can also take $G = \mathbb{Z}^r$ if we restrict the $\mathbb{Z}^r$-components in $\mathrm{Dih}(\mathbb{Z}^r)$ to $[0, n-1]^r$:

Theorem (SMALL 2022)

For all $n \gg 0$, more of the sets $A \subseteq \mathbb{Z}/2 \ltimes [0, n-1]^r \subseteq \mathrm{Dih}(\mathbb{Z}^r)$ of size m are MSTD than MDTS for $6 \leq m \leq c \cdot \sqrt{n}$ where c is a global constant.

Theorem (SMALL 2022)

For any $\epsilon > 0$, there exist m_ϵ and c_ϵ such that for all $n \gg 0$, if $m_\epsilon \leq m \leq c_\epsilon \sqrt{n}$, a proportion of at least $1 - \epsilon$ of the subsets are MSTD among $A \subseteq \mathbb{Z}/2 \ltimes [0, n-1]^r \subseteq \mathrm{Dih}(\mathbb{Z}^r)$ of size m.

Proof: Construct a bijection $\mathbb{Z}/2 \ltimes [0, n-1]^r \to D_{2n^r}$ that preserves collisions.

Future work

We would like to extend the bounds on m to show that for all n, D_{2n} has more MSTD sets than MDTS sets:

- $c\sqrt{n} < m \leq n$.
- Carefully count collisions.
- Analyze missed elements for m close to n.
- Construct injections from MDTS sets to MSTD sets in $\mathcal{S}_m$.

Any results we prove for D_{2n} may translate to generalized dihedral groups.

Expected size

Calculate expected sizes of $|A+A|$ and $|A-A|$.

Theorem (SMALL 2022)

For prime n and a random set $A \subseteq D_{2n}$ with $|A| = m$, we have that

$$\mathbb{E}(|A-A|) = 2n - n\frac{\binom{n}{m}}{\binom{2n}{m}}2^m - n^2(n-1)\sum_{k=1}^{m-1}\frac{\binom{n+k-m-1}{m-k-1}\binom{n-k-1}{k-1}}{\binom{2n}{m}k(m-k)} - \frac{(n-1)(2n)\binom{n-m-1}{m-1}}{(m)\binom{2n}{m}}.$$

Would also require understanding of variance.

Expected size for difference sets

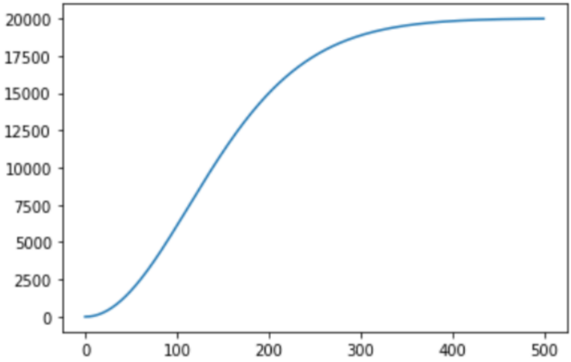


Figure: $\mathbb{E}(|A - A|)$ versus m for $n = 10007$.

Acknowledgments

Thank you!

This research was conducted as part of the **2022 SMALL REU program** at Williams College, and was funded by **NSF Grant DMS1947438**, **Harvey Mudd College**, and **Williams College** funds. We thank Steven J. Miller and our colleagues from the 2022 SMALL REU program for many helpful conversations.

References

[1] Peter Hegarty and Steven J Miller. "When almost all sets are difference dominated". In: *Random Structures & Algorithms* 35.1 (2009), pp. 118–136.

[2] Virginia Hogan and Steven J Miller. "When Generalized Sumsets are Difference Dominated". In: *arXiv preprint arXiv:1301.5703* (2013).

[3] Greg Martin and Kevin O'Bryant. "Many sets have more sums than differences". In: *arXiv preprint math/0608131* (2006).

[4] Steven J. Miller and Kevin Vissuet. *Most Subsets are Balanced in Finite Groups*. 2013. DOI: 10.48550/ARXIV.1308.2344. URL: https://arxiv.org/abs/1308.2344.