An Infinite Family of MSTD Sets

Steven J. Miller (Williams)
Daniel Scheinerman (Brown)

2008 AMS/MAA Conference

January 5, 2009

Introduction

Definition

$[a, b] = \{x \in \mathbb{Z} : a \leq x \leq b\}$
$[1, n] = \{1, 2, 3, \ldots, n\}$

Introduction

Definition

$[a,b] = \{x \in \mathbb{Z} : a \leq x \leq b\}$
$[1,n] = \{1,2,3,\dots,n\}$

Basic Operations

Given a set A, two basic operations:

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

Introduction

Definition

$[a,b] = \{x \in \mathbb{Z} : a \leq x \leq b\}$
$[1,n] = \{1,2,3,\dots,n\}$

Basic Operations

Given a set A, two basic operations:

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

$A = \{1,2,5\}$
$A + A = \{2,3,4,6,7,10\}$
$A - A = \{-4,-3,-1,0,1,3,4\}$

Introduction

Definition

$[a, b] = \{x \in \mathbb{Z} : a \leq x \leq b\}$
$[1, n] = \{1, 2, 3, \ldots, n\}$

Basic Operations

Given a set A, two basic operations:

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

$A = \{1, 2, 5\}$
$A + A = \{2, 3, 4, 6, 7, 10\}$
$A - A = \{-4, -3, -1, 0, 1, 3, 4\}$

Which is bigger?

In general which do we expect to be bigger $|A + A|$ or $|A - A|$?

Introduction

Subtraction is not commutative

In general $a - b \neq b - a$, so we expect to have more differences.

Definition

Given a set A, we will call More Sum Than Difference or MSTD if $|A + A| > |A - A|$.

Introduction

Subtraction is not commutative

In general $a - b \neq b - a$, so we expect to have more differences.

Definition

Given a set A, we will call More Sum Than Difference or MSTD if $|A+A| > |A-A|$.

Are there any MSTD sets?

Yes! The "smallest" is $A = \{1,3,4,5,8,12,13,15\}$. We then have $A+A = [2,30] \setminus \{3,22,29\}$ so $|A+A| = 26$, and $A-A = [-14,14] \setminus \{-13,-6,6,13\}$, so $|A-A| = 25$.

History of the Problem

What proportion of sets are MSTD?

Let $p(n)$ be the proportion of sets in $[1, n]$ that are MSTD. A natural question is does $p(n)$ tend to zero? The answer is no. Martin and O'Bryant (2006) showed that $p(n) > 2 \times 10^{-7}$ for $n \geq 15$.

History of the Problem

What proportion of sets are MSTD?

Let $p(n)$ be the proportion of sets in $[1, n]$ that are MSTD. A natural question is does $p(n)$ tend to zero? The answer is no. Martin and O'Bryant (2006) showed that $p(n) > 2 \times 10^{-7}$ for $n \geq 15$.

Conjecture: $p(n)$ is increasing.

History of the Problem

What proportion of sets are MSTD?

Let $p(n)$ be the proportion of sets in $[1, n]$ that are MSTD. A natural question is does $p(n)$ tend to zero? The answer is no. Martin and O'Bryant (2006) showed that $p(n) > 2 \times 10^{-7}$ for $n \geq 15$.

Conjecture: $p(n)$ is increasing.

Our Goal

Martin and O'Bryant proved that MSTDs have positive density by probabilistic techniques. Our goal was to prove the same result by direct construction of an infinite family of MSTD sets.

History of the Problem

What proportion of sets are MSTD?

Let $p(n)$ be the proportion of sets in $[1, n]$ that are MSTD. A natural question is does $p(n)$ tend to zero? The answer is no. Martin and O'Bryant (2006) showed that $p(n) > 2 \times 10^{-7}$ for $n \geq 15$.

Conjecture: $p(n)$ is increasing.

Our Goal

Martin and O'Bryant proved that MSTDs have positive density by probabilistic techniques. Our goal was to prove the same result by direct construction of an infinite family of MSTD sets.

Our Result

- Most families of MSTDs have density at most about $2^{-n/2}$.
- Ours has density at least $\frac{1}{n^4}$

Why should there be MSTD sets?

Since $a - b \neq b - a$, we have nearly twice as many "potential differences" as sums. The number of formal sums is

$$\binom{n}{2} + n = \frac{n^2 + n}{2}$$

and the number of formal differences is

$$2\binom{n}{2} + 1 = n^2 - n + 1.$$

Why should there be MSTD sets?

Since $a - b \neq b - a$, we have nearly twice as many "potential differences" as sums. The number of formal sums is

$$\binom{n}{2} + n = \frac{n^2 + n}{2}$$

and the number of formal differences is

$$2\binom{n}{2} + 1 = n^2 - n + 1.$$

Lots of collisions (especially in the middle)

If $A \subset [1, n]$ then $A + A \subset [2, 2n]$ and $A - A \subset [1 - n, n - 1]$. These two intervals have width $2n - 1$ which is much less than n^2.

Why should there be MSTD sets?

Since $a - b \neq b - a$, we have nearly twice as many "potential differences" as sums. The number of formal sums is

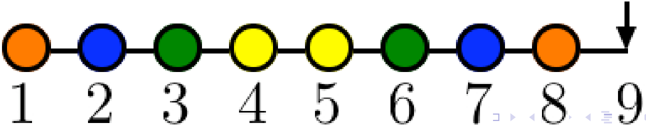
$$\binom{n}{2} + n = \frac{n^2 + n}{2}$$

and the number of formal differences is

$$2\binom{n}{2} + 1 = n^2 - n + 1.$$

Lots of collisions (especially in the middle)

If $A \subset [1, n]$ then $A + A \subset [2, 2n]$ and $A - A \subset [1 - n, n - 1]$. These two intervals have width $2n - 1$ which is much less than n^2.



History of the Problem

Expected size of $A+A$ and $A-A$

Let A be a uniformly randomly chosen subset of $[1,n]$ then

$$\Pr(k \in A+A) = \begin{cases} 1-(3/4)^{\frac{k-1}{2}} & 1 \le k \le n \text{ and } k \text{ is odd} \\ 1-(1/2)(3/4)^{\frac{k}{2}-1} & 2 \le k \le n \text{ and } k \text{ is even} \\ 1-(3/4)^{n-\frac{k-1}{2}} & n \le k \le 2n-1 \text{ and } k \text{ is odd} \\ 1-(1/2)(3/4)^{n-\frac{k}{2}} & n < k \le 2n \text{ and } k \text{ is even} \end{cases}$$

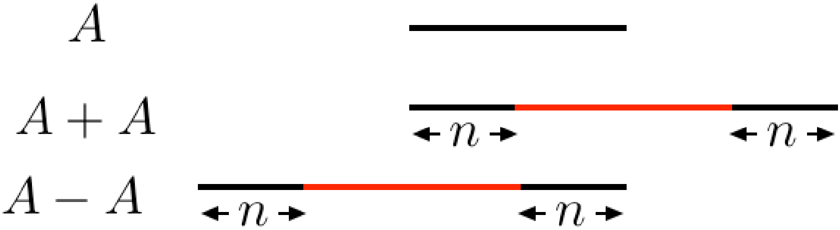
A similar expression holds for $\Pr(k \in A-A)$. It is fairly elementary to sum these and conclude:

$$\mathrm{E}(|A+A|) = (1/2^n) \sum_{A \subset [1,n]} |A+A| = 2n - 11 + o(1)$$

$$\mathrm{E}(|A-A|) = (1/2^n) \sum_{A \subset [1,n]} |A-A| = 2n - 7 + o(1).$$

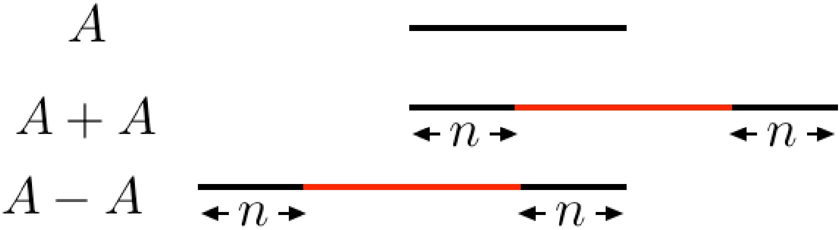
P_n sets

We say A is a P_n set if both its sum set and its difference set contain all but the n fringe elements on either side.



P_n sets

We say A is a P_n set if both its sum set and its difference set contain all but the n fringe elements on either side.

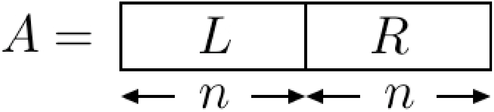


Most sets are P_n

A random set drawn from $[1, 2n]$ in the uniform model is P_n with probability approaching 1.

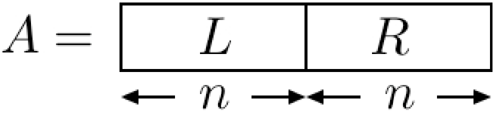
An Important Lemma

Let $A \subset [1, 2n]$ with $1, 2n \in A$ and suppose that A is P_n. Write $A = L \cup R$ where $L \subset [1, n]$ and $R \subset [n+1, 2n]$.

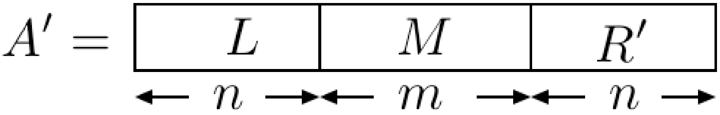


An Important Lemma

Let $A \subset [1, 2n]$ with $1, 2n \in A$ and suppose that A is P_n. Write $A = L \cup R$ where $L \subset [1, n]$ and $R \subset [n+1, 2n]$.



Construct a new set A' by "sticking in" a set M in the middle so that A' is P_n as well. So $A' = L \cup M \cup (R + m)$.



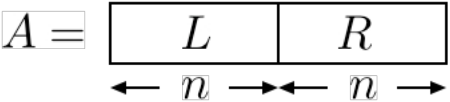
Then A' is P_n, and in general $|A' + A'| - |A + A| = |A' - A'| - |A - A|$.

Constructing new MSTD sets from old

Our objective is to take a P_n, MSTD set $A = L \cup R$ and find a family of sets M so that we can construct a family $A' = L \cup M \cup (R + m)$ of MSTD sets.

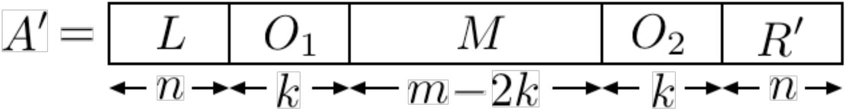
Constructing new MSTD sets from old

Our objective is to take a P_n, MSTD set $A = L \cup R$ and find a family of sets M so that we can construct a family $A' = L \cup M \cup (R + m)$ of MSTD sets.

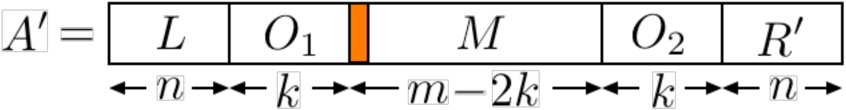


We add three pieces to the middle.

- $O_1 = [n+1, n+K]$ $\leftarrow$ contains k consecutive elements
- $O_2 = [n+m+k+1, n+m+2k]$ $\leftarrow$ k consecutive elements
- M has width $m - 2k$ and no run of k missing elements

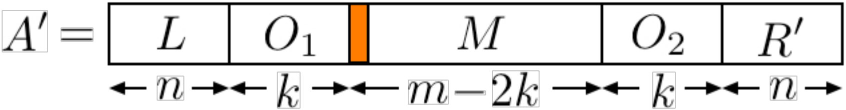


How large is this family of sets?



- L, O_1, O_2 and R' are fixed.
- We have added elements into a total width of m. So hopefully we have about 2^m possibilities.

How large is this family of sets?



- L, O_1, O_2 and R' are fixed.
- We have added elements into a total width of m. So hopefully we have about 2^m possibilities.

How many choices for M?

- Divide M into blocks of width $k/2$.
- Assume every block contains at least one element.
- Remove first element for uniqueness

$$\text{Lower bound for fixed k is } 2^{m-2k}\left(1-\frac{1}{2^{k/2}}\right)^{\frac{m-2k}{k/2}}$$

How large is this family of sets?

Number of possible M for a fixed k:

$$2^{m-2k}\left(1-\frac{1}{2^{k/2}}\right)^{\frac{m-2k}{k/2}}$$

Fix the width of the inserted interval, m. The number of possibilities can be found by summing over k:

$$S=\sum_{k=n}^{m/2} 2^{m-2k}\left(1-\frac{1}{2^{k/2}}\right)^{\frac{m-2k}{k/2}}$$

We compute the desired proportion by dividing by 2^m:

$$\frac{S}{2^m}=\sum_{k=n}^{m/2}\frac{1}{2^{2k}}\left(1-\frac{1}{2^{k/2}}\right)^{2m/k}$$

How large is this family of sets?

$$\frac{S}{2^m} = \sum_{k=n}^{m/2} \frac{1}{2^{2k}} \left(1 - \frac{1}{2^{k/2}}\right)^{2m/k}$$

We can achieve bounds:

$$\frac{1}{m^4} < \frac{S}{2^m} < \left(\frac{\log m}{m}\right)^4$$

This is not quite a positive percentage, but much denser than other families of MSTDs.

Questions?