On sum-free and solution-free sets of integers

Andrew Treglown
Joint work with József Balogh, Hong Liu, Maryam Sharifzadeh, and Robert Hancock

University of Birmingham

7th October 2016

Introduction

Definition

A set $S \subseteq [n]$ is sum-free if no solutions to $x + y = z$ in S.

Examples

- $\{1, 2, 4\}$ is not sum-free.
- Set of odds is sum-free.
- {n/2+1, n/2+2,. . . , n} is sum-free.

Introduction

What do sum-free subsets of $[n]$ look like?

Introduction

What do sum-free subsets of $[n]$ look like?

- Every sum-free subset of $[n]$ has size at most $\lceil n/2 \rceil$.

Introduction

What do sum-free subsets of $[n]$ look like?

- Every sum-free subset of $[n]$ has size at most $\lceil n/2 \rceil$.

Deshouillers, Freiman, Sós and Temkin (1999)

If $S \subseteq [n]$ is sum-free then at least one of the following holds:

(i) $|S| \leq 2n/5 + 1$;

(ii) S consists of odds;

(iii) $|S| \leq min(S)$.

Introduction

Examples of sum-free sets

- Set of odds is sum-free.
- {n/2+1, n/2+2,. . . , n} is sum-free.

These two examples show there are at least $2^{n/2}$ sum-free subsets of $[n]$.

Introduction

Cameron-Erdős Conjecture (1990)

The number of sum-free subsets of $[n]$ is $O(2^{n/2})$.

Introduction

Cameron-Erdős Conjecture (1990)

The number of sum-free subsets of $[n]$ is $O(2^{n/2})$.

Green; Sapozhenko c. 2003

There are constants c_e and c_o, s.t. the number of sum-free subsets of $[n]$ is

$$(1+o(1))c_e 2^{n/2},\ \textit{or}\ (1+o(1))c_o 2^{n/2}$$

depending on the parity of n.

Introduction

- The previous result doesn't tell us anything about the distribution of the sum-free sets in $[n]$.
- In particular, recall that $2^{n/2}$ sum-free subsets of $[n]$ lie in a single maximal sum-free subset of $[n]$.

Cameron-Erdős Conjecture (1999)

There is an absolute constant $c > 0$, s.t. the number of maximal sum-free subsets of $[n]$ is $O(2^{n/2-cn})$.

Lower bound construction

There are at least $2^{\lfloor n/4 \rfloor}$ maximal sum-free subsets of $[n]$.

- Suppose n is even. Let S consist of n together with precisely one number from each pair $\{x, n-x\}$ for odd $x < n/2$.
- Notice distinct S lie in distinct maximal sum-free subsets of $[n]$.
- Roughly $2^{n/4}$ choices for S.

Main sum-free result

Denote by $f_{\max}(n)$ the number of maximal sum-free subsets in $[n]$. Recall that $f_{\max}(n) \geq 2^{\lfloor n/4 \rfloor}$.

Cameron-Erdős Conjecture (1999)

$$\exists c > 0, \quad f_{\max}(n) = O(2^{n/2-cn}).$$

Łuczak-Schoen (2001)

$$f_{\max}(n) \leq 2^{n/2-2^{-28}n} \text{ for large } n$$

Wolfovitz (2009)

$$f_{\max}(n) \leq 2^{3n/8+o(n)}.$$

Balogh-Liu-Sharifzadeh-T. (2015)

$$f_{\max}(n) = 2^{n/4+o(n)}.$$

Main sum-free result

Balogh-Liu-Sharifzadeh-T. (2015)

$$f_{\max}(n) = 2^{n/4+o(n)}.$$

Main sum-free result

Balogh-Liu-Sharifzadeh-T. (2015)

$$f_{\max}(n) = 2^{n/4+o(n)}.$$

Balogh-Liu-Sharifzadeh-T. (2016+)

For each $1 \leq i \leq 4$, there is a constant C_i such that, given any $n \equiv i \mod 4$, $[n]$ contains $(C_i + o(1))2^{n/4}$ maximal sum-free sets.

Tools

From additive number theory:

- Container lemma of Green.
- Removal lemma of Green.
- Structure of sum-free sets by Deshouillers, Freiman, Sós and Temkin.

From extremal graph theory: upper bound on the number of maximal independent sets for

- all graphs by Moon and Moser.
- triangle-free graphs by Hujter and Tuza.
- Not too sparse and almost regular graphs.

Sketch of the proof

Balogh-Liu-Sharifzadeh-T. (2014)

$$f_{\max}(n) = 2^{n/4+o(n)}.$$

Container Lemma [Green]

There exists $\mathcal{F} \subseteq 2^{[n]}$, s.t.

(i) $|\mathcal{F}| = 2^{o(n)}$;

(ii) $\forall S \subseteq [n]$ sum-free, $\exists F \in \mathcal{F}$, s.t. $S \subseteq F$;

(iii) $\forall F \in \mathcal{F}$, $|F| \leq (1/2 + o(1))n$ and the number of Schur triples in F is $o(n^2)$.

Sketch of the proof

Balogh-Liu-Sharifzadeh-T. (2014)

$$f_{\max}(n) = 2^{n/4+o(n)}.$$

Container Lemma [Green]

There exists $\mathcal{F} \subseteq 2^{[n]}$, s.t.

(i) $|\mathcal{F}| = 2^{o(n)}$;

(ii) $\forall S \subseteq [n]$ sum-free, $\exists F \in \mathcal{F}$, s.t. $S \subseteq F$;

(iii) $\forall F \in \mathcal{F}$, $|F| \leq (1/2 + o(1))n$ and the number of Schur triples in F is $o(n^2)$.

By (i) and (ii), it suffices to show that for every container $A \in \mathcal{F}$,

$$f_{\max}(A) \leq 2^{n/4+o(n)}.$$

Constructing maximal sum-free sets

Removal+Structural lemmas $\Rightarrow$ classify containers $A \in \mathcal{F}$:

- Case 1: small container, $|A| \leq 0.45n$;
- Case 2: 'interval' container, 'most' of A in $[n/2+1, n]$.
- Case 3: 'odd' container, $|A \setminus O| = o(n)$.

Moreover, in all cases $A = B \cup C$ where B is sum-free and $|C| = o(n)$.

Constructing maximal sum-free sets

Removal+Structural lemmas $\Rightarrow$ classify containers $A \in \mathcal{F}$:

- Case 1: small container, $|A| \leq 0.45n$;
- Case 2: 'interval' container, 'most' of A in $[n/2+1, n]$.
- Case 3: 'odd' container, $|A \setminus O| = o(n)$.

Moreover, in all cases $A = B \cup C$ where B is sum-free and $|C| = o(n)$.

Crucial observation

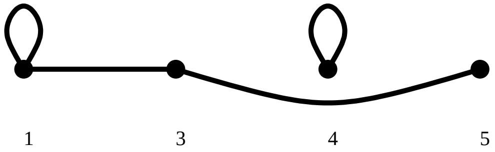
Every maximal sum-free subset in A can be built in two steps:
(1) Choose a sum-free set S in C;
(2) Extend S in B to a maximal one.

maximal sum-free sets $\Rightarrow$ maximal independent sets

Definition

Given $S, B \subseteq [n]$, the link graph of S on B is $L_S[B]$, where $V = B$ and $x \sim y$ iff $\exists z \in S$ s.t. $\{x, y, z\}$ is a Schur triple.

$L_2[1, 3, 4, 5]$



maximal sum-free sets $\Rightarrow$ maximal independent sets

Definition

Given $S, B \subseteq [n]$, the link graph of S on B is $L_S[B]$, where $V = B$ and $x \sim y$ iff $\exists z \in S$ s.t. $\{x, y, z\}$ is a Schur triple.

Lemma

Given $S, B \subseteq [n]$ sum-free and $I \subseteq B$, if $S \cup I$ is a maximal sum-free subset of $[n]$, then I is a maximal independent set in $L_S[B]$.

Case 1: small container, $|A| \leq 0.45n$.

Recall $A = B \cup C$, B sum-free, $|C| = o(n)$.

Crucial observation

Every maximal sum-free subset in A can be built in two steps:
(1) Choose a sum-free set S in C;
(2) Extend S in B to a maximal one.

Case 1: small container, $|A| \leq 0.45n$.

Recall $A = B \cup C$, B sum-free, $|C| = o(n)$.

Crucial observation

Every maximal sum-free subset in A can be built in two steps:
(1) Choose a sum-free set S in C;
(2) Extend S in B to a maximal one.

- Fix a sum-free $S \subseteq C$ (at most $2^{|C|} = 2^{o(n)}$ choices).
- Consider link graph $L_S[B]$.
- Moon-Moser: $\forall$ graphs G, $MIS(G) \leq 3^{|G|/3}$.
- So # extensions in (2) is at most $MIS(L_S[B])$,

$$MIS(L_S[B]) \leq 3^{|B|/3} \leq 3^{0.45n/3} \ll 2^{0.249n}.$$

- In total, A contains at most $2^{o(n)} \times 2^{0.249n} \ll 2^{n/4}$ maximal sum-free sets.

Cases 2 and 3.

- Now container A could be bigger than $0.45n$.
- This means crude Moon-Moser bound doesn't give accurate bound on $f_{\max}(A)$.
- Instead we obtain more structural information about the link graphs.

Cases 2 and 3.

- Now container A could be bigger than $0.45n$.
- This means crude Moon-Moser bound doesn't give accurate bound on $f_{\max}(A)$.
- Instead we obtain more structural information about the link graphs.

- For example, when A 'close' to interval $[n/2+1, n]$ link graphs are triangle-free
- Hujta-Tuza: $MIS(G) \leq 2^{|G|/2}$ for all triangle-free graphs G.
- Gives better bound on $f_{\max}(A)$.

Balogh-Liu-Sharifzadeh-T. (2016+)

For each $1 \leq i \leq 4$, there is a constant C_i such that, given any $n \equiv i \mod 4$, $[n]$ contains $(C_i + o(1))2^{n/4}$ maximal sum-free sets.

(i) Count by hand the maximal sum-free sets S that are 'extremal':
- S that contain precisely one even number.
- S where $\min(S) \approx n/4$, $\min_2(S) \approx n/2$.

(ii) Count remaining maximal sum-free sets using the container method.

Solution-free sets: Introduction

Let $\mathcal{L}$ denote the equation $a_1x_1 + \cdots + a_kx_k = b$ where $a_1, \ldots, a_k, b \in \mathbb{Z}$.

Solution-free sets: Introduction

Let $\mathcal{L}$ denote the equation $a_1x_1 + \cdots + a_kx_k = b$ where $a_1, \ldots, a_k, b \in \mathbb{Z}$.

Definitions:

1. $\mathcal{L}$ is translation-invariant if $\sum a_i = b = 0$.
2. A subset $A \subseteq [n]$ is $\mathcal{L}$-free if it does not contain any 'non-trivial' solutions to $\mathcal{L}$.
3. A subset $A \subseteq [n]$ is a maximal $\mathcal{L}$-free set if it is $\mathcal{L}$-free, and if the addition of any further $x \in [n] \setminus A$ would make it no longer $\mathcal{L}$-free.

Solution-free sets: Introduction

Fundamental Questions

- **Q1:** What is the size of the largest $\mathcal{L}$-free subset of $[n]$?
- **Q2:** How many $\mathcal{L}$-free subsets of $[n]$ are there?
- **Q3:** How many maximal $\mathcal{L}$-free subsets of $[n]$ are there?

Q1: What is the size of the largest $\mathcal{L}$-free subset of $[n]$?

Let $\mu_{\mathcal{L}}(n)$ be the size of the largest $\mathcal{L}$-free subset of $[n]$.

$\mathcal{L}$	$\mu_{\mathcal{L}}(n)$	Comment
$x+y=z$	$\lceil n/2 \rceil$	odds or interval
$x+y=2z$	$o(n)$	Roth's theorem (1953)
$p(x+y)=rz$, $r>2p$	$n-\lfloor 2n/r \rfloor$	union (Hegarty 2007)

Q1: What is the size of the largest $\mathcal{L}$-free subset of $[n]$?

Let $\mu_{\mathcal{L}}(n)$ be the size of the largest $\mathcal{L}$-free subset of $[n]$.

$\mathcal{L}$	$\mu_{\mathcal{L}}(n)$	Comment
$x + y = z$	$\lceil n/2 \rceil$	odds or interval
$x + y = 2z$	$o(n)$	Roth's theorem (1953)
$p(x + y) = rz$, $r > 2p$	$n - \lfloor 2n/r \rfloor$	union (Hegarty 2007)

In general...

$\mathcal{L}$	$\mu_{\mathcal{L}}(n)$
translation-invariant	$o(n)$
not translation-invariant	$\Omega(n)$

Q1: What is the size of the largest $\mathcal{L}$-free subset of $[n]$?

Hancock, T. 2015+

Let $\mathcal{L}$ be $px + qy = z$ where $p \geq q$ and $p \geq 2, p, q \in \mathbb{N}$. If n is sufficiently large then $\mu_{\mathcal{L}}(n) = n - \lfloor n/(p+q) \rfloor$.

- More recently, we have determined $\mu_{\mathcal{L}}(n)$ for a range of different equations $\mathcal{L}$ of the form $px + qy = rz$ where $p \geq q \geq r$.
- In each case, the extremal examples are 'intervals' or 'congruency classes'.

Q2: How many $\mathcal{L}$-free subsets of $[n]$ are there?

Let $f(n, \mathcal{L})$ be the number of $\mathcal{L}$-free subsets of $[n]$.
Clearly for any $\mathcal{L}$, we have $f(n, \mathcal{L}) \geq 2^{\mu_{\mathcal{L}}(n)}$.

Q2: How many $\mathcal{L}$-free subsets of $[n]$ are there?

Let $f(n, \mathcal{L})$ be the number of $\mathcal{L}$-free subsets of $[n]$.
Clearly for any $\mathcal{L}$, we have $f(n, \mathcal{L}) \geq 2^{\mu_{\mathcal{L}}(n)}$.

Green, Sapozhenko 2003

Let $\mathcal{L}$ be $x + y = z$. Then $\exists\, C_1, C_2$ s.t. given any $n \equiv i \mod 2$, $f(n, \mathcal{L}) = (C_i + o(1))2^{n/2}$.

Q2: How many $\mathcal{L}$-free subsets of $[n]$ are there?

Let $f(n, \mathcal{L})$ be the number of $\mathcal{L}$-free subsets of $[n]$.
Clearly for any $\mathcal{L}$, we have $f(n, \mathcal{L}) \geq 2^{\mu_{\mathcal{L}}(n)}$.

Green, Sapozhenko 2003

Let $\mathcal{L}$ be $x + y = z$. Then $\exists\, C_1, C_2$ s.t. given any $n \equiv i \bmod 2$, $f(n, \mathcal{L}) = (C_i + o(1))2^{n/2}$.

Observation (Cameron-Erdős 1990)

Let $\mathcal{L}$ be translation-invariant. Then it is not true that $f(n, \mathcal{L}) = \Theta(2^{\mu_{\mathcal{L}}(n)})$.

Q2: How many $\mathcal{L}$-free subsets of $[n]$ are there?

Green 2005

Let $\mathcal{L}$ be $a_1x_1 + \cdots + a_kx_k = 0$ where $a_1, \ldots, a_k \in \mathbb{Z}$.
Then $f(n, \mathcal{L}) = 2^{\mu_{\mathcal{L}}(n)+o(n)}$ (where $o(n)$ depends on $\mathcal{L}$).

Q2: How many $\mathcal{L}$-free subsets of $[n]$ are there?

Green 2005

Let $\mathcal{L}$ be $a_1x_1 + \cdots + a_kx_k = 0$ where $a_1, \ldots, a_k \in \mathbb{Z}$.
Then $f(n, \mathcal{L}) = 2^{\mu_\mathcal{L}(n)+o(n)}$ (where $o(n)$ depends on $\mathcal{L}$).

Hancock, T. 2015+

Fix $p, q \in \mathbb{N}$ where (i) $q \geq 2$ and $p > q(3q-2)/(2q-2)$ or (ii) $q = 1$ and $p \geq 3$. Let $\mathcal{L}$ be $px + qy = z$.
Then $f(n, \mathcal{L}) = \Theta(2^{\mu_\mathcal{L}(n)})$.

Q3: How many maximal $\mathcal{L}$-free subsets of $[n]$ are there?

Let $f_{\max}(n, \mathcal{L})$ be the number of maximal $\mathcal{L}$-free subsets of $[n]$.
We have already seen:

Balogh-Liu-Sharifzadeh-Treglown 2016+

Let $\mathcal{L}$ be $x + y = z$. For each $1 \leq i \leq 4$, there is a constant C_i s.t. given any $n \equiv i \bmod 4$, $f_{\max}(n, \mathcal{L}) = (C_i + o(1))2^{n/4}$.

Q3: How many maximal: An initial upper bound

Definition

- $\mathcal{L}$-triple: A solution to $\mathcal{L}$ when $\mathcal{L}$ is in three variables.
- $\mathcal{M}_{\mathcal{L}}(n)$: The set of $x \in [n]$ s.t. x does not lie in any $\mathcal{L}$-triple in $[n]$.
- $\mu^*_{\mathcal{L}}(n) := |\mathcal{M}_{\mathcal{L}}(n)|$.

Hancock, T. 2015+

Let $\mathcal{L}$ be $px + qy = rz$ where $p, q, r \in \mathbb{Z}$.
Then $f_{\max}(n, \mathcal{L}) \leq 3^{(\mu_{\mathcal{L}}(n) - \mu^*_{\mathcal{L}}(n))/3 + o(n)}$.

Q3: How many maximal: Further improvements?

Hancock, T. 2015+

Let $\mathcal{L}$ be $px + qy = z$ where $p \geq q \geq 2$ are integers s.t. $p \leq q^2 - q$ and $\gcd(p, q) = q$.
Then $f_{\max}(n, \mathcal{L}) \leq 2^{(\mu_{\mathcal{L}}(n) - \mu^*_{\mathcal{L}}(n))/2 + o(n)}$.

Q3: How many maximal: Further improvements?

Hancock, T. 2015+

Let $\mathcal{L}$ be $px + qy = z$ where $p \geq q \geq 2$ are integers s.t. $p \leq q^2 - q$ and $\gcd(p, q) = q$.
Then $f_{\max}(n, \mathcal{L}) \leq 2^{(\mu_{\mathcal{L}}(n) - \mu^*_{\mathcal{L}}(n))/2 + o(n)}$.

Hancock, T. 2016+

Let $\mathcal{L}$ be $qx + qy = z$ where $q \geq 2$ is an integer.
Then $f_{\max}(n, \mathcal{L}) = 2^{n/2q + o(n)}$.

Open problem

Given an abelian group G let $\mu(G)$ denote the size of the largest sum-free subset of G.

Green–Ruzsa (2005)

There are $2^{\mu(G)+o(|G|)}$ sum-free subsets of G.

Conjecture [Balogh-Liu-Sharifzadeh-T.]

There are at most $2^{\mu(G)/2+o(|G|)}$ maximal sum-free subsets of G.

- Easy to prove $3^{\mu(G)/3+o(|G|)}$ as an upper bound.