ADDITIVE AND MULTIPLICATIVE SIDON SETS

O. ROCHE-NEWTON and A. WARREN

Johann Radon Institute for Computational and Applied Mathematics (RICAM), Linz, Austria
e-mails: o.rochenewton@gmail.com, audie.warren@oeaw.ac.at

(Received March 24, 2021; accepted April 6, 2021)

Abstract. We give a construction of a set \( A \subset \mathbb{N} \) such that any subset \( A' \subset A \) with \( |A'| \gg |A|^{2/3} \) is neither an additive nor multiplicative Sidon set. In doing so, we refute a conjecture of Klurman and Pohoata.

1. Introduction

A set \( A \subset \mathbb{R} \) is a Sidon set if the only solutions to the energy equation

\[ a + b = c + d \quad \text{such that} \quad a, b, c, d \in A \]

are the trivial solutions whereby \( \{a, b\} = \{c, d\} \). Sidon sets are important and widely studied objects in Additive Combinatorics; see [14] for background on this topic and [7] for a thorough survey.

Given an arbitrary finite set \( A \subset \mathbb{R} \), we define \( s_+(A) \) to be the size of the largest Sidon set \( A' \) such that \( A' \subset A \). The problem of understanding the behaviour of \( s_+(\mathbb{N}) \) has attracted particular interest, and it is known that \( s_+(\mathbb{N}) = \Theta(N^{1/2}) \). See [4] for a nice introduction to this question.

Komlós, Sulyok and Szemerédi [6] proved (as part of a much more general result about sets avoiding certain linear configurations) that, up to constant factors, \( s_+(A) \) is minimised among sets of size \( N \) in the case when \( A = \{1, 2, \ldots, N\} \). That is, they proved that the bound

\[ s_+(A) \gg |A|^{1/2} \]

holds for any\(^1\) finite set \( A \subset \mathbb{R} \).

---

\(^1\) In fact, this was proved in [6] only for sets of integers, but it has later been established that the same result holds over the reals; see for instance [9, Lemma 2.2].

0236-5294/$20.00 © 2021 Akadémiai Kiadó, Budapest, Hungary
Similarly, a *multiplicative Sidon set* is a set $A \subseteq \mathbb{R}$ such that the only solutions to

$$ab = cd$$

are the trivial solutions with $\{a, b\} = \{c, d\}$. Given an arbitrary finite set $A \subseteq \mathbb{R}$, we define $s_*(A)$ to be the size of the largest multiplicative Sidon set $A'$ such that $A' \subseteq A$. Applying the estimate (1.1) to the set $\log A = \{\log a : a \in A\}$, it follows that

$$s_*(A) \gg |A|^{1/2}$$

for any finite $A \subseteq \mathbb{R}$.

In the spirit of the sum-product problem, one may ask the question of whether it is guaranteed that, for a fixed set $A \subseteq \mathbb{R}$, at least one of the inequalities (1.1) and (1.2) is improvable. This question was indeed asked by Klurman and Pohoata [8], who made the following conjecture.

**Conjecture 1.1.** There exists a constant $c > 0$ such that for all finite $A \subseteq \mathbb{R}$,

$$\max\{s_+(A), s_*(A)\} \gg |A|^{1/2+c}.$$  

Klurman and Pohoata in fact also made the stronger conjecture that, for all $\varepsilon > 0$,

$$\max\{s_+(A), s_*(A)\} \gg |A|^{1-\varepsilon}.$$  

The main result of this note is a construction disproving (1.4).

**Theorem 1.2.** For all $N \in \mathbb{N}$, there exists a set $A \subseteq \mathbb{N}$ with $|A| \geq N$ and such that $s_+(A) \ll |A|^{1/2} \log |A|$ and $s_*(A) \ll |A|^{2/3}$. In particular,

$$\max\{s_+(A), s_*(A)\} \ll |A|^{2/3}.$$  

**1.1. Notation.** Throughout the paper, the standard notation $\ll, \gg$ and respectively $O$ and $\Omega$ is applied to positive quantities in the usual way. That is, $X \gg Y, Y \ll X, X = \Omega(Y)$ and $Y = O(X)$ all mean that $X \geq cY$, for some absolute constant $c > 0$. If both $X \ll Y$ and $Y \ll X$ hold we write $X \approx Y$, or equivalently $X = \Theta(Y)$. All logarithms are in base $e$. For a set $A$, the notation $A(\cdot)$ is used for its characteristic function.
2. Construction

Proof of Theorem 1.2. For a sufficiently large integer $n$, define $P$ to be the set of all primes less than or equal to $n$, and define

$$Q = \left\{ n < q \leq \frac{n^2}{\log n} : q \text{ prime} \right\}.$$ 

The construction is simply the product set

$$A := P \cdot Q = \left\{ p \cdot q : p, q \text{ prime}, p \leq n < q \leq \frac{n^2}{\log n} \right\}.$$ 

$A$ is a fairly dense subset of the set $\{1, 2, \ldots, n^3/\log n\}$, which implies that $A + A$ is small and thus $s_+(A)$ is also small. Indeed, by the Fundamental Theorem of Arithmetic and the Prime Number Theorem,

$$|A| = |P||Q| \approx \frac{n^3}{\log^3 n}.$$ 

Since $A + A \subset \{1, \ldots, \frac{2n^3}{\log n}\}$, it follows that

$$|A + A| \ll |A|^2 |A|.$$ 

Furthermore, since a Sidon set of size $m$ contained in $A$ gives rise to $\binom{m}{2} + m$ distinct sums in $A + A$, it follows that

$$s_+(A) \ll |A + A|^{1/2} \ll |A|^{1/2} \log |A|.$$ 

It remains to prove the upper bound

$$s_+(A) \ll \frac{n^2}{\log^2 n} \approx |A|^{2/3}.$$ 

To do this, we appeal to an argument Erdős [3] used in order to establish the existence of a set with very small additive energy but still having fairly small $s_+(A)$. What follows can be viewed as an adaptation of this argument in the multiplicative setting.

We view the elements of $A$ as edges of a complete bipartite graph $G = (V, A)$ with vertex set $V = P \sqcup Q$, and with the edge between $p \in P$ and $q \in Q$ labelled by the product $pq$.

Suppose that $E \subset A$ is a multiplicative Sidon set. Then $E$ does not give rise to any copies of the cycle $C_4$ in $G$. Indeed, suppose that we have

Acta Mathematica Hungarica
four edges in $E$ which form a cycle. Then the cycle must be of the form $pq, qp', p'q', q'p$. But if all of these four edges are in $E$ then we have

$$(pq)(p'q') = (p'q)(pq'),$$

which contradicts the fact that $E$ is a multiplicative Sidon set.

On the other hand, any $E \subseteq A$ which does not contain a copy of $C_4$ has size $|E| \ll \frac{n^2}{\log^2 n}$ by the Kővári-Sós-Turán Theorem. For completeness, we give a quick proof below using an application of Cauchy–Schwarz:

$$|E|^2 = \left( \sum_{q \in Q} \sum_{p \in P} E(pq) \right)^2 \leq |Q| \sum_{q \in Q} \left( \sum_{p \in P} E(pq) \right)^2 \ll \frac{n^2}{\log^2 n} \sum_{p, p' \in P, : p \neq p'} \sum_{q \in Q} E(pq)E(p'q).
$$

If the sum $|E| + \frac{n^2}{\log^2 n}$ is dominated by the second term then we are done. Therefore we may assume otherwise, and the last inequality gives

$$|E|^2 \ll \frac{n^2}{\log^2 n} \left( |E| + \frac{n^2}{\log^2 n} \right) \ll \frac{n^2}{\log^2 n} |E|.$$

We have proved that any subset of size greater than $Cn^2 / \log^2 n$, for some constant $C$, contains a $C_4$, and thus is not a multiplicative Sidon set. This proves (2.2) and thus completes the proof of the theorem. \( \square \)

3. Connections to additive and multiplicative energy

3.1. The relationship between $s_+(A)$ and additive energy. The additive energy of $A$, denoted $E_+(A)$, is the number of ordered quadruples $(a, b, c, d) \in A \times A \times A \times A$ such that

$$a + b = c + d.$$

Note that $E_+(A)$ also counts the trivial solutions, and so we have $E_+(A) \geq |A|^2$. One may suspect that if $A$ has small additive energy then $s_+(A)$ must be large. This is true to some extent, as the following lemma shows.

**Lemma 3.1.** For any finite $A \subseteq \mathbb{R}$

$$s_+(A) \gg \frac{|A|^{4/3}}{(E_+(A))^{1/3}}.$$

*Acta Mathematica Hungarica*
The proof of Lemma 3.1 uses a simple probabilistic argument, and is implicit in the work of Alon and Erdős [1]. The multiplicative analogue of Lemma 3.1 (with $s_\ast(A)$ and $E_\ast(A)$ in place of $s_+(A)$ and $E_+(A)$ respectively) holds via the same reasoning. This lemma overtakes the bound (1.1) when $E_+(A) \ll |A|^{5/2}$. Combining these two results together, we have

$$s_+(A) \gg \max\left\{ \frac{|A|^{4/3}}{(E_+(A))^{1/3}}, |A|^{1/2} \right\}.$$ (3.1)

The purpose of this section is to make the observation that the combined bound (3.1) is in fact optimal. This is a consequence of the following result of Kohayakawa, Lee and Rödl [5] (we only state the parts of the result which are relevant to our analysis).

**Theorem 3.2** [5, Theorem 1.1]. Let $1/3 \leq a \leq 1$ be a fixed constant and $m = n^a(1 + o(1))$. Let $B \subseteq [n]$ be a random set of size $m$ (i.e. $B$ is chosen randomly from all sets of size $m$). Then, almost surely,

1. if $1/3 \leq a \leq 2/3$ then $s_+(B) = n^{1/3 + o(1)}$,
2. if $2/3 \leq a \leq 1$ then $s_+(B) = |B|^{1/2 + o(1)}$.

The $o(1)$ terms in the exponents suppress logarithmic factors. A more precise version of this statement is given in [5, Theorems 2.2-2.4].

For a random set $B \subseteq [n]$ of size $m$, the expected size of $E_+(B)$ is

$$cn^3 \frac{(n-4)^m}{m^n} = c' \frac{m^4}{n}.$$ (3.2)

Part (2) of Theorem 3.2 implies that, for a random set $B \subseteq [n]$ of size $n^{2/3 + o(1)} \leq |B| \leq n$, the estimate $s_+(B) \gg |B|^{1/2}$ contained in (3.1) is tight, up to the $o(1)$ factor. By (3.2) these sets will typically have far from maximal additive energy, and taking $m$ close to $n^{2/3}$ yields a set $B$ with $E_+(B) \approx |B|^{5/2}$ and $s_+(B)$ very small.

Furthermore we have, with high probability,

$$\frac{|B|^{4/3}}{E_+(B)^{1/3}} \approx n^{1/3},$$

and thus taking $B$ to be a random subset of $[n]$ with size $n^a$, $1/3 \leq a \leq 2/3$, the estimate from Lemma 3.1 matches the information $s_+(B) = n^{1/3 + o(1)}$ given by part (1) of Theorem 3.2, up to the $o(1)$ factor.

**3.2. A variant of the Klurman–Pohoata Conjecture.** Conjecture 1.1 could be viewed as a possible line of attack for the sum-product problem,
since the inequalities

\[ |A + A| \gg s_+(A)^2, \quad |AA| \gg s_*(A)^2 \]

imply that a positive answer to Conjecture 1.3 would give a non-trivial sum-product bound. Unfortunately, the construction in Theorem 1.2 shows that even the best possible result in this direction would not yield quantitative improvements to known sum-product inequalities.

However, it is not necessary that a set be additively/multiplicatively Sidon in order for it to determine many sums/products. A weaker assumption that the additive or multiplicative energy is small would suffice, in light of the usual Cauchy–Schwarz energy bounds

\[ E_+(A) \geq \frac{|A|^4}{|A + A|}, \quad E_*(A) \geq \frac{|A|^4}{|AA|}. \]

With this in mind, we propose a variant of Conjecture 1.1. For a set \( A \subset \mathbb{R} \), define\(^2\) \( t_+(A) \) to be the size of the largest subset \( A' \subset A \) such that

\[ E_+(A') < 2|A'|^2. \]

Similarly, \( t_*(A) \) denotes the size of the largest \( A' \subset A \) such that \( E_*(A') < 2|A'|^2. \)

**Question 3.3.** For what value of \( \kappa > 0 \) is it true that any finite set \( A \subset \mathbb{R} \) satisfies

\[ \max\{t_+(A), t_*(A)\} \geq |A|^{\kappa}? \]

The bound

\[ \max\{t_+(A), t_*(A)\} \gg |A|^{1/2} \]

follows immediately from the result of Komlós, Sulyok and Szemerédi [6], and can also be proved by much simpler means. For this question, an improvement of (3.3) follows from the Balog–Wooley Theorem. We will use the following result of Rudnev, Shkredov and Stevens [10].

**Theorem 3.4.** Let \( A \subset \mathbb{R} \). Then there exists \( A' \subset A \) with \( |A'| \geq |A|/2 \) and

\[ \min\{E_+(A'), E_*(A')\} \ll |A|^{11/4}. \]

We prove the following.

---

\(^2\) The multiplicative constant 2 in the definition of \( t_+(A) \) is not particularly important, but is chosen so that at least half of the contributions to \( E_+(A') \) come from the \( |A'|^2 \) trivial solutions.
Theorem 3.5. For any $A \subset \mathbb{R}$,
\[
\max\{t_+(A), t_+(A)\} \gg |A|^{5/8}.
\]

Proof. Apply Theorem 3.4 to obtain a subset $A'$ and suppose that $E_+(A') \leq C|A|^{11/4}$. Let $A''$ be a $p$-random subset of $A'$, with $p = \frac{1}{100C^{1/2}|A|^{3/8}}$. The expected size of $A''$ is at least $p|A|/2$. Let $E_0^+(A'')$ denote the number of non-trivial solutions to
\[
a_1 + a_2 = a_3 + a_4 \quad \text{such that } a_i \in A''.
\]
The expected value of $E_0^+(A'')$ is at most
\[
Cp^4|A|^{11/4} + p^3|A|^2.
\]
By Markov's inequality, the probability that
\[
E_0^+(A'') \geq \frac{1}{100} p^2|A|^2
\]
is at most $1/10$. By Chebyshev, the probability that
\[
|A''| < \frac{1}{10} p|A|
\]
is at most $1/10$. Therefore, with positive probability both events (3.4) and (3.5) do not occur. It therefore follows that there exists a set $A'' \subset A$ such that
\[
|A''| \gg |A|^{5/8} \quad \text{and} \quad E_0^+(A'') < \frac{1}{100} p^2|A|^2 \leq |A''|^2.
\]
In particular, $E_+(A'') < 2|A''|^2$, and it follows that $t_+(A) \gg |A|^{5/8}$.

If instead we have $E_*(A') \leq C|A|^{11/4}$ at the outset, we can run the same argument in the multiplicative setting and conclude that $t_+(A) \gg |A|^{5/8}$. □

Note that a small improvement to this result can be obtained by instead applying a quantitative improvement to Theorem 3.4, due to Shakan [11], but we have avoided this above in order to simplify the presentation. It can be calculated that this change to the proof results in the improved bound
\[
\max\{t_+(A), t_+(A)\} \gg |A|^{33/52-o(1)}.
\]
Note that $33/52 = 5/8 + 1/104$.

The construction in the proof of Theorem 1.2 does not give a non-trivial construction for Question 3.3. However, the original construction of Balog and Wooley 3.6 does yield the following.
THEOREM 3.6. There exists a finite set $A \subset \mathbb{R}$ such that
\[ \max\{t_+(A), t_*(A)\} \ll |A|^{5/6} \]

PROOF. Define
\[ A := \{(2i - 1)2^j : i \in [N^2], j \in [N]\}. \]
This set is a union of $N$ arithmetic progressions of length $N^2$, and since any solution to $(2i - 1)2^j = (2i' - 1)2^{j''}$ implies that $i = i'$ and $j = j'$, we have that $|A| = N^3$. Suppose that $|A'| \geq C|A|^{5/6} = CN^{5/2}$ for a sufficiently large constant $C$. Note that since
\[ AA \subseteq \{(2i - 1)2^j : i \in [2N^4], j \in [2N]\} \]
we have $|A'A'| \leq |AA| \leq 4N^5$. Applying the Cauchy Schwarz energy bound then yields
\[ E_*(A') \geq \frac{|A'|^4}{|A'A'|} \geq \frac{C^2}{4} |A'|^2. \]
Therefore, as long as $C$ is sufficiently large we must have $t_*(A) \leq C|A|^{5/6}$.

To show that $t_+(A) \ll |A|^{5/6}$, we define
\[ A_j = \{(2i - 1)2^j : i \in [N^2]\}, \]
noting that we have $A = \bigcup_{j \in [N]} A_j$, and that each $A_j$ is an arithmetic progression with $|A_j| = N^2$. Let us define
\[ I := \{j \in [N] : |A' \cap A_j| \geq 2|A'|^{1/3}N^{2/3}\}. \]
We then have
\[ |A'| = \sum_{j \in [N]} |A' \cap A_j| = \sum_{j \in I} |A' \cap A_j| + \sum_{j \notin I} |A' \cap A_j| \]
\[ \leq \sum_{j \in I} |A' \cap A_j| + 2|A'|^{1/3}N^{5/3}. \]
We now bound the energy of $A'$.
\[ E_+(A') \geq \sum_{j \in I} E(A' \cap A_j) \geq \sum_{j \in I} \frac{|A' \cap A_j|^4}{|A_j + A_j|} \gg \frac{1}{N^2} \sum_{j \in I} |A' \cap A_j|^4 \]
\[ \geq 8|A'| \sum_{j \in I} |A' \cap A_j| \geq 8|A'|^2 - 16|A'|^{4/3}N^{5/3}. \]
From the final inequality it follows that $E_+(A') \geq 2|A'|^2$ as long as the constant $C$ is sufficiently large. Therefore $t_+(A) \ll |A|^{5/6}$, proving the result. □

Note that the bound $|AA| \leq N^5$ above is slightly wasteful and can be improved by a logarithmic factor. Using this fact and slightly re-balancing the construction of $A$ above gives the small improvement

$$\max\{t_+(A), t_+(A)\} \ll \frac{|A|^{5/6}}{\log^c |A|}$$

for some constant $c > 0$.

### 3.3. Connection to the Balog–Wooley decomposition.

Conjecture (1.3) is also connected to a question of Balog and Wooley [2]: For which $\kappa > 0$ is it true that any finite set $A \subseteq \mathbb{R}$ can always be partitioned into two subsets $B$ and $C$, such that

$$\max\{E_+(B), E_+(C)\} \ll |A|^{3-\kappa}. \tag{3.6}$$

A construction is given in [2] (and which was repeated here in the proof of Theorem 3.6) proving that any such result must have $\kappa < 2/3$. Although the authors do not go as far as to conjecture that an exponent of $7/3$ is attainable in (3.6), a similar conjecture is stated in [10], where the authors conjecture that for all finite $A \subseteq \mathbb{R}$, there exists a subset $A' \subseteq A$ with $|A'| \geq \frac{|A|}{2}$ such that

$$\min\{E_+(A'), E_+(A')\} \ll |A|^{7/3+o(1)}. \tag{3.7}$$

A positive answer to this conjecture would imply a positive answer to Conjecture 1.1. Indeed, utilising Lemma 3.1 (or possibly its multiplicative analogue), the subset $A'$ satisfies

$$|A|^{5/9-o(1)} \ll |A'|^{5/9-o(1)} \ll \max\{s_+(A'), s_+(A')\} \leq \max\{s_+(A), s_+(A)\},$$

thus giving Conjecture 1.1 with $c = 1/18 - o(1)$. Weaker variants of conjecture (3.7) also imply Conjecture 1.1, as long as the exponent is at most $5/2 - \varepsilon$. Stating this in the contrapositive, any construction disproving Conjecture 1.1 also disproves strong Balog–Wooley type conjectures.

### 4. Further remarks

#### 4.1. Small sum set implies large $s_+(A)$?

Another conjecture of Klurman and Pohoata was the following: for all $\varepsilon > 0$, and any set $A \subset \mathbb{R}$

$$|A + A| \leq K|A| \Rightarrow s_+(A) \gg_{K, \varepsilon} |A|^{1-\varepsilon}. \tag{4.1}$$

*Acta Mathematica Hungarica*
By combining the multiplicative analogue of Lemma 3.1 with Solymosi’s [13] bound on the multiplicative energy via sumsets, they proved that

\[(4.2) \quad |A + A| \leq K|A| \Rightarrow s_*(A) \gg \frac{|A|^{2/3}}{K^{2/3}(\log |A|)^{1/3}}.\]

The construction given in the proof of Theorem 1.2 disproves the conjecture (4.1), at least in the range when \(K = \log^c |A|\). Indeed, we recorded in (2.1) that the construction satisfies \(|A + A| \ll |A| \log^2 |A|\). Moreover, since we have \(s_*(A) \ll |A|^{2/3}\), this construction also shows that the bound (4.2) is in fact optimal in this range, up to logarithmic factors.

4.2. Other constructions. Theorem 1.2 was obtained independently by Green and Peluse (private communication). We thank them for sharing their construction with us. The construction is similar to the set \(A\) defined in the proof of Theorem 1.2. Paraphrasing slightly, they define a set

\[B = \{p_1p_2p_3 : 1 \leq p_i \leq N, \ p_i \ \text{prime}\}.\]

The set \(B\) is a dense enough in \([N^3]\) to ensure that its sum set is small and thus \(s_+(B)\) is small. On the other hand, and similarly to our proof of Theorem 1.2, an application of the Cauchy–Schwarz inequality can be used to show that any subset of \(B\) larger than \(C|B|^{2/3}\) must contain four elements of the form

\[pqr, \ p'qr, \ pq'r', \ p'q'r',\]

which gives rise to a non-trivial multiplicative energy solution \((pqr)(p'q'r') = (p'qr)(pq'r')\).

In a forthcoming paper, Shkredov [12] gives another construction of a set with \(\max\{s_+(A), s_*(A)\} \ll |A|^{2/3}\). His construction is somewhat different in nature, and comes from taking \(A\) to be a sum set of carefully chosen geometric progressions.

Acknowledgements. We are very grateful to Cosmin Pohoata for bringing this problem to our attention, and for several helpful discussions. Thanks also to Ben Green, Oleksiy Klurman, Sarah Peluse, Ilya Shkredov and Sophie Stevens for helpful discussions.

References

[1] N. Alon and P. Erdős, An application of graph theory to additive number theory, *European J. Combin.*, 6 (1985), 201–203.

[2] A. Balog and T. D. Wooley, A low-energy decomposition theorem, *Q. J. Math.*, 68 (2017), 207–226.
ADDITIVE AND MULTIPLICATIVE SIDON SETS

[3] P. Erdős, Extremal problems in number theory, combinatorics and geometry, in: Proceedings of the International Congress of Mathematicians (Warsaw, 1983), Vols. 1,2, PWN (Warsaw, 1984), pp. 51–70.

[4] W. T. Gowers, What are dense Sidon subsets of \{1,2,\ldots,n\} like?, blog post https://gowers.wordpress.com/2012/07/13/what-are-dense-sidon-subsets-of-12-n-like/.

[5] Y. Kohayakawa, S. Lee and V. Rödl, The maximum size of a Sidon set contained in a sparse random set of integers, in: Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM (Philadelphia, PA, 2011), pp. 159–171.

[6] J. Komlós, M. Súlyok and E. Szemerédi, Linear problems in combinatorial number theory, Acta Math. Acad. Sci. Hungar., 26 (1975), 113–121.

[7] K. O'Bryant, A Complete Annotated Bibliography of Work Related to Sidon Sequences, Electron. J. Combin. (2004), #DS11.

[8] C. Pohoata, Sidon sets and sum-product phenomena, blog post https://pohoatza.wordpress.com/2021/01/23/sidon-sets-and-sum-product-phenomena/.

[9] O. Raz, A note on distinct distances, Combin. Probab. Comput., 29 (2020), 650–663.

[10] M. Rudnev, I. D. Shkredov and S. Stevens, On the energy variant of the sum-product conjecture, Rev. Mat. Iberoam., 36 (2020), 207–232.

[11] G. Shakan, On higher energy decompositions and the sum-product phenomenon, Math. Proc. Cambridge Phil. Soc., 167 (2019), 599–617.

[12] I. D. Shkredov, On an application of higher energies to Sidon sets, arXiv:2103.14670.

[13] J. Solymosi, Bounding multiplicative energy by the sumset, Adv. Math., 222 (2009), 402–408.

[14] T. Tao and V. Vu, Additive Combinatorics, Cambridge University Press (2006).