# On a problem of Erdős and Graham about consecutive sums in strictly increasing sequences

Adrian Beker

University of Zagreb

## The problem

Given a finite sequence of integers $a = (a_i)_{1 \le i \le k}$, wet let

$$S(a) = \left\{ \sum_{i=u}^{v} a_i \;\middle|\; 1 \le u \le v \le k \right\}.$$

be the set of its consecutive sums. Erdős and Graham [3] asked the following question, which also appears as problem #356 on Thomas Bloom's website *Erdős problems* [2].

### A question of Erdős and Graham

Is there some $c > 0$ such that, for all sufficiently large $n$, there exist integers $1 \le a_1 < \ldots < a_k \le n$ such that there are at least $cn^2$ distinct integers of the form $\sum_{i=u}^{v} a_i$ with $1 \le u \le v \le k$?

The obvious example in which $k = n$ and $a_i = i$ for all $1 \le i \le k$ only just fails, attaining

$$\Theta\Big(\frac{n^2}{(\log n)^{\delta}(\log\log n)^{3/2}}\Big)$$

distinct consecutive sums, where $\delta = 1 - \frac{1+\log\log 2}{\log 2} \approx 0.086$ is the Erdős-Ford-Tenenbaum constant (see [4] and [5]).

### Previous work

Erdős and Graham also asked what happens in the following cases:

- arbitrary sequences (without the monotonicity assumption): Hegyvári [6] constructed a sequence of length $\left(\frac{1}{3} + o(1)\right) n$ in $[n]$ with all consecutive sums distinct;
- permutations of $[n]$: Konieczny [7] estalished that, if $\pi \in S_n$ is chosen uniformly at random, then

$$|S(\pi)| \sim \Big(\frac{1+e^{-2}}{4}\Big)n^2$$

with high probability.

In [1], we gave an affirmative answer to the starting question. It turns out that there are many sequences that work:

### A probabilistic construction

There exists a constant $c_1 > 0$ such that the following holds for all positive integers $n$. Let $\varepsilon_1, \ldots, \varepsilon_n$ be independent Rademacher random variables and set $a_i = 3i + \varepsilon_i$ for $1 \le i \le n$. Then with positive probability, we have $|S(a)| \ge c_1 n^2$.

One can also write down an explicit example:

### A deterministic construction

There exists a constant $c_2 > 0$ such that the following holds for all positive integers $n$. Let $b$ be a positive integer such that $\log n \le b \le \frac{n}{(\log n)^2}$ and define

$$a_i = \begin{cases} 2i & \text{if } b \mid i \\ 2i - 1 & \text{otherwise} \end{cases}$$

for $1 \le i \le n$. Then $|S(a)| \ge c_2 n^2$.

## Verifying the constructions

The key to the verification of the constructions is the notion of *additive energy* and its relation to the size of sumsets/difference sets.

### Additive energy

Following Tao and Vu [8], the additive energy of a finite non-empty set $P \subseteq \mathbb{Z}$ is defined to be

$$E(P) = |\{(x, y, z, w) \in P^4 \mid x - y = z - w\}|.$$

Writing $r_P(t)$ for the number of representations of $t \in \mathbb{Z}$ as a difference of two elements of $P$, one obtains:

$$E(P) = \sum_{t \in P-P} r_P(t)^2.$$

Since $\sum_{t \in P-P} r_P(t) = |P|^2$, the Cauchy-Schwarz inequality implies that

$$E(P) \ge \frac{|P|^4}{|P-P|}.$$

In our setting, $S(a)$ is just the set of positive elements of $P(a) - P(a)$, where

$$P(a) = \left\{ \sum_{j=1}^{l} a_j \;\middle|\; 0 \le l \le k \right\}$$

is the set of partial sums of $a$. Therefore, in order to establish the probabilistic version, it suffices to prove the following:

### Expected additive energy is small

Let $n$ be a positive integer and let $\varepsilon_1, \ldots, \varepsilon_n$ be independent Rademacher random variables. Define $a_i = 3i + \varepsilon_i$ for $1 \le i \le n$. Then the expected value of $E(P(a))$ is $O(n^2)$.

By linearity of expectation, our task amounts to showing that

$$\sum_{\substack{i,j,k,l \in [0,n] \\ i<j,\ k<l}} \mathbb{P}\Bigg(\sum_{u=i+1}^{j} a_u = \sum_{v=k+1}^{l} a_v\Bigg) = O(n^2).$$

By taking symmetric differences, we may assume that we are summing over disjoint intervals, i.e. that $j \le k$ in the above sum. Upon rearranging, the probability in question becomes

$$g\Bigg(|[i+1,j]| + |[k+1,l]|, 3\Bigg(\sum_{v \in [k+1,l]} v - \sum_{u \in [i+1,j]} u\Bigg)\Bigg),$$

where $g(m, \cdot)$ is the probability mass function of the centred symmetric binomial distribution with parameter $m$. This can now be analysed by grouping the terms according to the lengths of the intervals and the difference of their sums. To accomplish this, one needs the following simple lemma:

### Modular anticoncentration for binomial distribution

Let $n, q$ be positive integers and let $X$ be a symmetric binomial random variable with parameter $n$. Then $\mathbb{P}(X \equiv 0 \pmod q) \le \frac{1}{q} + \frac{2}{\sqrt{n}}$.

The proof of the deterministic version is similar. Fixing the lengths of the intervals to be $s, t$, we have to estimate the number of pairs of left endpoints $(i, j)$ such that

$$2si - 2tj = t^2 - s^2 + \left\lfloor \frac{t}{b} \right\rfloor - \left\lfloor \frac{s}{b} \right\rfloor + \delta$$

for some $\delta \in \{-1, 0, 1\}$. If $q = \gcd(s, t)$ doesn't divide the right-hand side, there are no such $(i, j)$. Otherwise, there are at most

$$O\Big(\frac{n}{\max(s', t')}\Big) = O\Big(\frac{qn}{\max(s, t)}\Big) \qquad (*)$$

many solutions, where $s' = \frac{s}{q}$, $t' = \frac{t}{q}$. The divisibility condition implies

$$s' - t' \in \Big(-\frac{2b}{q}, \frac{2b}{q}\Big) + b\mathbb{Z}.$$

If $b$ is moderately large, this cuts down the number of admissible pairs $(s, t)$ by roughly a factor of $q$. This cancels out the factor of $q$ coming from $(*)$ and is enough to get the required $O(n^2)$ bound.

## An upper bound

In [1], we also establish the following upper bound on the size of $S(a)$:

### A non-trivial upper bound on $|S(a)|$

Let $n$ be a positive integer and let $1 \le a_1 < \ldots < a_k \le n$ be integers. Then

$$|S(a)| \le (c_3 + o(1))n^2,$$

where $c_3 = \frac{e^2-1}{2(e^2+1)} \approx 0.381$.

To prove this result, we use an idea of Konieczny [7]. Namely, we split $S(a)$ into the elements that are at most $\alpha n^2$ and those that are greater than $\alpha n^2$, for some constant parameter $\alpha$. The number of small sums is trivially at most $\alpha n^2$, whereas the number of large sums is bounded above by the number of pairs $(i, j)$ such that $1 \le i \le j \le k$ and

$$\sum_{u=i}^{j} u \ge \alpha n^2.$$

By a volume argument, the latter quantity can be related to the area of the region of the unit square $[0, 1]^2$ lying above the hyperbola $y^2 - x^2 = \alpha$. Optimising over $\alpha$ gives the desired bound.

## References

[1] A. Beker, *On a problem of Erdős and Graham about consecutive sums in strictly increasing sequences*, Bulletin of the London Mathematical Society, to appear.

[2] T. Bloom, *Erdős problems*, `https://www.erdosproblems.com/`.

[3] P. Erdős and R. Graham, *Old and new problems and results in combinatorial number theory*, Monographies de L'Enseignement Mathematique (1980).

[4] K. Ford, *The distribution of integers with a divisor in a given interval*, Ann. of Math. (2008).

[5] K. Ford, *Rough integers with a divisor in a given interval*, Journal of the Australian Mathematical Society (2021).

[6] N. Hegyvári, *On consecutive sums in sequences*, Acta Math. Hungar. (1986).

[7] J. Konieczny, *On consecutive sums in permutations*, Journal of Combinatorics (2021).

[8] T. Tao and V. Vu, *Additive Combinatorics*, volume 105 of Cambridge Studies in Advanced Mathematics, Cambridge University Press (2006).

Supported by HRZZ IP-2022-10-5116



adrian.beker@math.hr