Journal of Number Theory **83**, 194–201 (1999)
doi:10.1006/jnth.1999.2386, available online at http://www.idealibrary.com on IDEAL®

Convexity and Sumsets

György Elekes and Melvyn B. Nathanson

Eotvos Lorand University, Budapest, Hungary

and

Imre Z. Ruzsa

A. Rényi Math. Inst. of the Hungarian Academy of Sciences, Budapest, Hungary

Communicated by V. T. Sos

Received March 17, 1998; published online June 1, 2000

1. INTRODUCTION

Erdős and Szemerédi [3] proved that, for any set A of n real numbers either the set of sums

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

or the set of products

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

has at least cn^{α} elements, with positive constants c and $\alpha>1$. They did not specify the value of α. Nathanson [6] gave the value $\alpha=32/31$, and Elekes [2] improved it to 5/4.

Hegyvári [4] proved, confirming a conjecture of Erdős, that any set $B=\{b_1, ..., b_n\}$ of n integers such that $b_i<b_{i+1}$ and $b_i-b_{i-1}<b_{i+1}-b_i$ for all i satisfies

$$|B-B| \geqslant cn \frac{\log n}{\log \log n}.$$

The third author (unpublished) improved this estimate to $cn^{4/3}$, and as a common generalization of these problems formulated the following conjecture. Let $A \subset \mathbb{R}$ be a finite set, $|A|=n$, and let f be a strictly convex (or concave) function, defined on an interval containing A. Write

$$f(A)=\{f(a) : a \in A\}.$$

Then we have

$$\max\{|A+A|, |f(A)+f(A)|\} \geqslant cn^{\alpha}$$

with constants $c>0$ and $\alpha>1$.

0022-314X/00 $35.00

The result of Erdős and Szemerédi would follow by considering $f(x) = \log x$, and that of Hegyvári by placing the points (i, b_i) on a suitable convex curve. (Details will be given in Section 3.)

In this note we establish a general result which involves these as special cases and has further applications, among others to a problem of Erdős and Szemerédi from [3] about the number of the sums

$$a_1 + 1/a_2, \qquad a_1, a_2 \in A.$$

(See Corollaries 3.7 and 3.8.) The method is an extension of the one used by the first author [2].

2. THE MAIN RESULT

Let $A \subset \mathbb{R}$ be a finite set, $|A| = n$, and let f be a strictly convex or concave function, defined on an interval containing A. Write

$$f(A) = \{ f(a) : a \in A \}, \tag{2.1}$$

$$S = \{ (a, f(a)) : a \in A \} \subset \mathbb{R}^2. \tag{2.2}$$

THEOREM 1. *For every finite set* $T \subset \mathbb{R}^2$ *we have*

$$|S + T| \geqslant c \min(n\,|T|, n^{3/2}\,|T|^{1/2}) \tag{2.3}$$

with a certain absolute constant $0 < c < 1$.

For the proof we need some preparation.

DEFINITION 2.1. A system Γ of continuous plane curves (homeomorphic images of an interval) is a *pseudo-line system* if any two members of Γ share at most one point in common.

LEMMA 2.2. *Let n and N be positive integers. Let, moreover, P be a set of N distinct points in the plane and $e_1, e_2, ..., e_m$ some (also distinct) pseudo lines. If each of the e_i contains n or more points of P, then*

$$N \geqslant c \min(n^{3/2}, nm) \tag{2.4}$$

with a positive absolute constant c.

This is an immediate consequence of a result of Clarkson *et al.* [1] (see also [7]), which asserts that the number of incidences, which is in our case at least mn, is bounded from above by a constant multiple of $N + m + (Nm)^{3/2}$. For straight lines it was proved by Szemcrédi and Trotter [10]. (See also Pach and Sharir [8] for generalizations.) Recently, a simple proof was found by Székely [9].

Proof of Theorem 1. Let $I \supset A$ be an interval on which f is defined. We have

$$S \subset J = \{(x, f(x)) : x \in I\}.$$

Our system Γ of pseudo-lines will consist of the translated copies $J + t$ of this curve. The fact that each two intersect in at most one point is a direct consequence of convexity (or concavity, resp.). The number of curves is $m = |T|$. On these curves we consider the points

$$(a, f(a)) + t, \qquad a \in A, \quad t \in T.$$

On each curve the number of points is $n = |A|$, consistent with the notations in both the theorem and the lemma. The total number of points is $N = |S + T|$. We apply (2.4) to get

$$N = |S + T| \geqslant c \min(n^{3/2} m^{1/2}, nm)$$

as claimed. ∎

Remark 2.3. For $|T| \leqslant n$ the bound is $cn\,|T|$, and $n\,|T|$ is obviously an upper bound, thus we have the exact order of magnitude. The lower bound can also easily be obtained directly by observing that the $|T|$ sets of the form $S + t$, $t \in T$, each have n elements and the intersection of any two has at most one element.

For $n \leqslant |T| \leqslant n^3$ the bound is $cn^{3/2}\,|T|^{1/2}$, and in this generality this is also the best possible. Namely, let $A = \{1, 2, ..., n\}$ and $f(x) = x^2$. Now consider the set

$$T = \{(i, j) : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant nk\}. \tag{2.5}$$

We have $|T| = nk^2$, and an easy calculation yields (we assume $k \leqslant n$) that $|S + T| \ll kn^2$.

For $|T| > n^3$ the bound becomes worse than $|T|$. In this range nothing essentially better than $|T|$ can be claimed. This is shown by the previous example for $k > n$.

3. APPLICATIONS

The most interesting applications arise by substituting direct products into T. Put $T = C \times D$, and for simplicity assume that $|T| = |C|\,|D| \geqslant n$. Since $S \subset A \times f(A)$, we have

$$S + T \subset (A + C) \times (f(A) + D). \tag{3.1}$$

Thus from Theorem 1 we obtain the following.

COROLLARY 3.1. *Assume that $C, D \subset \mathbb{R}$ and that $|C|\,|D| \geqslant n$. We have*

$$|A + C|\,|f(A) + D| \geqslant cn^{3/2}(|C|\,|D|)^{1/2}. \tag{3.2}$$

The case $C = A$, $D = f(A)$ of (3.2) gives the following.

COROLLARY 3.2. *We have*

$$|A + A|\,|f(A) + f(A)| \geqslant cn^{5/2}. \tag{3.3}$$

In particular, either $A + A$ or $f(A) + f(A)$ has at least $cn^{5/4}$ elements.

We can use this to improve Hegyvári's result mentioned in the Introduction.

COROLLARY 3.3. *Let $B = \{b_1, ..., b_n\}$ be a set of n real numbers such that $b_i < b_{i+1}$ and $b_i - b_{i-1} < b_{i+1} - b_i$ for all i. We have*

$$|B + D| \geqslant c_1 n^{3/2},$$

for every set D with $|D| = n$, in particular

$$|B + B| \geqslant c_1 n^{3/2}$$

and

$$|B - B| \geqslant c_1 n^{3/2}.$$

Proof. This condition implies that we can find a convex function f such that $f(i) = b_i$ for $i = 1, ..., n$. We apply (3.2) to this function and the sets $A = C = \{1, 2, ..., n\}$, where we have $f(A) = B$. Since $|A + A| = 2n - 1$, we get $|B + D| \geqslant (c/2)\, n^{3/2}$. ∎

The case $C = f(A)$, $D = A$ of (3.2) gives the following.

COROLLARY 3.4. *For every set A and convex function f we have*

$$|A + f(A)| \geqslant cn^{5/4}. \tag{3.4}$$

We introduce the multiplicative analog of our notations

$$AB = \{ab : a \in A, b \in B\}.$$

By considering $f(x) = \log x$, after obvious transformations from Corollary 3.3 we obtain the following.

COROLLARY 3.5. *Let A, C, D be sets of positive reals. We have*

$$|A + C|\ |AD| \geqslant n^{3/2}(|C|\ |D|)^{1/2}. \tag{3.5}$$

With $C = D = A$ this reduces to the following result from [2].

COROLLARY 3.6. *Let A be a set of n positive real numbers. We have*

$$|A + A|\ |AA| \geqslant cn^{5/2}; \tag{3.6}$$

in particular, either $A + A$ or AA has at least $n^{5/4}$ elements.

By considering $f(x) = 1/x$, with the notation

$$1/A = \{1/a : a \in A\}, \tag{3.7}$$

Corollaries 3.2 and 3.4 yield the following results.

COROLLARY 3.7. *Let A be a set of n positive real numbers. We have*

$$|A + A|\ |1/A + 1/A| \geqslant cn^{5/2}; \tag{3.8}$$

in particular, either $A + A$ or $1/A + 1/A$ has at least $n^{5/4}$ elements.

COROLLARY 3.8. *We have always*

$$|A + 1/A| \geqslant n^{5/4}. \tag{3.9}$$

We conjecture that the exponents in (3.6) and (3.8) should have the value $3 - \varepsilon$. We think that the proper exponent in (3.9) is $2 - \varepsilon$. We have an example of a set where

$$|A + 1/A| \ll n^{2 - c/\log\log n},$$

and for every $\delta > 0$ we have an example where

$$|A + A| \ll n^{1+\delta}, \qquad |AA| \ll n^{2 - c(\delta)/\log\log n}.$$

4. MORE THAN TWO SUMMANDS

We write $Sk = S + S + \cdots + S$, for a repeated addition of k identical sets.

THEOREM 2. *Let S, T be as in Theorem 1. We have for every integer $k \geqslant 1$*

$$|Sk + T| \geqslant c' n^{3(1-2^{-k})} |T|^{2^{-k}} \tag{4.1}$$

if $|T| \geqslant n$, and

$$|Sk + T| \geqslant c' n^{3-2^{-(k-2)}} |T|^{2^{-k-1}} \tag{4.2}$$

if $|T| < n$, with a certain absolute constant $0 < c' < 1$.

Proof. This follows by a repeated application of Theorem 1 to the sets T, $S+T$, ..., $S(k-1)+T$. For the constant we obtain the value

$$c^{2-2^{1-k}}, \tag{4.3}$$

Since the exponent does not exceed 2, we may put $c' = c^2$. ▮

Remark 4.1. Unlike Theorem 1, Theorem 2 is probably not the best possible for $k \geqslant 2$.

By putting $T = S$ into Theorem 2 we obtain the following.

COROLLARY 4.2. *For every integer $k \geqslant 1$ we have*

$$|Sk| \geqslant cn^{3-2^{-(k-2)}}. \tag{4.4}$$

Remark 4.3. For the set $S = \{(i, i^2) : 1 \leqslant i \leqslant n\}$ we have

$$|S3| \asymp \frac{n^3}{\sqrt{\log n}}, \qquad |S4| \asymp n^3. \tag{4.5}$$

We conjecture that this example is nearly extremal, that is, we have

$$|S3| \gg n^{3-\varepsilon}, \qquad |S4| \gg n^3 \tag{4.6}$$

for every set S. This would also improve Theorem 2 for $k \geqslant 3$.

Here the second claim follows from Cauchy's lemma (see e.g. Nathanson [7, p. 31]). This yields that $(x, y) \in S4$ if x, y are odd, $1 \leqslant x \leqslant n$, and $x^2/4 \leqslant y \leqslant x^2/3$.

To get the first claim, consider the numbers

$$a = i + j + k, \qquad b = i^2 + j^2 + k^2.$$

These satisfy

$$d = 12b - 4a^2 = (2i - j - k)^2 + 3(j - k)^2.$$

This shows that d can contain only primes $\equiv 1 \pmod 6$ with an odd exponent. The number of admissible values of d up to x is $\sim c(x/\sqrt{\log x})$ and we obtain the upper estimate by grouping the pairs (a, b) according to the value of d. Conversely, let d be a number, divisible by 4 and free of primes $\equiv -1 \pmod 6$. T has a representation in the form $s^2 + 3t^2$, and we must have $s \equiv t \pmod 2$. For an arbitrary value of k the integers

$$j = t + k, \qquad i = \frac{t + s}{2} + k$$

yield a pair (a, b) with this value of d, and we have

$$a = i + j + k = \frac{3t + s}{2} + 3k.$$

The last equation shows that different values of k yield different values of a. Obviously, if we start with $d < (n/2)^2$, take nonnegative values of s, t and $0 < k < n/2$, then $i, j, k < n$ and we find a point in our set $S3$.

Put $T = C \times D$, and assume that $|T| = |C|\,|D| \geqslant n$. Since $S \subset A \times B$, we have

$$Sk + T \subset (Ak + C) \times (f(A)\,k + D). \tag{4.7}$$

Thus from Theorem 2 we obtain the following.

COROLLARY 4.4. *Assume that* $C, D \subset \mathbb{R}$, *and* $|C|\,|D| \geqslant n$. *For every integer* $k \geqslant 1$ *we have*

$$|Ak + C|\,|f(A)\,k + D| \geqslant c' n^{3(1 - 2^{-k})}(|C|\,|D|)^{2^{-k}}. \tag{4.8}$$

The case $C = A$, $D = f(A)$ of (4.8) gives the following.

COROLLARY 4.5. *For every integer* $k \geqslant 2$ *we have*

$$|Ak|\,|f(A)\,k| \geqslant cn^{3 - 2^{1-k}}. \tag{4.9}$$

We introduce the multiplicative analog of our notation

$$A^k = AA \cdots A, \qquad k \text{ factors}.$$

By considering $f(x) = \log x$, from Corollary 4.5 we obtain the following:

COROLLARY 4.6. *For every integer $k \geqslant 2$ we have*

$$|Ak|\,|A^k| \geqslant cn^{3-2^{1-k}}. \tag{4.10}$$

We conjecture that the exponent should go to infinity here rather than to 3. However, to obtain such a result we would need to find a more specific approach, since, as the previous examples indicate, the corresponding result with general convex function cannot yield exponents exceeding 3.

It would be interesting to find analogous results for different summands. We can iterate Theorem 1 with different sets, but because of its alternative nature a branching can occur at each place and the results are complicated and not satisfactory.

REFERENCES

1. K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl, Combinatorial complexity bounds for arrangements of curves and surfaces, *Discrete Comput. Geom.* **5** (1990), 99–106.
2. Gy. Elekes, On the number of sums and products, *Acta Arith.* **81** (1997), 365–367.
3. P. Erdős and E. Szemerédi, On sums and products of integers, *in* "Studies in Pure Math. to the Memory of P. Turán," pp. 213–218, Akadémiai Kiadó, Budapest, 1983.
4. N. Hegyvári, On consecutive sums in sequences, *Acta Math. Acad. Sci. Hungar.* **48** (1986), 193–200.
5. M. B. Nathanson, "Additive Number Theory: The Classical Bases," Graduate Texts in Mathematics, Vol. 164, Springer-Verlag, New York/Berlin, 1996.
6. M. B. Nathanson, On sums and products of integers, *Proc. Amer. Math. Soc.* **125** (1997), 9–16.
7. J. Pach and P. K. Agarwal, "Combinatorial Geometry," Wiley, New York/Chichester, 1995.
8. J. Pach and M. Sharir, Repeated angles in the plane and related problems, *J. Combin. Theory Ser. A* **59** (1990), 12–22.
9. L. A. Székely, Crossing numbers and hard Erdős problems in discrete geometry, *Combin. Probab. Comput.* **11** (1993), 1–10.
10. E. Szemerédi and W. T. Trotter, Jr., Extremal problems in discrete geometry, *Combinatorica* **3** (1983), 381–392.