

Sumsets without powerful numbers

by

ARTŪRAS DUBICKAS and ANDRIUS STANKEVIČIUS (Vilnius)

1. Introduction. Let $A = \{a_1 < a_2 < \dots\}$ be an infinite sequence of positive integers. There are many interesting and important problems concerning various arithmetical properties of sumsets $A + A := \{a + a' \mid a, a' \in A\}$ (see, e.g., a recent book [18]). An old problem of Erdős and Silverman is to determine the maximal density of a set A such that the sumset $A + A$ contains no squares. It is conjectured that the maximal density is $11/32$. Massias produced a corresponding set A as a union of eleven infinite arithmetic progressions with difference 32 each. It remains to prove that every set whose upper density is strictly greater than $11/32$ contains two (not necessarily distinct) elements which sum to a perfect square. See the papers of Lagarias, Odlyzko and Shearer [9], [10] and Schoen [17] for some progress concerning this problem. See also the papers of Alon and Freiman [1], Erdős and Freiman [4], Lipkin [12] and Sárközy [16], where various techniques are used to treat similar problems on sets whose subset sums contain or do not contain squares, powers, etc.

An interesting variation of this problem due to Luca [13] asks whether an infinite set obtained by adding any number of distinct elements of an infinite sequence A can be free of squares and, more generally, free of powers. The condition that the elements of A in the sum are distinct is natural and necessary in order to avoid the trivial situation when taking $a + \dots + a$, say, a times one gets a^2 . Luca proved that such a sequence A exists (see also [2] for a “finite” version of this problem). However, in his example the set A is very sparse. The n th element a_n is, roughly speaking, doubly exponential in n . Another construction which reduced Luca’s bound was given by the first author and Šarka in [3]. However, the bound in [3] is still exponential. More precisely, in [3] it was shown that, for any $\varepsilon > 0$, there is a positive constant $K = K(\varepsilon)$ and a sequence A whose distinct elements do not sum to a power and whose n th element is bounded as $a_n < K \exp(\varepsilon n)$ for each $n \in \mathbb{N}$.

2000 *Mathematics Subject Classification*: 11A41, 11A99, 11B05, 11B99.

Key words and phrases: powerful number, infinite sequence, prime number, sumset.

The aim of this note is to produce some sequences A that are even more dense. This allows us to improve the exponential bound $a_n < K \exp(\varepsilon n)$ to the polynomial bound $a_n < c_0 n^c$. In addition, we strengthen the corresponding arithmetical condition for the resulting sumset $\{a_{i_1} + \dots + a_{i_l} \mid l \in \mathbb{N}, a_{i_j} \in A, a_{i_u} \neq a_{i_v}\}$. This set of all possible sums of distinct a_n 's will be not only "free of powers" but also, more generally, "free of powerful numbers". A number $a > 1$ is called *powerful* if for every prime number p dividing a , p^2 also divides a . These numbers were studied already in 1934 by Erdős and Szekeres [5]. Sometimes powerful numbers (named "powerful" by Golomb in [7]) are called *squareful*, *square-full*, or *2-full*. Golomb [7] proved that the number of powerful numbers $\leq x$ is $\sim b\sqrt{x}$, where $b = \zeta(3/2)/\zeta(3) = 2.173\dots$. See [8] for more open problems concerning powerful numbers.

THEOREM 1. *There is an infinite sequence $A = \{a_1 < a_2 < \dots\}$ of positive integers satisfying $a_n \leq 2n^9$ for each $n \in \mathbb{N}$ such that the sum of any number of distinct elements of A is not a powerful number.*

For every $n \geq 2$, the factor 2 can be replaced by a smaller constant. However, in general, we do not know how to improve the exponent 9. In our next theorem, we slightly improve the bound by replacing $2n^9$ by $n^9 \exp(-f(n))$ for each n large enough, where $f(n)$ tends to infinity a little more slowly than $\log n$.

THEOREM 2. *There is an infinite sequence $A = \{a_1 < a_2 < \dots\}$ of positive integers satisfying*

$$a_n < n^9 \exp\left(-\frac{(\log 2)^2 \log n}{(\log \log n)^2}\right)$$

for each sufficiently large n such that the sum of any number of distinct elements of A is not a powerful number.

It seems likely that this statement "squeezes" almost everything from our method, so any further improvement seems to be of interest. (In fact, we do not give a formal proof that our choice of parameters in Section 3 is optimal.)

In our final theorem we improve the exponent 9 in Theorem 1 and replace it by $3 + \varepsilon$, but only for some infinite subsequence of \mathbb{N} rather than for all sufficiently large n .

THEOREM 3. *For any sequence of real numbers g_n , $n = 1, 2, \dots$, with $g_n \rightarrow \infty$ as $n \rightarrow \infty$, there is an infinite sequence $A = \{a_1 < a_2 < \dots\}$ of positive integers satisfying $a_n < g_n n^3$ for infinitely many $n \in \mathbb{N}$ such that the sum of any number of distinct elements of A is not a powerful number.*

Note that in Theorem 3 the sequence $g_n, n = 1, 2, \dots$, may tend to infinity at any prescribed rate.

In contrast to Theorem 3, a result of Lev (see Theorem 5' in [11]) states that if a sequence $A = \{a_1 < a_2 < \dots\}$ of positive integers satisfies $a_n < n^2/(50 \log n)$ for all sufficiently large $n \in \mathbb{N}$ then the sums of distinct elements of A contain an infinite arithmetic progression of the form $d, 2d, 3d, \dots$. This progression obviously contains infinitely many squares, so $g_n n^3$ in Theorem 3 cannot be replaced by $n^2/(50 \log n)$. See also [6] and [14] for some previous work on this problem.

In the next section, we present a general construction of sets whose subset sums avoid powerful numbers. An easy application of this construction leads to Theorem 1 immediately. In Sections 3 and 4, by a more elaborate choice of parameters, we prove Theorems 2 and 3, respectively.

2. Construction of A and the proof of Theorem 1. Let $p_1 < p_2 < \dots$ be an arbitrary sequence of prime numbers. Put $p_0 := 1$, and

$$A_k := \left\{ (jp_k^2 + p_k) \prod_{i=0}^{k-1} p_i^2 \mid j = 0, 1, \dots, p_k - 2 \right\}$$

for $k \geq 1$. The set A_k contains $p_k - 1$ distinct elements. Note that the smallest element of A_k is greater than the largest element of A_{k-1} , because $p_k p_{k-1}^2 > (p_{k-1} - 2)p_{k-1}^2 + p_{k-1}$. It follows that A_k and A_l are disjoint if $k \neq l$. Put

$$A := \bigcup_{k=1}^{\infty} A_k = \{a_1 < a_2 < \dots\}.$$

We claim that the sum of any number of distinct elements of A cannot be a powerful number. Indeed, by construction, the sum of any number of distinct elements of A_k is divisible by p_k , but not by p_k^2 . Furthermore, every element of A_l , where $l > k$, is divisible by p_k^2 . Suppose that S is a sum of several distinct elements of A . Let k be the least positive integer such that at least one element of A_k is one of the summands of S . Then, by the above, S is divisible by p_k , but not by p_k^2 . So S is not a powerful number.

Next, for every $n \in \mathbb{N}$, there is a unique positive integer k satisfying

$$p_{k-1} \leq n < p_k.$$

Clearly, the n th element of A , a_n , does not exceed the n th element of A_k . It follows that

$$a_n \leq ((n - 1)p_k^2 + p_k) \prod_{i=0}^{k-1} p_i^2 = (np_k^2 - p_k^2 + p_k) \prod_{i=0}^{k-1} p_i^2 < n(p_k p_{k-1} \dots p_1)^2.$$

In order to prove Theorems 1 and 2, we shall first choose certain sequences of prime numbers $p_1 < p_2 < \dots$ and then derive upper bounds for the corresponding products $p_k p_{k-1} \dots p_1$ in terms of n .

Proof of Theorem 1. Set $p_0 := 1, p_1 := 2, p_2 := 5$. For any real $x \geq 9$, the interval $(x, 4x/3)$ contains a prime number (see [15]). So, for each $k \geq 3$, let p_k be the least prime number satisfying

$$p_1 \dots p_{k-1} < p_k < 1.4p_1 \dots p_{k-1}.$$

We have $p_3 = 11$ and $a_1 = 2, a_2 = 20, a_3 = 120, \dots, a_{10} = (4 \cdot 11^2 + 11) \cdot 5^2 \cdot 2^2 = 49500$. One can check easily that $a_n \leq 2n^9$ for each $n \leq 10$, with equality for $n = 1$ only.

Suppose that $p_{k-1} \leq n < p_k$, where $k \geq 3$. On applying the above upper bound for p_k and then the above lower bound for p_{k-1} (the lower bound $p_k > p_{k-1} \dots p_1$ holds for every $k \in \mathbb{N}$), we obtain

$$p_k p_{k-1} \dots p_1 < 1.4p_{k-1}^2 (p_{k-2} \dots p_1)^2 < 1.4p_{k-1}^2 p_{k-1}^2 = 1.4p_{k-1}^4.$$

Using $p_{k-1} \leq n$, we deduce that $p_k p_{k-1} \dots p_1 < 1.4n^4$. Hence

$$a_n < n(p_k p_{k-1} \dots p_1)^2 < n(1.4n^4)^2 < 2n^9$$

for each $n \geq 11$.

3. Proof of Theorem 2. This time, by Bertrand’s postulate, for each $k \in \mathbb{N}$ and each real $\tau_k > 1$, there is a prime number p_k satisfying

$$\tau_k p_0 p_1 \dots p_{k-1} < p_k < 2\tau_k p_0 p_1 \dots p_{k-1}.$$

Fix any such p_k . Recall that $p_0 = 1$. (The τ_k will be chosen later.) Now,

$$\begin{aligned} p_k p_{k-1} \dots p_1 &< 2\tau_k p_{k-1}^2 (p_{k-2} \dots p_1)^2 < 2\tau_k p_{k-1}^2 (p_{k-1}/\tau_{k-1})^2 \\ &= 2\tau_k \tau_{k-1}^{-2} p_{k-1}^4. \end{aligned}$$

Using $p_{k-1} \leq n$, as above, we deduce that

$$a_n < n(p_k p_{k-1} \dots p_1)^2 < n(2\tau_k \tau_{k-1}^{-2} n^4)^2 = 4\tau_k^2 \tau_{k-1}^{-4} n^9.$$

It remains to give an upper bound for the factor $4\tau_k^2 \tau_{k-1}^{-4}$ in terms of n . Using $p_j > \tau_j p_{j-1} \dots p_1$ for $j = k - 1, k - 2, \dots, 1$, we obtain

$$\begin{aligned} n &\geq p_{k-1} > \tau_{k-1} p_{k-2} \dots p_1 > \tau_{k-1} \tau_{k-2} (p_{k-3} \dots p_1)^2 \\ &> \tau_{k-1} \tau_{k-2} \tau_{k-3}^2 (p_{k-4} \dots p_1)^4 > \tau_{k-1} \tau_{k-2} \tau_{k-3}^2 \tau_{k-4}^4 (p_{k-5} \dots p_1)^8 > \dots \\ &> \tau_{k-1} \tau_{k-2} \tau_{k-3}^2 \tau_{k-4}^4 \dots \tau_1^{2^{k-3}}. \end{aligned}$$

Similarly, using $p_j < 2\tau_j p_{j-1} \dots p_1$ for $j = k, k - 1, \dots, 1$, we obtain

$$\begin{aligned} n < p_k < 2\tau_k p_{k-1} \dots p_1 < 4\tau_k \tau_{k-1} (p_{k-2} \dots p_1)^2 < 8\tau_k \tau_{k-1} \tau_{k-2}^2 (p_{k-3} \dots p_1)^4 \\ < 16\tau_k \tau_{k-1} \tau_{k-2}^2 \tau_{k-3}^4 (p_{k-4} \dots p_1)^8 < \dots < 2^{2^{k-1}} \tau_k \tau_{k-1} \tau_{k-2}^2 \tau_{k-3}^4 \dots \tau_1^{2^{k-2}}. \end{aligned}$$

Therefore, putting

$$\mu_k := \tau_1^{2^{k-2}} \tau_2^{2^{k-3}} \dots \tau_{k-2}^2 \tau_{k-1},$$

one finds that

$$\tau_{k-1} \mu_{k-1} < n < 2^{2^{k-1}} \tau_k \mu_k.$$

Select

$$\tau_k := \exp(2^k / \sqrt{k+1})$$

for $k \in \mathbb{N}$. Since $\tau_j^{2^{k-j-1}} = \exp(2^{k-1} / \sqrt{j+1})$ for $1 \leq j \leq k-1$, we find that

$$\mu_k = \tau_1^{2^{k-2}} \tau_2^{2^{k-3}} \dots \tau_{k-2}^2 \tau_{k-1} = \exp\left(2^{k-1} \sum_{j=2}^k \frac{1}{\sqrt{j}}\right)$$

for $k \geq 2$. Below, we shall use the following standard estimate:

$$2(\sqrt{k+1} - \sqrt{2}) = \int_2^{k+1} \frac{dx}{\sqrt{x}} < \sum_{j=2}^k \frac{1}{\sqrt{j}} < \int_1^k \frac{dx}{\sqrt{x}} = 2(\sqrt{k} - 1).$$

Using the upper bound of this estimate, we get $\log \mu_k < 2^k(\sqrt{k} - 1)$. This implies that

$$\begin{aligned} n &< 2^{2^{k-1}} \tau_k \mu_k = \exp(2^{k-1} \log 2 + 2^k / \sqrt{k+1} + \log \mu_k) \\ &< \exp(2^k(0.35 + (k+1)^{-1/2} + \sqrt{k} - 1)) < \exp(2^k \sqrt{k}) \end{aligned}$$

for $k \geq 2$. Similarly, using $\sum_{j=2}^k 1/\sqrt{j} > 2(\sqrt{k+1} - \sqrt{2})$, we get $\log \mu_k > 2^k(\sqrt{k+1} - \sqrt{2})$, so $\log \mu_{k-1} > 2^{k-1}(\sqrt{k} - \sqrt{2})$. It follows that

$$n > \tau_{k-1} \mu_{k-1} > \mu_{k-1} > \exp(2^{k-1}(\sqrt{k} - \sqrt{2})) > \exp(2^{k-2} \sqrt{k-1})$$

for $k \geq 7$. Summarizing, we deduce that

$$2^{k-2} \sqrt{k-1} < \log n < 2^k \sqrt{k}$$

for each $k \geq 7$.

Next, by the choice of τ_k , we have

$$\tau_k^2 \tau_{k-1}^{-4} = \exp(2^{k+1}((k+1)^{-1/2} - k^{-1/2})).$$

The inequality

$$\begin{aligned} k^{-1/2} - (k+1)^{-1/2} &= \frac{\sqrt{k+1} - \sqrt{k}}{\sqrt{k(k+1)}} = \frac{1}{\sqrt{k(k+1)}(\sqrt{k+1} + \sqrt{k})} \\ &> \frac{1}{2(k+1)^{3/2}} \end{aligned}$$

implies that $\tau_k^2 \tau_{k-1}^{-4} < \exp(-2^k(k+1)^{-3/2})$. Thus

$$4\tau_k^2 \tau_{k-1}^{-4} < 4 \exp(-2^k(k+1)^{-3/2}) = \exp(\log 4 - 2^k(k+1)^{-3/2}).$$

Recall that $a_n < 4\tau_k^2\tau_{k-1}^{-4}n^9$. Hence, in order to prove the required inequality

$$a_n < n^9 \exp\left(-\frac{(\log 2)^2 \log n}{(\log \log n)^2}\right),$$

it remains to verify that

$$2^k(k+1)^{-3/2} - \log 4 > (\log 2)^2 \log n / (\log \log n)^2.$$

For this, we shall use the estimates $2^{k-2}\sqrt{k-1} < \log n < 2^k\sqrt{k}$ for $\log n$. Firstly,

$$2^k(k+1)^{-3/2} = \frac{2^k\sqrt{k}}{\sqrt{(k+1)^3k}} > \frac{\log n}{\sqrt{(k+1)^3k}} > \frac{\log n}{(k+1)^2},$$

because $2^k\sqrt{k} > \log n$. Secondly, for $k \geq 257$, the inequality $2^{k+2} \leq 2^{k-2}\sqrt{k-1} < \log n$ implies that $k+1 < (\log \log n)/\log 2 - 1$. Hence

$$2^k(k+1)^{-3/2} - \log 4 > \frac{\log n}{(k+1)^2} - \log 4 > \frac{\log n}{((\log \log n)/\log 2 - 1)^2} - \log 4.$$

This is greater than $(\log 2)^2 \log n / (\log \log n)^2$ for each sufficiently large n .

4. Proof of Theorem 3. Choose any sequence of prime numbers p_k , $k = 1, 2, \dots$, satisfying $p_k > p_{k-1}$ and

$$(p_{k-1}p_{k-2} \dots p_1)^2 < g_{p_k}.$$

Let A be the set defined in Section 2. We shall prove that $a_n < g_n n^3$ for every n of the form $n = p_k$, where $k \geq 2$. Put

$$m := (p_1 - 1) + \dots + (p_k - 1) = p_1 + \dots + p_k - k.$$

The inequality $p_1 + \dots + p_{k-1} \geq 2(k-1) \geq k$ implies that $n = p_k \leq m$. By the choice of m , the m th element of A is equal to the largest element of A_k . Using the condition imposed on the growth of p_k 's, we find that

$$a_n \leq a_m = ((p_k - 2)p_k^2 + p_k) \prod_{i=1}^{k-1} p_i^2 < p_k^3 \prod_{i=1}^{k-1} p_i^2 < p_k^3 g_{p_k} = g_n n^3.$$

Acknowledgements. We thank the referee of this paper for pointing out several useful references. This research was supported in part by the Lithuanian State Studies and Science Foundation.

References

[1] N. Alon and G. Freiman, *On sums of subsets of a set of integers*, *Combinatorica* 8 (1988), 297–306.
 [2] J. Cilleruelo, *Solution of Problem 38*, *Gaceta de la Real Sociedad Matemática Española* 9 (2006), 455–460.

- [3] A. Dubickas and P. Šarka, *Infinite sets of integers whose distinct elements do not sum to a power*, J. Integer Sequences 9 (2006), Article 06.4.8, 9 pp.
- [4] P. Erdős and G. Freiman, *On two additive problems*, J. Number Theory 34 (1990), 1–12.
- [5] P. Erdős und G. Szekeres, *Über die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisches Problem*, Acta Litt. Sci. Szeged 7 (1934), 95–102.
- [6] J. Folkman, *On the representation of integers as sums of distinct terms from a fixed sequence*, Canad. Math. J. 18 (1966), 643–655.
- [7] S. W. Golomb, *Powerful numbers*, Amer. Math. Monthly 77 (1970), 848–852.
- [8] R. K. Guy, *Unsolved Problems in Number Theory*, 3rd ed., Springer, 2004.
- [9] J. C. Lagarias, A. M. Odlyzko and J. B. Shearer, *On the density of sequences of integers the sum of no two of which is a square. I: Arithmetic progressions*, J. Combin. Theory Ser. A 33 (1982), 167–187.
- [10] —, —, —, *On the density of sequences of integers the sum of no two of which is a square. II: General sequences*, *ibid.* 35 (1983), 123–139.
- [11] V. F. Lev, *Blocks and progressions in subset sum sets*, Acta Arith. 106 (2003), 123–142.
- [12] E. Lipkin, *On representation of r th powers by subset sums*, *ibid.* 52 (1989), 353–365.
- [13] F. Luca, *Infinite sets of positive integers whose sums are free of powers*, Rev. Colombiana Mat. 36 (2002), 67–70.
- [14] T. Łuczak and T. Schoen, *On the maximal density of sum-free sets*, Acta Arith. 95 (2000), 225–229.
- [15] J. Nagura, *On the interval containing at least one prime number*, Proc. Japan Acad. 28 (1952), 177–181.
- [16] A. Sárközy, *Finite addition theorems. II*, J. Number Theory 48 (1994), 197–218.
- [17] T. Schoen, *On sets of natural numbers whose sumset is free of squares*, J. Combin. Theory Ser. A 88 (1999), 385–388.
- [18] T. Tao and V. H. Vu, *Additive Combinatorics*, Cambridge Univ. Press, 2006.

Department of Mathematics and Informatics
Vilnius University
Naugarduko 24, Vilnius LT-03225, Lithuania
E-mail: arturas.dubickas@mif.vu.lt
andrius.stankevicius@mif.vu.lt

Received on 7.5.2007
and in revised form on 3.9.2007

(5440)