Product representation of perfect cubes*

Zsigmond György Fleiner[†1], Márk Hunor Juhász[‡1], Blanka Kövér[§1], Péter Pál Pach[¶2,3], and Csaba Sándor[‖2,3,4]

[1]ELTE Eötvös Loránd University Faculty of Science, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary
[2]Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[3]MTA-BME Lendület Arithmetic Combinatorics Research Group, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[4]Department of Stochastics, Institute of Mathemetics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary

Abstract

Let $F_{k,d}(n)$ be the maximal size of a set $A \subseteq \{1, 2, \dots, n\}$ such that the equation

$$a_1 a_2 \dots a_k = x^d,\ a_1 < a_2 < \dots < a_k$$

has no solution with $a_1, a_2, \dots, a_k \in A$ and integer x. Erdős, Sárközy and T. Sós studied $F_{k,2}$, and gave bounds when $k = 2, 3, 4, 6$ and also in the general case. We study the problem for $d = 3$, and provide bounds for $k = 2, 3, 4$ and 6, furthermore, in the general case as well. In particular, we refute an 18-year-old conjecture of Verstraëte.

We also introduce another function $f_{k,d}$ closely related to $F_{k,d}$: While the original problem requires $a_1, \dots, a_k$ to all be distinct, we can relax this and only require that the multiset of the a_i's cannot be partitioned into d-tuples where each d-tuple consists of d copies of the same number.

1 Introduction

The problem of the solvability of equations of the form

$$a_1 a_2 \dots a_k = x^2,\ a_1 < a_2 < \dots < a_k$$

in a set $A \subseteq [n] = \{1, 2, \dots, n\}$ first appeared in a 1995 paper of Erdős, Sárközy and T. Sós [3]. They investigated the maximal size of a set A such that the equation cannot be solved in A, that is, there are no distinct $a_1, \dots, a_k \in A$ whose product is a perfect square. This motivates the following definitions:

Let $F_{k,d}(n)$ be the maximal size of a set $A \subseteq [n]$ such that

$$a_1 a_2 \dots a_k = x^d,\ a_1 < a_2 < \dots < a_k \tag{1}$$

*The full version of this work can be found in [4] and will be published elsewhere.

†Email: zsgyfleiner@gmail.com
‡Email: markh.shepherd@gmail.com
§Email: koverblanka@gmail.com
¶Email: pach.peter@vik.bme.hu. Research of P. P. P. supported by the Lendület program of the Hungarian Academy of Sciences (MTA) and by the National Research, Development and Innovation Office NKFIH (Grant Nr. K146387)
‖Email: sandor.csaba@ttk.bme.hu. Research of C. S. supported by the Lendület program of the Hungarian Academy of Sciences (MTA) and by the National Research, Development and Innovation Office NKFIH (Grant Nr. K146387)

has no solution with $a_1, a_2, \dots, a_k \in A$ and integer x. Similarly, let $f_{k,d}(n)$ be the maximal size of a set $A \subseteq [n]$ such that

$$a_1 a_2 \dots a_k = x^d \tag{2}$$

has no solution with $a_1, a_2, \dots, a_k \in A$ and integer x, except *trivial* solutions that we specify below. If we allow some of the a_i's in equation (2) to coincide, some trivial solutions do arise: It is clear, for instance, that $a_1 = \dots = a_d$ will yield a solution to the equation $a_1 \dots a_d = x^d$. Let us call a solution trivial if the multiset of the a_i's can be partitioned into d-tuples where each d-tuple consists of d copies of the same number: see for example $(a_1a_1a_1)(a_2a_2a_2)(a_3a_3a_3) = x^3$ for $k = 9$, $d = 3$. Note that trivial solutions arise only if $d \mid k$. Let $f_{k,d}(n)$ be the maximal size of a set $A \subseteq [n]$ such that the equation $a_1a_2 \dots a_k = x^d$ does not have any nontrivial solution with $a_1, a_2, \dots, a_k \in A$. Note that $f_{k,d} \le F_{k,d}$.

With our notation, Erdős, Sárközy and T. Sós [3] proved the following results (and also gave bounds for $F_{k,2}$ for every k):

Theorem 1 (Erdős, Sárközy, T. Sós)**.** *For every $\ell \in \mathbb{Z}^+$, we have*

(i) $F_{2,2}(n) = \left(\frac{6}{\pi^2} + o(1)\right) n$;

(ii) $\frac{n^{3/4}}{(\log n)^{3/2}} \ll F_{4,2}(n) - \pi(n) \ll \frac{n^{3/4}}{(\log n)^{3/2}}$;

(iii) $\frac{n^{2/3}}{(\log n)^{4/3}} \ll F_{6,2}(n) - \left(\pi(n) + \pi\left(\frac{n}{2}\right)\right) \ll n^{7/9} \log n$.

Later Győri [5] and the fourth named author [7] improved the upper bound for $F_{6,2}(n) - \left(\pi(n) + \pi\left(\frac{n}{2}\right)\right)$. The current best upper bound is

$$F_{6,2}(n) - \left(\pi(n) + \pi\left(\frac{n}{2}\right)\right) \ll n^{2/3} (\log n)^{2^{1/3} - 1/3 + o(1)}.$$

For general cases, the current best lower bound estimates have been proved recently by the fourth named author and Vizer [8].

Note that the case $2 \mid k$ is closely related to (generalized) multiplicative Sidon sets, as a solution to the (multiplicative) Sidon equation $a_1 \dots a_k = b_1 \dots b_k$ provides a solution $a_1 \dots a_k b_1 \dots b_k = x^2$. However, the case $2 \nmid k$ seems to be much more difficult. Erdős, Sárközy and T. Sós proved the following results:

Theorem 2 (Erdős, Sárközy, T. Sós)**.** *For every $\ell \in \mathbb{Z}^+$ and $\varepsilon > 0$, we have*

(i) $\frac{n}{(\log n)^{1+\varepsilon}} \ll n - F_{3,2}(n) \le n - f_{3,2}(n) \ll n (\log n)^{\frac{e \log 2}{2} - 1 + \varepsilon}$;

(ii) $\liminf_{n \to \infty} \frac{F_{2\ell+1,2}(n)}{n} \ge \log 2 = 0.69\dots$;

(iii) $\frac{n}{(\log n)^2} \ll n - F_{2\ell+1,2}(n)$.

Note that similar bounds can be proved for the functions $f_{k,2}(n)$.

It remained an interesting problem to find the right shape of the function $F_{2\ell+1,2}$ for $\ell \ge 2$. Very recently, Tao [10] proved that for every $k \ge 4$ there exists some constant $c_k > 0$ such that $F_{k,2}(n) \le (1 - c_k + o(1))n$ as $n \to \infty$.

Based on the work of Erdős, Sárközy, and T. Sós, Verstraëte [11] studied a similar problem: He aimed to find the maximal size of a set $A \subseteq [n]$ such that no product of k distinct elements of A is in the value set of a given polynomial $f \in \mathbb{Z}[x]$. He showed that for a certain class of polynomials the answer is $\Theta(n)$, for another class it is $\Theta(\pi(n))$, and conjectured that these are the only two possibilities:

Conjecture 3. *Let $f \in \mathbb{Z}[x]$ and let k be a positive integer. Then, for some constant $\rho = \rho(k, f)$ depending only on k and f, the maximal size of a set $A \subseteq [n]$ such that no product of k distinct elements of A is in the value set of f is either $(\rho + o(1))n$ or $(\rho + o(1))\pi(n)$ as $n \to \infty$.*

For further related results, see [6, 9].

2 Our results

We investigated the original problem in the case $d = 3$, and provided bounds for both $F_{k,3}$ and $f_{k,3}$. As expected, several additional difficulties arise compared to the case $d = 2$. To overcome these, various new ideas are needed of combinatorial and number theoretic nature. We summarize our results below.

For $k = 2$, the following bounds hold:

Theorem 4. *There exist positive constants c_1 and c_2 such that*

$$c_1 n^{2/3} < n - F_{2,3}(n) \leq n - f_{2,3}(n) < c_2 n^{2/3}.$$

For the case $k = 3$ we prove that $f_{3,3}(n)/n$ converges to a constant $c_{3,3} \in (0, 1)$, which we can approximate (theoretically to arbitrary precision):

Theorem 5. *There exists a constant $0.6224 \leq c_{3,3} \leq 0.6420$ such that*

$$f_{3,3}(n) = (c_{3,3} + o(1))n.$$

(An analogous result holds for $F_{3,3}(n)$, as well.)

In the case $k = 4$ we show that for large n, the following bounds hold. Our proofs generalize and extend ideas from [3] used for the estimation of $F_{3,2}(n)$.

Theorem 6. *Let $\varepsilon > 0$. There exists some $n_0(\varepsilon)$ such that for every $n \geq n_0(\varepsilon)$ we have*

$$\frac{n}{(\log n)^{2+\varepsilon}} < n - F_{4,3}(n) \leq n - f_{4,3}(n) < \frac{n}{(\log n)^{1-\frac{e \log 3}{2\sqrt{3}}-\varepsilon}}.$$

For $k = 6$ we obtained the following results:

Theorem 7. *There exist positive constants c_1 and c_2 such that*

$$c_1 \frac{n^{3/4}}{(\log n)^{3/2}} < f_{6,3}(n) - \pi(n) < c_2 \frac{n^{3/4}}{(\log n)^{3/2}}.$$

Theorem 8. *For $F_{6,3}(n)$ the following holds:*

$$F_{6,3}(n) = (1 + o(1)) \frac{n \log \log n}{\log n}.$$

Note that Theorem 8 refutes Conjecture 3 of Verstraëte [11].

We also give bounds for larger values of k, all the results and proofs are contained in the preprint [4].

3 Proof ideas

The different behaviours of the function $f_{k,3}$ (and $F_{k,3}$) can be illustrated by the cases $k = 2, 3, 4, 6$. Here we give a brief outline of the proof ideas in these cases.

3.1

For proving Theorem 4 we shall notice that $a_1 a_2 = x^3$ holds if and only if the product of the cubefree parts of a_1 and a_2 is a perfect cube. That is, if the cubefree part of a_1 is uv^2 (where uv is squarefree), then in a solution the cubefree part of a_2 has to be $u^2 v$. With the help of this observation one can show the exact result

$$f_{2,3}(n) = n - \sum_{\substack{1 \leq u < v \\ \gcd(u,v)=1 \\ uv^2 \leq n \\ u,v \text{ squarefree}}} \left\lfloor \sqrt[3]{\frac{n}{uv^2}} \right\rfloor,$$

for getting the claimed bound we have to estimate this sum. (Also, note that $F_{2,3}(n) = f_{2,3}(n) + 1$.)

3.2

For getting the bound in Theorem 5 let r be a fixed positive integer and let p_i denote the ith prime. Each cubefree positive integer a can be written as

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} a',$$

where $\alpha_1, \dots, \alpha_r \in \{0,1,2\}$ and a' is cubefree satisfying $\gcd(a', p_1 p_2 \dots p_r) = 1$. Here $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ is the p_r*-smooth* and a' is the p_{r+1}*-rough* part of the number a. Observe that the product of three integers is a perfect cube if and only if so are the product of their p_r-smooth parts and the product of their p_{r+1}-rough parts. In particular, for a fixed a' there cannot be three elements in A with p_{r+1}-rough part a' such that the product of their p_r-smooth parts is a perfect cube. Note that the product of three p_r-smooth numbers is a cube if and only if the sum of their exponent vectors $(\alpha_1, \alpha_2, \dots, \alpha_r)$ add up to $(0,0,\dots,0)$ calculating coordinate-wise modulo 3. Alternatively, if we consider the exponent vectors as elements of $\mathbb{F}_3^r$, they form a nontrivial 3-term arithmetic progression (3AP). Let $L_r(i)$ be the set of p_r-smooth cubefree integers up to i:

$$L_r(i) := \{p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} : \ \alpha_1, \dots, \alpha_r \in \{0,1,2\}\} \cap [i],$$

and let $s_r(i)$ denote the largest possible size of a subset of $L_r(i)$ avoiding nontrivial solutions to $a_1 a_2 a_3 = x^3$. Note that $s_r(i)$ is the size of the largest 3AP-free subset of

$$\{(\alpha_1, \dots, \alpha_r) \in \{0,1,2\}^r : \ \alpha_1 \log p_1 + \dots + \alpha_r \log p_r \le \log i\},$$

if we consider this set as a subset of $\mathbb{F}_3^r$. Clearly, for every $i \ge p_1^2 \dots p_r^2$, we have $s_r(i) = s_r(p_1^2 \dots p_r^2)$ (whose common value is $r_3(\mathbb{F}_3^r)$, the largest possible size of a 3AP-free subset of $\mathbb{F}_3^r$). For getting good numerical bounds we shall calculate these $s_r(i)$ values, for which we used IP solvers. Note that the exact value of $r_3(\mathbb{F}_3^r)$ is known only for $r \le 6$, thus significantly improving our numerical bounds is a very difficult task.

3.3

First we sketch the proof of the lower bound in Theorem 6 (which provides upper bounds for $f_{4,3}$ and $F_{4,3}$).

Let $A \subseteq \{1,2,\dots,n\}$ be a subset such that $a_1 a_2 a_3 a_4 \ne x^3$ if $a_i \in A$, $a_1 < a_2 < a_3 < a_4$ and let $D = \{d_1, \dots, d_t\}$ be the set of all positive integers d such that $d \le n^{1/3}$ and $\Omega(d) \le \frac{1}{3} \log\log n$, where $\Omega(d)$ denotes the number of prime factors of d (counted by multiplicity). A calculation yields that

$$t = |D| > \frac{n^{1/3}}{(\log n)^{1 + \frac{1}{3}\log\frac{1}{3} - \frac{1}{3} + \frac{\varepsilon}{3}}}.$$

Let H be the 3-uniform hypergraph on the vertex set $\{P_1, \dots, P_t\}$ such that $\{P_i, P_j, P_k\}$ is an edge in H if and only if $d_i d_j d_k \in A$. Let M be the set of those $m \in [n]$ such that $m \notin A$ and $m = d_i d_j d_k$ for some $1 \le i < j < k \le t$, then $|A| \le n - |M|$.

For a fixed $m \in M$ let $h(m)$ denote the number of triples (d_i, d_j, d_k) such that $m = d_i d_j d_k$, $1 \le i < j < k \le t$. If $m = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r} \in M$, then

$$\Omega(m) = \Omega(d_i) + \Omega(d_j) + \Omega(d_k) \le \log\log n,$$

hence

$$h(m) \le \tau_3(m) = \prod_{i=1}^{r} \binom{k_i + 2}{2} \le \prod_{i=1}^{r} 3^{k_i} = 3^{\Omega(m)} \le 3^{\log\log n} = (\log n)^{\log 3},$$

where $\tau_3(m)$ denotes the number of triples (a,b,c) with $a,b,c \in \mathbb{Z}^+$ such that $m = abc$.

If H contains a K_4^3 (a subhypergraph G with vertex set $V = \{P_{i_1}, P_{i_2}, P_{i_3}, P_{i_4}\}$ such that $V \setminus \{P_{i_j}\}$ is an edge in G for every $j \in \{1,2,3,4\}$), then for some $d_{i_1} < d_{i_2} < d_{i_3} < d_{i_4}$ and

$$a_1 = d_{i_1}d_{i_2}d_{i_3}, \quad a_2 = d_{i_1}d_{i_2}d_{i_4}, \quad a_3 = d_{i_1}d_{i_3}d_{i_4}, \quad a_4 = d_{i_2}d_{i_3}d_{i_4}$$

we have $a_1 < a_2 < a_3 < a_4$, $a_1, a_2, a_3, a_4 \in A$ and $a_1a_2a_3a_4 = (d_{i_1}d_{i_2}d_{i_3}d_{i_4})^3$. Therefore, H does not contain any K_4^3. Hence, by a result of de Caen [1] there exists a constant $\delta > 0$ such that there at least δt^3 triples (i,j,k), $1 \le i < j < k \le t$ such that $\{P_i, P_j, P_k\}$ is not an edge in H.

Let $h = \max_{m \in M} h(m) \le (\log n)^{\log 3}$. If $\{P_i, P_j, P_k\} \notin H$, $1 \le i < j < k \le t$, then $m = d_id_jd_k$ has at most h decompositions as a product of three positive integers, which gives the following bound on M:

$$|M| \ge \frac{\delta t^3}{h} \gg \frac{n}{(\log n)^{3+\log\frac{1}{3}-1+\varepsilon} \cdot (\log n)^{\log 3}} = \frac{n}{(\log n)^{2+\varepsilon}},$$

which completes the proof of the lower bound.

The construction providing the upper bound is the set of the integers a such that

(i) $\frac{n}{\log n} \le a \le n$,

(ii) $d^2 \mid a$ implies $d \le \log n$, and

(i)ii a cannot be written in the form $a = uvw$ with integers u, v, w such that $\frac{\sqrt[3]{n}}{(\log n)^{16}} \le u, v, w \le \sqrt[3]{n}(\log n)^{16}$.

Here, we omit the details.

3.4

The proof of Theorem 7 is a modification of the similar bounds for multiplicative 3-Sidon sets, that is, for sets avoiding solutions to the equation $a_1a_2a_3 = b_1b_2b_3$. (Note that the main term for multiplicative 3-Sidon sets is larger, $\pi(n) + \pi(n/2)$, so neither bound is a corollary, instead the methods should be adapted to this slightly different setting.)

The set achieving the asymptotically largest possible size for Theorem 8 is

$$A = \left\{ m : \ m = pq, \ \frac{n}{\log n} < m \le n, \ p, q \text{ primes}, \ p < \frac{q}{\log n} \right\}.$$

The upper bound is a consequence of [2, Theorem 3], since, according to this result, if n is large enough, there exist distinct $a_1, a_2, \ldots, a_6 \in A$ such that

$$a_1a_2 = a_3a_4 = a_5a_6,$$

however, then $a_1a_2a_3a_4a_5a_6$ is a perfect cube.

4 Concluding remarks and open problems

We gave bounds for the functions $F_{k,3}(n)$ and $f_{k,3}(n)$.

Finally, we pose some problems for further research.

Problem 1. *Let us suppose that* $1 < k < d$. *Is it true that*

$$n^{k/d} \ll n - F_{k,d}(n) \le n - f_{k,d}(n) \ll n^{k/d}?$$

Problem 2. *Is it true that there exists a constant* c *such that*

$$f_{2,3}(n) = n - (c + o(1))n^{2/3}?$$

Problem 3. *Let $d \geq 4$. Is it true that*

$$f_{d+1,d}(n) = (1 - o(1))n?$$

As a corollary of the above theorems we get the following result:

Corollary 9. *For $d = 2, 3$ and $k > d$, $d \mid k$, there exist constants $c_{k,d} > 0$ and $C_{k,d} \in \mathbb{Z}^+$ such that*

$$F_{k,d}(n) = (c_{k,d} + o(1))\pi_{C_{k,d}}(n),$$

where $\pi_r(n)$ denotes the number of positive integers up to n which have exactly r prime factors (counted with multiplicity)

Problem 4. *Is it true that for any $d \geq 4$ and $k > d$, $d \mid k$, there exist constants $c_{k,d} > 0$ and $C_{k,d} \in \mathbb{Z}^+$ such that*

$$F_{k,d}(n) = (c_{k,d} + o(1))\pi_{C_{k,d}}(n)?$$

References

[1] D. de Caen, Extensions of a theorem of Moon and Moser on complete subgraphs, *Ars Combin.* **16** (1983), 5–10.

[2] P. Erdős, On the multiplicative representation of integers, *Israel Journal of Mathematics*, **2** (4) (1964) 251–261.

[3] P. Erdős, A. Sárközy, V. T. Sós, On product representations of powers, I., *European Journal of Combinatorics* **16 (6)** (1995) 567–588.

[4] Zs. Gy. Fleiner, M. H. Juhász, B. Kövér, P. P. Pach, Cs. Sándor, Product representation of perfect cubes, arXiv: 2405.12088

[5] E. Győri, C_6-free bipartite graphs and product representation of squares, *Discrete Mathematics* **165** (1997) 371–375.

[6] P. P. Pach, Generalized multiplicative Sidon sets, *Journal of Number Theory* **157** (2015) 507–529.

[7] P. P. Pach, An improved upper bound for the size of the multiplicative 3-Sidon sets, *Int. J. Number Theory* **15 (8)** (2019) 1721–1729.

[8] P. P. Pach, M. Vizer, Improved lower bounds for multiplicative square-free sequences, *Electronic Journal of Combinatorics* **30 (4)** (2023) Article Number P4.31.

[9] G. N. Sárközy, Cycles in bipartite graphs and an application in number theory, *Journal of Graph Theory* **19 (3)** (1995) 323–331.

[10] T. Tao, On product representations of squares, arXiv: 2405.11610

[11] J. Verstraëte, Product representations of polynomials, *E*uropean Journal of Combinatorics, **27 (8)** (2006) 1350–1361.