GENERALIZED MULTIPLICATIVE SIDON SETS

PÉTER PÁL PACH

ABSTRACT. Let us call a set of positive integers a multiplicative k-Sidon set, if the equation $a_1a_2\dots a_k = b_1b_2\dots b_k$ does not have a solution consisting of distinct elements of this set. Let $G_k(n)$ denote the maximal size of a multiplicative k-Sidon subset of $\{1, 2, \dots, n\}$. In this paper we prove that $\pi(n) + \pi(n/2) + c_1n^{2/3}/(\log n)^{4/3} \le G_3(n) \le \pi(n)+\pi(n/2)+c_2n^{2/3}\frac{\log n}{\log\log n}$ for some constants $c_1, c_2 > 0$. It is also shown that $\pi(n)+n^{3/5}/(\log n)^{6/5} \le G_4(n) \le \pi(n) + (10 + \varepsilon)n^{2/3}$. Furthermore, for every k the order of magnitude of $G_k(n)$ is determined and an upper bound, similar to the previously mentioned ones, is given. This problem is related to a problem of Erdős-Sárközy-T. Sós and Győri: They examined how many elements of the set $\{1, 2, \dots, n\}$ can be chosen in such a way that none of the $2k$-element products is a perfect square. The maximal size of such a subset is denoted by $F_{2k}(n)$. As a consequence of our upper estimates for $G_k(n)$ the upper estimates for $F_{2k}(n)$ are strengthened because $G_k(n) \ge F_{2k}(n)$. Moreover, by a new construction we also sharpen their lower bound for $F_8(n)$.

1. INTRODUCTION

A set $A \subseteq \mathbb{N}$ is called a Sidon set, if for every s the equation $x+y = s$ has at most one solution with $x, y \in A$. A multiplicative Sidon set is analogously defined by requiring that the equation $xy = s$ has at most one solution in A. To emphasize the difference, throughout the paper the first one will be called an additive Sidon set. There are many results on the maximal size of an additive Sidon set in $\{1, 2, \dots, n\}$ and about the infinite case, as well. Moreover, a natural generalization of additive Sidon sets is also studied, they are called $B_h[g]$ sequences: A sequence A of positive integers is called a $B_h[g]$ sequence, if every integer n has at most g representations $n = a_1 + a_2 + \dots + a_h$ with all a_i in A and $a_1 \le a_2 \le \dots \le a_h$. Note that an additive Sidon sequence is a $B_2[1]$ sequence.

2011 *Mathematics Subject Classification*: Primary 11B83, Secondary 11B05.
Key words and phrases: multiplicative Sidon set, C_{2k}-free graph.
Date: May 1, 2015.

In this paper our aim is to generalize the multiplicative Sidon sequences and give some bounds on the maximal size of them. A set $A \subseteq \mathbb{N}$ is going to be called a multiplicative k-Sidon sequence, if the equation $a_1a_2\dots a_k = b_1b_2\dots b_k$ does not have a solution in A consisting of distinct elements. With other words, A is k-Sidon, if the equation $a_1a_2\dots a_k = b_1b_2\dots b_k$ does not have a "nontrivial solution" in A.

In [10] I investigated the equation $a_1a_2\dots a_k = b_1b_2\dots b_l$, and proved that for $k \neq l$ there is no density-type theorem, which means that a subset of $\{1, 2, \dots, n\}$ not containing a "nontrivial solution", that is, a solution consisting of distinct elements, can have size $c \cdot n$. However, a Ramsey-type theorem can be proved: if we colour the integers by r colours, then the equation $a_1a_2\dots a_k = b_1b_2\dots b_l$ has a nontrivial monochromatic solution. The case when $k = l$ is much more interesting, in this paper this is going to be investigated.

Let $G_k(n)$ denote the maximal size of a multiplicative k-Sidon sequence in $\{1, 2, \dots, n\}$. Erdős studied the case $k = 2$. In [3] he gave a construction with $\pi(n) + c_1n^{3/4}/(\log n)^{3/2}$ elements, and proved that the maximal size of such a set is at most $\pi(n) + c_2n^{3/4}$. 31 years later Erdős [4] himself improved this upper bound to $\pi(n)+c_2n^{3/4}/(\log n)^{3/2}$. Hence, in the lower- and upper bounds of $G_2(n)$ not only the main terms are the same, but the error terms only differ in a constant factor. In this paper our aim is to asymptotically determine $G_k(n)$, and give lower- and upper bounds on the error term, as well.

Our question about the solvability of $a_1a_2\dots a_k = b_1b_2\dots b_k$ is not only a natural generalization of the multiplicative Sidon sequences, but it is also strongly connected to the following problem from combinatorial number theory: Erdős, Sárközy and T. Sós [5] examined how many elements of the set $\{1, 2, \dots, n\}$ can be chosen in such a way that none of the $2k$-element products from this set is a perfect square. The maximal size of such a subset is denoted by $F_{2k}(n)$. Note that the functions F and G satisfy the inequality $F_{2k}(n) \leq G_k(n)$ for every n and k because if the equation $a_1 \dots a_k = b_1 \dots b_k$ has a solution of distinct elements, then the product of these $2k$ numbers is a perfect square. Erdős, Sárközy and T. Sós proved the following estimates for $k = 3$:

$$\pi(n) + \pi(n/2) + c\frac{n^{2/3}}{(\log n)^{4/3}} \leq F_6(n) \leq \pi(n) + \pi(n/2) + cn^{7/9}\log n.$$

Besides, they noted that by improving their graph theoretic lemma used in the proof the upper bound $\pi(n) + \pi(n/2) + cn^{2/3}\log n$ could be obtained, so the lower and upper bounds would only differ in a log-power factor in the error term. Later Győri [7] improved this graph theoretic lemma, and gained the desired bound. Furthermore, Erdős, Sárközy and T. Sós gave the following estimates for $k = 4$:

$$\pi(n) + c_1 n^{4/7}/(\log n)^{8/7} \le F_8(n) \le \pi(n) + c_2 n^{3/4} \log n.$$

Moreover, they proved the upper bound $F_{2k}(n) \le \pi(n)+cn^{3/4}/(\log n)^{3/2}$ for even $k \ge 2$ and $F_{2k}(n) \le \pi(n) + \pi(n/2) + cn^{7/9} \log n$ for odd $k \ge 3$. In this paper these bounds are going to be improved as a consequence of my upper estimates for $G_k(n)$. For $k = 3$ Győri's previously mentioned upper bound's error term is strengthened by a $\log \log n$ factor, and for $k = 4$ the exponent of n is decreased from $3/4$ to $2/3$ in the error term of the estimate of Erdős, Sárközy and T. Sós. For $k = 4$ the lower bound $F_8(n) \ge \pi(n) + cn^{4/7}/(\log n)^{8/7}$ given by Erdős, Sárközy and T. Sós is also improved with the help of a new construction, it is going to be proved that $F_8(n) \ge \pi(n) + n^{3/5}/(\log n)^{6/5}$.

2. Preliminary lemmas

Throughout the paper the maximal number of edges of a graph not containing a cycle of length k is conventionally denoted by $ex(n, C_k)$, and let us use the notation $ex(u, v, C_{2k})$ for the maximal number of edges of a C_{2k}-free bipartite graph, where the number of vertices of the two classes are u and v. (Note that every graph appearing in this paper is simple.)

Lemma 2.1. *Let $n \in \mathbb{N}$. Then*

$$\frac{1}{3}n^{3/2} < ex(n, C_4) < \frac{n}{4}(1 + \sqrt{4n - 3}),$$

if n is large enough.

Proof. Reiman [11] proved the upper bound, and he also constructed a C_4-free graph with $n = p^2 + p + 1$ vertices and $\frac{1}{2}p(p+1)^2 \sim \frac{1}{2}n^{3/2}$ edges for any prime p. From this the lower bound can be derived easily by looking the largest prime p such that $p^2 + p + 1 \le n$, taking the C_4-free graph for $p^2 + p + 1$ and adding $n - p^2 - p - 1$ isolated vertices to it. □

Lemma 2.2. *Let $n \in \mathbb{N}$. Then*

$$ex(n, C_6) < 0.6272n^{\frac{4}{3}},$$

if n is large enough.

Proof. This is the second statement of Theorem 1.1 in [6]. □

Lemma 2.3. *Let $n \in \mathbb{N}$. Then*

$$ex(n, C_{2k}) < 100kn^{\frac{k+1}{k}}.$$

Proof. This is a special case of Theorem 1. (setting $l = k$) in [2]. □

Lemma 2.4. *Let* $u, v \in \mathbb{N}$. *Then*

$$ex(u, v, C_6) \le 2^{1/3}(uv)^{2/3} + 16(u + v).$$

Proof. This is Theorem 1.2 in [6]. □

Lemma 2.5. *Let* $u, v \in \mathbb{N}$ *satisfying* $v \le u$. *Then*

$$ex(u, v, C_6) < 2u + v^2/2.$$

Proof. This is Theorem 1. in [7]. □

Lemma 2.6. *Let* $u, v \in \mathbb{N}$. *Then for every* $k \ge 2$

$$ex(u, v, C_{2k}) \le (2k-3)[(uv)^{\frac{k+1}{2k}} + u + v], \text{ if } k \text{ is odd},$$

and

$$ex(u, v, C_{2k}) \le (2k-3)[u^{\frac{k+2}{2k}} v^{\frac{1}{2}} + u + v], \text{ if } k \text{ is even}.$$

Proof. This is Corollary 2. in [9]. □

Lemma 2.7. *There exists some* $c > 0$ *constant such that for large enough* n *there exists a graph with* n *vertices and girth 8 having at least* $cn^{4/3}$ *edges.*

Proof. This is a consequence of Theorem 1. in [1]. In the previously mentioned theorem it is proved that for each prime power q there exists a $(q+1)$-regular graph of girth 8 having $n = 2(q^3+q^2+q+1)$ vertices. Therefore, the number of edges in this graph is $(q^3+q^2+q+1)(q+1) \sim 2^{-4/3}n^{4/3}$. Moreover, the prime powers are dense enough to guarantee the existence of a graph with n vertices and girth 8 having at least $cn^{4/3}$ edges for all large enough n. □

Lemma 2.8. *Let us denote by* $N_i(x)$ *the number of positive integers* $n \le x$ *satisfying* $\Omega(n) \le i$. *(Here,* $\Omega(n)$ *denotes the number of prime factors of* n *with multiplicity.) For every* $\delta > 0$ *there exists some constant* $C = C(\delta)$ *such that for* $1 \le i \le (1-\delta)\log\log x$ *we have*

$$N_i(x) < C(\delta) \cdot \frac{x}{\log x} \cdot \frac{(\log\log x)^{i-1}}{(i-1)!}.$$

Proof. Let $\pi_i(x) = |\{n : n \le x, \Omega(n) = i\}|$. Landau [8] proved that for every $\eta > 0$ there exists some $D = D(\delta)$ such that for every $1 \le i \le (1-\eta)\log\log x$ the following inequality holds:

$$\pi_i(x) < D(\eta) \cdot \frac{x}{\log x} \cdot \frac{(\log\log x)^{i-1}}{(i-1)!}.$$

Let $\delta > 0$ be arbitrary and $1 \le i \le (1-\delta)\log\log x$. By using the result of Landau an upper bound for $N_i(x)$ can be given:

$$N_i(x) = \sum_{j=0}^{i} \pi_j(x) \le \sum_{j=0}^{i} D(1+\delta)\cdot\frac{x}{\log x}\cdot\frac{(\log\log x)^{j-1}}{(j-1)!} =$$

$$= D(1+\delta)\cdot\frac{x}{\log x}\cdot\frac{(\log\log x)^{i-1}}{(i-1)!}\sum_{j=0}^{i}\frac{j(j+1)\dots(i-1)}{(\log\log x)^{i-j}} \le$$

$$\le D(1+\delta)\cdot\frac{x}{\log x}\cdot\frac{(\log\log x)^{i-1}}{(i-1)!}\sum_{j=0}^{i}(1-\delta)^{i-j} \le \frac{2D(1+\delta)}{\delta}\cdot\frac{x}{\log x}\cdot\frac{(\log\log x)^{i-1}}{(i-1)!},$$

hence for constant $C(\delta) = \frac{2D(1+\delta)}{1-\delta}$ the required inequality holds. □

Lemma 2.9. *Let $n \in \mathbb{N}$. Every $m \le n$ positive integer can be written in the form*

$$m = uv, \ \ v \le u,$$

where $u \le n^{2/3}$, or u is a prime.

Proof. This is Lemma I. in [3]. □

Similarly, an even sharper statement can be proved.

Lemma 2.10. *Let n be a positive integer and $1 < g < n^{1/6}$ an arbitrary real number. Every $m \le n$ can be written in the form*

$$m = uv \ (u, v \in \mathbb{N}),$$

where one of the following conditions holds:

(a) $v \le u \le \sqrt{n}\cdot g$,
(b) $\sqrt{n}\cdot g < u \le n^{2/3}$ *such that* $\Omega(u) \le \frac{\log n}{2\log g}$,
(c) $n^{2/3} < u$ *is a prime.*

Proof. Let the prime factorization of m be $m = q_1 q_2 \dots q_r$. We may suppose that $n^{2/3} > q_1 \ge q_2 \ge \dots \ge q_r$, otherwise (c) holds. Starting with q_1 we make two products out of the prime factors in such a way that we always add the next prime to the product which is smaller. Accordingly, at first q_1 forms one of the products, and the value of the other (empty) product is 1. In the next step the other product is going to be q_2, then q_3 goes to the product containing q_2 because $q_1 \ge q_2$, so the two products are going to be q_1 and $q_2 q_3$. Hereafter, we continue dividing the prime factors in the above described way. If we manage to adject all the q_i in such a way that none of the obtained products are bigger than $\sqrt{n}\cdot g$, then (a) holds. Otherwise, let i be the smallest index such that by adjecting q_i one of the products would be bigger than $\sqrt{n}\cdot g$. It was possible to divide the primes $q_1,\dots,q_{i-1}$ into two parts in such a way that in both parts the product of the primes is at most $\sqrt{n}\cdot g$. Let us call the two products A and B, then the inequality

$$A \le B \le \sqrt{n}\cdot g$$

holds. It is known that $Aq_i > \sqrt{n}\cdot g$, that is, $A > \frac{\sqrt{n}\cdot g}{q_i}$.

Since

$$A^2 \le AB \le \frac{m}{q_i} \le \frac{n}{q_i},$$

we have that

$$\frac{n\cdot g^2}{q_i^2} < A^2 \le \frac{n}{q_i},$$

which yields $q_i > g^2$. As q_i is the ith biggest prime divisor

$$n \ge m \ge q_1 q_2 \dots q_i \ge g^{2i},$$

so

$$i \le \frac{\log n}{2\log g}.$$

Hence, (b) holds with $u = Aq_i$, if $Aq_i \le n^{2/3}$. If $Aq_i > n^{2/3}$, then

$$q_i \ge \frac{ABq_i^2}{n} \ge \frac{(Aq_i)^2}{n} > n^{1/3},$$

so the value of i can be only 1 or 2. Since $A \le B$, so $A = 1$, that is, the inequality $Aq_i > n^{2/3}$ yields that $q_i > n^{2/3}$ is the biggest prime divisor of the number n. Therefore, $i = 1$ and $q_1 > n^{2/3}$, so (c) holds.

□

Let us denote by $G_k(n)$ the possible maximal size of a subset of $\{1,2,\dots,n\}$ such that no $2k$ distinct elements taken from this subset satisfy the equation $a_1a_2\dots a_k = b_1b_2\dots b_k$.

3. The equation $s_1s_2s_3 = t_1t_2t_3$

Theorem 3.1. *For every $\varepsilon > 0$ there exists an $N = N(\varepsilon)$ such that if $n > N = N(\varepsilon)$, then*

$$(1)\quad \pi(n) + \pi(n/2) + cn^{2/3}/(\log n)^{4/3} \le G_3(n) \le$$
$$\le \pi(n) + \pi(n/2) + \left(\frac{2^{4/3}e}{3} + \varepsilon\right)\cdot n^{2/3}\cdot\frac{\log n}{\log\log n},$$

where $c > 0$ is a constant.

Proof. At first the lower bound is going to be proved. By Lemma 2.7. there exists a graph G such that the vertices of G are the odd primes not greater than $\sqrt{n}$, the girth of G is at least 8 and for the number of edges of G we have $l_G \geq c(\pi(\sqrt{n}))^{4/3}$. Let $A = \{p \mid \sqrt{n} < p \leq n, p \text{ is prime}\} \cup \{2q \mid \sqrt{n} < q \leq n/2, q \text{ is prime}\} \cup \{uv \mid uv \in E(G)\}$. Now $A \subseteq \{1, 2, \ldots, n\}$, and we show that the equation

$$s_1 s_2 s_3 = t_1 t_2 t_3 \ (s_1, s_2, s_3, t_1, t_2, t_3 \in A) \tag{2}$$

has no solution consisting of distinct elements. We will refer to the edge uv of G by the product uv. At first assume that one of the variables in a solution of (2) is an edge of G, for instance, $s_1 = uv \in E(G)$. Then v is a prime, so it divides the right hand side as well, so it can be assumed that $t_1 = vw \in E(G)$, where $w \neq u$. Now w divides the left hand side, therefore it can be assumed that $s_2 = wz \in E(G)$, and so on... By continuing this method, we get a cycle of length at most 6, which is a contradiction. So in a solution of $s_1 s_2 s_3 = t_1 t_2 t_3$ only odd primes and odd primes multiplied by 2 can occur. In this case exactly 3 of the 6 variables would be divisible by 2 and none of them by 4, which is contradiction again. Furthermore, for the size of the set A we have

$$|A| \geq \pi(n) - \pi(\sqrt{n}) + \pi(n/2) - \pi(\sqrt{n}) + c(\pi(\sqrt{n}))^{4/3} \geq$$
$$\geq \pi(n) + \pi(n/2) + cn^{2/3}/(\log n)^{4/3}.$$

For the upper bound assume that for $A \subseteq \{1, 2, \ldots, n\}$ equation (2) has no solution consisting of distinct elements.

Let $g(n) = e^{\frac{\log n}{\log \log n}}$. Let $A = \{a_1, \ldots, a_l\}$, where $1 \leq a_1 < a_2 < \cdots < a_l \leq n$. Applying Lemma 2.10. for n and $g = g(n)$ we obtain that the elements of the set A can be written in the form $a_i = u_i v_i$, where u_i and v_i are positive integers, and one of the following conditions holds:

(i) $n^{2/3} < u_i$ is a prime,
(ii) $\sqrt{n} \cdot g(n) \leq u_i \leq n^{2/3}$ and $\Omega(u_i) \leq \frac{\log n}{2 \log g(n)}$,
(iii) $v_i \leq u_i \leq \sqrt{n} \cdot g(n)$.

If any $1 \leq i \leq l$ can be written as $u_i v_i$ in more than one way, then we choose such an u_i and v_i that v_i is minimal. The number of elements of A for which $u_i = v_i$ can be estimated from above by the number of square numbers in $\{1, 2, \ldots, n\}$, hence

$$|\{i | 1 \leq i \leq l, u_i = v_i\}| \leq \sqrt{n}. \tag{3}$$

For proving the upper estimate let us assume that $v_i \neq u_i$ for every $a_i \in A$. Then adding $\sqrt{n}$ to the obtained upper bound we gain an upper estimate for an arbitrary set A. Assume that (2) has no such solution where $s_1, s_2, s_3, t_1, t_2, t_3$ are distinct. Let $G = (V, E)$ be a graph where the vertices are the integers not greater than $n^{2/3}$ and the primes from the interval $(n^{2/3}, n]$:

$$V(G) = \{a \in \mathbb{N} | a \le n^{2/3}\} \cup \{p | n^{2/3} < p \le n, p \text{ is a prime}\}.$$

Then the number of the vertices of G is $|V(G)| = \pi(n) + [n^{2/3}] - \pi(n^{2/3})$. The edges of G will correspond to the elements of A: For each $1 \le i \le l$ let $u_i v_i$ be an edge, and denote it by a_i: $E(G) = \{u_i v_i | 1 \le i \le l\}$. In this way distinct edges are assigned to distinct elements of A. In the graph G there are no loops because we have omitted the elements where $u_i = v_i$, moreover $|E(G)| = |A| = l$. Furthermore, G contains no hexagons. Indeed, if $x_1x_2x_3x_4x_5x_6x_1$ is a hexagon in G, then

$$s_1 = x_1x_2, t_1 = x_2x_3, s_2 = x_3x_4, t_2 = x_4x_5, s_3 = x_5x_6, t_3 = x_6x_1$$

would be a solution of (2), contradicting our assumption.

Now our aim is to estimate from above the number of edges of G. At first let us partition the edges of G into some parts. Let G_0 be the subgraph that contains such $u_i v_i$ edges of G for which $\max(u_i, v_i) \le \sqrt{n}$:

$$E(G_0) = \{u_i v_i | u_i \le \sqrt{n}\}.$$

Let K_1 be a positive integer, which is going to be determined later, and for every $1 \le h \le K_1$ let G_h be the subgraph which contains those $u_i v_i$ edges of G for which the inequality $\sqrt{n} \cdot g(n)^{\frac{h-1}{K_1}} < u_i \le \sqrt{n} \cdot g(n)^{\frac{h}{K_1}}$ holds:

$$E(G_h) = \{u_i v_i | \sqrt{n} \cdot g(n)^{\frac{h-1}{K_1}} < u_i \le \sqrt{n} \cdot g(n)^{\frac{h}{K-1}}\}.$$

The graphs $G_0, G_1, \dots, G_{K_1}$ contain all of the edges of G that satisfy (iii).

Out of the remaining edges those are divided into K_2 parts which satisfy (ii), where K_2 will also be determined later. For these $u_i v_i$ edges $\sqrt{n} \le u_i \le n^{2/3}$ and $\Omega(u_i) \le \frac{\log n}{2 \log g(n)}$ hold. For $1 \le h \le K_2$ let G_{K_1+h} be the subgraph which contains such $u_i v_i$ edges of the graph $G \setminus (G_0 \cup \dots \cup G_{K_1})$ which satisfy the inequality $n^{\frac{1}{2} + \frac{h-1}{6K_2}} \le u_i < n^{\frac{1}{2} + \frac{h}{6K_2}}$:

$$E(G_{K_1+h}) = \{u_i v_i | n^{\frac{1}{2} + \frac{h-1}{6K_2}} \le u_i < n^{\frac{1}{2} + \frac{h}{6K_2}}\} \setminus \cup_{j=0}^{K_1} E(G_j).$$

Finally, let $G_{K_1+K_2+1}$ be the graph which is obtained by deleting the edges of $G_0, G_1, \dots, G_{K_1+K_2}$ from G. For the edges $u_i v_i$ in $G_{K_1+K_2+1}$ we have $n^{2/3} < u_i$. That is, u_i is a prime, and these edges satisfy (i):

$$E(G_{K_1+K_2+1}) = \{u_i v_i | n^{2/3} \le u_i, u_i \text{ is prime}\}.$$

So we divided the graph G into $K_1 + K_2 + 2$ parts.

Denote by l_h the number of edges of G_h $(0 \le h \le K_1 + K_2 + 1)$. In the remaining part of the proof we estimate the l_h number of edges separately, and at the end we add up these estimates. There are at most $[n^{1/2}]$ vertices of G_0 that are endpoints of some edges because $u_i v_i \in E(G_0)$ implies $v_i < u_i \le n^{1/2}$. Hence, by Lemma 2.2. for large enough n

$$l_0 \le 0.6272(n^{1/2})^{4/3} = 0.6272 n^{2/3} \tag{4}$$

holds.

Now let $1 \le h \le K_1$. If any $a_i = u_i v_i$ is an edge of the graph G_h, then $\sqrt{n} \cdot g(n)^{\frac{h-1}{K_1}} < u_i \le \sqrt{n} \cdot g(n)^{\frac{h}{K_1}}$, and so $v_i = \frac{a_i}{u_i} \le \frac{n}{u_i} \le \frac{\sqrt{n}}{g(n)^{\frac{h-1}{K_1}}}$. Thus G_h is a bipartite graph with bipartition U_h and V_h, where

$$U_h \subseteq \left\{ [\sqrt{n} \cdot g(n)^{\frac{h-1}{K_1}}] + 1, \dots, [\sqrt{n} \cdot g(n)^{\frac{h}{K_1}}] \right\},$$

and

$$V_h \subseteq \left\{ 1, 2, \dots, \left[\sqrt{n}/g(n)^{\frac{h-1}{K_1}} \right] \right\}.$$

(We delete those vertices of G_h which are not endpoints of any edge.) By Lemma 2.4. the following inequality holds for the number of edges of G_h:

$$(5) \quad l_h \le 2^{1/3}(|U_h||V_h|)^{2/3} + 16(|U_h| + |V_h|) \le$$
$$\le 2^{1/3} n^{\frac{2}{3}} g(n)^{\frac{2}{3K_1}} + 16([\sqrt{n} \cdot g(n)^{\frac{h}{K_1}}] - [\sqrt{n} \cdot g(n)^{\frac{h-1}{K_1}}]) + 16\sqrt{n}/g(n)^{\frac{h-1}{K_1}}.$$

By adding up the upper estimates of l_h for $1 \le h \le K_1$:

(6)
$$\sum_{h=1}^{K_1} l_h \le 2^{1/3} K_1 n^{\frac{2}{3}} g(n)^{\frac{2}{3K_1}} + 16 \sum_{h=1}^{K_1} ([\sqrt{n} \cdot g(n)^{\frac{h}{K_1}}] - [\sqrt{n} \cdot g(n)^{\frac{h-1}{K_1}}]) +$$
$$+16 \sum_{h=1}^{K_1} \frac{\sqrt{n}}{g(n)^{\frac{h-1}{K_1}}} \le 2^{1/3} K_1 n^{\frac{2}{3}} g(n)^{\frac{2}{3K_1}} + 16\sqrt{n} \cdot g(n) + 16 \cdot \frac{1 - \frac{1}{g(n)}}{1 - \frac{1}{g(n)^{1/K_1}}} \cdot \sqrt{n}$$

because one of the summas is a telescopic sum and the other is the sum of the members of a geometric series of K_1 elements. Furthermore, we get the asymptotically best estimate, if we choose the value of K_1 in such a way that $K_1 g(n)^{\frac{2}{3K_1}}$ is minimal. Examining the function $K_1 \to K_1 g(n)^{\frac{2}{3K_1}}$ we get that it attains the smallest value for $K_1 = \frac{2 \log g(n)}{3}$, where its value is $\frac{2e}{3} \log g(n)$. Therefore, let $K_1 = \left\lceil \frac{2 \log g(n)}{3} \right\rceil$, and note that the ceiling gives us an error of neglectable size:

$$K_1 g(n)^{\frac{2}{3K_1}} < \left(\frac{2\log g(n)}{3}+1\right) g(n)^{1/\log g(n)} = \frac{2e}{3}\cdot \log g(n) + e. \tag{7}$$

Since $K_1 \le \log g(n)$, so

$$16 \cdot \frac{1-\frac{1}{g(n)}}{1-\frac{1}{g(n)^{1/K_1}}}\cdot\sqrt{n} \le \frac{16}{1-1/e^{3/2}}\cdot\sqrt{n}. \tag{8}$$

Therefore, from (6) with the choice of $K_1 = \left\lceil \frac{2\log g(n)}{3} \right\rceil$ by considering (7) and (8) we obtain the following upper bound:

(9)

$$\sum_{h=1}^{K_1} l_h \le \frac{2^{4/3}e}{3}\cdot n^{2/3}\log g(n) + 2^{1/3}e\cdot n^{2/3} + 16\sqrt{n}\cdot g(n) + \frac{16}{1-1/e^{3/2}}\cdot\sqrt{n} \le$$

$$\le \frac{2^{4/3}e}{3}\cdot n^{2/3}\cdot\frac{\log n}{\log\log n} + c_1 n^{2/3},$$

where c_1 is an arbitrary constant bigger than $2^{1/3}e$.

Now let $1 \le h \le K_2$. If any $a_i = u_i v_i$ is an edge of G_{K_1+h}, then

$$n^{\frac{1}{2}+\frac{h-1}{6K_2}} < u_i \le n^{\frac{1}{2}+\frac{h}{6K_2}},$$

and so

$$v_i = \frac{a_i}{u_i} \le \frac{n}{u_i} \le n^{\frac{1}{2}-\frac{h-1}{6K_2}}.$$

This means that G_h is such a bipartite graph where the two independent classes of vertices U_{K_1+h} and V_{K_1+h} satisfy the following conditions:

$$U_{K_1+h} \subseteq \left\{\left[n^{\frac{1}{2}+\frac{h-1}{6K_2}}\right]+1, \dots, \left[n^{\frac{1}{2}+\frac{h}{6K_2}}\right]\right\},$$

and

$$V_{K_1+h} \subseteq \left\{1, 2, \dots, \left[n^{\frac{1}{2}-\frac{h-1}{6K_2}}\right]\right\},$$

furthermore for every u_i element of U_{K_1+h}

$$\Omega(u_i) \le \frac{\log n}{2\log g(n)} = \frac{1}{2}\cdot\log\log n \tag{10}$$

also holds. (We delete those vertices of G_{K_1+h} which are not endpoints of any edge.)

Let us denote by $N_{s+1}(x)$ the number of the numbers which are less or equal than x and can be written as the product of at most $s+1$ primes:

$$N_{s+1}(x) = |\{a \in \mathbb{N} | a \le x \text{ and } \Omega(a) \le s+1\}|.$$

Let $s = \left\lfloor \frac{1}{2} \cdot \log\log n \right\rfloor - 1$. By Lemma 2.8. there exists such a c' constant depending on c with which the following inequality holds:

$$N_{s+1}(x) \le c' \cdot \frac{x}{\log x} \cdot \frac{(\log\log x)^s}{s!}. \tag{11}$$

Applying inequality (11) for $x = n^{\frac{1}{2}+\frac{h}{6K_2}}$ we have

(12)

$$|U_{K_1+h}| \le N_s(n^{\frac{1}{2}+\frac{h}{6K_2}}) \le c' \cdot \frac{n^{\frac{1}{2}+\frac{h}{6K_2}}}{(\frac{1}{2}+\frac{h}{6K_2})\log n} \cdot \frac{(\log\log n^{\frac{1}{2}+\frac{h}{6K_2}})^s}{s!} \le$$

$$\le 2c' \cdot \frac{n^{\frac{1}{2}+\frac{h}{6K_2}}}{\log n} \cdot \frac{(\log\log n)^s}{s!}.$$

To estimate the obtained expression we give an upper bound for $\frac{\log\log n}{s}$. Let $\eta > 0$ be arbitrary. If n is large enough, then

$$\frac{\log\log n}{s} = \frac{\log\log n}{\left\lfloor \frac{1}{2} \cdot \log\log n \right\rfloor - 1} \le 2 + \eta.$$

Using this and the $s! \ge (s/e)^s$ inequalities we have

$$\frac{(\log\log n)^s}{s!} \le \frac{(\log\log n)^s}{(s/e)^s} = ((2+\eta)e)^{(1/2)\log\log n} = \tag{13}$$

$$= (\log n)^{\frac{1}{2}\log((2+\eta)e)} < (\log n)^{9/10},$$

if $0 < \eta$ is chosen to be sufficiently small, because for $\eta = 0$ the value of the exponent of $\log n$ is smaller than 0.9. Substituting $\frac{(\log\log n)^s}{s!} < (\log n)^{9/10}$ into (12) we get

$$|U_{K_1+h}| \le 2c' \cdot \frac{n^{\frac{1}{2}+\frac{h}{6K_2}}}{(\log n)^{1/10}}.$$

Furthermore, it is clear that

$$|V_{K_1+h}| \le n^{\frac{1}{2}-\frac{h-1}{6K_2}}.$$

By Lemma 2.4. for the number of edges of G_{K_1+h} the following inequality holds:

$$l_{K_1+h} \le 2^{1/3}(|U_{K_1+h}||V_{K_1+h}|)^{2/3} + 16(|U_{K_1+h}| + |V_{K_1+h}|) \le \tag{14}$$

$$\le 2(c')^{2/3} n^{\frac{2}{3}+\frac{1}{9K_2}}/(\log n)^{1/15} + 16\left(2c' \cdot \frac{n^{\frac{1}{2}+\frac{h}{6K_2}}}{(\log n)^{1/10}} + n^{\frac{1}{2}-\frac{h-1}{6K_2}}\right).$$

Summing up the upper bounds of l_h for $1 \le h \le K_2$:

$$(15)\quad \sum_{h=1}^{K_2} l_{K_1+h} \le 2(c')^{2/3} K_2 n^{\frac{2}{3}+\frac{1}{9K_2}}/(\log n)^{1/15}+ \\ + 16\sum_{h=1}^{K_2}\left(2c'\cdot\frac{n^{\frac{1}{2}+\frac{h}{6K_2}}}{(\log n)^{1/10}} + n^{\frac{1}{2}-\frac{h-1}{6K_2}}\right).$$

In this expression summing the geometric progression $n^{\frac{1}{2}+\frac{h}{6K_2}}$ $(1 \le h \le K_2)$ we have

$$(16)\qquad \frac{32c'}{(\log n)^{1/10}}\cdot\sum_{h=1}^{K_2} n^{\frac{1}{2}+\frac{h}{6K_2}} = \frac{32c'}{(\log n)^{1/10}}\cdot\frac{n^{\frac{2}{3}+\frac{1}{6K_2}} - n^{\frac{1}{2}+\frac{1}{6K_2}}}{n^{\frac{1}{6K_2}}-1}.$$

In the estimate (15) the largest term is $2(c')^{2/3}K_2 n^{\frac{2}{3}+\frac{1}{9K_2}}/(\log n)^{1/15}$, therefore in order to obtain the best upper bound we have to choose the value of K_2 in such a way that $K_2 n^{\frac{1}{9K_2}}$ is minimal. Examining the function $K_2 \to K_2 n^{\frac{1}{9K_2}}$ we get that it obtains the smallest value for $K_2 = \frac{\log n}{9}$, where the value of the function is $\frac{e\log n}{9}$. Accordingly, let $K_2 = \lceil \frac{\log n}{9} \rceil$, and note that the upper integer part gives us an error of neglectable size:

$$K_2 n^{\frac{1}{9K_2}} < \left(\frac{\log n}{9}+1\right) n^{\frac{1}{9K_2}} \le \frac{e\log n}{9} + e.$$

With this choice of K_2 the value of (16):

$$(17)\qquad \frac{32c'}{(\log n)^{1/10}}\cdot\frac{n^{\frac{2}{3}+\frac{1}{6K_2}} - n^{\frac{1}{2}+\frac{1}{6K_2}}}{n^{\frac{1}{6K_2}}-1} \le c_2\cdot n^{2/3},$$

where $c_2 > 0$ is arbitrary. The sum of the other geometric progression appearing in (15) is less than $n^{1/2}\log n$, hence with this choice of c_2 the inequality (15) yields that

$$(18)\qquad \sum_{h=1}^{K_2} l_{K_1+h} \le \frac{2e(c')^{2/3}}{9} n^{2/3}(\log n)^{14/15} + c_3\cdot n^{2/3},$$

where $c_3 > c_2$ is arbitrary.

Finally, $G_{K_1+K_2+1}$ is also a bipartite graph, the two independent vertex classes are the primes from the interval $(n^{2/3}, n]$ and the positive integers less than $n^{1/3}$. (We delete again the vertices of degree 0.) If $p \in (n/2, n]$, then the vertex corresponding to p is the endpoint of at most one edge: The one corresponding to $p\cdot 1$ because $2p > n$, so p cannot be connected with an integer bigger than 1. Delete the $1p$

edges and the p vertices for $n/2 < p \leq n$ from the graph $G_{K_1+K_2+1}$, and let the remaining graph be $G'_{K_1+K_2+1}$. Note that the number of deleted edges is at most $\pi(n) - \pi(n/2)$. The graph $G'_{K_1+K_2+1}$ does not contain any hexagons either, and all of its edges join a prime from $(n^{2/3}, n/2]$ with a positive integer less than $n^{1/3}$. Therefore, it is a bipartite graph whose independent vertex classes R and S satisfy the following conditions:

$$R \subseteq \{p | n^{2/3} < p \leq n/2, p \text{ is a prime}\} \text{ and}$$

$$S \subseteq \{a \in \mathbb{N} | a < n^{1/3}\}.$$

By Lemma 2.5. for the number of edges of $G'_{K_1+K_2+1}$ the inequality

$$l'_{K_1+K_2+1} \leq 2|R| + |S|^2/2 \leq 2(\pi(n/2) - \pi(n^{2/3})) + n^{2/3}/2$$

holds. Accordingly,

$$l_{K_1+K_2+1} \leq \pi(n) - \pi(n/2) + l'_{K_1+K_2+1} \leq \pi(n) + \pi(n/2) + n^{2/3}/2. \tag{19}$$

Adding up the inequalities (4), (9), (18), (19):

$$l = \sum_{h=0}^{K_1+K_2+1} l_h \leq 0.6272 n^{2/3} + \frac{2^{4/3}e}{3} \cdot n^{2/3} \cdot \frac{\log n}{\log\log n} + c_1 n^{2/3} + \tag{20}$$

$$+ \frac{2e(c')^{2/3}}{9} n^{2/3} (\log n)^{14/15} + c_3 \cdot n^{2/3} + \pi(n) + \pi(n/2) + n^{2/3}/2 \leq$$

$$\leq \pi(n) + \pi(n/2) + \left(\frac{2^{4/3}e}{3} + \varepsilon \right) \cdot n^{2/3} \cdot \frac{\log n}{\log\log n},$$

where $\varepsilon > 0$ is arbitrary and n is large enough. Remember that the error coming from the square numbers is $O(n^{1/2})$ by (3), so this upper bound holds for any set A, if n is large enough. Consequently, the statement of the theorem is proved.

□

4. The equation $s_1 s_2 s_3 s_4 = t_1 t_2 t_3 t_4$

Now we give an upper bound for $G_4(n)$, moreover for $G_{2k}(n)$ for every $k \geq 2$.

Theorem 4.1. *For every $k \geq 2$ and $\varepsilon > 0$ there exists some $N = N(k, \varepsilon)$ such that for $n > N$ we have*

$$G_{2k}(n) \leq \pi(n) + (c + \varepsilon) n^{2/3},$$

where $c = 10$ for $k = 2$, $c = 18$ for $k = 3$ and $c = 4k - 3$ for $k > 3$.

Proof. Let

$$A = \{a_1, \dots, a_l\}, \text{ where } 1 \le a_1 < a_2 < \cdots < a_l \le n.$$

Assume that in A the equation

$$s_1 s_2 \dots s_{2k} = t_1 t_2 \dots t_{2k} \ (s_1, \dots, s_{2k}, t_1, \dots, t_{2k} \in A) \tag{21}$$

does not have a solution consisting of distinct elements. By applying Lemma 2.9. for n we get that the elements of A can be written in the form

$$a_i = u_i v_i,$$

where u_i and v_i are positive integers for which one of the following conditions holds:

(i) $n^{2/3} < u_i$ is a prime,
(ii) $v_i \le u_i \le n^{2/3}$.

If for some $1 \le i \le l$ there are more possibilities for a_i to be written as a product satisfying the above conditions, then choose u_i and v_i in such a way that v_i is minimal. Similarly as in the proof of Theorem 3.1., the number of elements of A such that $u_i = v_i$ can be estimated from above by the number of square numbers in $\{1, 2, \dots, n\}$, hence

$$|\{i \mid 1 \le i \le l, u_i = v_i\}| \le \sqrt{n}. \tag{22}$$

At first for the upper estimate we shall assume that $v_i \ne u_i$ for every $a_i \in A$. Then adding $\sqrt{n}$ to the obtained upper bound we gain an upper estimate for an arbitrary set A.

Assume that (21) has no such solution where s_1, ..., s_{2k}, t_1, ..., t_{2k} are distinct. Let $G = (V, E)$ be a graph where the vertices are the integers not greater than $n^{2/3}$ and the primes from the interval $(n^{2/3}, n]$:

$$V(G) = \{a \in \mathbb{N} | a \le n^{2/3}\} \cup \{p | n^{2/3} < p \le n, p \text{ is a prime}\}.$$

The number of the vertices of G is $|V(G)| = \pi(n) + [n^{2/3}] - \pi(n^{2/3})$. The edges of G correspond to the elements of A. For each $1 \le i \le l$ let $u_i v_i$ be an edge. This edge will be denoted by $a_i = u_i v_i$:

$$E(G) = \{u_i v_i | 1 \le i \le l\}.$$

This way distinct edges are assigned to distinct elements of A. The graph G has no loops because we have omitted the elements where $u_i = v_i$, moreover $|E(G)| = |A| = l$. From the assumption that (21) has no solution consisting of distinct elements, it follows that there is no cycle of length $4k$ in the graph G.

Since if $x_1 x_2 \dots x_{4k} x_1$ is a cycle in G, then

$$s_i = x_{2i-1} x_{2i}, \ t_i = x_{2i} x_{2i+1} \ (1 \le i \le 2k)$$

would be a solution of (21) ($x_{4k+1} := x_1$), contradicting our assumption.

Now our aim is to estimate the number of edges of G from above. For this we partition the edges of G into some parts.

Let G_0 be the subgraph that contains such u_iv_i edges of G for which $v_i \le u_i \le \sqrt{n}$:

$$E(G_0) = \{u_iv_i | u_i \le \sqrt{n}\}.$$

Let G_1 be the subgraph which contains the u_iv_i edges satisfying $\sqrt{n} < u_i \le n^{2/3}$. In the case when $k = 2$ the edges of G_1 have to be split into two parts in order to obtain a good estimate: Let G_1' and G_1'' be the subgraphs which contain such u_iv_i edges of G_1 that satisfy $\sqrt{n} < u_i \le n^{7/12}$ and $n^{7/12} < u_i \le n^{2/3}$, respectively:

$$E(G_1') = \{u_iv_i \mid \sqrt{n} < u_i \le n^{7/12}\}$$

and

$$E(G_1'') = \{u_iv_i \mid n^{7/12} < u_i \le n^{2/3}\}.$$

The graphs G_0 and G_1 contain all the edges satisfying (ii).

Let G_2 be the graph that we get after deleting the edges of G_0 and G_1 from G. For the elements of A corresponding to the edges of the graph G_2 we have $n^{2/3} < u_i$, hence u_i is a prime number, and these edges satisfy (i):

$$E(G_2) = \{u_iv_i \mid n^{2/3} \le u_i,\ u_i \text{ is a prime}\}.$$

So we divided the graph G into 3 (4 in the case $k = 2$) parts.

Denote by l_h the number of edges of G_h $(0 \le h \le 2)$. In the remaining part of the proof we estimate the l_h number of edges separately, then we add them up.

The graph G_0 has at most $[\sqrt{n}]$ vertices of positive degree, since for $u_iv_i \in E(G_0)$ we have $v_i < u_i \le \sqrt{n}$. Therefore, by Lemma 2.3. the number of edges of G_0 satisfies the inequality

$$l_0 \le 200k \cdot n^{\frac{1}{2}+\frac{1}{4k}}. \tag{23}$$

If u_iv_i is an edge of the graph G_1, then

$$v_i = \frac{n}{u_i} \le \frac{n}{\sqrt{n}} = \sqrt{n}.$$

This means that the sizes of the independent vertex classes of the bipartite graph G_1 are at most $n^{2/3}$ and $n^{1/2}$. By Lemma 2.6. for the number of edges of G_1 we obtain the upper bound:

$$\begin{aligned} l_1 &\le (4k-3)(n^{\frac{2}{3}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{2k+2}{4k}} + n^{\frac{2}{3}} + n^{\frac{1}{2}}) = \\ &= (4k-3)n^{\frac{1}{3}+\frac{k+1}{4k}} + (4k-3)n^{2/3} + (4k-3)n^{1/2}. \end{aligned} \tag{24}$$

When $k=2$ this estimate is not sharp enough, so we give upper bounds for the number of edges of G_1' and G_1'' separately by using Lemma 2.6.:

$$l_1' \le 5(n^{\frac{7}{12}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{3}{4}} + n^{\frac{7}{12}} + n^{\frac{1}{2}}) = 5n^{2/3} + 5n^{7/12} + 5n^{1/2},$$

$$l_1'' \le 5(n^{\frac{2}{3}\cdot\frac{1}{2}+\frac{5}{12}\cdot\frac{3}{4}} + n^{\frac{2}{3}} + n^{\frac{5}{12}}) = 5n^{2/3} + 5n^{31/48} + 5n^{5/12}.$$

Here, when l_1'' was estimated, we used the observation that if $u_i v_i$ is an edge of G_1'', then $v_i \le n/u_i \le n^{5/12}$. So in the case $k=2$ we get that

$$l_1 = l_1' + l_1'' \le 10n^{2/3} + 5n^{7/12} + 5n^{1/2} + 5n^{31/48} + 5n^{5/12}. \tag{25}$$

Finally, let us look at the graph G_2, which is bipartite, as well and the two independent vertex classes are the set of the primes in $(n^{2/3}, n]$ and the set of the positive integers less than $n^{1/3}$. (We omit the vertices with degree 0.) So G_2 is a bipartite graph with independent vertex classes R and S satisfying

$$R \subseteq \{p \mid n^{2/3} < p \le n,\ p \text{ is a prime}\} \text{ and}$$

$$S \subseteq \{a \in \mathbb{N} \mid a < n^{1/3}\}.$$

The graph G_2 does not contain a cycle of length $4k$, and it can be shown that it does not contain k pairwise edge-disjoint 4-cycles either. Assume to the contrary that $y_{i,1}y_{i,2}y_{i,3}y_{i,4}y_{i,1}$ $(1 \le i \le k)$ are edge-disjoint 4-cycles in G_2. Then the product of the numbers $y_{i,1}y_{i,2}$ and $y_{i,3}y_{i,4}$ is equal to the product of the numbers $y_{i,2}y_{i,3}$ and $y_{i,4}y_{i,1}$ for every $1 \le i \le k$. Therefore, the equation $s_1 \dots s_{2k} = t_1 \dots t_{2k}$ has a solution consisting of distinct elements of A, which contradicts our assumption. So G_2 does not contain k edge-disjoint 4-cycles, so after deleting at most $4(k-1)$ edges it can be guaranteed that there are no more 4-cycles in the graph at all. (If it contains a 4-cycle, we delete the edges of it, if it still contains a 4-cycle, we delete those edges too, and so on. After the $(k-1)$-th step it will not contain any 4-cycle.) Let us denote the remaining graph by G_2'. For the number of edges in G_2' we have $l_2' \ge l_2 - 4(k-1)$.

Now we define a graph H on S. The edges of H are obtained in the following way: Take the points of R one by one, and for every vertex $v \in R$ take a maximal matching of the neighbours of v. Let these be the edges of H. If the degree of v is 0 or 1, then we do not get any edge, if the degree of v is even, then we get $d(v)/2$ edges and if it is odd, then we get $(d(v)-1)/2 = \lfloor d(v)/2 \rfloor$ edges. If ab is an edge in H, then this edge is drawn for a common $G_2'-$neighbour of a and b. This common neighbour is unique, since in G_2' there is no 4-cycle. So, by this process, for different vertices of R we have different edges in H. If $d(v) \ge 2$, then $d(v)/3 \le \lfloor d(v)/2 \rfloor$, so the number of edges of H is at least 1/3 times the number of such edges of G_2' which have an endpoint in R with degree at least 2. Hence,

$$l_2' \leq |R| + 3l_H,$$

where l_H denotes the number of edges of H. We show that H does not contain a $2k$-cycle: Suppose to the contrary that $u_1u_2\dots u_{2k}u_1$ is a cycle in H. Then, by the definition of H, there exist vertices $v_1, v_2, \dots, v_{2k} \in R$ for which u_iv_i, v_iu_{i+1} (where $u_{2k+1} = u_1$) are all edges of G_2. Hence, the numbers $s_i = u_iv_i$, $t_i = v_iu_{i+1}$ form a solution of equation (21) consisting of distinct elements of A, which contradicts our assumption. So H is a C_{2k}-free graph having $[n^{1/3}]$ vertices, hence by Lemma 2.3. we obtain that

$$l_H \leq (100k)n^{\frac{1}{3}\left(1+\frac{1}{k}\right)}.$$

Therefore,

$$l_2 \leq |R| + 3l_H + 4(k-1) \leq \pi(n) + (300k)n^{\frac{1}{3}\left(1+\frac{1}{k}\right)} + 4(k-1). \tag{26}$$

Summarizing the results, namely, adding up the inequalities (23), (24) and (26):

$$\begin{aligned} l = l_0 + l_1 + l_2 \leq (200k \cdot n^{\frac{1}{2}+\frac{1}{4k}}) + ((4k-3)n^{\frac{1}{3}+\frac{k+1}{4k}} + (4k-3)n^{2/3} + \\ + (4k-3)n^{1/2}) + (\pi(n) + (300k)n^{\frac{1}{3}\left(1+\frac{1}{k}\right)} + 4(k-1)) \leq \\ \leq \pi(n) + (4k-3+\varepsilon)n^{2/3} \end{aligned}$$

holds for every $k \geq 4$, if $\varepsilon > 0$ and n is sufficiently large. If $k = 3$, then we get the upper bound $k \leq \pi(n) + (18+\varepsilon)n^{2/3}$. If $k = 2$, then for estimating k_1 we use (25):

$$\begin{aligned} l = l_0 + l_1 + l_2 \leq (400 \cdot n^{\frac{1}{2}+\frac{1}{8}}) + (10n^{2/3} + 5n^{7/12} + 5n^{1/2} + 5n^{31/48} + \\ 5n^{5/12}) + (\pi(n) + (300 \cdot 2)n^{\frac{1}{3}\left(1+\frac{1}{2}\right)} + 4) \leq \\ \leq \pi(n) + (10+\varepsilon)n^{2/3}, \end{aligned}$$

where $\varepsilon > 0$ and n is sufficiently large. These upper bounds are valid for any A, since the error term coming from (22) is negligible. Therefore, we proved the desired statement. □

Now we give a lower estimate for $G_4(n)$.

Theorem 4.2. *If n is large enough, then the inequality*

$$G_4(n) \geq \pi(n) + n^{3/5}/(\log n)^{6/5}$$

holds.

Proof. Let $n \in \mathbb{N}$,

$$S = \{p \mid p \leq n^{2/5}(\log n)^{1/5},\ p \text{ is a prime}\} \text{ and}$$

$$T = \{p \mid n^{2/5}(\log n)^{1/5} < p \leq n,\ p \text{ is a prime}\}.$$

At first we construct a bipartite graph G_0, where the two independent vertex classes are S and T, so the set of the vertices is $V(G_0) = S \cup T$. In order to do this, let us take take a C_4-free graph H on S, whose number of edges satisfies the following inequality:

$$\frac{1}{3}\pi(n^{2/5}(\log n)^{1/5})^{3/2} \leq l_H \leq \frac{2}{5}\pi(n^{2/5}(\log n)^{1/5})^{3/2}.$$

Note that such a graph exists according to Lemma 2.1. Now, we make the edges of H correspond injectively to such vertices of T which are in the interval $\left(n^{2/5}(\log n)^{1/5}, n^{3/5}/(\log n)^{1/5}\right]$. It can be done, since

$$\left|T \cap \left(n^{2/5}(\log n)^{1/5}, n^{3/5}/(\log n)^{1/5}\right]\right| =$$
$$= \pi(n^{3/5}/(\log n)^{1/5}) - \pi(n^{2/5}(\log n)^{1/5}) \geq \frac{2}{5}\pi(n^{2/5}(\log n)^{1/5})^{3/2},$$

if n is sufficiently large. If the edge $uv \in E(H)$ corresponds to the vertex $w \in T$, then displace the uv edge with the uwv cherry. To different uv edges different $w \in T$ vertices belong, moreover the inequalities $uw \leq n$ and $vw \leq n$ hold because $u, v \leq n^{2/5}(\log n)^{1/5}$ and $w \leq n^{3/5}/(\log n)^{1/5}$. Let us call the obtained bipartite graph G_0. In G_0 two vertices from S have at most one common neighbour, and they have exactly one, if there is an edge between them in H. Accordingly, the number of edges of G_0 is

$$|E(G_0)| = 2|E(H)| \geq \frac{2}{3}\pi(n^{2/5}(\log n)^{1/5})^{3/2}.$$

We claim that there is no cycle of length 4 and 8 in G_0. Every second vertex of a 4-cycle would be in S and every second in T. However, in this case the two vertices from S would have two common neighbours from T, which is not possible by the construction of this graph. On the other hand, if $x_1x_2x_3x_4x_5x_6x_7x_8x_1$ would be a 8-cycle in G_0, where $x_1, x_3, x_5, x_7 \in S,\ x_2, x_4, x_6, x_8 \in T$, then $x_1x_3x_5x_7x_1$ would be a 4-cycle in H because for every $i \in \{1, 3, 5, 7\}$ the vertex x_{i+1} is the common neighbour of x_i and x_{i+2} in G_0 ($x_9 := x_1$).

Now, let us start to examine the number of edges of G_0. In the graph G_0 the degree of every vertex of T is 0 or 2. Denote by $T_1 \subseteq T$ the set of vertices of degree 0 and by $T_2 \subseteq T$ the set of vertices of degree 2. Because of the bijective correspondence between the edges of H and the vertices of T_2 we have

$$|T_2| = |E(H)| \geq \frac{1}{3}\pi(n^{2/5}(\log n)^{1/5})^{3/2}.$$

Let G be the bipartite graph which is obtained from G_0 by adding 1 to S and connecting it with all of the vertices of T_1. That is, the two independent vertex classes are going to be $S \cup \{1\}$ and T: $V(G) = S \cup \{1\} \cup T$, and the set of the edges of the graph is $E(G) = E(G_0) \cup \{1x \mid x \in T_1\}$. We claim that the set $A = \{xy \mid xy \in E(G)\}$ satisfies the conditions: $A \subseteq \{1, 2, \ldots, n\}$ and the equation $s_1s_2s_3s_4 = t_1t_2t_3t_4$ does not have a solution consisting of distinct elements from A.

From the construction it follows that $A \subseteq \{1, 2, \ldots, n\}$, moreover if n is large enough, then

$$\begin{aligned}|A| = |E(G)| = |T| + |T_2| \geq \\ \geq \pi(n) - \pi(n^{2/5}(\log n)^{1/5}) + \frac{1}{3}\pi(n^{2/5}(\log n)^{1/5})^{3/2} \geq \\ \geq \pi(n) + n^{3/5}/(\log n)^{6/5},\end{aligned}$$

since for different xy edges of G the product xy is also different. Now, it is going to be proved that the equation

$$s_1s_2s_3s_4 = t_1t_2t_3t_4 \tag{27}$$

does not have a solution of distinct elements of A. The set A has only one element which is divisible by the prime $p \in T_1$, namely p. This means that if p would occur on one of the sides, then it would have to occur on the other side as well, which is impossible. Therefore, the primes of T_1 cannot occur on either of the sides of the equation, that is, the numbers $s_1, s_2, s_3, s_4, t_1, t_2, t_3, t_4$ all correspond to some edges of G_0, so each of them can be written as the product of a prime of S and one of T_2. Moreover, if the equation (27) would hold, then the set of edges corresponding to the variables would be a union of cycles. Since the graph is bipartite, this would be only possible, if they would form two cycles of length 4 or one of length 8. However, G_0 does not contain either C_4 or C_8, so this is impossible, as well. Therefore, the desired statement is proved.

□

Summing up the lower- and upper bounds of Theorems 4.1. and 4.2 obtained for G_4 we get the following result:

Corollary 4.3. *For arbitrary $\varepsilon > 0$ there exists such an $N = N(\varepsilon)$ that for every $n > N$ the following inequality holds:*

$$\pi(n) + n^{3/5}/(\log n)^{6/5} \leq G_4(n) \leq \pi(n) + (10 + \varepsilon)n^{2/3}.$$

5. Corollaries

Erdős proved the following theorem about the size of the multiplicative 2-Sidon sequences:

Theorem (Erdős, [4]). There exist such c_1 and c_2 positive constants for which the inequality

$$\pi(n) + c_1 \frac{n^{3/4}}{(\log n)^{3/2}} \le G_2(n) \le \pi(n) + c_2 \frac{n^{3/4}}{(\log n)^{3/2}}$$

holds.

Now, by using Erdős's previously mentioned theorem and with the help of Theorems 3.1. and 4.1. some estimates about $G_k(n)$ standing for arbitrary k are going to be proved.

Corollary 5.1. *Let $3 \le k$ be a positive integer and $\varepsilon > 0$ be arbitrary. Then there exists such an $N = N_k(\varepsilon)$ with which for every $N < n$ the inequality*

$$G_k(n) \le \pi(n) + (c_k + \varepsilon) n^{2/3}$$

holds, if k is even and

$$G_k(n) \le \pi(n) + \pi(n/2) + (c_k + \varepsilon) \cdot n^{2/3} \cdot \frac{\log n}{\log \log n},$$

if k is odd.

Here $c_4 = 10$, $c_6 = 18$, $c_k = 2k - 3$ for even $6 < k$ and $c_k = \frac{2^{4/3}e}{3}$ for odd $3 \le k$.

Proof. According to Theorem 4.1. the statement holds, if k is even.

For odd k the inequality is going to be proved by induction.

By Theorem 3.1. the statement stands for $k = 3$. Let us assume that the inequality is already proved for an odd k bigger than 3. That is, for every $\varepsilon > 0$ there exists such an $N_k = N_k(\varepsilon)$ bound that if $n > N_k$ and for a set $A \subseteq \{1, 2, \dots, n\}$

$$|A| \ge \pi(n) + \pi(n/2) + \left(\frac{2^{4/3}e}{3} + \varepsilon\right) n^{2/3} \frac{\log n}{\log \log n}$$

holds, then $2k$ distinct elements of A can be chosen for which $s_1 \dots s_k = t_1 \dots t_k$. Now let $n > N_k$, $A \subseteq \{1, 2, \dots, n\}$, and assume that

$$|A| \ge \pi(n) + \pi(n/2) + \left(\frac{2^{4/3}e}{3} + \varepsilon\right) n^{2/3} \frac{\log n}{\log \log n}.$$

If n is large enough, then this yields that

$$|A| \geq \pi(n) + \pi(n/2) + \left(\frac{2^{4/3}e}{3} + \varepsilon\right) n^{2/3} \cdot \frac{\log n}{\log\log n} \geq$$
$$\geq \pi(n) + C_2 n^{3/4}/(\log n)^{3/2},$$

therefore according to the result of Erdős about the 2-Sidon sequences the equation

$$s_{k+1}s_{k+2} = t_{k+1}t_{k+2}$$

has a solution of distinct elements in A. Let us fix one such solution. Applying the induction hypothesis for the set $A\setminus\{s_{k+1}, s_{k+2}, t_{k+1}, t_{k+2}\}$, if n is large enough, then $2k$ pairwise distinct elements can be chosen for which

$$s_1 \dots s_k = t_1 \dots t_k.$$

The numbers $s_1, \dots, s_{k+2}, t_1, \dots, t_{k+2}$ are pairwise distinct, and

$$s_1 \dots s_{k+2} = t_1 \dots t_{k+2},$$

so we proved the statement for $k+2$. Therefore, the theorem is proved. □

Remark. *It is easy to check that for even k the set $\{p \mid 1 \leq p \leq n, p$ is a prime$\}$ and for odd k the set $\{p \mid \sqrt{n} < p \leq n, p$ is a prime$\} \cup \{2q \mid \sqrt{n} < q \leq n/2, q$ is a prime$\}$ is a multiplicative k-Sidon sequence. This means that Corollary 5.1 implies that $G_k(n)$ is asymptotically $\pi(n)$ for even k and $\pi(n) + \pi(n/2)$ for odd k.*

Erdős, Sárközy and T. Sós examined that at most how many elements of a set can be chosen in such a way that the product of any $2k$ of them is not a square. They proved the following theorem about the maximal size, $F_{2k}(n)$, of such sets:

Theorem (Erdős, Sárközy, T. Sós, [5]**).** Let $1 < k$ be a positive integer. There exists such a constant $c > 0$ that the following inequalities hold:

$$F_{2k}(n) \leq \pi(n) + cn^{3/4}/(\log n)^{3/2},$$

if k is even and n is large enough, and respectively

$$F_{2k}(n) \leq \pi(n) + \pi(n/2) + cn^{7/9}\log n,$$

if k is odd and n is large enough.

For $k = 3$ Győri strengthened this result by proving the following theorem:

Theorem (Győri, [7]). There exists such a constant $c > 0$ that the following inequality holds:

$$F_6(n) \leq \pi(n) + \pi(n/2) + cn^{2/3} \log n.$$

Moreover, this result implies that a similar upper bound can be given for $F_{2k}(n)$, when n is odd. However, by using Corollary 5.1. we can prove a stronger statement than the previously quoted one of Erdős, Sárközy and T. Sós and note that for odd k it is even slightly stronger than the result of Győri:

Corollary 5.2. *Let $3 \leq k$ be a positive integer and $\varepsilon > 0$ be arbitrary. Then there exists such an $N = N_k(\varepsilon)$ with which for every $N < n$ one of the following inequalities holds depending on the parity of k:*

$$F_{2k}(n) \leq \pi(n) + (c_k + \varepsilon)n^{2/3}, \textit{ if } k \textit{ is even,}$$

and

$$F_{2k}(n) \leq \pi(n) + \pi(n/2) + (c_k + \varepsilon) \cdot n^{2/3} \cdot \frac{\log n}{\log \log n}, \textit{ if } k \textit{ is odd.}$$

Here $c_4 = 10$, $c_6 = 18$, $c_k = 2k - 3$ for even $6 < k$ and $c_k = \frac{2^{4/3}e}{3}$ for odd $3 \leq k$.

Proof. If the equation

$$s_1 \dots s_k = t_1 \dots t_k \ (s_1 \dots, s_k, t_1, \dots, t_k \in A)$$

has a solution of distinct elements, then $x = s_1 \dots s_k$ and $s_{k+i} = t_i$ give a solution of the equation

$$s_1 \dots s_{2k} = x^2.$$

Therefore, $F_{2k}(n) \leq G_k(n)$ holds for every n. So, Corollary 5.1. yields the desired statement.

□

Moreover, the lower bound of $F_8(n)$ given by Erdős, Sárközy és T. Sós is also developed in this paper. They showed that $F_8(n) \geq \pi(n) + cn^{4/7}/(\log n)^{8/7}$, and we increase the exponent of n to $3/5$ in the error term.

Corollary 5.3. *If n is sufficiently large, then the following inequality holds:*

$$F_8(n) \geq \pi(n) + n^{3/5}/(\log n)^{6/5}.$$

Proof. The construction occuring in the proof of Theorem 4.2. is also appropriate for proving this problem. That proof can also be applied here with some little changes. □

6. Acknowledgements

This research was realized in the frame of TÁMOP 4.2.4. A/1-11-1-2012-0001 National Excellence Program – Elaborating and operating an inland student and researcher personal support system. The project was subsidized by the European Union and co-financed by the European Social Fund. The Hungarian National Foundation for Scientific Research Grant K108947 has also contributed to this research.

References

[1] C. T. Benson: *Minimal regular graphs of girths eight and twelve*, Canad. J. Math. 18 (1966) 1091–1094.

[2] J. A. Bondy, M. Simonovits: *Cycles of even length in graphs*, J. Combinatorial Theory Ser. B 16 (1974) 97–105.

[3] P. Erdős: *On sequences of integers no one of which divides the product of two others and some related problems*, Tomsk. Gos. Univ. Učen. Zap. 2 (1938) 74–82.

[4] P. Erdős: *On some applications of graph theory to number theoretic problems*, Publ. Ramanujan Inst. 1 (1969) 131–136.

[5] P. Erdős, A. Sárközy, V. T. Sós: *On the product representations of powers, I*, European J. Comb. 16 (1995) 567–588.

[6] Z. Füredi, A. Naor, J. Verstraëte: *On the Turán number for the hexagon*, Adv. Math. 203(2) (2006) 476–496.

[7] E. Győri: *C_6-free bipartite graphs and product representation of squares*, Discrete Math. 165/166 (1997) 371–375.

[8] E. Landau: *Handbuch der Lehre von der Verteilung der Primzahlen*, vol. 1, Leipzig (1909).

[9] A. Naor, J. Verstraëte: *A note on bipartite graphs without $2k$-cycles*, Probability, Combinatorics and Computing 14(5-6) (2005) 845–849.

[10] P. P. Pach: *The Ramsey-type version of a problem of Pomerance and Schinzel*, Acta Arithmetica, 156(1) (2012) 1–5.

[11] I. Reiman: *Über ein Problem von K. Zarankiewicz*, Acta. Math. Acad. Sci. Hung. 9 (1959) 269–279.

Department of Computer Science and Information Theory, Budapest University of Technology and Economics, 1521 Budapest, Magyar tudósok körútja 2., Hungary

E-mail address: ppp@cs.bme.hu