# On products of integers. II

P. ERDŐS and A. SÁRKÖZY

**1.** Throughout this paper, $c_1, c_2, \dots$ denote absolute constants; $k_0(\alpha, \beta, \dots)$, $k_1(\alpha, \beta, \dots), \dots, x_0(\alpha, \beta, \dots), \dots$ denote constants depending only on the parameters $\alpha, \beta, \dots$; $v(n)$ denotes the number of the prime factors of the positive integer $n$, counted according to their multiplicity. The number of the elements of a finite set $S$ is denoted by $|S|$.

Let $k, n$ be any positive integers, $A=\{a_1, a_2, \dots, a_n\}$ any finite, strictly increasing sequence of positive integers satisfying

$$a_1 = 1,\ a_2 = 2, \dots, a_k = k \tag{1}$$

(consequently, $|A|=n \geqq k$). Let us denote the number of integers which can be written in form

$$\prod_{i=1}^{n} a_i^{\varepsilon_i} \quad (\varepsilon_i = 0 \text{ or } 1) \tag{2}$$

or

$$a_i a_j \quad (1 \leqq i, j \leqq n),$$

respectively by $f(A, n, k)$ and $g(A, n, k)$. Let us write

$$F(n, k) = \min_A f(A, n, k) \quad \text{and} \quad G(n, k) = \min_A g(A, n, k)$$

where the minimums are extended over all sequences $A$ satisfying (1) and $|A|=n$.

Starting out from a conjecture of G. Halász, the second author showed in the first part of this paper (see [4]) that

$$G(n, k) > n \cdot \exp\left(c_1 \frac{\log k}{\log \log k}\right).$$

Note that to get many distinct products of form $a_i a_j$, we need a condition of type (1); otherwise e.g. the sequence $A=\{1, 2, 2^2, \dots, 2^{n-1}\}$ is a counterexample, namely for this sequence the number of the distinct products is $2n-1=O(n)$.

Received June 10, 1977.

Furthermore, $G(n, k)/n$ is not much greater for fixed $k$ and large $n$ than for $n=k$, i.e. for $A=B_k$ where

$$B_k = \{1, 2, \dots, k\}.$$

This can be shown by the following construction: let $A^*=\{a_1^*, a_2^*, \dots, a_n^*\}$ be the sequence of the integers of form $p^i j$ where $p$ is a fixed prime number greater than $k$, $i=1, 2, \dots, m$, $j=1, 2, \dots, k$, and $m$ is any positive integer. Clearly,

$$\frac{g(A^*, n, k)}{n} < 2\frac{g(B_k, k, k)}{k} = 2\frac{G(k, k)}{k}$$

thus

$$\frac{G(n, k)}{n} < 2\frac{G(k, k)}{k} \quad \text{for} \quad k/n,$$

hence

$$\frac{G(n, k)}{n} < 4\frac{G(k, k)}{k} \quad (= o(k)) \quad \text{for every } n.$$

The authors conjectured that

$$\frac{G(n, k)}{n} > c_2\frac{G(k, k)}{k} \tag{3}$$

for every $n \geqq k$, and furthermore, that for any $\omega>0$, $k>k_0(\omega)$ and $n \geqq k$, we have

$$F(n, k) > n^2 k^\omega$$

or perhaps

$$n^2 \exp\left(c_3\frac{k}{\log k}\right) < F(n, k) < n^2 \exp\left(c_4\frac{k}{\log k}\right) \tag{4}$$

for large $k$ and $n \geqq k$. (See [4], also Problem 9 in [3].)

The aim of this paper is to disprove (3) (Theorem 1) and to prove a slightly weaker form of (4) (Theorem 2).

**2.** In this section, we will disprove (3).

P. Erdős showed in [1] (see Theorem 1) that for any $\varepsilon>0$ and $k>k_0(\varepsilon)$,

$$\frac{k^2}{(\log k^2)^{1+\varepsilon}}(e\log 2)^{\frac{\log\log k^2}{\log 2}} = g(B_k, k, k) = \sum_{\substack{m\leqq k^2\\ m=xy\\ x\leqq k,\, y\leqq k}} 1 < \frac{k^2}{(\log k^2)^{1-\varepsilon}}(e\log 2)^{\frac{\log\log k^2}{\log 2}}.$$

This inequality can be written in the equivalent form

$$\frac{k^2}{(\log k)^{c_5+\varepsilon}} < G(k, k) < \frac{k^2}{(\log k)^{c_5-\varepsilon}}$$

where

$$c_5 = 1 - \frac{1+\log\log 2}{\log 2}.$$

An easy computation shows that

$$0{,}086 < c_5 < 0{,}087.$$

Hence, for large $k$,

$$\frac{k}{(\log k)^{0,087}} < \frac{G(k,k)}{k} < \frac{k}{(\log k)^{0,086}}. \tag{5}$$

Thus to disprove (3), it is sufficient to show that for large $k$, there exist a positive integer $n$ $(\geqq k)$ and a sequence $A$ such that $|A|=n$, (1) holds and

$$\frac{g(A,n,k)}{n} < \frac{k}{(\log k)^{c_6}} \tag{6}$$

where

$$c_6 > 0{,}087. \tag{7}$$

In fact, by (5) and the definition of the function $G(n,k)$, this would imply

$$\frac{G(n,k)}{n} < \frac{k}{(\log k)^{c_6}} < \frac{1}{(\log k)^{c_7}} \cdot \frac{G(k,k)}{k} \tag{8}$$

where

$$c_7 = c_6 - 0{,}087 > 0$$

by (7).

Let us write $\varphi(x) = 1 + x\log x - x$ and let $z$ denote the single real root of the equation

$$\varphi(x) = \varphi(1+x). \tag{9}$$

A simple computation shows that

$$0{,}54 < z < 0{,}55. \tag{10}$$

Theorem 1. *For any* $\varepsilon > 0$ *and* $k > k_1(\varepsilon)$, *there exist a positive integer* $n(\geqq k)$ *and a sequence* $A$ *such that* $|A|=n$, (1) *holds and*

$$\frac{g(A,n,k)}{n} < \frac{k}{(\log k)^{c_8-\varepsilon}} \tag{11}$$

*where*

$$c_8 = \varphi(z). \tag{12}$$

(The function $\varphi(x)$ is decreasing for $0<x<1$. Thus with respect to (10), we obtain by a simple computation that

$$c_8 = \varphi(z) > \varphi(0{,}55) > 0{,}121.$$

Hence, Theorem 1 yields that for large $k$, (6) holds with $c_6 = 0{,}121$ which satisfies (7). Thus in fact, (8) holds with $c_7 = 0{,}121 - 0{,}087 = 0{,}034$ which disproves (3).)

Proof. Let $k$ be a positive integer which is sufficiently large (in terms of $\varepsilon$) and let $m$ be any positive integer satisfying

$$m > k^2. \tag{13}$$

Let $D_k$ denote the set of those integers $d$ for which

$$1 \leqq d \leqq k \tag{14}$$

and

$$\nu(d) > \log\log k \tag{15}$$

hold. Let $p$ be a prime number satisfying

$$p > k. \tag{16}$$

Let $E_k$ denote the set of those integers $e$ which can be written in form $p^\alpha d$ where

$$1 \leqq \alpha \leqq m \tag{17}$$

and

$$d \in D_k. \tag{18}$$

Finally, let

$$A = E_k \cup B_k.$$

We are going to show that for large enough $k$, this sequence $A$ satisfies (11).

Obviously,

$$n = |A| = |E_k| + |B_k| \leqq mk + k < 2mk. \tag{19}$$

Furthermore, by a theorem of P. Erdős and M. Kac [2], we have

$$|D_k| > \frac{1}{3}k.$$

Thus (with respect to (16))

$$n = |A| > |E_k| = m \cdot |D_k| > \frac{1}{3}mk. \tag{20}$$

To estimate the number of the distinct products of form $a_i a_j$, we have to distinguish four cases.

*Case 1.* Assume at first that $a_i \in B_k$, $a_j \in B_k$. Since $B_k$ consists of $k$ elements, the pair $a_i, a_j$ can be chosen in at most

$$k^2 < m < n$$

ways (with respect to (13) and (20)).

*Case 2.* Assume now that $a_i = p^\alpha d \in E_k$ (where (14), (15) and (16) hold),

$$a_j \in B_k \tag{21}$$

and

$$\nu(a_j) \leqq z \log\log k. \tag{22}$$

Then

$$a_i a_j = p^\alpha d a_j. \tag{23}$$

Let $\pi_i(x)$ denote the number of those integers $u$ for which $u \leqq x$ and $v(u)=i$ hold. By a theorem of Hardy and Ramanujan, for any $\omega>0$ there exists a constant $c_9=c_9(\omega)$ such that for large $x$ and $1 \leqq i \leqq \omega \log x$, we have

$$\pi_i(x) < c_9 \frac{x}{\log x} \frac{(\log\log x)^{i-1}}{(i-1)!}. \tag{24}$$

Choosing here $\omega=1$ and using Stirling's formula, we obtain that for $k>k_2(\omega)$, the number of the integers $a_j$ satisfying (21) and (22) is at most

$$\sum_{0 \leqq i \leqq z \log\log k} \pi_i(k) < \tag{25}$$

$$< 1+\sum_{1 \leqq i \leqq z \log\log k} c_9 \frac{k}{\log k} \frac{(\log\log k)^{i-1}}{(i-1)!} <$$

$$< 1+c_9 \frac{k}{\log k} \sum_{1 \leqq i \leqq z \log\log k} \frac{(\log\log k)^{[z \log\log k]-1}}{([z \log\log k]-1)!} \leqq$$

$$< 1+c_9 \frac{k}{\log k}\, z \log\log k \frac{(\log\log k)^{[z \log\log k]-1}}{([z \log\log k]-1)!} <$$

$$< 1+c_{10} \frac{k}{\log k} \frac{(\log\log k)^{[z \log\log k]}}{([z \log\log k]-1)^{[z \log\log k]-1/2} e^{-[z \log\log k]-1}} <$$

$$< 1+c_{11} \frac{k}{\log k} \frac{(\log\log k)^{[z \log\log k]}}{(z \log\log k)^{[z \log\log k]-1/2} e^{-z \log\log k}} <$$

$$< c_{12} \frac{k}{\log k} \frac{1}{(\log k)^{z \log z} (\log\log k)^{-1/2} (\log k)^{-z}} < \frac{k}{(\log k)^{c_8-\varepsilon/3}}$$

(where $c_8$ is defined by (12)) since $\dfrac{(\log\log k)^{i-1}}{i-1!}$ is increasing for $1 \leqq i \leqq \log\log k$.

By (14), (17) and (18), $\alpha$ and $d$ can be chosen in at most $m$ and $k$ ways, respectively. Thus the number of the products of form (23) is less than

$$m \cdot k \cdot \frac{k}{(\log k)^{c_8-\varepsilon/3}} < n \frac{k}{(\log k)^{c_8-\varepsilon/2}}$$

(with respect to (20)).

*Case 3.* Assume that $a_i=p^\alpha d \in E_k$ (where (14), (15) and (16) hold),

$$a_j \in B_k \tag{26}$$

and

$$v(a_j) > z \log\log k. \tag{27}$$

Then

$$a_i a_j = (p^\alpha d) a_j = p^\alpha (d a_j). \tag{28}$$

By (14), (15), (18), (26) and (27),

$$d a_j \leqq k \cdot k = k^2$$

and

$$v(d a_j) = v(d) + v(a_j) > \log\log k + z \log\log k = (1+z)\log\log k.$$

Thus applying (24) with $\omega = 100$, we obtain that for any $0 < \delta < z/2$ and $k > k_3(\delta)$, and writing $r = [(1+z-\delta)\log\log k^2]$, the number of the distinct products of form $d a_j$ is at most

$$\sum_{(1+z)\log\log k < i} \pi_i(k^2) < \sum_{(1+z-\delta)\log\log k^2 < i} \pi_i(k^2) = \tag{29}$$

$$= \sum_{r < i \leqq 100 \log\log k^2} \pi_i(k^2) + \sum_{100 \log\log k^2 < i} \pi_i(k^2) <$$

$$< \sum_{r < i \leqq 100 \log\log k^2} c_9 \frac{k^2}{\log k^2} \frac{(\log\log k^2)^{i-1}}{(i-1)!} + R(k^2) <$$

$$< c_{13} \frac{k^2}{\log k} \frac{(\log\log k^2)^r}{r!} \sum_{j=0}^{+\infty} \left( \frac{\log\log k^2}{r} \right)^j + R(k^2) <$$

$$< c_{14} \frac{k^2}{\log k} \frac{(\log\log k^2)^r}{r!} \sum_{j=0}^{+\infty} \left( \frac{1}{1+z-\delta} \right)^j + R(k^2) <$$

$$< c_{15} \frac{k^2}{\log k} \frac{(\log\log k^2)^r}{r!} + R(k^2)$$

where

$$R(x) = \sum_{100 \log\log x < i} \pi_i(x).$$

Applying Stirling's formula, we obtain that for $k > k_4(\delta)$,

$$\frac{k^2}{\log k} \frac{(\log\log k^2)^r}{r!} < \tag{30}$$

$$< c_{16} \frac{k^2}{\log k} \frac{(\log\log k^2)^{[(1+z-\delta)\log\log k^2]}}{([(1+z-\delta)\log\log k^2])^{[(1+z-\delta)\log\log k^2]+1/2} e^{-[(1+z-\delta)\log\log k^2]}} <$$

$$< c_{17} \frac{k^2}{\log k} \frac{(\log\log k^2)^{[(1+z-\delta)\log\log k^2]}}{((1+z-\delta)\log\log k^2)^{[(1+z-\delta)\log\log k^2]+1/2} e^{-(1+z-\delta)\log\log k}} <$$

$$< c_{18} \frac{k^2}{\log k} \frac{1}{e^{(1+z-\delta)\log(1+z-\delta)\log\log k} (\log\log k)^{1/2} (\log k)^{-(1+z-\delta)}} <$$

$$< c_{18} \frac{k^2}{(\log k)^{\varphi(1+z-\delta)}}.$$

The function $\varphi(x)$ is continuous at $x=1+z$. Thus if $\delta$ is sufficiently small in terms of $\varepsilon$ then for $k>k_5(\delta)=k_5(\delta(\varepsilon))=k_6(\varepsilon)$, we obtain from (30) that

$$\frac{k^2}{\log k}\frac{(\log\log k^2)^r}{r!}<\frac{k^2}{(\log k)^{\varphi(1+z)-\varepsilon/3}}=\frac{k^2}{(\log k)^{c_8-\varepsilon/3}} \tag{31}$$

(since $\varphi(1+z)=\varphi(z)=c_8$ by the definition of $z$).

Furthermore, P. Erdős proved in [1] (see formulae (5) and (6)) that for large $x$,

$$R(x)<2\frac{x}{(\log x)^2}. \tag{32}$$

(29), (31) and (32) yield that the number of the distinct products of form $da_j$ is at most

$$\sum_{(1+z)\log\log k<i}\pi_i(k^2)<c_{15}\frac{k^2}{(\log k)^{c_8-\varepsilon/3}}+2\frac{k^2}{(\log k^2)^2}<c_{19}\frac{k^2}{(\log k)^{c_8-\varepsilon/3}}. \tag{33}$$

Finally, by (17), $\alpha$ in (28) can be chosen in $m$ ways. Thus with respect to (20), we obtain that the number of the distinct products of form (28) is less than

$$m\cdot c_{19}\frac{k^2}{(\log k)^{c_8-\varepsilon/3}}<n\frac{k}{(\log k)^{c_8-\varepsilon/2}}.$$

*Case 4.* Assume that $a_i=p^\alpha d_1\in E_k$, $a_j=p^\beta d_2\in E_k$ where

$$1\leqq\alpha,\ \beta\leqq m \tag{34}$$

and

$$d_1,\ d_2\in D_k. \tag{35}$$

Then the product $a_ia_j$ can be written in form

$$a_ia_j=(p^\alpha d_1)(p^\beta d_2)=p^{\alpha+\beta}d_1d_2=p^\gamma d \tag{36}$$

where by (34) and (35),

$$2\leqq\gamma\leqq 2m \tag{37}$$

and

$$d=d_1d_2\leqq k\cdot k=k^2,\quad \nu(d)=\nu(d_1)+\nu(d_2)>2\log\log k. \tag{38}$$

By (37), $\gamma$ can be chosen in at most $2m-1<2m$ ways, while in view of (33), at most

$$\sum_{2\log\log k<i}\pi_i(k^2)<\sum_{(1+z)\log\log k<i}\pi_i(k^2)<c_{19}\frac{k^2}{(\log k)^{c_8-\varepsilon/3}}$$

integers $d$ satisfy (38). Thus the number of the distinct products $a_ia_j$ of form (36) is less than

$$2m\cdot c_{19}\frac{k^2}{(\log k)^{c_8-\varepsilon/3}}<n\frac{k}{(\log k)^{c_8-\varepsilon/2}}.$$

Summarizing the results obtained above, we get that for $k > k_7(\varepsilon)$,

$$g(A, n, k) < n + 3 \cdot n \cdot \frac{k}{(\log k)^{c_8 - \varepsilon/2}} < n \cdot \frac{k}{(\log k)^{c_8 - \varepsilon}}$$

which completes the proof of Theorem 1.

**3.** In this section, we will estimate $F(n, k)$.

Theorem 2. *There exist absolute constants* $c_{20}, c_{21}$ *such that for* $k > k_8$ *and* $n \geqq k$,

$$n^2 \exp\left(c_{20} \frac{k}{\log^2 k}\right) < F(n, k) < n^2 \exp\left(c_{21} \frac{k}{\log k}\right). \tag{39}$$

Proof. First we prove the upper estimate. We will show at first that

$$F(k, k) = f(B_k, k, k) < \exp\left(c_{22} \frac{k}{\log k}\right). \tag{40}$$

In case $A = B_k = \{1, 2, \dots, k\}$ (and $n = k$), all the products of form (2) are divisors of $k!$. Thus applying Legendre's formula and the prime number theorem (or a more elementary theorem), we obtain that

$$F(k, k) \leqq d(k!) = \prod_{p \leqq k} \left(1 + \sum_{\alpha=1}^{+\infty} \left[\frac{k}{p^\alpha}\right]\right) \leqq$$

$$\leqq \prod_{p \leqq k} \left(2 \sum_{\alpha=1}^{+\infty} \left[\frac{k}{p^\alpha}\right]\right) < \prod_{p \leqq k} \left(\sum_{\alpha=1}^{+\infty} \frac{2k}{p^\alpha}\right) = \prod_{p \leqq k} \frac{2k}{p-1} \leqq \prod_{p \leqq k} \frac{4k}{p} =$$

$$= \prod_{j=1}^{\left[\frac{\log k}{\log 2}\right]} \prod_{\frac{k}{2^j} < p \leqq \frac{k}{2^{j-1}}} \frac{4k}{p} < \prod_{j=1}^{\left[\frac{\log k}{\log 2}\right]} \prod_{\frac{k}{2^j} < p \leqq \frac{k}{2^{j-1}}} 4k \cdot \frac{2^j}{k} \leqq$$

$$\leqq \prod_{j=1}^{\left[\frac{\log k}{\log 2}\right]} (4 \cdot 2^j)^{\pi\left(\frac{k}{2^{j-1}}\right)} < \exp\left\{c_{23} \left(\sum_{j=1}^{\left[\frac{\log k}{\log 2}\right]} \frac{k}{2^{j-1}} \cdot \frac{1}{\log \frac{k}{2^{j-1}}} \cdot \log 4 \cdot 2^j\right)\right\} <$$

$$< \exp\left\{c_{24} \left(\sum_{j=1}^{\left[\frac{1}{2} \cdot \frac{\log k}{\log 2}\right]} \frac{k}{2^j} \cdot \frac{1}{\log \sqrt{k}} \cdot j + \sum_{j=\left[\frac{1}{2} \cdot \frac{\log k}{\log 2}\right]+1}^{\left[\frac{\log k}{\log 2}\right]} \frac{k}{2^j} \cdot j\right)\right\} <$$

$$< \exp\left\{c_{25} \left(\frac{k}{\log k} + \sqrt{k}\right)\right\} < \exp\left(c_{26} \frac{k}{\log k}\right)$$

which proves (40).

Assume now that $n > k$. Let $p$ denote a prime number satisfying $p > k$ and let

$$A = \{1, 2, \dots, k, p, p^2, \dots, p^{n-k}\}.$$

For this sequence $A$, $|A|=n$, and the products (2) can be written in form

$$\prod_{i=1}^{k} i^{\varepsilon_i} \prod_{j=1}^{n-k} p^{j\delta_j} = a \cdot p^{\beta} \tag{41}$$

where $\varepsilon_i=0$ or 1 and $\delta_j=0$ or 1. Here $a$ may assume $F(k, k)$ different values, and obviously, $\beta$ may assume any integer value (independently of $\alpha$) from the interval

$$0 \leqq \alpha \leqq \sum_{j=1}^{n-k} 1 = \frac{(n-k)(n-k+1)}{2}$$

of length $\frac{(n-k)(n-k+1)}{2}$. Furthermore, the prime factors of $a$ are less than $p$, thus for different pairs $a$, $\beta$, we obtain different products of form (41). Thus with respect to (40),

$$F(n, k) \leqq f(A, n, k) = F(k, k) \cdot \frac{(n-k)(n-k+1)}{2} <$$

$$< \exp\left(c_{22}\frac{k}{\log k}\right) \cdot \frac{n^2}{2} < n^2 \exp\left(c_{22}\frac{k}{\log k}\right)$$

which completes the proof of the second inequality in (39).

Now we are going to prove that the first inequality in (39) holds with $c_{20}=\frac{1}{92}$, in other words,

$$F(n, k) > n^2 \exp\left(\frac{1}{92}\frac{k}{\log^2 k}\right). \tag{42}$$

Let us assume at first that

$$n \leqq \exp\left(\frac{1}{3}\frac{k}{\log k}\right).$$

Then for large $k$, the right hand side of (42):

$$\begin{aligned} n^2 \exp\left(\frac{1}{92}\frac{k}{\log^2 k}\right) &\leqq \exp\left(\frac{2}{3}\frac{k}{\log k} + \frac{1}{92}\frac{k}{\log^2 k}\right) < \\ &< \exp\left(\frac{2}{3}\frac{k}{\log k} + \frac{1}{100}\frac{k}{\log k}\right) = \exp\left(\frac{68}{100}\frac{k}{\log k}\right). \end{aligned} \tag{43}$$

On the other hand, let $A$ denote any sequence satisfying (1). Let us form all those products of form (2) for which

$$\varepsilon_i = \begin{cases} 0 & \text{or 1 if } a_i \text{ is a prime numbes and } a_i \leqq k, \\ 0 & \text{otherwise.} \end{cases}$$

By (1), A contains all the $\pi(k)$ prime numbers $p \leqq k$, thus the number of these products is $2^{\pi(k)}$. Hence, by the prime number theorem, we have

(44) $$\bigl(F(n,k)\geqq\bigr)f(A,n,k)\geqq 2^{\pi(k)}=\exp\bigl(\log 2\pi(k)\bigr)>$$
$$>\exp\left(\frac{69}{100}\pi(k)\right)>\exp\left(\frac{68}{100}\frac{k}{\log k}\right).$$

(43) and (44) yield (42) in this case.

Let us assume now that

(45) $$n>\exp\left(\frac{1}{3}\frac{k}{\log k}\right).$$

Let
$$l=\left[\frac{1}{7}\frac{k}{\log^2 k}\right].$$

Denote the $i^{\text{th}}$ prime number by $p_i$ ($p_1=2, p_2=3, \dots$) and let $q_i=p_{i+1}$ for $i=1,2,\dots,l$, $Q=\{q_1,q_2,\dots,q_l\}$, $R=\{q_1,2q_1,q_2,2q_2,\dots,q_l,2q_l\}$. Obviously, (45) implies that $R\subset\{a_1,a_2,\dots,a_{[n/2]}\}$. Let us define the sequence $E=\{e_1,e_2,\dots,e_m\}$ by
$$\{a_1,a_2,\dots,a_{[n/2]}\}=E\cup R,\quad E\cap R=\emptyset.$$

For $s=1,2,\dots,\left[\frac{n}{4}\right]+1$, we denote the interval $[n-2[n/4]-1+2s,n]$ by $I_s$, and let $F_s$ denote the set of those products of form (2) for which
$$\varepsilon_i=0\quad\text{if}\quad a_i\in R,\quad \sum_{i:\,a_i\in B}\varepsilon_i=2,$$
$$\varepsilon_i=0\quad\text{if}\quad \left[\frac{n}{2}\right]<i\leqq n-2[n/4]-2+2s,$$
and
$$\varepsilon_l=1\quad\text{if}\quad i\in I_s\quad(\text{i.e. } n-2[n/4]-1+2s\leqq i\leqq n).$$

In other words, $F_s$ denotes the set of those numbers which can be written in form
$$\Bigl(\prod_{\mu\in I_s}a_\mu\Bigr)\cdot e_ie_j$$
where $1\leqq i$, $j\leqq m$, $i\neq j$. Let $F$ denote the set of those numbers which can be written in form
$$e_ie_j\quad\text{where}\quad 1\leqq i,\quad j\leqq m,\quad i\neq j.$$
Then obviously,

(46) $$|F_s|=|F|,$$

independently of $s$.

Furthermore, for $s=1, 2, \dots, \left[\frac{n}{4}\right]+1$, let $G_s$ denote the set of those products of form (2) for which

$$\varepsilon_i = 0 \text{ or } 1 \quad \text{if} \quad a_i \in R, \quad \sum_{i: a_i \in B} \varepsilon_i = 1,$$

$$\varepsilon_i = 0 \quad \text{if} \quad \left[\frac{n}{2}\right] < i \leqq n-2[n/4]-2+2s$$

and

$$\varepsilon_i = 1 \quad \text{if} \quad i \in I_s \quad (\text{i.e. } n-2[n/4]-1+2s \leqq i \leqq n).$$

In other words, $G_s$ denotes the set of those numbers which can be written in form

$$\Big(\prod_{\mu \in I_s} a_\mu\Big) \cdot e_i \prod_{j=1}^{l} q_j^{\varepsilon_j} \prod_{t=1}^{l} (2q_t)^{\varphi_t}$$

(where $\varepsilon_j = 0$ or 1, $\varphi_t = 0$ or 1). Then $|G_s|$ is equal to the number of the products of form

$$e_i \prod_{j=1}^{l} q_j^{\varepsilon_j} \prod_{t=1}^{l} (2q_t)^{\varphi_t} = 2^\alpha e_i \prod_{j=1}^{l} q_j^{\delta_j} \tag{47}$$

where

$$\delta_j = 0, 1 \quad \text{or} \quad 2 \tag{48}$$

and

$$0 \leqq \alpha \leqq l. \tag{49}$$

Let $G$ denote the set of those numbers which can be written in form

$$e_i \prod_{j=1}^{l} q_j^{\delta_j}$$

where (48) holds. Obviously, for any product of this form, there exist exponents $\varepsilon_j$, $\varphi_t$ and $\alpha$, satisfying (47), (49), $\varepsilon_j = 0$ or 1 and $\varphi_t = 0$ or 1. A product of form (47) can be obtained from at most $l+1$ distinct elements of $G$; namely, by (49), $\alpha$ may assume only at most $l+1$ distinct values. Thus

$$|G_s| \geqq \frac{|G|}{l+1} \tag{50}$$

(again, independently of $s$).

We are going to show that for $s \neq t$,

$$(F_s \cup G_s) \cap (F_t \cup G_t) = 0. \tag{51}$$

In fact, assume that $s > t$. Then for $y \in F_t \cup G_t$,

$$\begin{aligned} y &\geqq \prod_{\mu \in I_t} a_\mu = \prod_{n-2[n/4]-1+2t \leqq \mu < n-2[n/4]-1+2s} a_\mu \cdot \prod_{\mu \in I_s} a_\mu \geqq \\ &\geqq a_{n-2[n/4]-1+2t}\, a_{n-2[n/4]+2t} \cdot \prod_{\mu \in I_s} a_\mu > (a_{[n/2]})^2 \prod_{\mu \in I_s} a_\mu \quad (\text{for } y \in F_t \cup G_t). \end{aligned} \tag{52}$$

On the other hand, for $z \in F_s$,

(53) $$z = e_i e_j \prod_{\mu \in I_s} a_\mu \leqq (a_{[n/2]})^2 \prod_{\mu \in I_s} a_\mu \quad (\text{for } z \in F_s).$$

Finally, if $v \in G_t$, then we have

(54) $$v \leqq e_i \prod_{j=1}^{l} q_j \prod_{t=1}^{l} 2q_t \cdot \prod_{\mu \in I_s} a_\mu \leqq a_{[n/2]} \cdot 2^l \left( \prod_{j=1}^{l} q_j \right)^2 \cdot \prod_{\mu \in I_s} a_\mu .$$

By the prime number theorem,

$$\log \left( \prod_{i=1}^{x} p_i \right) \sim x \log x.$$

Thus if $k$ (and consequently $l$) are sufficiently large then with respect to (45) we have

$$2^l \left( \prod_{j=1}^{l} q_j \right)^2 = 2^l \left( \prod_{i=2}^{l+1} p_i \right)^2 < 2^l \left( \exp \left\{ \frac{35}{34} (l+1) \log (l+1) \right\} \right)^2 <$$

$$< \exp \left( \frac{1}{7} \frac{k}{\log^2 k} \cdot \log 2 \right) \exp \left\{ \frac{35}{17} \left( \frac{1}{7} \frac{k}{\log^2 k} + 1 \right) \log \left( \frac{1}{7} \frac{k}{\log^2 k} + 1 \right) \right\} <$$

$$< \exp \left( \frac{k}{\log^2 k} \right) \exp \left( \frac{5}{16} \frac{k}{\log^2 k} \log k \right) =$$

$$= \exp \left( \frac{k}{\log^2 k} + \frac{5}{16} \frac{k}{\log k} \right) < \frac{1}{3} \exp \left( \frac{5}{15} \frac{k}{\log k} \right) < \frac{1}{3} n < \left[ \frac{n}{2} \right] \leqq a_{[n/2]}.$$

Putting this into (54), we obtain that

(55) $$v \leqq (a_{[n/2]})^2 \prod_{\mu \in I_s} a_\mu \quad (\text{for } v \in G_s);$$

(52), (53) and (55) yield (51).

By (46), (50) and (51), we have

(56) $$f(A, n, k) \geqq \left| \bigcup_{s=1}^{[n/4]+1} (F_s \cup G_s) \right| = \sum_{s=1}^{[n/4]+1} |F_s \cup G_s| \geqq$$

$$\geqq \sum_{s=1}^{[n/4]+1} \max \{|F_s|, |G_s|\} \geqq \sum_{s=1}^{[n/4]+1} \max \left\{ |F|, \frac{|G|}{l+1} \right\} =$$

$$= ([n/4]+1) \max \left\{ |F|, \frac{|G|}{l+1} \right\} > \frac{n}{4} \frac{1}{l+1} \max \{|F|, |G|\}.$$

Thus to complete the proof of Theorem 2, we need a lower estimate for $\max \{|F|, |G|\}$. In the next section, we will prove the following lemma (using the same method as in [4]):

Lemma 1. *Let* $Q = \{q_1, q_2, \dots, q_l\}$ *be any set consisting of* $l$ *(distinct) prime numbers. Let* $E = \{e_1, e_2, \dots, e_m\}$ *(where* $e_1 < e_2 < \dots < e_m$*) be any sequence of positive* *integers. Let F and G denote the sets consisting of those integers which can be respectively written in form*

$$e_i e_j \quad (1 \leqq i, j \leqq m, \ i \neq j) \quad \textit{and} \quad e_i \prod_{j=1}^{l} q_j^{\delta_j} \quad (\delta_j = 0, 1 \text{ or } 2).$$

*Then for*

$$l > l_0, \tag{57}$$

*we have*

$$\max\{|F|, |G|\} > m \exp\left(\frac{2}{25} l\right). \tag{58}$$

Let us suppose now that Lemma 1 has been proved. Then the proof of Theorem 2 can be completed in the following way:

For large $k$, (57) holds by the definition of $l$. Thus we may apply Lemma 1. We obtain that (58) holds. Putting this into (56), we get that for large $k$ and any sequence $A$ (satisfying (1) and $|A|=n$),

$$f(A, n, k) > \frac{n}{4} \frac{1}{l+1} m \exp\left(\frac{2}{25} l\right). \tag{59}$$

With respect to (45),

$$m = |E| = [n/2] - |R| = [n/2] - 2l = \left[\frac{n}{2}\right] - 2\left[\frac{1}{7} \frac{k}{\log^2 k}\right] >$$

$$> \frac{n}{3} - \frac{2}{7} \frac{k}{\log^2 k} > \frac{n}{3} - \frac{1}{3} \frac{k}{\log k} > \frac{n}{3} - \log n > \frac{n}{4}.$$

Thus we obtain from (59) that for large $k$,

$$f(A, n, k) > \frac{n}{4} \frac{1}{l+1} \frac{n}{4} \exp\left(\frac{2}{25} l\right) > \frac{n^2}{16} \exp\left(\frac{2}{26} l\right) =$$

$$= \frac{n^2}{16} \exp\left\{\frac{1}{13}\left[\frac{1}{7} \frac{k}{\log^2 k}\right]\right\} > n^2 \exp\left(\frac{1}{92} \frac{k}{\log^2 k}\right)$$

which proves (42) and thus also Theorem 2.

**4.** To complete the proof of Theorem 2, we still have to give a

Proof of lemma 1. Let us write every $e \in E$ in form

$$e = (rs^2)(q_1^{\varepsilon_1} q_2^{\varepsilon_2} \dots q_l^{\varepsilon_l}) = bd \tag{60}$$

where $r$, $s$ are positive integers, $\varepsilon_i = 0$ or 1 (for $i = 1, 2, \dots, l$), $p/r$ implies that $p \notin Q$, $p/s$ implies that $p \in Q$ (also $r=1$ and $s=1$ may occur) and $b = rs^2$, $d = q_1^{\varepsilon_1} q_2^{\varepsilon_2} \dots q_l^{\varepsilon_l}$. Let us denote the occuring values of $b$ by $b_1, b_2, \dots, b_z$ ($b_i \neq b_j$

for $i \neq j$), let $B=\{b_1, b_2, \dots, b_z\}$ and let us denote the set of those numbers $e \in E$ for which $b=b_i$ in (60) (for fixed $i$, $1 \leqq i \leqq z$), by $E(b_i)$. Then obviously,

$$E = \bigcup_{i=1}^{z} E(b_i) \quad \text{and} \quad E(b_i) \cap E(b_j) = \emptyset \quad \text{for} \quad i \neq j,$$

thus

$$m = |E| = \sum_{i=1}^{z} |E(b_i)|. \tag{61}$$

For $b \in B$, let $F(b)$ denote the set of those numbers which can be written in form

$$e_x e_y \quad \text{where} \quad e_x \in E(b), \quad e_y \in E(b), \quad e_x \neq e_y.$$

Furthermore, for fixed $b \in B$ and for each $e_x = b q_1^{\varepsilon_1} q_2^{\varepsilon_2} \dots q_l^{\varepsilon_l}$, let us form all the products of form

$$e_x(q_1^{\gamma_1} q_2^{\gamma_2} \dots q_l^{\gamma_l}) = (b q_1^{\varepsilon_1} q_2^{\varepsilon_2} \dots q_l^{\varepsilon_l})(q_1^{\gamma_1} q_2^{\gamma_2} \dots q_l^{\gamma_l}) \tag{62}$$

where

$$\gamma_i = \begin{cases} 0 \ \text{ or } \ 1 & \text{if } \ \varepsilon_i = 1 \\ 1 \ \text{ or } \ 2 & \text{if } \ \varepsilon_i = 0 \end{cases}$$

and let us denote the set of these products by $G(b)$.

Obviously,

$$F \supset \bigcup_{i=1}^{z} F(b_i) \tag{63}$$

and

$$G \supset \bigcup_{i=1}^{z} G(b_i). \tag{64}$$

We are going to show that

$$F(b_i) \cap F(b_j) = \emptyset \quad \text{for} \quad i \neq j \tag{65}$$

and

$$G(b_i) \cap G(b_j) = \emptyset \quad \text{for} \quad i \neq j. \tag{66}$$

In fact, let us assume that

$$b_i = r_i s_i^2 \neq b_j = r_j s_j^2, \tag{67}$$

$$e_x = b_i q_1^{\varepsilon_1} q_2^{\varepsilon_2} \dots q_l^{\varepsilon_l} \in E(b_i), \quad e_y = b_i q_1^{\varphi_1} q_2^{\varphi_2} \dots q_l^{\varphi_l} \in E(b_i),$$

$$e_u = b_j q_1^{\alpha_1} q_2^{\alpha_2} \dots q_l^{\alpha_l} \in E(b_j) \quad \text{and} \quad e_v = b_j q_1^{\beta_1} q_2^{\beta_2} \dots q_l^{\beta_l} \in E(b_j).$$

Then

$$e_x e_y = r_i^2 s_i^4 q_1^{\varepsilon_1+\varphi_1} q_2^{\varepsilon_2+\varphi_2} \dots q_l^{\varepsilon_l+\varphi_l} \quad (\in F(b_i)) \tag{68}$$

and

$$e_u e_v = r_j^2 s_j^4 q_1^{\alpha_1+\beta_1} q_2^{\alpha_2+\beta_2} \dots q_l^{\alpha_l+\beta_l} \quad (\in F(b_j)). \tag{69}$$

If $r_i \neq r_j$ then there exists a prime power $p^\gamma$ such that $p \notin Q$ and $p^\gamma / e_x e_y$ but $p^\gamma \nmid e_x e_y$, or conversely; this implies that $e_x e_y \neq e_u e_v$. If $r_i = r_j$ then by (67), $s_i \neq s_j$ must hold. Thus there exists a prime power $q_t^\mu$ such that $q_t \in Q$ and $q_t^\mu / s_i$ but $q_t^\mu \nmid s_j$ (or conversely). Then the exponent of $q_t$ is at least $4\mu + \varepsilon_i + \varphi_i \geqq 4\mu$ in the canonical form of $e_x e_y$ and at most $4(\mu - 1) + \alpha_i + \beta_i \leqq 4\mu - 2$ in the canonical form of $e_u e_v$, thus $e_x e_y \neq e_u e_v$ holds also in this case, which proves (65).

In order to prove (66), note that we may write the product (62) in form

$$r(s^2 q_1 q_2 \dots q_l) q_1^{\alpha_1} q_2^{\alpha_2} \dots q_l^{\alpha_l} \quad \text{where} \quad \alpha_i = 0 \text{ or } 1 \quad \text{for} \quad i = 1, 2, \dots, l.$$

Obviously, a number of this form uniquely determines each of the factors $r, s, q_1^{\alpha_1}, \dots, q_l^{\alpha_l}$, which proves (66).

(63), (64), (65) and (66) imply that

$$\max\{|F|, |G|\} \geqq \max\left\{\left|\bigcup_{i=1}^{z} F(b_i)\right|, \left|\bigcup_{i=1}^{z} G(b_i)\right|\right\} = \tag{70}$$

$$= \max\left\{\sum_{i=1}^{z} |F(b_i)|, \sum_{i=1}^{z} |G(b_i)|\right\} \geqq \frac{1}{2}\left(\sum_{i=1}^{z} |F(b_i)| + \sum_{i=1}^{z} |G(b_i)|\right) =$$

$$= \frac{1}{2} \sum_{i=1}^{z} \left(|F(b_i)| + |G(b_i)|\right) \geqq \frac{1}{2} \sum_{i=1}^{z} \max\{|F(b_i)|, |G(b_i)|\}.$$

Thus in order to prove (58), it suffices to show that for $b \in B$, $\max\{|F(b)|, |G(b)|\}$ is large.

Let us assume that $b \in B$. We have to distinguish two cases.

*Case 1:*

$$(0 <) |E(b)| \leqq 2^{\frac{7}{8}l - 1}. \tag{71}$$

We are going to show that in this case $|G(b)|$ is large (in terms of $|E(b)|$). Let us fix an element $e_x$ of $E(b)$ and for this $e_x$, form all the products of form (62). Obviously, the factor $q_1^{\gamma_1} q_2^{\gamma_2} \dots q_l^{\gamma_l}$ can be chosen in $2^l$ ways thus the number of these products is $2^l$. Hence, with respect to (71),

$$|G(b)| \geqq 2^l = 2^{\frac{1}{8}l + 1} \cdot 2^{\frac{7}{8}l - 1} = 2^{\frac{1}{8}l + 1} |E(b)|. \tag{72}$$

*Case 2:*

$$|E(b)| > 2^{\frac{7}{8}l - 1}. \tag{73}$$

In this case, we shall need the following lemma:

Lemma 2. *Let $\varrho$ be any real number, satisfying*

$$0 < \varrho < \frac{1}{2} \tag{74}$$

*and*

$$f(\varrho) \overset{\text{def}}{=} -\varrho \log \varrho - (1-\varrho)\log(1-\varrho) - \left(1-\frac{\varrho}{2}\right)\log 2 < 0, \tag{75}$$

*and let $l$ be any integer, sufficiently large depending on $\varrho$:*

$$l > l_1(\varrho). \tag{76}$$

*Put*

$$\varphi(l) = 2^{-\frac{\varrho}{2}l-1}.$$

*Let $S$ denote the set of the $2^l$ $l$-tuples $(\mu_1, \mu_2, \dots, \mu_l)$, satisfying $\mu_h = 0$ or 1 for $h = 1, 2, \dots, l$. Let $R$ be any subset of $S$ for which*

$$|R| > \varphi(l)2^l. \tag{77}$$

*Then the number of the distinct sums of form*

$$(\mu_1+\nu_1, \dots, \mu_l+\nu_l) = (\mu_1, \dots, \mu_l) + (\nu_1, \dots, \nu_l), \tag{78}$$

*where $(\mu_1, \dots, \mu_l) \in R$ and $(\nu_1, \dots, \nu_l) \in R$, is greater than $(\varphi(l))^{-1}|R|$.*

This lemma is identical with Lemma 2 in [4].

Using Lemma 2, we are going to show that (73) implies that $|F(b)|$ is large.

Let us choose $\varrho = \frac{1}{4}$ in Lemma 2. Then (74) holds trivially, and a simple computation shows that

$$f\left(\frac{1}{4}\right) = \frac{3}{8}(\log 8 - \log 9) < 0,$$

thus $\varrho$ satisfies also (75). Furthermore, we choose $R$ as the set of those $l$-tuples $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l)$ (where $\varepsilon_i = 0$ or 1) for which $bq_1^{\varepsilon_1}q_2^{\varepsilon_2}\dots q_l^{\varepsilon_l} \in E(b)$ holds. Then by (73), also (77) holds:

$$|R| = |E(b)| > 2^{\frac{7}{8}l-1} = 2^{-\frac{1}{8}l-1}\cdot 2^l = \varphi(l)2^l.$$

Thus we may apply Lemma 2. We obtain that the number of the distinct sums of form (78) (where $(\mu_1, \dots, \mu_l) \in R$ and $(\nu_1, \dots, \nu_l) \in R$) is greater than $(\varphi(l))^{-1}|R|$. But distinct sums of form (78) determine distinct products of form

$$e_x e_y = (bq_1^{\mu_1}\dots q_l^{\mu_l})(bq_1^{\nu_1}\dots q_l^{\nu_l}) = b^2 q_1^{\mu_1+\nu_1}\dots q_l^{\mu_l+\nu_l},$$

and with at most $|E(b)|$ exception, also $e_x \neq e_y$ holds. Thus

$$\begin{aligned}|F(b)| &> (\varphi(l))^{-1}|R| - |E(b)| = \left(2^{-\frac{1}{8}l-1}\right)^{-1}|E(b)| - |E(b)| = \\ &= \left(2^{\frac{1}{8}l+1}-1\right)|E(b)| > 2^{\frac{1}{8}l}|E(b)|.\end{aligned} \tag{79}$$

(72) and (79) yield that for any $b \in B$.

$$\max\{|F(b)|, |G(b)|\} > 2^{\frac{1}{8}l}|E(b)|.$$

Putting this into (70), we obtain (with respect to (61)) that

$$\max\{|F|, |G|\} \geqq \frac{1}{2}\sum_{i=1}^{z} \max\{|F(b_i)|, |G(b_i)|\} >$$

$$> \frac{1}{2}\sum_{i=1}^{z} 2^{\frac{1}{8}l}|E(b_i)| = 2^{\frac{1}{8}l-1}\sum_{i=1}^{z}|E(b_i)| = m2^{\frac{1}{8}l-1} =$$

$$= m\exp\left\{\log 2\left(\frac{1}{8}l-1\right)\right\} > m\exp\left\{\left(\frac{\log 2}{8}-\frac{1}{1000}\right)l\right\} > m\exp\left(\frac{2}{25}l\right)$$

which completes the proof of Lemma 1.

## References

[1] P. Erdős, An asymptotic inequality in the theory of numbers, *Vestnik Leningrad. Univ.*, **15**: 13 (1960), 41—49 (Russian).

[2] P. Erdős and M. Kac, The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.*, **62** (1940), 738—742.

[3] P. Erdős and A. Sárközy, Some solved and unsolved problems in combinatorial number theory, *Mat. Slovaca*, to appear.

[4] A. Sárközy, On products of integers, *Studia Sci. Math. Hung.*, **9** (1974), 161—171.

MATH. INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
REÁLTANODA U. 13—15.
1053 BUDAPEST, HUNGARY