ON THE PRIME FACTORS OF $\binom{n}{k}$ AND OF CONSECUTIVE INTEGERS

P. Erdös and A. Sárközy

1. First we introduce some notations.

$c_1, c_2, \ldots$ denote positive absolute constants. The number of elements of the finite set S is denoted by $|S|$. We write $e^x = \exp x$. We denote the i^{th} prime number by p_i. $\nu(n)$ denotes the number of distinct prime factors of n. m_k is the smallest integer m for which

$$\nu\left(\binom{m}{k}\right) > k .$$

$P(n)$ is the greatest, $p(n)$ the smallest prime factor of n. n_k is the smallest integer n for which

$$P(n+i) > k \text{ for all } 1 \le i \le k .$$

In [3], Erdös, Gupta, and Khare proved that

$$m_k > c_1 k^2 \log k . \tag{1}$$

We prove

THEOREM 1. *For* $k > k_0$ *we have*

$$m_k > c_2 k^2 (\log k)^{4/3} (\log\log k)^{-4/3} (\log\log\log k)^{-1/3} . \tag{2}$$

It seems certain that $m_k > k^{2+c}$. It follows from results of Selfridge and the first author [4] that, by an averaging process, $m_k < k^{e+\varepsilon}$. The exact determination of, or even good inequalities for, m_k seems very difficult.

Denote by m'_k the least integer such that for every $m \ge m'_k$

$$\nu\left(\binom{m}{k}\right) \ge k .$$

Szemerédi and the first author proved in [5] that $m'_k < (e+\varepsilon)^k$ for every $\varepsilon > 0$ if $k > k_0$. Very likely $m'_k < k^c$ but we could not even prove $m'_k < (e-\varepsilon)^k$. Schinzel conjectured that $\nu\left(\binom{m}{k}\right) = k$ holds for

UTILITAS MATHEMATICA Vol. 16 (1979), pp. 197-215.

infinitely many values of m. This is almost certainly true, but will be hopelessly difficult to prove. Let $g(n)$ be the smallest integer $t \ge 1$ (if it exists) for which $n-t$ has more than one prime factor greater than t (or the same definition for $n+t$); surely $g(n)$ exists for both definitions but it is not known (see [4]).

In a previous paper [2, p.273], P. Erdös outlined the proof of

$$n_k < k^{\log k / \log\log k} . \tag{3}$$

No reasonable lower bound for n_k was known. We prove

THEOREM 2. *If* $k > k_0$, *then* $n_k > k^{5/2}/16$.

It seems certain that for $k > k_0(t)$, $n_k > k^t$ and, in fact, perhaps the upper bound (3) is close to the "truth".

Some other related problems and results can be found in [1].

2. *Proof of Theorem 1.*

We need two lemmas.

LEMMA 1. *If* $x \ge 0$, $y \ge 2$ *and* d *is a positive integer, then the number of solutions of*

$$p-q = d,\ x < p \le x+y ,$$

where p, q *are prime numbers, is less than*

$$c_3 \prod_{p/d} \left(1 + \frac{1}{p}\right) \frac{y}{\log^2 y} < c_4 \log\log (d+2) \frac{y}{\log^2 y} .$$

This lemma can be proved by Brun's sieve (see e.g. [6] or [7]).

LEMMA 2. *Let* $N \ge 2$ *be a positive integer and* $s_1 < s_2 < \ldots < s_N$ *be any integers. Put* $s_N - s_1 = S$ *and, for fixed* d, *denote the number of solutions of*

$$s_x - s_y = d$$

by $F(d)$. *Then there exists a positive integer* D *such that*

$$D \le \frac{4S}{N} \ \ \textit{and} \ \ F(D) \ge \frac{N^2}{48S} .$$

Proof of Lemma 2.

For $j = 1, 2, \ldots, [N/4] + 1$, let $I_j = [s_1+(j-1)y, s_1+jy]$. Then

$$s_1 + \left(\left|\frac{S}{y}\right| + 1\right)y > s_1 + \frac{S}{y}\, y = s_1+S = s_N ,$$

and thus

$$\bigcup_{j=1}^{[S/y]+1} I_j \supset [s_1, s_N] ;$$

hence

$$(4) \qquad \sum_{\substack{1\le j\le [N/4]+1 \\ N_j\ge 2}} N_j = \sum_{1\le j\le [N/4]+1} N_j - \sum_{\substack{1\le j\le [N/4]+1 \\ N_j=1}} N_j$$

$$= N - ([N/4] + 1) \ge N - \frac{N}{2} = \frac{N}{2}$$

since

$$[\tfrac{N}{4}] + 1 \le \frac{N}{4} + \frac{N}{4} = \frac{N}{2} \quad \text{for} \quad N \ge 4$$

and

$$[\tfrac{N}{4}] + 1 \le \frac{N}{2}$$

holds also for $N = 2, 3$.

Let $N_j \ge 2$ for some j and define z by $s_z \notin I_j$, $s_{z+1} \in I_j$. Then for all the $\binom{N_j}{2}$ pairs, $1 \le x < y \le N_j$, we have

$$1 \le s_{z+y} - s_{z+x} \le s_z + N_j - s_{z+1} \le 4S/N .$$

Thus, with respect to (4) and using Cauchy's inequality, we obtain that

$$\sum_{1\le d\le 4S/N} F(d) \ge \sum_{\substack{1\le j\le [N/4]+1 \\ N_j\ge 2}} \binom{N_j}{2}$$

$$= \sum_{\substack{1\le j\le [N/4]+1 \\ N_j\ge 2}} N_j \frac{N_j - 1}{2} \ge \sum_{\substack{1\le j\le [N/4]+1 \\ N_j\ge 2}} N_j \frac{N_j/2}{2}$$

$$= \frac{1}{4} \sum_{\substack{1\le j\le [N/4]+1 \\ N_j \ge 2}} N_j^2 \ge \frac{1}{4} \left(\sum_{\substack{1\le j\le [N/4]+1 \\ N_j \ge 2}} N_j \right)^2 \Bigg/ \sum_{\substack{1\le j\le [N/4]+1 \\ N_j \ge 2}} 1$$

$$\ge \frac{1}{4}\left(\frac{N}{2}\right)^2 \Bigg/ \sum_{1\le j\le [N/4]+1} 1 \ge \frac{N^2}{16} \Bigg/ \left(\frac{N}{4}+1\right) \ge \frac{N^2}{16} \Bigg/ \frac{3N}{4} = \frac{N}{12} ;$$

hence

$$\max_{1\le d\le 4S/N} F(d) \ge \frac{\sum_{1\le d\le 4S/N} F(d)}{\sum_{1\le d\le 4S/N} 1} \ge \frac{N/12}{4S/N} = \frac{N^2}{48S} ,$$

which completes the proof of the lemma.

In order to prove Theorem 1, it suffices to show that if ε is sufficiently small then the assumptions

(5) $$\nu\left(\binom{m}{k}\right) > k ,$$

(6) $$m \ge \varepsilon k^2 (\log k)^{4/3} (\log \log k)^{-4/3} (\log \log \log k)^{-1/3} ,$$

lead to a contradiction for all $k \ge k_0(\varepsilon)$. By (1) we may assume that

(7) $$m > c_1 k^2 \log k$$

also holds.

For all $0 \le i \le k-1$, we write

$$m-i = a_i b_i ,$$

where $P(a_i) \le k$ and $p(b_i) > k$ (if, for example, all the prime factors of $m-i$ are less than or equal to k then we write $b_i = 1$). Furthermore, let S_1 denote the set consisting of the integers $0 \le i \le k-1$ such that

$$P(m-i) < \frac{k \log k}{\log \log k}$$

and let $S_2 = \{0,1,\dots,k-1\} - S_1$ (i.e. $i \in S_2$ if and only if $||0 \le i \le k-1$ and $P(m-i) > k \log k (\log \log k)^{-1}$).

First we give a lower estimate for $\prod_{i \in S_2} a_i \cdot \binom{m}{k}$ is an integer; thus

$$k! \;\Big|\; m(m-1)\dots(m-k+1) = \prod_{i=0}^{k-1} a_i \prod_{i=0}^{k-1} b_i ,$$

and, obviously, $\left(k!, \prod_{i=0}^{k-1} b_i\right) = 1$. Thus we have $k! \;\Big|\; \prod_{i=0}^{k-1} a_i$

and hence by Stirling's formula,

$$\prod_{i=0}^{k} a_i \ge k! > \left(\frac{k}{3}\right)^k \tag{8}$$

(for sufficiently large k).

From (5), we have

$$k < \nu\left(\binom{m}{k}\right) \le (\nu(m(m-1)\dots(m-k+1)) \tag{9}$$

$$= \sum_{\substack{p \le k \\ p | m(m-1)\dots(m-k+1)}} 1 + \sum_{\substack{k < p \le k \log k (\log\log k)^{-1} \\ p | m(m-1)\dots(m-k+1)}} 1$$

$$+ \sum_{\substack{k \log k (\log\log k)^{-1} < p \\ p | m(m-1)\dots(m-k+1)}} 1 .$$

Here the first term is less than or equal to $\pi(k)$. Furthermore, if $k \log k (\log\log k)^{-1} < p$ and $p \mid m(m-1)\dots(m-k+1)$ then $p | m-i$ for some $i \in S_2$, and by (6), we have

$$\frac{m-i}{p} \le \frac{m}{k \log k (\log\log k)^{-1}}$$

$$\le \frac{\varepsilon k^2 (\log k)^{4/3} (\log\log k)^{-4/3} (\log\log\log k)^{-1/3}}{k \log k (\log\log k)^{-1}}$$

$$= \varepsilon k (\log k)^{1/3} (\log\log k)^{-1/3} (\log\log\log k)^{-1/3}$$

$$< k \log k (\log\log k)^{-1} .$$

Thus for all $i \in S_2$, $m-i$ has only one prime factor greater than $k \log k (\log \log k)^{-1}$. This implies that the third term in (9) is less than or equal to $|S_2|$. Finally, we write

$$\sum_{\substack{k<p\le k \log k (\log \log k)^{-1} \\ p | m(m-1)\dots(m-k+1)}} 1 = R .$$

We obtain from (9) that

$$k \le \pi(k)+R+|S_2| ;$$

hence

$$R \ge (k-|S_2|)- \pi(k) = |S_1| - \pi(k) . \tag{10}$$

Obviously, we have

$$\prod_{i \in S_1} a_i = \prod_{i \in S_1} \frac{m-i}{b_i} \le \frac{m^{S_1}}{\prod_{\substack{k<p\le k \log k (\log \log k)^{-1} \\ p | m(m-1)\dots(m-k+1)}} p} \tag{11}$$

$$\le m^{|S_1|} \Big/ \prod_{i=\pi(k)+1}^{i=\pi(k)+R} p_i$$

By the prime number theorem, we have $p_i \sim i \log i$. Furthermore, by the prime number theorem and the definition of R,

$$R \le \pi(k \log k (\log \log k)^{-1}) < k$$

(for large k). Thus by (10) and using the prime number theorem and Stirling's formula, we obtain, for sufficiently large k and $|S_1| > c_8$, that

$$\prod_{i=\pi(k)+1}^{\pi(k)+R} p_i > \prod_{i=\pi(k)+1}^{\pi(k)+R} \left(1 - \frac{1}{11}\right) i \log i \tag{12}$$

$$\ge \prod_{i=\pi(k)+1}^{\pi(k)+R} \left(1 - \frac{1}{11}\right) i \log (\pi(k)+1)$$

$$> \prod_{i=\pi(k)+1}^{\pi(k)+R} \left(1 - \frac{1}{10}\right) i \log k = \left(\frac{9}{10}\right)^R (\log k)^R \frac{(\pi(k)+R)!}{(\pi(k))!}$$

$$> \left(\frac{9}{10}\right)^k (\log k)^{|S_1|-\pi(k)} \frac{|S_1|!}{\pi(k)^{\pi(k)}}$$

$$> \left(\frac{9}{10}\right)^k (\log k)^{|S_1|} (\pi(k) \log k)^{-\pi(k)} \frac{|S_1|^{|S_1|}}{3^{|S_1|}}$$

$$> \left(\frac{9}{10}\right)^k (|S_1| \log k)^{|S_1|} e^{-\pi(k) \log (\pi(k) \log k)} 3^{-k}$$

$$> \left(\frac{9}{10}\right)^k (|S_1| \log k)^{|S_1|} e^{-2k} 3^{-k}$$

$$> \left(\frac{9}{10}\right)^k (|S_1| \log k)^{|S_1|} 9^{-k} 3^{-k} = 30^{-k} (|S_1| \log k)^{|S_1|} .$$

(11) and (12) yield that, for $|S_1| > c_8$,

$$\prod_{i\in S_1} a_i < \frac{m^{|S_1|}}{30^{-k}(|S_1|\log k)^{|S_1|}} = k^{|S_1|} 30^k \left(\frac{m}{|S_1| k \log k}\right)^{|S_1|} . \tag{13}$$

By (10) and the prime number theorem, we have

$$|S_1| \le R+\pi(k) = \sum_{\substack{k<p\le k \log k(\log \log k)^{-1} \\ p|m(m-1)\dots(m-k+1)}} 1 + \pi(k) \tag{14}$$

$$\le \sum_{k<p\le k \log k(\log \log k)^{-1}} 1 + \pi(k)$$

$$= (\pi(k \log k (\log \log k)^{-1}) - \pi(k)) + \pi(k)$$

$$= \pi(k \log k (\log \log k)^{-1}) < 2k(\log \log k)^{-1} .$$

If $a > 0$ then an easy discussion shows that the function $f(x) = (a/x)^x$ is increasing in the interval $0 < x < a/e$. If we put $a = m/(k \log k)$ and $x = |S_1|$ then from (7) and (14) we have

$$\frac{a}{e} = \frac{m}{e k \log k} > \frac{c_1 k^2 \log k}{e k \log k} = \frac{c_1}{e} k > 2k(\log \log k)^{-1} > |S_1| = x$$

for sufficiently large k. Thus, for large k and $|S_1| > c_8$, (13) and (14) yield that

$$(15)\qquad \prod_{i\in S_1} a_i < k^{|S_1|}\, 30^k \left(\frac{m}{|S_1| k \log k}\right)^{|S_1|}$$

$$< k^{|S_1|}\, 30^k \left(\frac{m}{2k(\log\log k)^{-1}\, k \log k}\right)^{2k(\log\log k)^{-1}}$$

$$< k^{|S_1|} 30^k \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^{2k(\log\log k)^{-1}}.$$

For large k, this holds also for $0 \le |S_1| \le c_8$ since in this case, from (6) and (7) we have

$$k^{|S_1|} 30^k \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^{2k(\log\log k)^{-1}}$$

$$> 30^k \left(\frac{c_1 k^2 \log k}{k^2 \log k\, (\log\log k)^{-1}}\right)^{2k(\log\log k)^{-1}} > 30^k ,$$

while

$$\prod_{i\in S_1} a_i < \prod_{i\in S_1} m = m^{|S_1|} < (k^3)^{|S_1|} \le k^{3c_8} < 30^k$$

for sufficiently large k.

(8) and (15) yield that

$$(16)\qquad \prod_{i\in S_2} a_i = \prod_{i=0}^{k-1} a_i \left(\prod_{i\in S_1} a_i\right)^{-1}$$

$$> \left(\frac{k}{3}\right)^k \left(k^{|S_1|} 30^k \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^{2k(\log\log k)^{-1}}\right)^{-1}$$

$$= k^{k-|S_1|}\, 90^{-k} \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^{-2k(\log\log k)^{-1}}$$

$$= k^{|S_2|}\, 90^{-k} \left(\frac{m}{k^2 \log k(\log\log k)^{-1}}\right)^{-2k(\log\log k)^{-1}} .$$

Let S_3 denote the set consisting of the integers i satisfying $0 \le i \le k-1$, $i \in S_2$, and $a_i \le 10^{-6}k$. Then we have

$$(17) \qquad \prod_{i\in S_2} a_i = \prod_{i\in S_2, i\notin S_3} a_i \prod_{i\in S_3} a_i$$

$$\le \prod_{i\in S_2, i\notin S_3} 10^{-6}k \prod_{i\in S_3} \frac{m-i}{b_i}$$

$$\le \prod_{i\in S_2, i\notin S_3} 10^{-6}k \prod_{i\in S_3} k\,\frac{m}{kP(m-i)}$$

$$\le k^{|S_2|} \prod_{i\in S_2, i\notin S_3} 10^{-6} \prod_{i\in S_3} \frac{m}{k^2 \log k\, (\log\log k)^{-1}}$$

$$= k^{|S_2|} (10^{-6})^{|S_2|-T} \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^T$$

$$\le k^{|S_2|} (10^{-6})^{k-T} \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^T,$$

where

$$T = \sum_{i\in S_3} 1 .$$

(16) and (17) yield that

$$k^{|S_2|} 90^{-k} \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^{-2k(\log\log k)^{-1}}$$

$$< k^{|S_2|} (10^{-6})^{k-T} \left(\frac{m}{k^2 \log k\, (\log\log k)^{-1}}\right)^T ;$$

hence

$$(18) \qquad (10^6)^{k-T} 90^{-k} < \left(\frac{m}{k^2 \log k (\log\log k)^{-1}}\right)^{T+2k(\log\log k)^{-1}}$$

We are going to show that this inequality implies that

$$(19) \qquad T > k(\log\log k)^{-1} .$$

If $T \ge k/2$ then this inequality holds trivially for large k; thus we may assume that $T < k/2$. Then we obtain from (18) that

$$\left(\frac{m}{k^2 \log k(\log\log k)^{-1}}\right)^{T+2k(\log\log k)^{-1}}$$

$$> (10^6)^{k-k/2}90^{-k} > 10^{3k}\, 10^{-2k} = 10^k .$$

Thus, from (6), we have

$$T+2k(\log\log k)^{-1} > \frac{k \log 10}{\log \dfrac{m}{k^2 \log k(\log\log k)^{-1}}}$$

$$\geq \frac{k \log 10}{\log \dfrac{\varepsilon k^2 (\log k)^{4/3}(\log\log k)^{-4/3}(\log\log\log k)^{-1/3}}{k^2 \log k\,(\log\log k)^{-1}}}$$

$$= \frac{k \log 10}{\log\,(\varepsilon(\log k)^{1/3}(\log\log k)^{-1/3}(\log\log\log k)^{-1/3})}$$

$$> \frac{k \log 10}{\log\,(\log k)^{1/3}} = \frac{3k \log 10}{\log\log k} > 3k(\log\log k)^{-1}$$

for sufficiently large k, which yields (19).

By the definition of a_i, we have $a_i | m-i$. If $i \in S_3$, then $a_i > 10^{-6}k$ also holds. But, if $a > 10^{-6}k$, then a may have at most

$$\left|\frac{k}{a}\right|+1 \leq \frac{k}{a} + 1 < 10^6 + 1 < 2\cdot 10^6$$

multiples amongst the numbers $m, m-1, \ldots, m-k+1$. Thus, for fixed a,

$$a_i = a, \quad 1 \in S_3 ,$$

has less than $2\cdot 10^6$ solutions. From (19), this implies that the number of the distinct a_i's with $i \in S_3$ is greater than

$$\frac{\sum_{i\in S_3} 1}{2\cdot 10^6} = \frac{T}{2\cdot 10^6} > k(2\cdot 10^6 \log\log k)^{-1} .$$

By the definitions of the sets S_2 and S_3, $i \in S_3(\subset S_2)$ implies that

$$k \log k\,(\log\log k)^{-1} < P(m-i) \leq b_i ;$$

hence, from (6),

$$(20)\qquad (10^{-6}k <)a_i = \frac{m-i}{b_i} \le \frac{m}{k \log k (\log\log k)^{-1}}$$

$$\le \frac{\varepsilon k^2(\log k)^{4/3}(\log\log k)^{-4/3}(\log\log\log k)^{-1/3}}{k \log k (\log\log k)^{-1}}$$

$$= \varepsilon k (\log k)^{1/3}(\log\log k)^{-1/3}(\log\log\log k)^{-1/3}.$$

Let us write

$$t = k (\log k)^{-5/3} (\log\log k)^{2/3}(\log\log\log k)^{2/3}$$

and for all

$$i \le j \le \left[\frac{\varepsilon k(\log k)^{1/3}(\log\log k)^{-1/3}(\log\log\log k)^{-1/3}}{t}\right] + 1,$$

let us form the interval $I_j = ((j-1)t, jt]$. By (20), these intervals cover all the a_i's (with $i \in S_3$), and the number of these intervals is

$$\left[\frac{\varepsilon k(\log k)^{1/3}(\log\log k)^{-1/3}(\log\log\log k)^{-1/3}}{t}\right] + 1$$

$$= \left[\frac{\varepsilon k(\log k)^{1/3}(\log\log k)^{-1/3}(\log\log\log k)^{-1/3}}{k(\log k)^{-5/3}(\log\log k)^{2/3}(\log\log\log k)^{2/3}}\right] + 1$$

$$< \varepsilon(\log k)^2(\log\log k)^{-1}(\log\log\log k)^{-1} + 1$$

$$< 2\varepsilon(\log k)^2(\log\log k)^{-1}(\log\log\log k)^{-1}$$

(for large k). Thus the matchbox principle yields that there exists an interval I_j which contains more than

$$\frac{k(2\cdot 10^6 \log\log k)^{-1}}{2\varepsilon(\log k)^2(\log\log k)^{-1}(\log\log\log k)^{-1}}$$

$$= (4\cdot 10^6)^{-1}\varepsilon^{-1} k (\log k)^{-2} \log\log\log k$$

distinct a_i's. In other words, there exist indices $i_1, i_2, \dots, i_N$ satisfying $i_\ell \in S_3$ for $i \le \ell \le N$,

(21) $$(j-1)t < a_{i_1} < a_{i_2} < \ldots < a_{i_N} \le jt$$

and

(22) $$N > (4\cdot 10^6)^{-1}\varepsilon^{-1} k (\log k)^{-2} \log\log\log k .$$

Let us apply Lemma 2 with the set $a_{i_1}, a_{i_2}, \ldots, a_{i_N}$ in place of the set $s_1, s_2, \ldots, s_N$. (By (21), $N \ge 2$ holds trivially.) Then from (21), we have

$$S = s_N - s_1 = a_{i_N} - a_{i_1}$$
$$< t = k(\log k)^{-5/3}(\log\log k)^{2/3}(\log\log\log k)^{2/3} .$$

Thus, from (22), we obtain that there exists an integer

(23) $$D \le \frac{4S}{N} < \frac{4k(\log)\ k)^{-5/3}(\log\log k)^{2/3}(\log\log\log k)^{2/3}}{(4\cdot 10^6)^{-1}\ \varepsilon^{-1}\ k\ (\log k)^{-2}\ \log\log\log k}$$
$$= 16\cdot 10^6 \varepsilon(\log k)^{1/3}(\log\log k)^{2/3}(\log\log\log k)^{-1/3}$$

for which

(24) $$a_{i_x} - a_{i_y} = D$$

has at least

(25) $$F(D) \ge \frac{N^2}{48S} > \frac{(4\cdot 10^6)^{-2}\ \varepsilon^{-2}\ k^2(\log k)^{-4}\ (\log\log\log k)^2}{48k(\log k)^{-5/3}(\log\log k)^{2/3}(\log\log\log k)^{2/3}}$$
$$> 10^{-15}\ \varepsilon^{-2}\ k(\log k)^{-7/3}(\log\log k)^{-2/3}(\log\log\log k)^{4/3}$$

solutions.

By the definition of the set S_3, $i_\ell \in S_3$ implies, from (6), that for large k

$$b_{i_\ell} = \frac{m-1_\ell}{a_{i_\ell}} < \frac{m}{10^{-6}k}$$

$$\le 10^6\ \varepsilon\ k(\log k)^{4/3}(\log\log k)^{-4/3}(\log\log\log k)^{-1/3}$$

$$< k(\log k)^{4/3}$$

and, on the other hand,

$$b_{i_\ell} \ge P(m-i_\ell) > k \log k\ (\log\log k)^{-1}\ . \tag{26}$$

By the definition of b_{i_ℓ}, we have $p(b_{i_\ell}) > k$; thus these inequalities imply that $b_{i_\ell} = q_\ell$ must be a prime number. Then $q_1, q_2, \ldots, q_N$ are distinct since by (26), $q_\ell > k \log k\ (\log\log k)^{-1}$ (for all ℓ). Thus q_ℓ may have at most one multiple amongst the numbers $m, m-1, \ldots, m-k+1$. Furthermore, we have

$$q_\ell = \frac{m-i_\ell}{a_{i_\ell}} < \frac{m}{(j-1)t}$$

(it can be shown easily that $j > 1$) and

$$q_\ell = \frac{m-i_\ell}{a_{i_\ell}} = \frac{m}{a_{i_\ell}} - \frac{i_\ell}{a_{i_\ell}} > \frac{m}{jt} - \frac{k}{10^{-6}k} = \frac{m}{jt} - 10^6\ .$$

Thus all the primes q_ℓ belong to the interval

$$I = \left(\frac{m}{jt} - 10^6,\ \frac{m}{(j-1)t}\right). \tag{27}$$

From $a_{i_\ell} > 10^{-6}k$, (5), and the definition of t, the length of this interval is

$$(28)\quad |I| = \frac{m}{(j-1)t} - \left(\frac{m}{jt} - 10^6\right) = \frac{m}{(j-1)jt} + 10^6$$

$$< \frac{mt}{(jt-t)jt} < \frac{mt}{(a_{i_\ell}-t)a} < \frac{mt}{(10^{-6}k-t)10^{-6}k} < 2\cdot 10^{-12}mtk^{-2}$$

$$< 2\cdot 10^{-12}\varepsilon k^2(\log k)^{4/3}(\log\log k)^{-4/3}(\log\log\log k)^{-1/3}$$

$$\cdot k(\log k)^{-5/3}(\log\log k)^{2/3}(\log\log\log k)^{2/3}\, k^{-2}$$

$$= 2\cdot 10^{-12}\varepsilon k(\log k)^{-1/3}(\log\log k)^{-2/3}(\log\log\log k)^{1/3}\ .$$

Furthermore, if a_{i_x} and a_{i_y} satisfy (24) then we have

$$q_y - q_x = \frac{m-i_y}{a_{i_y}} - \frac{m-i_x}{a_{i_x}} = m\,\frac{a_{i_x}-a_{i_y}}{a_{i_x}a_{i_y}} + \frac{i_x}{a_{i_x}} - \frac{i_y}{a_{i_y}}$$

$$= \frac{mD}{a_{i_y}(a_{i_y}+D)} + \frac{i_x}{a_{i_x}} - \frac{i_y}{a_{i_y}}\ .$$

Hence

$$(29)\qquad \left|(q_y-q_x) - \frac{mD}{a_{i_y}(a_{i_y}+D)}\right| \le \frac{i_y}{a_{i_x}} + \frac{i_y}{a_{i_y}} \le 2\cdot\frac{k}{10^{-6}k} = 2\cdot 10^6\ .$$

On the other hand, by (6) and (23), we have

$$(30)\qquad \left|\frac{mD}{a_{i_y}(a_{i_y}+D)} - \frac{mD}{jt(jt+D)}\right| = mD\,\frac{\left|\left((jt)^2-(a_{i_y})^2\right)+D(jt-a_{i_y})\right|}{a_{i_y}(a_{i_y}+D)jt(Jt+D)}$$

$$= mD\frac{\left|jt-a_{i_y}\right|\ jt+a_{i_y}+D}{a_{i_y}(a_{i_y}+D)jt(jt+D)} = \frac{mD\left|jt-a_{i_y}\right|(jt+a_{i_x})}{a_{i_y}a_{i_x}jt(jt+D)}$$

$$< \frac{mD\cdot t\cdot 2jt}{a_{i_y}a_{i_x}jt\ a_{i_x}} < \frac{2mDt}{(10^{-6}k)^3} = 2\cdot 10^{18}mDtk^{-3}$$

$$< 2\cdot 10^{18}\varepsilon k^2(\log k)^{4/3}(\log\log k)^{-4/3}(\log\log\log k)^{-1/3}$$

$$\cdot 16\cdot 10^6\varepsilon(\log k)^{1/3}(\log\log k)^{2/3}(\log\log\log k)^{-1/3}$$

$$\cdot k(\log k)^{-5/3}(\log\log k)^{2/3}(\log\log\log k)^{2/3}\, k^{-3}$$

$$= 32\cdot 10^{24}\varepsilon^2.$$

For small ε, (29) and (30) yield that

$$\left|(q_y-q_x) - \frac{mD}{jt(jt+D)}\right| \le \left|(q_y-q_x) - \frac{mD}{a_{i_y}(a_{i_y}+D)}\right| + \left|\frac{mD}{a_{i_y}(a_{i_y}+D)} - \frac{mD}{jt(jt+D)}\right|$$

$$< 2\cdot 10^6+32\cdot 10^{24}\varepsilon^2 < 3\cdot 10^6 .$$

Thus for all the $F(D)$ solutions x, y of (24), q_y-q_x is in the interval

$$\frac{mD}{jt(jt+D)} - 3\cdot 10^6 < q_y-q_x < \frac{mD}{jt(jt+d)} + 3\cdot 10^6 .$$

The length of this interval is $6\cdot 10^6$; thus, by the matchbox principle, there exists an integer d such that (from (6) and (23))

$$(31)\qquad d < \frac{mD}{jt(jt+d)} + 3\cdot 10^6 < \frac{mD}{(10^{-6}k)^2} + 3\cdot 10^6$$

$$< 10^{12}\varepsilon k^2(\log k)^{4/3}(\log\log k)^{-4/3}(\log\log\log k)^{-1/3}$$

$$\cdot 16\cdot 10^6\varepsilon(\log k)^{1/3}(\log\log k)^{2/3}(\log\log\log k)^{-1/3}\cdot k^{-2}$$

$$= 16\cdot 10^{18}\varepsilon^2(\log k)^{5/3}\log\log k)^{-2/3}(\log\log\log k)^{-2/3}$$

and, denoting the number of solutions of

$$(32)\qquad q_x-q_y = d,\quad q_x \in I,\quad q_y \in I$$

by $G(d)$, we have from (25)

$$(33)\qquad G(d) \ge \frac{F(D)}{6\cdot 10^6+1}$$

$$> \frac{10^{-15}\varepsilon^{-2}k(\log k)^{-7/3}(\log\log k)^{-2/3}(\log\log\log k)^{4/3}}{6\cdot 10^6+1}$$

$$> 10^{22}\varepsilon^{-2}k(\log k)^{-7/3}(\log\log k)^{-2/3}(\log\log\log k)^{4/3} .$$

On the other hand, by (28), Lemma 1 yields that the number of solutions of (32) is

(34) $G(d) < c_4 \log\log(d+2)\dfrac{|I|}{\log^2|I|}$

$$< c_4 \log\log(16\cdot 10^{18}\varepsilon^2(\log k)^{5/3}(\log\log k)^{-2/3}(\log\log\log k)^{-2/3}+2)$$

$$\cdot \frac{2\cdot 10^{-12}\varepsilon k(\log k)^{-1/3}(\log\log k)^{-2/3}(\log\log\log k)^{1/3}}{\log^2(2\cdot 10^{-12}\varepsilon k(\log k)^{-1/3}(\log\log k)^{-2/3}(\log\log\log k)^{1/3})}$$

$$< 2c_4 \log\log\log k \frac{2\cdot 10^{12}\varepsilon k(\log k)^{-1/3}(\log\log k)^{-2/3}(\log\log\log k)^{1/3}}{2^{-1}\log^2 k}$$

$$= 8\cdot 10^{-12}\varepsilon_4\, k(\log k)^{-7/3}(\log\log k)^{-2/3}(\log\log\log k)^{4/3} .$$

(33) and (34) yield that

$$10^{-22}\varepsilon^{-2}k(\log k)^{-7/3}(\log\log k)^{-2/3}(\log\log\log k)^{4/3}$$

$$< 8\cdot 10^{-12}\varepsilon\, c_4\, k(\log k)^{-7/3}(\log\log k)^{-2/3}(\log\log\log k)^{4/3};$$

hence

$$8^{-1}10^{-10}c_4^{-1} < \varepsilon^3 .$$

But for sufficiently small ε, this inequality cannot hold, and this contradiction proves Theorem 1. (In fact, as this inequality shows, Theorem 1 holds, for example, with $c_2 = 10^{-4}\, c_4^{-1/3}$.)

3. *Proof of Theorem 2.*

We have to prove that if $k > k_0$ and

(35) $$0 \le n \le k^{5/2}/16 ,$$

then there exists an integer u such that

(36) $$n < u \le n+k$$

and

(37) $$P(u) \le k .$$

We are going to show that there exists an integer u satisfying (36) and of the form

$$u = d_1 d_2 d_3 ,$$

where the integers d_i satisfy

(38) $$0 < d_i \le k \quad \text{for} \quad i = 1, 2, 3 .$$

Obviously, an integer u of this form also satisfies (37).

Let us define the positive integers x and z by

$$z = [k^{1/2}/3]$$

and

$$x = [(n/z)^{1/2}] + 1 .$$

Then

$$x^2 > ((n/z)^{1/2})^2 = n/z ;$$

hence

(39) $$x^2 - \frac{n}{z} > 0 .$$

Let us define the non-negative integer y by

(40) $$y < \left(x^2 - \frac{n}{z}\right)^{1/2} \le y+1 .$$

Finally, let

$$d_1 = z,$$

$$d_2 = x-y,$$

and

$$d_3 = x+y.$$

Then $d_1 > 0$ and $d_3 > 0$ hold trivially while $d_2 > 0$ will follow from $d_1 > 0$, $d_3 > 0$ and $u = d_1 d_2 d_3 > n$. Furthermore, $d_1 = z = [k^{1/2}/3] < k$ holds trivially. Finally, in view of (35) and (39), we have

$$
\begin{aligned}
d_2 = x-y \le x+y = d_3 &< \left(x + x^2 - \frac{n}{z}\right)^{1/2} \\
&< (n/z)^{1/2}+1 + \left(\left((n/z)^{1/2}+1\right)^2 - \frac{n}{z}\right)^{1/2} \\
&= (n/z)^{1/2}+1+(2(n/z)^{1/2}+1)^{1/2} \\
&< \left(\frac{k^{5/2}}{16[k^{1/2}/3]}\right)^{1/2} + 1 + \left(2\left(\frac{k^{5/2}}{16[k^{1/2}/3]}\right)^{1/2} + 1\right)^{1/2} \\
&= (1+o(1))3^{1/2}4^{-1}k+1+O(k^{1/2}) = (1+o(1))3^{1/2}4^{-1}k
\end{aligned}
$$

for $k \to +\infty$. Here $3^{1/2}4^{-1} < 1$; thus also $d_2 \le d_3 < k$ holds for large k which completes the proof of (38) (provided that $u > n$).

By (40), we have

$$
(41) \quad u = d_1 d_2 d_3 = z(x-y)(x+y) = z(x^2-y^2) > z\left(x^2 - \left(x^2 - \frac{n}{z}\right)^{1/2}\right)^2 = z \cdot \frac{n}{z} = n .
$$

Finally, by (40),

$$
x^2 - \frac{n}{z} \le (y+1)^2 ;
$$

hence

$$
z(x^2-(y+1)^2) \le n .
$$

Thus, from (35), we have

$$
\begin{aligned}
(42) \qquad u = d_1 d_2 d_3 &= z(x-y)(x+y) = z(x^2-y^2) \\
&= n + (z(x^2-y^2)-n) \le n+(z(x^2-y^2)-z(x^2-(y+1)^2)) \\
&= n+(2y+1)z < n+\left(2\left(x^2-\frac{n}{z}\right)^{1/2} + 1\right)z \\
&< n+\left(2\left(\left((n/z)^{1/2}+1\right)^2 - \frac{n}{z}\right)^{1/2} + 1\right)z \\
&= n+\left(2(2(n/z)^{1/2}+1)^{1/2}+1\right)z \\
&= n+\left(2\left(2\left(\frac{n}{[k^{1/2}/3]}\right)^{1/2} + 1\right)^{1/2} + 1\right)\frac{k^{1/2}}{3} \\
&< n+\left(2\left(2\left(\frac{k^{5/2}/16}{k^{1/2}/4}\right)^{1/2} + 1\right)^{1/2} + 1\right)\frac{k^{1/2}}{3} \\
&= n+\left(2(k+1)^{1/2}+1\right)\frac{k^{1/2}}{3} < n+3k^{1/2} \cdot \frac{k^{1/2}}{3} = n+k
\end{aligned}
$$

for sufficiently large k.

(41) and (42) yield (36) and this completes the proof of Theorem 2.

REFERENCES

[1] P. Erdös, *Über die Anzahl der Primfaktoren von* $\binom{n}{k}$, Archiv der Math. *24* (1973), 53-56.

[2] P. Erdös, *Problems and results on consecutive integers*, Publ. Math. Debrecen. *23* (1976), 271-282.

[3] P. Erdös, H. Gupta and S. P. Khare, *On the number of distinct prime factors of* $\binom{n}{k}$, Utilitas Math. *10* (1976), 51-60.

[4] P. Erdös and J. Selfridge, *Some problems on the prime factors of consecutive integers*, Illinois J. Math. *11* (1967), 428-430.

[5] P. Erdös and E. Szemerédi, *Problem 192*, Mat. Lapok *25* (1974), 182-183.

[6] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press, London - New York - San Francisco, 1974.

[7] H. Halberstam and K. F. Roth, *Sequences, Vol. I*, Oxford, Clarendon Press, 1966.

Mathematical Institute
Hungarian Academy of Sciences
Budapest

Received February, 1979.