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

P. ERDŐS
Hungarian Academy of Sciences, Budapest, Hungary
and
R. L. GRAHAM
Bell Laboratories, Murray Hill, New Jersey

A well known theorem of Sylvester and Schur (see [5]) states that for $n > 2k$, the binomial coefficient $\binom{n}{k}$ always has a prime factor exceeding $k$. This can be considered as a generalization of the theorem of Chebyshev: There is always a prime between $m$ and $2m$. Set

$$\binom{n}{k} = u_n(k)\nu_n(k)$$

with

$$u_n(k) = \prod_{p < k} p^\alpha \quad \nu_n(k) = \prod_{\rho \geq k} \rho^\alpha.$$ 

In [4] it is proved that $\nu_n(k) > u_n(k)$ for all but a finite number of cases (which are tabulated there).

In this note, we continue the investigation of $u_n(k)$ and $\nu_n(k)$. We first consider $\nu_n(k)$, the product of the large prime divisors of $\binom{n}{k}$.

**Theorem.**

$$\max_{1 \leq k \leq n} \nu_n(k) = \frac{n}{2}(1 + o(1)).$$

**Proof.** For $k < \epsilon n$ the result is immediate since in this case $\binom{n}{k}$ itself is less than $e^{n/2}$. Also, it is clear that the maximum of $\nu_n(k)$ is not achieved for $k > n/2$. Hence, we may assume $\epsilon n < k < n/2$. Now, for any prime

$$p \in \left\{ \frac{n - k}{r}, \frac{n}{r} \right\}$$

with $p > k$ and $r > 1$, we have $p | \nu_n(k)$. Also, if $k^2 > n$ then $p^2 | \nu_n(k)$ so that in this case the contribution to $\nu_n(k)$ of the primes

$$p \in \left\{ \frac{n - k}{r}, \frac{n}{r} \right\}$$

is (by the Prime Number Theorem (PNT)) just $\frac{n}{r}(1 + o(1))$. Thus, letting $\frac{n}{r + 1} < k < \frac{n}{r}$, we obtain

$$\nu_n(k) = \exp \left[ \left( \sum_{r=1}^{k} \frac{k}{r} + \left( \frac{n}{r} - k \right) \right) \left( 1 + o(1) \right) \right] = \exp \left[ \left( \frac{n}{r} \sum_{r=1}^{k-1} \frac{1}{r} \right) \left( 1 + o(1) \right) \right]$$

$$\leq \frac{n}{2}(1 + o(1)).$$

and the theorem is proved.

It is interesting to note that since

$$\frac{n}{r} \sum_{r=1}^{k-1} \frac{1}{r} = \frac{1}{2}$$

348
for both $t = 2$ and $t = 3$ then
\[
\lim_{n \to \infty} \nu_n(k)^{1/n} = e^{1/2}
\]
for any $k \in \left( \frac{n}{3}, \frac{n}{2} \right)$.

In Table 1, we tabulate the least value $k^*(n)$ of $k$ for which $\nu_n(k)$ achieves its maximum value for selected values of $n \leq 200$. It seems likely that infinitely often $k^*(n) = \frac{n}{2}$ but we are at present far from being able to prove this.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$k^*(n)$</th>
<th>$n$</th>
<th>$k^*(n)$</th>
<th>$n$</th>
<th>$k^*(n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>10</td>
<td>2</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>11</td>
<td>3</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>12</td>
<td>6</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>13</td>
<td>4</td>
<td>50</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>14</td>
<td>4</td>
<td>100</td>
<td>42</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>15</td>
<td>5</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>16</td>
<td>6</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>17</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that

$\nu_1(0) < \nu_1(1) < \nu_1(2) < \nu_1(3)$.

It is easy to see that for $n > 7$, the $\nu_n(k)$ cannot increase monotonically for $0 < k < \frac{n}{2}$.

Next, we mention several results concerning $u_n(k)$. To begin with, note that while $u_1(k) = 1$ for $0 < k < \frac{n}{2} = \frac{7}{2}$, this behavior is no longer possible for $n > 7$. In fact, we have the following more precise statement.

**Theorem.** For some $k < (2 + o(1)) \log n$, we have $u_n(k) > 1$.

**Proof.** Suppose $u_n(k) = 1$ for all $k < (2 + \varepsilon) \log n$. Choose a prime $p < (1 + \varepsilon) \log n$ which does not divide $n + 1$. Such a prime clearly exists (for large $n$) by the PNT. Since $p | n + 1$ then for some $k$ with $p < k < 2p$,

\[
p^2 | n(n - 1) \cdots (n - k + 1), \quad p^2 \nmid k!
\]

Thus, $p \nmid u_n(k)$ and since

\[
k < 2p < (2 + 2\varepsilon) \log n,
\]

the theorem is proved.

In the other direction we have the following result.

**Fact.** There exist infinitely many $n$ so that for all $k < (1/2 + o(1)) \log n$, $u_n(k) = 1$.

**Proof.** Choose $n + 1 = \text{l.c.m.} \left\{ 1, 2, \cdots, \left\lfloor \frac{n}{2} \right\rfloor \right\}$. By the PNT, $n = e^{(2 + o(1)) \frac{n}{2}}$. Clearly, if $m < t$ then $m! \left( \frac{n}{t} \right)$. Thus,

\[
u_n(k) = 1 \quad \text{for} \quad k < \left( \frac{1}{2} + o(1) \right) \log n
\]

as claimed.

In Table 2 we list the least value $n^*(k)$ of $n$ such that $u_n(i) = 1$ for $1 < i < k$

<table>
<thead>
<tr>
<th>$k$</th>
<th>$n^*(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>71</td>
</tr>
</tbody>
</table>
Of course, for \( k \leq 2 \), \( v_n(k) = 1 \) is automatic. By a theorem of Mahler [11], it follows that

\[
u_n(k) < n^{1+\epsilon}
\]

for \( k \geq 3 \) and large \( n \). It is well known that if \( \rho \alpha \mid \binom{n}{k} \) then \( \rho \alpha < n \). Consequently,

\[
u_n(k) < n^{\pi(k)},
\]

where \( \pi(k) \) denotes the number of primes not exceeding \( k \). It seems likely that the following stronger estimate holds:

\[
(*) \quad u_n(k) < n^{(1+o(1))(1-\gamma)\pi(k)}, \quad k > 5,
\]

where \( \gamma \) denotes Euler's constant. It is easy to prove \((*)\) for certain ranges of \( k \). For example, suppose \( k \) is relatively large compared to \( n \), say, \( k = n/t \) for a large fixed \( t \). Of course, any prime \( p \in (n - n/t, n) \) divides \( v_n(k) \) and by the PNT

\[
\prod_{n(1-1/t) < p < n} p = e^{(1+o(1))n/t}.
\]

More generally, if \( n \in (n/t, n) \) with \( r < t \) then \( p > k \) and \( p \mid v_n(k) \) so that again by the PNT

\[
\prod_{n/r < p < n/r} p = e^{(1+o(1))n/rt}.
\]

Thus

\[
v_n(k) > \prod_{1 < r < t} \frac{n}{r} \left( \frac{1}{r} \right) < p < \frac{n}{r} \prod_{1 < r < t} \frac{p}{r} = e^{(1+o(1))\sum_{1 < r < t} \frac{r}{t}} \frac{n}{r} = e^{(1+o(1))(1-\gamma)n/\pi(k)}.
\]

But by Stirling's formula we have

\[
\binom{n}{n/t} = e^{\frac{n}{2} \log t + o(1)} \binom{n}{t}.
\]

Thus,

\[
u_n(k) = \binom{n}{k} / v_n(k) < e^{\frac{n}{2} \log t + o(1)} \binom{n}{t} = n^{(1+o(1))(1-\gamma)\pi(k)} = e^{(1+o(1))(1-\gamma)n/\pi(k)}.
\]

which is just \((*)\).

In contrast to the situation for \( v_n(k) \), the maximum value of \( u_n(k) \) clearly occurs for \( k > n/2 \). Specifically, we have the following result.

**Theorem.** The value \( k(n) \) of \( k \) for which \( u_n(k) \) assumes its maximum value satisfies

\[
k(n) = (1+o(1)) \left( \frac{e-1}{e+1} \right) n.
\]

**Proof.** Let \( k = (1 - cn) \). For \( c < \frac{1}{2} \),

\[
v_n(k) = \prod_{n-k < p < n} p = e^{(1+o(1))cn}.
\]

Since

\[
\binom{n}{k} = \binom{n}{cn} = e^{-c \log c^{1-c}(1-c)(1+c)(1+o(1))n}
\]

then

\[
u_n(k) = \binom{n}{k} / v_n(k) = e^{-(1+o(1))(c+1+c)(1-c)n}.
\]

A simple calculation shows that the exponent is maximized by taking \( c = \frac{1}{e+1} = 0.2689 \ldots \).
Concluding remarks. We mention here several related problems which were not able to settle or did not have time to investigate. One of the authors [8] previously conjectured that \( \binom{2n}{n} \) is never squarefree for \( n > 4 \) (at present this is still open). Of course, more generally, we expect that for all \( \alpha \), \( \binom{2n}{n} \) is always divisible by an \( \alpha \)th power of a prime > \( k \) if \( n > n_\alpha (\alpha, k) \). We can show the much weaker result that \( n = 23 \) is the largest value of \( n \) for which all \( \binom{n}{k} \) are squarefree for \( 0 \leq k < n \). This follows from the observation that if \( p \) is prime and \( p^\alpha \mid \binom{n}{k} \) for any \( k < n \), then \( p^{\beta \mid n+1} \), where

\[
p^{\beta} \geq \frac{n+1}{p^{\alpha}-1}.
\]

Thus, \( 2^\beta \mid \binom{n}{k} \) for any \( k \) implies \( 2^{\beta \mid n+1} \) where \( 2^{\beta} \geq \frac{n+1}{3} \). Also, \( 3^\gamma \mid \binom{n}{k} \) for any \( k \) implies \( 3^{\gamma \mid n+1} \) where \( 3^{\gamma} \geq \frac{n+1}{8} \). Together they imply that \( d = 2^33^{\gamma \mid n+1} \) where \( d > (n+1)^3/24 \). Since \( d \) cannot exceed \( n+1 \) then \( n+1 < 24 \) is forced, and the desired result follows.

For given \( n \) let \( f(n) \) denote the largest integer such that for some \( k \), \( \binom{n}{k} \) is divisible by \( f(n)^{th} \) power of a prime. We can prove that \( f(n) \to \infty \) as \( n \to \infty \) (this is not hard) and very likely \( f(n) > c \log n \) but we are very far from being able to prove this. Similarly, if \( F(n) \) denotes the largest integer so that for all \( k \), \( 1 < k < n \), \( \binom{n}{k} \) is divisible by the \( F(n)^{th} \) power of some prime, then it is quite likely that \( \lim F(n) = \infty \), but we have not proved this.

Let \( P(x) \) and \( p(x) \) denote the greatest and least prime factors of \( x \), respectively. Probably

\[
p\left(\binom{n}{k}\right) > \max \left(n-k, k^{1+\epsilon}\right)
\]

but this seems very deep (for related results see the papers of Ramachandra and others [11], [12]).

J. L. Selfridge and P. Erdős conjectured and Ecklund [1] proved that \( p\left(\binom{n}{k}\right) < \frac{n}{2} \) for \( k > 1 \), with the unique exception of \( p\left(\binom{2}{3}\right) = 5 \). Selfridge and Erdős [9] proved that

\[
p\left(\binom{n}{k}\right) < \frac{c_1n}{k^{\epsilon_2}}
\]

and they conjecture

\[
p\left(\binom{n}{k}\right) < \frac{n}{k} \text{ for } n > k^2.
\]

Finally, let \( d\left(\binom{n}{k}\right) \) denote the greatest divisor of \( \binom{n}{k} \) not exceeding \( n \). Erdős originally conjectured that \( d\left(\binom{n}{k}\right) > n - k \) but this was disproved by Schinzel and Erdős [13]. Perhaps it is true however, that \( d_n > cn \) for a suitable constant \( c \).

For problems and results of a similar nature the reader may consult [2], [3], [6], [7], [10] or [11].

REFERENCES


*****