On the greatest prime factor of \( \prod_{k=1}^{x} f(k) \)

by

P. Erdős (Budapest) and A. Schinzel (Warszawa)

In memory of Trygve Nagell

Let \( P(n) \) denote the greatest prime factor of \( n \). T. Nagell was the first to give a non-trivial lower bound for \( P(\prod_{k=1}^{x} f(k)) \), where \( f \) is an arbitrary irreducible polynomial of degree greater than 1. In [5] he proved

\[ P(\prod_{k=1}^{x} f(k)) > c(f, \varepsilon) x^{(\log x)^{1+\varepsilon}} \text{ for all } \varepsilon > 0. \]

In 1951 the first named author improved considerably the above inequality by proving that for \( x > x_0(f) \)

\[ P(\prod_{k=1}^{x} f(k)) > x(\log x)^{c(f) \log \log x} \quad \text{with } c(f) > 0. \]

In the same paper [1] he has also claimed that

\[ P(\prod_{k=1}^{x} f(k)) > x \exp((\log x)^{\delta(f)}) \quad \text{with } \delta(f) > 0. \]

Our efforts to reconstruct the proof of the latter estimate have been unsuccessful. Instead we have proved the following

**Theorem 1.** Let \( f \in \mathbb{Z}[x] \) be an irreducible polynomial of degree \( l > 1 \). There exists an absolute constant \( c_1 > 0 \) such that for \( x > x_1(f) \)

\[ P(\prod_{k=1}^{x} f(k)) > x \exp(\exp(c_1(\log \log x)^{1/3})). \]

In the sequel we shall denote the \( n \)th iterate of \( \log \) by \( \log^{(n)} \), the number of solutions of the congruence \( f(k) \equiv 0 \pmod{m} \) in the interval \( 1 \leq k \leq x \) by
\( q_n(m) \), the number of divisors of an integer \( n \) in a set \( S \) by \( d(n, S) \) and we shall put:

\[
q_n(m) = \varphi(m).
\]

\( c_1, c_2, \ldots, x_1, x_2, \ldots \) will denote positive constants, in general depending on \( f \).

Theorem 1 is an immediate consequence of the following two theorems.

**Theorem 2.** Under the assumptions of Theorem 1 the number \( N(x) \) of positive integers \( k \leq x \) such that

\[
d\left( f(k), \left\lfloor \frac{x}{2^r} \right\rfloor \right) \geq 1
\]

satisfies for \( x > x_2 \)

\[
N(x) > \frac{x}{\log x} \exp(c_3 (\log x)^{1/3}),
\]

where \( c_3 \) is an absolute constant.

**Theorem 3.** Under the assumptions of Theorems 1 and 2

\[
P(\prod_{k=1}^{x} f(k)) = x \exp\left(\frac{\log x}{x} - N(x)\right)
\]

for \( x > x_3 \).

The proof of Theorem 3 follows closely the proof of (1) given in \([1]\). It is clear from this theorem that in order to prove (2) it would be enough to show that

\[
2x \frac{\varphi(m)}{m} \geq \varphi_4(m) \geq \frac{1}{2} \frac{x}{m} \varphi(m).
\]

**Lemma 1.** If \( x \geq m \) we have

\[
2x \frac{\varphi(m)}{m} \geq \varphi_4(m) \geq \frac{1}{2} x \frac{\varphi(m)}{m}.
\]

**Proof.** We have for \( x \geq m \)

\[
\frac{x}{2m} \varphi(m) \leq \frac{x}{m} \varphi(m) \leq \varphi_4(m) \leq \frac{x}{m} \varphi(m) + \varphi(m) \leq \frac{2x}{m} \varphi(m).
\]

**Lemma 2.** If \( z \geq 2y, y > y_1 \) we have

\[
2 \log \frac{\log z}{\log y} \geq \sum_{\gamma < p \leq x} \frac{\varphi(p)}{p} > \frac{1}{2} \log \frac{\log z}{\log y}.
\]

**Proof.** We shall use the prime ideal theorem in the form

\[
\sum_{\gamma < p \leq x} \varphi(p) \log p = y + O(\gamma^{-c_y \sqrt{\log \gamma}})
\]

(see \([4]\), Satz 190).

By partial summation we obtain

\[
\sum_{\gamma < p \leq x} \varphi(p) \log p = \log \log y + c_4 + O(\gamma^{-c_y \sqrt{\log \gamma}}),
\]

hence

\[
\sum_{\gamma < p \leq x} \varphi(p) \log p = \log \log y + O(\gamma^{-c_y \sqrt{\log \gamma}})
\]

and since for \( z \geq 2y \) the main term dominates the error we get the desired bounds.

**Lemma 3.** Assume that \( f \) is primitive. If \( P \) runs through all integers composed of \( n \) distinct prime factors we have for \( y \geq y_2 \)

\[
\sum_{\gamma < p \leq x} \varphi(P) \leq \frac{c_5 \log y + c_6 y^{n-1}}{(n-1)! \log y}
\]

**Proof.** Since \( \varphi(m) \) is multiplicative, we have \( \varphi(P) \leq P \). On the other hand, for the number \( \pi_n(x) \) of positive integers \( \leq x \) composed of \( n \) distinct prime factors we have the inequality (see \([3]\))

\[
\pi_n(x) \leq \frac{c_7 x \log x + c_8 y^{n-1}}{(n-1)! \log x}
\]

Hence

\[
\sum_{\gamma < p \leq x} \frac{1}{y} \pi_n(y) \leq \frac{c_9 \log y + c_8 y^{n-1}}{(n-1)! \log y}.
\]
Remark. The formulation of Lemma 3 and its proof have been corrected following a suggestion from G. Tenenbaum.

**Lemma 4.** Let \( c > 0 \), \( r = \lceil c \log_2 x \rceil^{1/3} \), \( A(c) \) be the set of all integers in the interval \([x/2, x]\) of the form
\[
pq_1 \cdots q_r,
\]
where \( p, q_1, \ldots, q_r \) are primes and for \( i = 1, 2, \ldots, r \)
\[
\exp \left( \frac{1}{2(2r + 7)} \log x \right)^{1/2r} < q_i < \exp \left( \frac{1}{2(2r + 7)} \log x \right)^{1 - \frac{1}{2r}}.
\]
The number \( N_0(x) \) of positive integers \( k \leq x \) such that
\[
d(f(k), A(c)) > 2lr! (2r + 7)^{r + 1}
\]
is \( o \left( x/(\log x)^{r + 2} \right) \).

**Proof.** We shall assume throughout that \( x \) is sufficiently large and without loss of generality that \( f \) is primitive. Then if \( p, q_1, \ldots, q_r \in [x/2, x] \) and \( \forall i \) satisfy the inequalities (3) we have \( p > x^{1/2} \). On the other hand for \( k \leq x \)
\[
|f(k)| < c_k x^l,
\]
hence \( f(k) \) can have at most 2\( l \) prime factors greater than \( x^{1/2} \). Therefore, (4) implies that \( f(k) \) has more than
\[
R = r! (2r + 7)^{r + 1}
\]
divisors in \( A(c) \), of the form \( pq_1 q_2 \cdots q_r \), where \( p \) is fixed.

Consider the family of sets \( \{q_1^{\alpha_1}, \ldots, q_r^{\alpha_r}\} \) (\( 1 \leq \alpha \leq R \). By the theorem of Erdős and Rado [2] the family contains a \( \alpha \)-system of cardinality \( 2r + 7 \). Let the common intersection of any two distinct sets of this \( \alpha \)-system be \( \{p_1, \ldots, p_s\} \), where \( 0 \leq \delta < r \). Let \( s \) be the integer defined by
\[
2^s - 1 < \frac{x}{p_1 \cdots p_s} < 2^s.
\]
By the condition \( pq_1 q_2 \cdots q_r \in A(c) \) \( f(k) \) has at least \( 2r + 7 \) pairwise coprime divisors in the interval \([2^{-2}, 2^2]\), each divisor consisting of \( r - \delta \) distinct prime factors all in the interval
\[
\left( \exp \left( \frac{1}{2(2r + 7)} \log x \right)^{1/2r}, \exp \left( \frac{1}{2(2r + 7)} \log x \right)^{1 - \frac{1}{2r}} \right)
\]
and all but one less than
\[
\exp \left( \frac{1}{2(2r + 7)} \log x \right)^{1 - \frac{1}{2r}}.
\]

Hence
\[
s \geq \frac{(\log x)^{1/2}}{2(2r + 7) \log 2} = s_0
\]
and
\[
N_0(x) \leq \sum_{s = 0}^{r - 1} \sum_{\delta = 0}^{s - 1} q_0 \left( P_1 P_2 \cdots P_{2r + 7} \right) \,
\]
where the sum \( q_0 \) is taken over all sets of \( 2r + 7 \) pairwise coprime integers \( P_1, P_2, \ldots, P_{2r + 7} \) in the interval \([2^{-2}, 2^2]\), each consisting of \( r - \delta \) distinct prime factors of the size described above. For every such set we have
\[
P_1 \cdots P_{2r + 7} < x^{1/2} \exp \left( \frac{r - \delta - 1}{2} (\log x)^{1 - 1/2r} \right) < x,
\]
thus by Lemma 1
\[
q_0 \left( P_1 P_2 \cdots P_{2r + 7} \right) < 2 e^{2 r - 2} P_{2r + 7}^{r - 1} P_1^{2r + 7} \frac{q_0(P)}{P},
\]
and by (6)
\[
N_0(x) \leq \sum_{s = 0}^{r - 1} \sum_{\delta = 0}^{s - 1} 2 x \sum_{P_1, \ldots, P_{2r + 7}} \frac{q_0(P)}{P} \leq 2 x \sum_{s = 0}^{r - 1} \sum_{\delta = 0}^{s - 1} \frac{q_0(P)}{P}^{2r + 7},
\]
where \( P \) runs through all integers \( P \) in the interval \([2^{-2}, 2^2]\) consisting of \( r - \delta \) distinct prime factors. By Lemma 3 we obtain
\[
\sum_{P} q_0(P) \leq c_0 \frac{r - \delta}{(r - \delta - 1)!} \log 2^s \leq c_0 \frac{r - \delta - 1}{(r - \delta - 1)!}.
\]
Hence
\[
N_0(x) \leq 2 x \sum_{\delta = 0}^{r - 1} \frac{c_0 (r - \delta - 1)!}{s^{2r + 7}} \sum_{s = 0}^{r - 1} \frac{(\log x)^{r - 1}(2r + 7)}{s^{2r + 7}}.
\]
For \( s > s_0 \) we have
\[
\log s > \log (s_0 - 1) > 2r + 7 > 2r + 5 \quad (r - 1) > r - 1.
\]
Therefore, on this halfline \( (\log s)^{r - 1}(2r + 7)/s^{2r + 7} \) is decreasing, since
\[
\frac{(\log s)^{r - 1}(2r + 7)}{s^{2r + 7}} < \frac{d}{ds} \left( \frac{(\log s)^{r - 1}(2r + 7)}{s^{2r + 7}} \right).
\]
Since
\[
\log s > \frac{2r + 7}{2r + 5} (r - 1),
\]
It follows that
\[
\sum_{a \geq a_0} \frac{(\log 2)^{\nu - 1}(2r + 7)}{3^{2r + 7}} \leq \int_{s_0 - 1}^{\infty} \frac{(\log s)^{\nu - 1}(2r + 7)}{s^{2r + 7}} ds 
\]
and by (7)
\[
N_0(x) \leq 2x(c_9 \log x)^{2r + 7} \exp O\left(\frac{r^2 \log x}{(\log x)^{r - 1}}\right) = o\left(\frac{x}{(\log x)^{r + 2}}\right).
\]

Proof of Theorem 2. For \(k \leq x\) by (5) \(f(k)\) has less than \(c_{10} \log x\) prime factors. Thus we have in the notation of Lemma 4
\[
d(f(k), A(c)) < \left(\frac{c_{10} \log x}{r + 1}\right)^{2r + 7} = \frac{c_{10}^{2r + 7}}{(r + 1)!} (\log x)^{r + 1}.
\]
From Lemma 4 and (8) we obtain
\[
\sum k d(f(k), A(c)) = o\left(\frac{x}{(\log x)^{r + 1}}\right),
\]
where in \(\sum\) \(k\) runs through all positive integers \(k \leq x\) with \(d(f(k), A(c)) > 2r!(2r + 7)^{r + 1}\).

On the other hand, by Lemma 1
\[
\sum_{k=1}^{x} d(f(k), A(c)) = \sum_{a \in A(c)} \frac{\varphi(a)}{a} = \frac{x}{2} \sum_{a \in A(c)} \varphi(a).
\]
We evidently have
\[
\sum_{a \in A(c)} \frac{\varphi(a)}{a} = \sum_{q_1} \frac{\varphi(q_1)}{q_1} \sum_{q_2} \frac{\varphi(q_2)}{q_2} \ldots \sum_{q_r} \frac{\varphi(q_r)}{q_r} \sum_{q_{r + 1}}^{\prime} \frac{\varphi(p)}{p},
\]
where the sum \(\sum_{q_i}\) is taken over all primes \(q_i\) in the interval (3) \((1 \leq i \leq r)\) and the sum \(\sum_{q_{r + 1}}^{\prime}\) is taken over all primes \(p\) in the interval
\[
\frac{x}{2q_1 \ldots q_r} \leq p \leq \frac{x}{q_1 \ldots q_r}.
\]
It follows from Lemma 2 that
\[
\sum_{i=1}^{r} \frac{\varphi(q_i)}{q_i} > \frac{1}{4r} \log_2 x \quad (1 \leq i \leq r),
\]
\[
\sum_{r + 1}^{\prime} \frac{\varphi(p)}{p} > \frac{1}{2} \log \left(1 + \frac{\log 2}{\log(x/2q_1 \ldots q_r)}\right) > \frac{\log 2}{2} \log x.
\]
Therefore,
\[
\sum_{a \in A(c)} \frac{\varphi(a)}{a} > \left(\frac{\log_2 x}{4r}\right)^r \frac{\log 2}{2} \log x
\]
and by (10)
\[
\sum_{k=1}^{x} d(f(k), A(c)) > \frac{2}{4r} \log x \left(\frac{\log_2 x}{4r}\right)^r.
\]
Since \(r = o(\log_2 x)\), it follows from (9) that
\[
\sum d(f(k), A(c)) > \frac{x}{6 \log x} \left(\frac{\log_2 x}{4r}\right)^r,
\]
where \(\sum\) is taken over all positive integers \(k \leq x\) such that \(d(f(k), A(c)) \leq 2r!(2r + 7)^{r + 1}\).

From (11) we obtain
\[
N(x) > \frac{1}{12r!(2r + 7)^{r + 1}} \log x \left(\frac{\log_2 x}{4r}\right)^r
\]
\[
> \frac{x}{\log x} \exp \left(r \left(3 \log x - 3 \log r + 1 - \log 8 + O\left(\frac{\log r}{r}\right)\right)\right)
\]
\[
> \frac{x}{\log x} \exp \left(c(\log_2 x)^{1/3} \left(-3 \log c + 1 - \log 8 + O(\log x)\right)\right).
\]
Choosing \(c = \sqrt[3]{e}/8\) (the choice \(c = \sqrt[3]{1/(8e^2)}\) is optimal) we obtain the theorem.

Remark. If instead of the theorem of Erdős and Révész we use their conjecture \(r!(2r + 7)^{r + 1}\) is replaced throughout by \(2(2r + 7)^{r + 1}\) and the above proof for \(r = \left[\frac{1}{2} \log_2 x^{1/3}\right]\) gives
\[
N(x) > \frac{x}{\log x} \exp \left(c_{11} \left(\log_2 x\right)^{1/2} \log_4 x\right)
\]
for \(x > x_4\),
where \(c_{11}\) is an absolute constant.

We proceed to the proof of Theorem 3. Denote by \(U\) the set of all integers \(u\) of the interval \([x/\log x, x]\) for which \(f(u)\) has no prime factor satisfying
\[
x < p \leq c_{12} x, \quad \text{where} \quad c_{12} = 2c_8.
\]

**Lemma 5.** \(\text{card } U > x - c_{13} x/\log x\).
Proof. Clearly
\[
\text{card } U = [x] - \left\lfloor \frac{x}{\log x} \right\rfloor - \sum_{p \leq \sqrt{x}} (\varphi(p) - \frac{x}{p \log x}) > x \frac{x}{\log x} - 1 - \ln(c_{12} x) > x - c_{13} \frac{x}{\log x}.
\]
For \( k \leq x \) put
\[
|f(k)| = A_k B_k, \quad \text{where } A_k = \prod_{p \mid f(k)} p^k, \quad B_k = |f(k)|/A_k
\]
and let
\[
P \left( \prod_{k=1}^x f(k) \right) = P_x.
\]

**Lemma 6.** For all \( u \in U \)
\[
A_k > \frac{x^t}{2(\log x)^t P_x^{t-1}}.
\]
Proof. Since by the definition of \( U \): \( x/\log x < u \leq x \) we have for \( x > x_s \)
\[
\frac{1}{2} \left( \frac{x}{\log x} \right)^t < |f(u)| < c_8 x^t.
\]
Further, \( f(u) \) has no prime factor in the interval \( (x, c_{12} x) \). Therefore by (12) and the choice of \( c_{12} \), \( B_k \) can have at most \( l-1 \) prime factors, multiple factors counted multiply. By (13) all prime factors of \( f(u) \) are at most \( P_x \), thus
\[
B_k \leq P_x^{l-1}.
\]
Hence
\[
A_k = |f(u)|/B_k > \frac{x^t}{2(\log x)^t P_x^{t-1}}.
\]

**Lemma 7.** Let \( u \in U \) be such that \( f(u) \) has a divisor in \([x/2, x]\). Then
\[
A_k > \frac{x^t}{2(\log x)^t P_x^{t-2}}.
\]
Proof. By the definition of \( U \) all prime factors of \( B_k \) are greater than \( c_{12} x \). Since \( f(u) \equiv 0 \pmod{d} \) for some \( d \in [x/2, x] \) we have by (12), (14) and the choice of \( c_{12} \)
\[
B_k < 2c_8 x^{t-1} = (c_{12} x)^{t-1}.
\]
Thus \( B_k \) can have at most \( l-2 \) prime factors, multiple factors counted multiply. Thus by (12) and (13)
\[
A_u = \frac{|f(u)|}{B_u} > \frac{x^t}{2(\log x)^t P_x^{t-2}}.
\]

**Lemma 8.**
\[
\sum_{k=1}^x \log A_k < x \log x + c_{14} x.
\]
Proof, see Nagell [6], pp. 180-182.

**Proof of Theorem 3.** The number of \( u \in U \) for which \( f(u) \) has a divisor in \([x/2, x]\) is at least equal to \( N(x) - (x - \text{card } U) \), hence by Lemma 5 is at least \( N(x) - c_9 x/\log x \). From Lemmata 5, 6, 7, and (8) we now obtain
\[
x \log x + c_{14} x
\]
\[
\geq \sum_{n \leq x} \log A_n > \left( x - c_{13} \frac{x}{\log x} \right) (\log x - \log_3 x - (l-1) \log P_x - \log 2)
\]
\[
+ \left( N(x) - c_{13} \frac{x}{\log x} \right) \log P_x
\]
\[
> \log x - \log_3 x - (l-1) \log P_x - c_{13} \log P_x
\]
\[
+ c_{13} (l-1) \frac{x}{\log x} \log P_x + N(x) \log P_x - c_{13} \frac{x}{\log x} \log P_x
\]
\[
> \log x - \log_3 x - (l-1) \log P_x - (c_{13} l + 1) x + N(x) \log P_x.
\]
Hence
\[
(l-1) x \log \frac{P_x}{x} > N(x) \log P_x - \log_3 x - (c_{13} l + c_{14} l + 1) x
\]
By Lemma 2 for \( x > x_\gamma \) there is at least one prime \( p \in [x/2, x] \) with \( \varphi(p) > 0 \), hence \( P_x > x/2 \). On the other hand, by Theorem 1 \( x \log_2 x = o(N(x) \log x) \).
Thus for \( x > x_\gamma \)
\[
\log x - \log_3 x + (c_{13} l + c_{14} l + 1) x < \frac{1}{7} N(x) \log x - N(x) \log 2
\]
and the inequality (15) gives
\[
(l-1) x \log \frac{P_x}{x} > \frac{l-1}{l} N(x) \log x
\]
and the inequality (15) gives
\[
P_x > x \exp \left( \frac{\log x}{lx} N(x) \right),
\]
which was to be proved.
References


Received on 21.2.1989
and in revised form on 28.4.1989