ON THE GREATEST PRIME FACTOR OF $\prod_{k=1}^{x} f(k)$

P. Erdös*.

[*Extracted from the Journal of the London Mathematical Society, Vol.* 27, 1952.]

Tchebycheff† proved that the greatest prime factor of $\prod_{k=1}^{x}(1+k^2)$ tends to infinity faster than any constant multiple of x. Later Nagell‡ proved the following sharper and more general theorem:

Let $f(x)$ be any polynomial with integer coefficients which is not the product of linear factors with integral coefficients. Denote by P_x the greatest prime factor of $\prod_{k=1}^{x} f(k)$. Then

$$P_x > c_1 x \log x. \tag{1}$$

Throughout this paper c_1, c_2, ... will denote positive constants depending only on the polynomial, p, q will denote primes, and x will be sufficiently large. In the present paper we shall obtain the following improvement on Nagell's result:

Theorem. *There exists a $c_2 = c_2\,(f)$ such that*

$$P_x > x(\log x)^{c_2 \log\log\log x}. \tag{2}$$

Clearly we can assume without loss of generality that $f(x)$ is irreducible in the rational field and of degree $l > 1$. (2) is very far from being best possible. I can prove in a much more complicated way that

$$P_x > x\, e^{(\log x)^{c_3}}. \tag{3}$$

---

* Received 3 October, 1951; read 15 November, 1951.

† E. Landau, *Handbuch über die Lehre von der Verteilung der Primzahlen*, 1 (1909), 559–561.

‡ *Abhandlungen aus dem Math. Seminar Hamburg*, 1 (1922), 179–194. See also G. Ricci, *Annali de Mat.* (4), 12 (1934), 295–303.

(3) will not be proved in the present paper. It seems likely that $P_x > c_4 x^l$, but this if true must be very deep.

Denote by $\rho(k)$ the number of solutions of $f(u) \equiv 0 \pmod{k}$, $0 \leqslant u < k$, and by $\rho_x(k)$ the number of solutions of $f(u) \equiv 0 \pmod{k}$, $0 < u \leqslant x$. We evidently have, for $k \leqslant x$,

$$\frac{x}{2k}\rho(k) \leqslant \left[\frac{x}{k}\right]\rho(k) \leqslant \rho_x(k) \leqslant \left[\frac{x}{k}\right]\rho(k)+\rho(k) \leqslant \frac{2x}{k}\rho(k). \tag{4}$$

We shall make use of the prime ideal theorem in the form*

$$\sum_{p \leqslant y} \rho(p) = \big(1+o(1)\big)\, y/\log y. \tag{5}$$

From (5) and $\rho(p) \leqslant l$ [l is the degree of $f(x)$] it follows that

$$\big(1+o(1)\big)\, y/\log y \geqslant \sum_{\substack{y \leqslant p \leqslant 2y \\ \rho(p)>0}} 1 \geqslant \big(1+o(1)\big)\, y/l\log y.$$

Hence
$$c_5'/\log y > \sum_{\substack{y \leqslant p \leqslant 2y \\ \rho(p)>0}} 1/p > c_5/\log y. \tag{6}$$

Denote by $a_1 < a_2 < \ldots$ the integers of the interval $(x/\log\log x,\, x)$ of the form pq, where

$$p > x^{\frac{1}{2}}, \quad \exp\left[(\log x)^{\frac{1}{2}}\right] < q < x^{\frac{1}{100}}, \quad \rho(a_i) > 0. \tag{7}$$

[The condition $\rho(a_i) > 0$ means $\rho(p) > 0$, $\rho(q) > 0$]. $d^+(n)$ denotes the number of divisors of n amongst the a's.

LEMMA 1. *The number of integers $t \leqslant x$ for which $f(t)$ is divisible by one of the a's is greater than*

$$c_6 x (\log\log x)(\log\log\log x)/\log x.$$

We prove Lemma 1 in several steps. We have by (4)

$$\sum_{k=1}^{x} d^+\big(f(k)\big) = \sum_i \rho_x(a_i) \geqslant \frac{x}{2}\sum_i \frac{\rho(a_i)}{a_i} \geqslant \frac{x}{2}\sum \frac{1}{a_i}. \tag{8}$$

We evidently have
$$\sum_i \frac{1}{a_i} = \Sigma_1 \frac{1}{q}\, \Sigma_2 \frac{1}{p},$$

where in Σ_1 $\exp\left[(\log x)^{\frac{1}{2}}\right] < q < x^{\frac{1}{100}}$, in Σ_2 $x/(q\log\log x) < p < x/q$ and $\rho(p) > 0$, $\rho(q) > 0$. From (6) we obtain

$$\Sigma_1 1/q > c_6 \log\log x, \quad \Sigma_2 1/p > c_7 \log\log\log x/\log x.$$

* If p does not divide the discriminant of $f(x)$, the number $\rho(p)$ of solutions of $f(x) \equiv 0 \pmod{p}$ is the same as the number of prime ideal factors of p of the first degree in the field generated by a zero of $f(x)$ (see Dedekind, *Abh. K. Ges. Wiss. Göttingen*, 1878). Thus the sum in (5) is essentially the same as the number of prime ideals $\mathfrak{p}$ with $N\mathfrak{p} < y$.

Thus
$$\sum_i 1/a_i > c_8 \log\log x \log\log\log x/\log x. \tag{9}$$

Hence, from (8) and (9),
$$\sum_{k=1}^{x} d^+\big(f(k)\big) > \tfrac{1}{2}c_8\, x \log\log x \log\log\log x/\log x. \tag{10}$$

Next we show that the number $N(x)$ of integers $k \leqslant x$ satisfying $d^+\big(f(k)\big) > 20l$ is $o\big(x/(\log x)^3\big)$.

First of all, for $k \leqslant x$, $f(k) < c_9 x^l$; thus $f(k)$ can have at most $2l$ prime factors greater than $x^{\frac{1}{2}}$. Thus it follows from (7) that if $d^+\big(f(k)\big) > 20l$, then $f(k)$ must have at least 10 factors pq_j, satisfying
$$q_1 < q_2 < \ldots < q_{10},\ \frac{x}{p\log\log x} < q_j < \frac{x}{p},\ p > x^{\frac{1}{2}},\ \exp[(\log x)^{\frac{1}{2}}] < q_j < x^{\frac{1}{100}}, \tag{11}$$
since pq_j, being an a_i, must lie in the interval $(x/\log\log x,\ x)$. Let s be the integer defined by
$$2^{s-1} < \frac{x}{p\log\log x} \leqslant 2^s.$$

Then, by (11), $f(k)$ has at least 10 distinct prime factors in the interval $(2^{s-1},\ 2^s \log\log x)$. Further*, by (11),
$$(\log x)^{\frac{1}{2}} < s < \tfrac{1}{10}\log x. \tag{12}$$
The number of integers $k \leqslant x$ for which $f(k)$ has at least 10 distinct prime factors in $(2^{s-1},\ 2^s \log\log x)$, s satisfying (12), is clearly less than
$$\sum_s \Sigma_3\, \rho_x(q_1 q_2 \cdots q_{10}), \tag{13}$$
where $(\log x)^{\frac{1}{2}} < s < \frac{1}{10}\log x$ and, in Σ_3, $2^{s-1} < q_j < 2^s \log\log x$ and the q's are distinct.

Clearly $N(x)$ is not greater than the sum (13); thus to prove
$$N(x) = o\big(x/(\log x)^3\big)$$
it will suffice to prove that the sum (13) is $o\big(x/(\log x)^3\big)$. We have, by (7),
$$q_1 q_2 \cdots q_{10} < x^{\frac{1}{10}} < 2^{\log x}.$$
Thus by (4) and $\rho(q) \leqslant l$ we have
$$N(x) \leqslant \sum_s \Sigma_3\, \rho_x'(q_1 \cdots q_{10}) < 2l^{10} x \sum_s \Sigma_3 \frac{1}{q_1 \cdots q_{10}}.$$

* x is sufficiently large.

From (6) we have

$$\Sigma_3 \frac{1}{q_1 \cdots q_{10}} < \left(\Sigma_3 \frac{1}{q_i}\right)^{10} < c_{10}(\log\log\log x)^{10}/s^{10}.$$

Thus finally

$$\sum_s \Sigma_3 \rho_x(q_1 \cdots q_{10}) < c_{10}\, x \sum_{s > (\log x)^{\frac{1}{2}}} (\log\log\log x)^{10}/s^{10} = o\left(\frac{x}{(\log x)^3}\right),$$

as was to be proved.

Since, for $k \leqslant x$, $f(k) < c_9 x^l$, $f(k)$ has less than $c_{11} \log x$ prime factors. Thus we have

$$d^+\left(f(k)\right) < c_{11}^2 (\log x)^2. \tag{14}$$

From (14) and $N(x) = o\left(x/(\log x)^3\right)$, we have

$$\Sigma_4\, d^+\left(f(k)\right) = o(x/\log x), \tag{15}$$

where in Σ_4, $k \leqslant x$ and $d^+\left(f(k)\right) \geqslant 20l$. From (10) and (15) we have

$$\Sigma_5\, d^+\left(f(k)\right) > c_{12}\, x \log\log x \log\log\log x/\log x, \tag{16}$$

where in Σ_5, $k \leqslant x$ and $d^+\left(f(k)\right) < 20l$. From (16) we finally obtain

$$\sum_{\substack{k \leqslant x \\ d^+ f(k) > 0}} 1 > \frac{c_{12}}{20l}\, x \log\log x \log\log\log x/\log x,$$

which proves Lemma 1 with $c_6 = c_{12}/20l$.

Denote by $u_1 < u_2 < \ldots$ the integers of the interval $(x/\log x, x)$ for which $f(u_i)$ has no prime factor p satisfying

$$x \leqslant p \leqslant c_{13}\, x \log\log x, \quad c_{13}^{l-1} > c_9.$$

Denote by $U(x)$ the number of the u's not exceeding x.

Lemma 2. $$U(x) > x - c_{14}\, x \log\log x/\log x.$$

Clearly

$$U(x) \geqslant x - \frac{x}{\log x} - \sum_{x \leqslant p \leqslant c_{13} x \log\log x} \rho_x(p) > x - \frac{x}{\log x} - l\pi(c_{13}\, x \log\log x)$$

$$> x - c_{14}\, x \log\log x/\log x,$$

as stated.

Assume now that the greatest prime factor P_x of $\prod_{k=1}^{x} f(k)$ is less than $x(\log x)^{c_2 \log\log\log x}$. This assumption will lead to a contradiction. Put*

$$f(k) = A_k B_k, \quad \text{where} \quad A_k = \prod_{\substack{p^\alpha \| f(x) \\ p \leqslant x}} p^\alpha, \quad B_k = f(k)/A_k.$$

LEMMA 3. $$A_{u_j} > x/(\log x)^{c_2 l \log\log\log x}.$$

Since by definition $x/\log x \leqslant u_j \leqslant x$, we have

$$c_{15}\, x^l/(\log x)^l < f(u_i) < c_9\, x^l. \tag{17}$$

Further, by the definition of the u_j, $f(u_j)$ has no prime factor in the interval $(x,\, c_{13}\, x \log\log x)$. Therefore, by (17), B_{u_j} $(= f(u_j)/A_{u_j})$ can have at most $l-1$ prime factors, multiple factors counted multiply. By assumption all prime factors of $f(u_i)$ are less than $x(\log x)^{c_2 \log\log\log x}$. Thus

$$B_{u_j} < x^{l-1}(\log x)^{(l-1)c_2 \log\log\log x}.$$

Hence by (17)

$$A_{u_j} = \frac{f(u_j)}{B_{u_j}} > \frac{c_{15}\, x^l}{B_{u_j}(\log x)^l} > \frac{c_{15}\, x^l}{x^{l-1}(\log x)^{l+(l-1)c_2 \log\log\log x}} > \frac{x}{(\log x)^{lc_2 \log\log\log x}},$$

as stated.

LEMMA 4. *The number of u's for which $f(u_j)$ is a multiple of an a_i is greater than* $c_{16}\, x \log\log x \log\log\log x/\log x$.

From Lemmas 1 and 2, the number of these u's is greater than

$$c_6\, x \log\log x \log\log\log x/\log x - \left(x - U(x)\right) > c_{16}\, x \log\log x \log\log\log x/\log x,$$

as stated.

LEMMA 5. *Let u_j be such that $f(u_j)$ is a multiple of one of the a's. Then*

$$A_{u_j} > x^{\frac{3}{2}}.$$

By definition of the u's all prime factors of B_{u_j} are greater than $c_{13}\, x \log\log x$. Thus since $f(u_j) \equiv 0 \pmod{a_i}$ we have from (17),

$$B_{u_j} < c_9\, x^l/(x/\log\log x) = c_9\, x^{l-1}\log\log x < (c_{13}\, x \log\log x)^{l-1}$$

if $c_{13}^{l-1} > c_9$. Thus B_{u_j} can have at most $l-2$ prime factors, multiple factors counted multiply. Thus by (17) and our assumption on P_x

$$A_{u_j} = \frac{f(u_j)}{B_{u_j}} > c_{15}\frac{x^l}{(\log x)^l}\left(x^{l-2}(\log x)^{(l-2)c_2 \log\log\log x}\right)^{-1} > x^{\frac{3}{2}},$$

as stated.

* $p^\alpha \| z$ means that $p^\alpha \mid z$, $p^{\alpha+1} \nmid z$.

LEMMA 6. $$\sum_{k=1}^{x} \log A_k < x \log x + c_{17} x.$$

This is a result of Nagell*.

Proof of the theorem. From Lemmas 2, 3, 4 and 5,

$$\sum_{k=1}^{x} \log A_k \geqslant \sum_i \log A_{u_i}$$
$$> (x - c_{14} x \log\log x/\log x)(\log x - lc_2 \log\log x \log\log\log x)$$
$$+ (c_{16} x \log\log x \log\log\log x/\log x)(\tfrac{1}{2} \log x). \quad (18)$$

The first summand of (18) is given by Lemmas 2 and 3, the second summand is given by Lemmas 4 and 5, *i.e.* by the u's satisfying $f(u_j) \equiv 0 \pmod{a_i}$. Thus from (18)

$$\sum_{k=1}^{x} \log A_k > x \log x - c_{14} x \log\log x - lc_2 x \log\log x \log\log\log x$$
$$+ \tfrac{1}{2} c_{16}\, x \log\log x \log\log\log x.$$

But this contradicts Lemma 6 for $c_2 < c_{16}/2l$. This contradiction proves

$$P_x > x(\log x)^{c_2 \log\log\log x},$$

and so completes the proof of the theorem.

Department of Mathematics,
University College, London.

* *Ibid.* (footnote ‡, p. 379), 180–182. Nagell does not state the result explicitly, but proves it on the above-mentioned pages [see in particular equation (7), p. 182].

Printed by C. F. Hodgson & Son, Ltd., Pakenham Street, London, W.C.1.