ON THE GREATEST AND LEAST PRIME FACTORS OF $n!+1$

P. ERDŐS AND C. L. STEWART

1. *Introduction*

Much work has been done on obtaining estimates from below for the greatest prime factors of the terms of certain sequences of integers. Let $P(m)$ denote the greatest prime factor of m and let $f(x)$ be any irreducible polynomial with degree > 1 and integer coefficients.

It can easily be deduced from Siegel's work [*cf.* 12] that $P(f(x)) \to \infty$ as $x \to \infty$ and recently Sprindzhuk and Kotov [7], using deep techniques of Baker, have shown that indeed

$$P(f(x)) > c \log\log x$$

for all integers x where $c = c(f) > 0$; the case of quadratic and cubic f was in fact covered by earlier works of Schinzel [10] and Keates [6].

In another context Birkhoff and Vandiver [1] proved, by elementary methods, that for distinct positive integers a, b, $P(a^n - b^n) \geqslant n+1$ for all integers $n > 6$. Recently, again using techniques of Baker, the second author [14] obtained some new results in this connexion; for instance, for the Fermat numbers we have (see [15])

$$P(2^{2^n}+1) > cn2^n,$$

for all positive integers n where c is a positive absolute constant. Related work has been carried out on the Fibonacci and, more generally, the Lucas and Lehmer numbers; moreover non-trivial lower bounds have been established (see [13]) for $P(u_n)$ where u_n denotes the n-th term of a general recurrence sequence.

In yet another direction it follows from the work of Tchebychev, Nagell and Ricci (see [2]) that

$$P(f(1) \dots f(x)) < cx \log x$$

for all sufficiently large integers x, where c is a positive number depending only on the degree of f. The first author [2] improved the exponent of $\log x$ here to $c \log\log\log x$ where $c = c(f) > 0$ and, indeed, stated without proof that the number on the right could be replaced by $xe^{(\log x)^\alpha}$ for all sufficiently large x where $0 < \alpha < 1$. Hooley [5], using sieve techniques, showed that if f is a quadratic polynomial then $\log x$ can be replaced by x^ε for some $\varepsilon > 0$. It seems probable in fact that if f is irreducible of degree $d > 1$ then

$$P(f(1) \dots f(x)) > cx^d$$

for $c = c(f) > 0$ and for all integers x, but a verification seems hopeless at present.

As regards estimates from below for least prime factors of sequences of the above kind very little of significance is apparently known. In fact the only substantial result of which we are aware is that $p(f(x)) > x^c$ infinitely often, where $p(m)$ denotes the least prime factor of m, f is a polynomial as above such that $f(1), \dots, f(k)$ have no common divisor for some integer k, and $c > 0$ depends only on the degree of f; this can be deduced from Brun's sieve.

Received 29 October, 1975.

[J. LONDON MATH. SOC. (2), 13 (1976), 513–519]

In the present paper we shall discuss the sequences $n!+1$ and $p_1 \dots p_n+1$ where p_k denotes the k-th prime; for these it would seem that no significant lower estimates for either the least or the greatest prime factors have hitherto been recorded. Perhaps our ignorance can best be illustrated by observing that we cannot exclude, even with recourse to the deep results of Baker and others, the possibility that infinitely often $n!+1$ is composed solely of the two smallest primes exceeding n. Nevertheless, we can prove by elementary arguments a few results which are, perhaps, not entirely trivial.

We begin by noting that $p(n!+1) \geqslant n+1$ for all n and by Wilson's theorem, equality holds here when $n+1$ is a prime. We shall show first that, by contrast, a better estimate obtains when $n+1$ is composite.

THEOREM 1. *For any positive integer n such that $n+1$ is not a prime we have*

$$p(n!+1) > n+(1-o(1)) \log n/\log\log n. \tag{1}$$

Further, for almost all integers n, we have

$$p(n!+1) > n+\varepsilon(n)n^{1/2}, \tag{2}$$

where $\varepsilon(n)$ is any positive function that decreases to 0 as $n \to \infty$.

We believe that the estimate (1) is best possible; this is certainly the case, see the proof of (1), if $i!+1$ is the product of a small number, less than e^i say, and a prime for infinitely many odd integers i. A similar estimate to (1) can be obtained for the greatest prime factor, without restriction. We prove

THEOREM 2. *For all positive integers n we have*

$$P(n!+1) > n+(1-o(1)) \log n/\log\log n. \tag{3}$$

Furthermore

$$\limsup_{n\to\infty} P(n!+1)/n > 2+\delta \tag{4}$$

where δ is an effectively computable positive constant.

It should be noted that the proof of (4) also suffices to establish that $p(n!+1) < (2-\delta')n$, where $\delta' = (\delta/(1+\delta))$, for infinitely many composite integers $n+1$.

Theorems 1 and 2 follow from a generalisation of Wilson's theorem; for the proof of (3) the well known result of Liouville that $(p-1)!+1$ is not a power of the prime $p(> 5)$ is invoked. Both theorems hold with $n!-1$ in place of $n!+1$ and the proofs apply virtually unchanged; the exception is the proof of (3) where the result of Liouville is not applicable and a theorem of Erdős and Oblàth [3] to the effect that the equation $n!-1 = x^m$ has no solution in integers $n > 1$, $x > 1$ and $m > 1$, is required instead.

We turn now to the sequence $p_1 \dots p_n+1$ and prove:

THEOREM 3. *For infinitely many integers n (> 0)*

$$P(p_1 \dots p_n+1) > p_{n+k}$$

where $k > c \log n/\log\log n$ for some positive absolute constant c.

The main reason for the discrepancy in the relative strengths of Theorems 1 and 2 as compared with Theorem 3 is that we do not have an analogue of Wilson's theorem for the product of the first n primes.

It is certainly not known, witness Theorems 2 and 3, whether $n!+1$ or $p_1 \dots p_n+1$ is a prime infinitely often and curiously enough it is not even known if $p_1 \dots p_n+1$ is infinitely often not a prime. From Wilson's theorem and a slight generalisation it follows that there are infinitely many even and infinitely many odd integers n for which $n!+1$ is composite and Schinzel [11] has shown that for any rational number c, $cn!+1$ is infinitely often composite. We would guess, of course, that both $n!+1$ and $p_1 \dots p_n+1$ are composite for almost all values of n.

Finally we shall prove the following result, a special case of which is needed for the proof of Theorem 3.

THEOREM 4. *The equations*

$$\prod_{p \leqslant n} p = x^m - y^m$$

and

$$\prod_{p \leqslant n} p = x^m + y^m$$

have no solutions in positive integers x, y, $n(> 2)$ and m (> 1).

For the proof of Theorem 4 we shall adopt a similar approach to that used by Erdős and Obláth [3]. They showed that for primes p and integers n the equations $n! = x^p - y^p$ $(p > 2,\ n > 1)$ and $n! = x^p + y^p$ $(p > 1, n > 2)$ have no solutions in positive integers x and y.

We close with the curious observation, related to the first of the above equations with $m = 2$ and $y = 1$, that $4p_1 \dots p_n+1$ is a square for $n = 1, 2, 3, 4$ and 7. C. Bach has checked the expression for all $n < 28$ with the aid of a computer and has found no more squares; and we would guess that $n = 7$ gives the last such square.

2. *Proof of Theorem* 1

For the proof of (1), we shall use the result, which follows immediately from Wilson's theorem, that for a prime p

$$(p-i-1)!\, i! \equiv (-1)^{i+1} \pmod{p} \quad 0 \leqslant i \leqslant p-1. \tag{5}$$

Now if $n+1$ is not a prime and a prime p divides $n!+1$ then since $p > n+1$ we may write $p-i-1 = n$ for some positive integer i. From the above identity it then follows that p divides $i!+(-1)^{i+1}$ and thus that $i! \geqslant p-1$. But now since $p-1 = n+i$ we have $i! - i \geqslant n$ and (1) therefore follows.

We shall now prove (2). We assume, therefore, that there exists a decreasing function ε for which (2) fails to hold on a set of integers of upper density $\delta > 0$ and we shall show that this leads to a contradiction. Accordingly, we can find arbitrarily large integers n such that between n and $2n$ there are δn integers t for which $p(t!+1) < t+\varepsilon(t)t^{1/2}$.

We first remove those integers t for which $t+1$ or $t+2$ is a prime, leaving, by the prime number theorem, a set T of at least $(\delta/2)n$ integers t for which $2 < p_t - t < \varepsilon(t)t^{1/2}$ where $p_t = p(t!+1)$. Now from (5) p_t divides $(p_t-t-1)!+(-1)^{t+1}$ and since

$2 < p_t - t < r$ where $r = [\varepsilon(n)(2n)^{1/2}]$ we have that either $p_t \mid b!+1$ or $p_t \mid b!-1$ for some integer $b = b_t$ $(= p_t - t - 1)$ with $1 < b < r$. Further, if $p_t = p_{t'}$ for $t \neq t'$ then $b_t \neq b_{t'}$. Therefore

$$\prod_{t \in T} p_t \Big| \prod_{b=2}^{r} (b!+1)(b!-1) \qquad (\neq 0)$$

and thus

$$\prod_{b=2}^{r} (b!)^2 > n^{\delta n/2}. \tag{6}$$

On the other hand, since $b! < b^b$, we have

$$\prod_{b=1}^{r} (b!)^2 < \exp\left(\sum_{b=1}^{r} 2b \log b\right).$$

But $\varepsilon \to 0$ as $n \to \infty$, and so, by definition, $r < n^{1/2}$ for n sufficiently large; whence the exponential expression on the right above is less than $n^{1+\dots+r}$ and so also less than $n^{2\varepsilon^2 n}$. This contradicts (6) for n sufficiently large and the theorem follows.

3. *Proof of Theorem* 2

We observe that (3) follows from (1) if $n+1$ is not a prime. If $n+1$ is a prime p (> 5) then, from a result of Liouville [8], $(p-1)!+1$ is divisible by a prime other than p. We can therefore conclude, by the same argument used in the proof of (1), that (3) is valid.

We next observe that we can easily establish that the expression on the left-hand side of (4) is $\geqslant 2$. For if $p \mid i!+1$ for some odd integer i (> 1) then from (5) $p \mid (p-i-1)!+1$. In particular we see that as i tends to infinity so also do p and $p-i-1$; furthermore, the maximum of p/i and $p/(p-i-1)$, $0 < i < p-1$, is clearly $\geqslant 2$. The proof that the aforementioned inequality is strict, however, is considerably more involved.

We shall now assume that (4) does not hold and show that this leads to a contradiction. Accordingly, given a positive constant δ $(< 1/2)$ we can find an integer $N = N(\delta)$, satisfying $(\log N)^{-1} < \delta^2/2$, such that for all integers $n > N$, $P(n!+1) < (2+\delta)n$. Now by the prime number theorem we can find an integer $n > N!+1$ such that the interval $\{(2-\delta)n, (2+4\delta)n\}$ contains at most $6\delta n/\log n$ prime numbers. Furthermore, if a prime p divides $m!+1$ for some odd integer $m \geqslant n$, then from (5) $p \mid (p-m-1)!+1$ and as $n > N!+1$ we have $p > N!+1$ and thus $p-m-1 > N$. Therefore, by assumption, both p/m and $p/(p-m-1)$ are less than $2+\delta$ and thus if $n \leqslant m \leqslant n+\delta n$ we plainly have $(2-\delta)n < p < (2+4\delta)n$.

Let $m_1 < \dots < m_t$ denote the odd integers in $\{n, (1+\delta)n\}$. If $p^l \mid m!+1$ then, upon observing that $p > (2-\delta)n$, $m < (1+\delta)n$, $m! < m^m$ and $\delta < 1/2$, we see than $l < 3n/2$. Further, if $p^{l_1} \mid m_i!+1$ and $p^{l_2} \mid m_j!+1$ $(m_i > m_j)$ then, on setting $l = \min\{l_1, l_2\}$, we find that $p^l \mid m_i m_{i-1} \dots m_{j+1} - 1$ and thus $l < m_i - m_j$. Accordingly, there can be at most 2^k integers $m_i!+1$ divisible by p^l where $l \geqslant \delta n/2^k$; hence p divides $(m_1!+1) \dots (m_t!+1)$ to at most the power $3n/2 + \delta n \log n/\log 2$. But now, as there are less than $6\delta n/\log n$ primes in $\{(2-\delta)n, (2+4\delta)n\}$ and as $\delta < 1/2$ we have

$$(m_1!+1)\dots(m_t!+1) < (4n)^{(3n/2+\delta n \log n/\log 2)\, 6\delta n/\log n}$$

which is $< n^{c'((\delta n)^2 + n^2/\log n)}$ for some absolute constant c'; whence the above product is less than $n^{c\delta^2 n^2}$ with $c = 3c'/2$, since N was chosen such that $(\log N)^{-1} < \delta^2/2$.

On the other hand, as $(m_t!+1) > n! > (n/e)^n$ and as t is certainly $> \delta n/3$ we have

$$(m_1!+1) \ldots (m_t!+1) > (n/e)^{\delta n^2/3}$$

which is $> n^{c_1\delta n^2}$ for some positive absolute constant c_1. This, however, contradicts our earlier estimate if $\delta < c_1/c$ and the result now follows.

4. *Proof of Theorem* 3

By the prime number theorem we can find, for arbitrarily large integers n, an interval $\{n, n+t\}$, with $t = c(\log n)^2/\log\log n$ for some small positive constant c, which contains $r = (1+o(1))\, t/\log n$ primes. We shall prove that for some prime p_m from the interval $\{n, n+t\}$, $p_1 \ldots p_m+1$ has a prime factor greater than $p_{m+r/3}$. This will then prove the theorem, for we have chosen r to be $(1+o(1))\, c \log n/\log\log n$ and it follows from the prime number theorem that $m = (1+o(1))\, n/\log n$.

Denote by p_k and p_{k+r} the least and the greatest prime numbers in $\{n, n+t\}$ and observe that $p_1 \ldots p_{k+a}+1$ and $p_1 \ldots p_{k+b}+1$ cannot both be divisible by the same prime factor p_{k+d}, $1 \leqslant a < b < d \leqslant r$. For if they are we have

$$p_{k+a+1} \cdots p_{k+b}-1 \equiv 0 (\operatorname{mod} p_{k+d}),$$

and thus, on writing $t_u = p_{k+d}-p_{k+u}$,

$$(-1)^{b-a} \prod_{u=a+1}^{b} t_u - 1 \equiv 0 (\operatorname{mod} p_{k+d}).$$

But now as the t_u in the above product are all less than t and as there are fewer than r of them we have $t^r > p_{k+d} > n$ and this contradicts our choice of t for c sufficiently small.

We now consider the integers $p_1 \ldots p_{k+u}+1$, $u = 1, \ldots, 2r/3$. Each of these integers has at least two distinct prime factors; for, from Theorem 4 it follows that the integers cannot be a power of a prime and if one of them is a prime then the theorem clearly holds. Not all of these primes can be less than $n+t$, for we would then have, by the preceding paragraph, $4r/3$ primes between n and $n+t$ and this contradicts our original assumption. Thus for some integer m, $k+1 \leqslant m \leqslant k+2r/3$, $p_1 \cdots p_m+1$ is divisible by a prime greater than $n+t$ and hence greater than $p_{m+r/3}$. The theorem now follows by our earlier remarks.

5. *Proof of Theorem* 4

We shall give a detailed proof that the first equation of Theorem 4 has no solution. For the proof that the second equation, involving a sum of powers, has no solution we shall merely need to make some minor adjustments to this proof.

We note that for both equations we may assume x and y are relatively prime positive integers and further that m is a prime q.

The proof of the first assertion now proceeds in three stages.

We observe that for $q = 2$ the equation cannot hold since we would then have that 2 divides $(x-y)(x+y)$ to exactly the first power, which is clearly impossible.

Next we observe that for primes $q > 2$ we have, on denoting $(x^q-y^q)/(x-y)$ by A, that $(x-y)^2 < A$ and thus that $(x^q-y^q)^2 < A^3$. But now q divides A to at most the first power, while all other prime factors of A are congruent to 1 (mod q) by a classical result of Euler [4]. Thus, if the equation is to hold, we must have

$$\left(\prod_{p\leqslant n} p\right)^{2/3} < q \prod_{\substack{p\equiv 1 \,(\mathrm{mod}\, q)\\ p\leqslant n}} p \tag{7}$$

for some integer n (> 2).

It follows from a result of Rosser and Schoenfeld (see 3.14 of [9]) that the logarithm of the left-hand side of (7) is greater than $\cdot 60n$ for $n \geqslant 563$. The final stage of the proof involves obtaining a good upper bound for the right-hand side of (7) for comparison with this estimate.

To this end we note that the expression

$$B(k) = q^{[k/(q-1)]}\frac{(q+1)\ldots((k-1)q+1)}{k!}$$

is an integer for all positive integers k. Indeed, the exponent of a prime p in the prime decomposition of $k!$ is $[k/p]+[k/p^2]+\ldots$ and, if $(p,q)=1$, this is clearly at least the exponent to which p divides the product in the numerator of $B(k)$; further, it is plain that the sum of the above series is $\leqslant [k/(p-1)]$.

Now since $(jq+1)/(j+1) < q$, we have $B(k) < q^{[k/(q-1)]+k-1}$. Furthermore, we certainly have

$$\begin{aligned} q \prod_{\substack{p\equiv 1 \,(\mathrm{mod}\, q)\\ p\leqslant n}} p &\leqslant qB([n/q]+1) \\ &< q^{n/q(q-1)+n/q+2}, \end{aligned}$$

which, on taking logarithms, is less than $n((\log 3)/2+2(\log q)/n)$ since $q \geqslant 3$.

From the original equation and the result of Euler [4], it follows that $n \geqslant q$. Accordingly it follows from the above paragraph that for $n \geqslant 563$, the logarithm of the expression on the right-hand side of (7) is less than $n((\log 3)/2+(2 \log 563)/563)$, which in turn is less than $\cdot 58n$. This contradicts our lower bound of $\cdot 60n$, however, and therefore the original equation cannot hold for $n \geqslant 563$. It is now a straightforward task to check that (7) and hence also the equation, cannot hold for $1 < n < 563$.

To prove that the product of the primes $\leqslant n$ cannot be the sum of q-th powers we first observe that the equation cannot hold for $q = 2$ and $n > 2$, since if it could we would have $3 \mid x^2+y^2$ for relatively prime positive integers x and y, which is, of course, impossible. For primes $q > 2$, on writing $(x^q+y^q)/(x+y)$ as A and noting that

$$(x+y)^2 \leqslant (x+y)^2+3(x-y)^2 \leqslant 4(x^3+y^3)/(x+y),$$

we find that $(x^q+y^q)^2 < 4A^3$. But now from a slight generalisation of the result of Euler [4] and from the preceding remark, if our original equation has a solution then (7) holds with the product on the right-hand side of the inequality now multiplied by the cube root of 4. We may now proceed as before to conclude that the equation cannot hold for $n \geqslant 563$ and a simple check extends this to $n > 2$. For $n = 2$ we note that we have the trivial solution $2 = 1^q+1^q$.

References

1. G. D. Birkhoff and H. S. Vandiver, "On the integral divisors of a^n-b^n", *Ann. of Math.*, 5 (1904), 173–180.

2. P. Erdős, "On the greatest prime factor of $\prod_{k=1}^{x} f(k)$", *J. London Math. Soc.*, 27 (1952), 379–384.

3. ——— and R. Obláth, " Über diophantische Gleichungen der Form $n! = x^p \pm y^p$ und $n! \pm m! = x^p$ ". *Acta Litt. Sci. Szeged*, 8 (1937), 241–255.

4. L. Euler "Theoremata circa divisores numerorum", *Commentationes Arithmeticae Collectae, Tom. I., Petropoli* (1849), 50–61.

5. C. Hooley, " On the greatest prime factor of a quadratic polynomial ", *Acta Mathematica*, 117 (1967), 281–299.

6. M. Keates, "On the greatest prime factor of a polynomial", *Proc. Edinburgh Math. Soc.*, 16 (1969), 301–303.

7. S. V. Kotov, " The greatest prime factor of a polynomial ", (Russian) *Mat. Zametki*, 13 (1973), 515–522.

8. J. Liouville, " Sur l'équation $1 \cdot 2 \cdot 3 \, .. (p-1)+1 = p^m$ ", *J. Math. Pure Appl.*, 1 (1856), 351–352.

9. J. Barkley Rosser and L. Schoenfeld, " Approximate formulas for some functions of prime numbers ", *Illinois J. Math.*, 6 (1962), 64–94.

10. A. Schinzel, " On two theorems of Gelfond and some of their applications ", *Acta Arith.*, 13 (1967), 177–236.

11. A. Schinzel, " On the composite integers of the form $c(ak+b)! \pm 1$ ", *Nordisk Mat. Tidskr.*, 10 (1962), 8–10.

12. C. L. Siegel, (under the pseudonym X). " The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + ... + k$ ", *J. London Math. Soc.*, 1 (1926), 66–68.

13. C. L. Stewart, " The greatest prime factor of a general linear recurrence sequence ", to appear.

14. ———, " The greatest prime factor of $a^n - b^n$ ", *Acta Arith.*, 26 (1975), 427–433.

15. ———, " Divisors of Fermat, Fibonacci, Lucas and Lehmer numbers ", to appear.

Mathematical Institute of the Hungarian Academy of Sciences,
Budapest,
Hungary

and

Trinity College,
Cambridge