On the greatest and least prime factors of $n! + 1$, II

By C. L. STEWART (Waterloo)

In memory of Béla Brindza

Abstract. Let ε be a positive real number. We prove that for infinitely many odd integers n the least prime factor of $n! + 1$ is at most $(\frac{\sqrt{145}-1}{8} + \varepsilon)n$ and that for infinitely many positive integers n the greatest prime factor of $n! + 1$ exceeds $(\frac{11}{2} - \varepsilon)n$.

1. Introduction

In 1856, LIOUVILLE [6] proved that $(p-1)! + 1$ is not a power of p for any prime p larger than 5. More than a century later ERDŐS and GRAHAM [2] asked if the equation

$$(p-1)! + a^{p-1} = p^k \tag{1}$$

has only finitely many solutions in positive integers a, k, p with p an odd prime. In 1991 BRINDZA and ERDŐS [1] resolved the question by proving that all solutions of (1) are smaller than an effectively computable number. A few years later YU and LIU [10] and then LE [5] determined the complete list of solutions.

In 1976 we investigated with ERDŐS [3] the arithmetical character of integers of the form $n! + 1$ where n is a positive integer. For any integer m larger than 1 let $P(m)$ denote the greatest prime factor of m and let $p(m)$ denote the least prime factor of m. By Wilson's theorem p divides $(p-1)!+1$ whenever p is a prime. Since all prime factors of $n!+1$ exceed n we see that $p(n! + 1) = n + 1$ whenever $n + 1$ is a prime. We showed with ERDŐS [3] that if $n + 1$ is not a prime then

$$p(n! + 1) > n + (1 - o(1))\frac{\log n}{\log\log n}. \tag{2}$$

Further, for almost all integers n,

$$p(n! + 1) > n + \varepsilon(n)n^{\frac{1}{2}}, \tag{3}$$

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

In [3] we indicated how to prove that for infinitely many integers n for which $n + 1$ is not a prime $p(n! + 1)$ is less than $2n$. We observed, as a direct consequence of Wilson's theorem, that if p is a prime then

$$(p - 1 - i)!\, i! \equiv (-1)^{i+1} \pmod{p}, \quad 0 \le i \le p - 1. \tag{4}$$

Thus if $p \mid i!+1$ for some odd integer i (> 1) then, from (4), $p \mid (p-i-1)!+1$. For any positive real numbers θ and t with $t > \frac{1}{\theta}$ we have

$$\max\left(\theta, \frac{1}{t - \theta^{-1}}\right) \ge \frac{2}{t} \tag{5}$$

and

$$\min\left(\theta, \frac{1}{t - \theta^{-1}}\right) \le \frac{2}{t}. \tag{6}$$

Note that $\frac{p}{p-i-1} = \frac{1}{\frac{p-1}{p} - \frac{i}{p}}$ and so on taking $\theta = \frac{p}{i}$ and $t = \frac{p-1}{p}$ we see from (5) and (6) that

$$\max\left(\frac{p}{i}, \frac{p}{p - i - 1}\right) \ge \frac{2p}{p - 1} = 2 + \frac{2}{p - 1},$$

and

$$\min\left(\frac{p}{i}, \frac{p}{p - i - 1}\right) \le 2 + \frac{2}{p - 1},$$

for $0 < i < p - 1$. As i tends to infinity so also do p and $p - i$. Thus for each $\varepsilon > 0$,

$$p(n! + 1) < (2 + \varepsilon)n, \tag{7}$$

for infinitely many composite integers $n + 1$. Further

$$P(n! + 1) > 2n, \tag{8}$$

for infinitely many positive integers n. We indicated in [3] that $2 + \varepsilon$ in (7) could be replaced by $2 - \delta$ for some positive number δ. Our first result will be of this character.

Theorem 1. *Let $\varepsilon > 0$. There are infinitely many odd integers n for which*

$$p(n! + 1) < \left(\frac{\sqrt{145} - 1}{8} + \varepsilon \right) n. \tag{9}$$

Observe that $\frac{\sqrt{145} - 1}{8} = 1.38019\ldots.$

With ERDŐS [3] we proved that (2) holds with $P(n! + 1)$ in place of $p(n! + 1)$ for all positive integers n. Of course this only is an improvement on (2) for those integers n for which $n + 1$ is a prime. Recently LUCA and SHPARLINSKI [7] sharpened this result by proving that

$$P(n! + 1) > n + \left(\frac{1}{4} + o(1) \right) \log n; \tag{10}$$

indeed they established (10) with $P(n! + 1)$ replaced by $P(n! + f(n))$ where f is any non-zero polynomial with integer coefficients. In 2002 MURTY and WONG [8] showed that if the *abc* conjecture holds, then

$$P(n! + 1) > (1 + o(1))n \log n.$$

In 1976 we improved on (8) with ERDŐS [3] by proving that there is a positive number δ such that

$$P(n! + 1) > (2 + \delta)n, \tag{11}$$

for infinitely many integers n. LUCA and SHPARLINSKI [7] established (11) with $(2 + \delta)n$ replaced by $(\frac{5}{2} + o(1))n$ and showed that their result applies with $n! + f(n)$ in place of $n! + 1$ where f is any non-zero polynomial with integer coefficients. Our next result gives a further improvement on (11).

For any set X we denote the cardinality of X by $|X|$. For any set A of positive integers and any positive integer n we put $A(n) = A \cap \{1, \ldots, n\}$. The lower asymptotic density of A is $\liminf \frac{|A(n)|}{n}$.

Theorem 2. *Let $\varepsilon > 0$. The set of positive integers n for which*

$$P(n! + 1) > \left(\frac{11}{2} - \varepsilon\right) n \tag{12}$$

has positive lower asymptotic density.

As we remarked in [3] estimates (2), (3), (7), (8) and (11) hold with $n! + 1$ replaced by $n! - 1$ and the same comment applies to the estimates (9) and (12). Further, the same techniques used to prove Theorems 1 and 2 allow one to prove, for instance, that for each positive real number ε there exist infinitely many positive integers n for which $P((2n)! + 1) > \left(\frac{17 + \sqrt{145}}{8} - \varepsilon\right) 2n$ and there exist infinitely many positive integers n for which $P((n! + 1)(n! - 1)) > \left(\frac{11 + \sqrt{85}}{2} - \varepsilon\right) n$.

2. Preliminary lemmas

Let $p_1, p_2, \ldots$ denote the sequence of prime numbers and put $d_k = p_{k+1} - p_k$ for $k = 1, 2, \ldots$. Our first lemma, due to HEATH-BROWN, gives a bound on the frequency of large differences between consecutive prime numbers.

Lemma 1. *Let ε be a positive real number. There is a positive number c, which depends on ε, such that*

$$\sum_{p_k \leq x} d_k^2 < c x^{\frac{23}{18} + \varepsilon}.$$

PROOF. This is Theorem 1 of [4].   □

Our next result gives a bound for the size of the greatest common divisor of a collection of terms of the form $k! + 1$.

Lemma 2. *Let n and t be positive integers with $t \geq 2$ and let $i_1, \ldots, i_t$ be distinct positive integers from a subinterval of $[1, n]$ of length ℓ. Then*

$$\gcd(i_1! + 1, \ldots, i_t! + 1) < n^{\frac{\ell}{t-1}}. \tag{13}$$

Further, there exists a positive number c_1 such that if n exceeds c_1 and $t \geq 3$ then

$$\gcd(i_1! + 1, \ldots, i_t! + 1) < e^n n^{\frac{2\ell}{(t-1)^2}}. \tag{14}$$

Furthermore, let δ and ε be positive real numbers. There exists a positive number c_2, which depends on ε and δ, such that if n exceeds c_2,

$$3 \leq t < n^{\frac{13}{18} - \delta}, \tag{15}$$

and

$$\ell > \frac{n}{(\log n)^{\frac{1}{2}}}, \tag{16}$$

then

$$\gcd(i_1! + 1, \ldots, i_t! + 1)$$
$$< \exp\left((1 + \varepsilon)\ell\left(\frac{\log t}{t} + \frac{\log\left(\frac{ne}{\ell}\right)}{t} + \frac{2\log n \max(1, \log\log t)}{(t-1)^2}\right)\right). \tag{17}$$

We remark that it is possible to replace the term $\max(1, \log\log t)$ on the right hand side of inequality (17) by $f(t)$ where f is any real valued function to the real numbers of size at least 1 for which $\lim_{t \to \infty} f(t) = \infty$ provided that c_2 is modified to depend on f.

PROOF OF LEMMA 2. Let A be a positive integer and let (k_1, k_2) and (k_3, k_4) be distinct pairs of integers with $n \geq k_1 > k_2 \geq 1$ and $n \geq k_3 > k_4 \geq 1$ for which

$$A \mid k_i! + 1, \tag{18}$$

for $i = 1, 2, 3, 4$. Suppose, without loss of generality, that

$$k_1 - k_2 \geq k_3 - k_4. \tag{19}$$

(Note that $\{k_1, k_2, k_3, k_4\}$ may only contain 3 elements.) Then, by (18),

$$A \mid k_i! - k_{i+1}!, \quad \text{for } i = 1, 3,$$

and so

$$A \mid k_{i+1}!(k_i(k_i - 1) \cdots (k_{i+1} + 1) - 1), \quad \text{for } i = 1, 3.$$

But, by (18), $\gcd(A, k_{i+1}!) = 1$ for $i = 1, 3$ hence

$$A \mid k_i \cdots (k_{i+1} + 1) - 1, \tag{20}$$

for $i = 1, 3$. Therefore

$$A \mid k_1 \cdots (k_2 + 1) - k_3 \cdots (k_4 + 1).$$

Since

$$k_1 \cdots (k_2 + 1) - k_3 \cdots (k_4 + 1) = (k_3 - k_4)! \left(\frac{(k_1 - k_2)!}{(k_3 - k_4)!} \binom{k_1}{k_2} - \binom{k_3}{k_4} \right),$$

$\gcd(A, (k_3 - k_4)!) = 1$ and $k_1 - k_2 \geq k_3 - k_4$, we find that

$$A \mid ((k_3 - k_4)!)^{-1}(k_1 \cdots (k_2 + 1) - k_3 \cdots (k_4 + 1)). \tag{21}$$

Notice that

$$|k_1 \cdots (k_2 + 1) - k_3 \cdots (k_4 + 1)| < n^{k_1 - k_2}. \tag{22}$$

Thus, provided that

$$k_1 \cdots (k_2 + 1) \neq k_3 \cdots (k_4 + 1), \tag{23}$$

we deduce from (21), (22) and the fact that $m! \geq \left(\frac{m}{e} \right)^m$, when m is a positive integer, that

$$A < n^{(k_1 - k_2) - (k_3 - k_4)} \left(\frac{ne}{k_3 - k_4} \right)^{k_3 - k_4}. \tag{24}$$

Since $g(x) = \left(\frac{ne}{x} \right)^x$ attains its maximum value at $x = n$ we see that, when (23) holds,

$$A < n^{(k_1 - k_2) - (k_3 - k_4)} e^n. \tag{25}$$

We now put

$$A = \gcd(i_1! + 1, \ldots, i_t! + 1).$$

We shall prove (13) first. Put

$$\mu = \min\{i_a - i_b \mid i_a > i_b, \ a, b \in \{1, \ldots, t\}\}$$

and let k_1, k_2 be elements of $\{i_1, \ldots, i_t\}$ with $k_1 - k_2 = \mu$. Since $\mu \leq \frac{\ell}{(t-1)}$ we see from (20) with $i = 1$ that

$$A < n^{k_1 - k_2} \leq n^{\frac{\ell}{(t-1)}},$$

as required.

We shall prove (14) next. There are $\binom{t}{2}$ pairs of integers (i, i') with $i > i'$ which can be chosen from $\{i_1, \ldots, i_t\}$. Associated to each such pair is the difference $i - i'$ and $0 < i - i' \leq \ell$. Therefore there are two such pairs, (k_1, k_2) and (k_3, k_4) say, for which

$$0 \leq (k_1 - k_2) - (k_3 - k_4) \leq \frac{\ell}{\left(\binom{t}{2} - 1\right)}. \qquad (26)$$

Thus, provided that (23) holds, by (25) and (26),

$$A < n^{\frac{\ell}{\left(\binom{t}{2} - 1\right)}} e^n,$$

hence, since $t \geq 3$ and $\binom{t}{2} - 1 \geq \frac{(t-1)^2}{2}$,

$$A < e^n n^{\frac{2\ell}{(t-1)^2}}.$$

It remains only to ensure that (23) holds. We may assume that A exceeds e^n since otherwise the result is immediate. Note that if $k_1 = k_3$ then, since the pairs (k_1, k_2) and (k_3, k_4) are distinct, $k_2 \neq k_4$ and so (23) holds. Thus we may assume that $k_1 > k_3$; a similar argument applies if $k_3 < k_1$. Further, we may assume, after renumbering k_2, k_3, k_4 if necessary, that $k_2 \geq k_3 > k_4$. Since, as in (20), A divides $k_1 \cdots (k_2 + 1) - 1$ we see that $A \leq n^{k_1 - k_2}$. But A exceeds e^n and so

$$k_1 - k_2 \geq \frac{n}{\log n}.$$

By a version of the prime number theorem with an explicit error term, for n sufficiently large there is a prime p between k_1 and k_2. As a consequence p divides $k_1 \cdots (k_2 + 1)$ and not $k_3 \cdots (k_4 + 1)$ and so (23) holds and (14) follows.

Finally, we shall prove (17). Let $c_3, c_4, \ldots$ denote positive numbers which are effectively computable in terms of ε and δ. Without loss of generality we may suppose that

$$n \geq i_1 > i_2 > \cdots > i_t \geq 1.$$

Note that we may also suppose that

$$t - 1 \geq (\log n)^{\frac{1}{8}}, \qquad (27)$$

since otherwise, by (16),

$$\frac{2\ell \log n}{(t-1)^2} \geq 2\ell(\log n)^{\frac{3}{4}} \geq 2n(\log n)^{\frac{1}{4}}$$

and therefore, by (14),

$$A < \exp\left((1+\varepsilon)\frac{2\ell(\log n)}{(t-1)^2}\right),$$

for n larger than c_3, whence (17) holds.

We consider the consecutive integers $i_{j+1}+1, \ldots, i_j$ for $j = 1, \ldots, t-1$. Notice that A divides $i_j!+1$ and $i_{j+1}!+1$ hence A divides $i_j \cdots (i_{j+1}+1)-1$. Therefore A is at most $n^{i_j - i_{j+1}}$. If for some j, with $1 \leq j \leq t-1$,

$$i_j - i_{j+1} < \frac{n}{t(\log n)^{\frac{3}{2}}},$$

then

$$A \leq \exp\left(\frac{n}{t(\log n)^{\frac{1}{2}}}\right),$$

and, by (16), (17) holds. Thus we may suppose that

$$i_j - i_{j+1} \geq \frac{n}{t(\log n)^{\frac{3}{2}}},$$

for $j = 1, \ldots, t-1$. Let m denote the number of the intervals $[i_{j+1}+1, i_j]$ for $j = 1, \ldots, t-1$ which do not contain a prime number. Let $p_1, p_2, \ldots$ denote the sequence of prime numbers and put $d_k = p_{k+1} - p_k$ for $k = 1, 2, \ldots$. Then

$$\sum_{p_k \leq n} d_k^2 \geq m\left(\frac{n}{t(\log n)^{\frac{3}{2}}}\right)^2.$$

But, by Lemma 1,

$$\sum_{p_k \leq n} d_k^2 < n^{\frac{23}{18}+\frac{\delta}{2}},$$

for $n > c_4$. In particular

$$m < \frac{t^2(\log n)^3}{n^{\frac{13}{18}-\frac{\delta}{2}}},$$

and by (15), since $t \leq n$,

$$m < tn^{-\frac{\delta}{2}} (\log n)^3 < t^{1-\frac{\delta}{3}}, \tag{28}$$

for $n > c_5$.

Put

$$t_1 = t - 1 - m, \tag{29}$$

and order the differences $i_j - i_{j+1}$ with $1 \leq j \leq t - 1$ for which there is a prime in the interval $[i_{j+1} + 1, i_j]$ according to size. Let us denote these differences by $\gamma_1, \ldots, \gamma_{t_1}$ so that

$$\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_{t_1}. \tag{30}$$

Observe that

$$\gamma_1 + \cdots + \gamma_{t_1} \leq i_1 - i_t \leq \ell. \tag{31}$$

For any real number x let $[x]$ denote the largest integer of size at most x. Put

$$t_2 = \left[t_1 / (\log \log t_1)^{\frac{1}{2}} \right]. \tag{32}$$

Then, by (30) and (31),

$$\gamma_{t_2}(t_1 - t_2) \leq \ell.$$

Thus, by (27), (28), (29), (30) and (32), for $n > c_6$,

$$\gamma_h < \left(1 + \frac{\varepsilon}{2} \right) \frac{\ell}{t}, \tag{33}$$

for $h = 1, \ldots, t_2$.

Next note that

$$(\gamma_{t_2} - \gamma_{t_2-1}) + (\gamma_{t_2-1} - \gamma_{t_2-2}) + \cdots + (\gamma_1 - 0) = \gamma_{t_2}$$
$$(\gamma_{t_2-1} - \gamma_{t_2-2}) + \cdots + (\gamma_1 - 0) = \gamma_{t_2-1}$$
$$\ddots \qquad\qquad \vdots$$
$$\ddots \qquad \vdots$$
$$(\gamma_1 - 0) = \gamma_1$$

hence

$$(\gamma_{t_2} - \gamma_{t_2-1}) + 2(\gamma_{t_2-1} - \gamma_{t_2-2}) + \cdots + t_2\gamma_1 = \gamma_{t_2} + \cdots + \gamma_1. \tag{34}$$

Put

$$\theta = \min(\gamma_{t_2} - \gamma_{t_2-1}, \ldots, \gamma_2 - \gamma_1, \gamma_1).$$

Then, by (31) and (34),

$$\frac{t_2(t_2+1)}{2}\theta \le \ell.$$

Therefore, by (27), (28), (29) and (32), for $n > c_7$,

$$\theta < (1+\varepsilon)\frac{2\ell \log\log t}{t^2}. \tag{35}$$

We have $\gamma_1 = i_r - i_{r+1}$ for an integer r with $1 \le r \le t-1$. Then

$$A \mid i_r \cdots (i_{r+1} + 1) - 1.$$

If $\theta = \gamma_1$, we see that

$$A < n^\theta,$$

and by (35) our result follows.

Thus we may suppose that $\theta = \gamma_s - \gamma_{s-1}$ for some integer s from $\{2, \ldots, t_2\}$. In particular,

$$\theta = (i_a - i_{a+1}) - (i_b - i_{b+1})$$

with a and b distinct integers from $\{1, \ldots, t-1\}$. Put $k_1 = i_a$, $k_2 = i_{a+1}$, $k_3 = i_b$ and $k_4 = i_{b+1}$. By construction there is a prime among the integers $k_2 + 1, \ldots, k_1$ and another prime among the integers $k_4 + 1, \ldots, k_3$. Thus the larger of the two primes divides one of $k_1 \cdots (k_2 + 1)$ and $k_3 \cdots (k_4 + 1)$ and not the other whence (23) holds. Note also that $\left(\frac{ne}{x}\right)^x$ is an increasing function of x for x positive and less than n. Therefore, by (24), (27), (33) and (35), we find that

$$A < n^{(1+\varepsilon)\frac{2\ell \log\log t}{t^2}} \left(\frac{net}{(1+\varepsilon)\ell}\right)^{(1+\varepsilon)\frac{\ell}{t}},$$

hence that

$$A < \exp\left((1+\varepsilon)\ell\left(\frac{2\log n \log\log t}{t^2} + \frac{\log t}{t} + \frac{\log\left(\frac{ne}{(1+\varepsilon)\ell}\right)}{t}\right)\right) \tag{36}$$

for $n > c_8$, as required.   □

For any prime p let $t(p)$ denote the number of positive integers k for which $p \mid k! + 1$. In [9, Theorem 7.5] we noted that

$$t(p) < \frac{(m+1)(m+2)}{2} \quad \text{where } m = (3p)^{\frac{1}{3}}. \tag{37}$$

To see this observe that if n and s are positive integers and p divides both $n! + 1$ and $(n+s)! + 1$ then p divides $(n+s)! - n!$ hence $(n+s) \cdots (n+1) \equiv 1$ (mod p). In particular, n is a solution of the polynomial congruence $(x+s) \cdots (x+1) \equiv 1$ (mod p), and by Lagrange's theorem the number of such solutions is at most s. Let I be an interval of length ℓ (≥ 1) and let $n_1 < n_2 < \cdots < n_k$ denote all the solutions of $x! + 1 \equiv 0$ (mod p) in I. Plainly

$$\sum_{i=1}^{k-1}(n_{i+1} - n_i) \leq \ell, \tag{38}$$

and by our earlier observation at most s of the terms in brackets in the above sum are equal to s. Therefore

$$\sum_{i=1}^{k-1}(n_{i+1} - n_i) \geq \sum_{s=1}^{u} s^2, \tag{39}$$

where u is defined by the inequalities

$$\sum_{s=1}^{u} s \leq k - 1 < \sum_{s=1}^{u+1} s.$$

Thus k is at most $\frac{(u+1)(u+2)}{2}$ and by (38) and (39)

$$\ell \geq \frac{u(u+1)(2u+1)}{6} > \frac{u^3}{3}. \tag{40}$$

Since all integers n for which $p \mid n! + 1$ lie in the interval $[1, p-1]$, (37) follows from (40) with $\ell = p$. Further, from (40) we obtain Lemma 3 below, a result which is the special case of Lemma 2 of LUCA and SHPARLINSKI [7] with $f(x)$ equal to 1.

Lemma 3. *There exists a positive number c such that if p is a prime number and I is an interval of the positive real numbers of length ℓ with $\ell \geq 1$ then the number of integers k in I for which p divides $k! + 1$ is at most $c\ell^{\frac{2}{3}}$.*

For any non-zero integer m and any prime p we denote by $\mathrm{ord}_p\, m$ the exponent of the largest power of p which divides m. As usual $|m|_p$ is the p-adic absolute value of m normalized so that

$$|m|_p = p^{-\,\mathrm{ord}_p\, m}.$$

Lemma 4. *There exists a positive number c_1 such that if p is a prime number, n a positive integer and I a subinterval of $[1,n]$ of length $\ell \geq 2$ then*

$$(\log p)\,\mathrm{ord}_p\left(\prod_{i\in I}(i!+1)\right) < \frac{2}{3}\ell\log\ell\log n + c_1\ell\log n + n\log n. \qquad (41)$$

Further, for each pair of positive real numbers ε and ε_1 there exist positive numbers c_2 and c_3 such that if ℓ exceeds $\varepsilon_1 n$ and n exceeds c_3 then

$$(\log p)\,\mathrm{ord}_p\left(\prod_{i\in I}(i!+1)\right) < (1+\varepsilon)\frac{2}{9}\ell(\log\ell)^2 + c_2 n\log n. \qquad (42)$$

PROOF. Let $i_1,\ldots,i_u$ be the integers i in I for which p divides $i!+1$. By Lemma 3 there is a positive number c such that

$$u \leq c\ell^{\frac{2}{3}}. \qquad (43)$$

Put $h_t = \mathrm{ord}_p(i_t!+1)$ for $t = 1,\ldots,u$. We may suppose that

$$h_1 \geq \cdots \geq h_u.$$

Then, by (13) of Lemma 2,

$$p^{h_t} < n^{\frac{\ell}{(t-1)}}, \qquad (44)$$

for $t = 2,\ldots,u$. In particular by (43) and (44),

$$(\log p)(h_2 + \cdots + h_u) < \ell\log n\left(1 + \int_2^{c\ell^{\frac{2}{3}}} \frac{1}{t-1}dt\right)$$

$$< \frac{2}{3}\ell\log\ell\log n + c_4\ell\log n, \qquad (45)$$

where c_4 is a positive number. Since

$$h_1\log p \leq n\log n \qquad (46)$$

and

$$\operatorname{ord}_p \left(\prod_{i \in I} (i! + 1) \right) = h_1 + \cdots + h_u, \tag{47}$$

(41) follows from (45) and (46).

Let $c_5, c_6, \ldots$ denote positive numbers which depend on ε and ε_1 and suppose that ℓ exceeds $\varepsilon_1 n$. Then by (43),

$$u \le cn^{\frac{2}{3}} < n^{\frac{13}{18} - \frac{1}{36}},$$

for $n > c_5$. Thus for $n > c_6$, (15) of Lemma 2 holds with $\delta = \frac{1}{36}$ and, since $\ell > \varepsilon_1 n$, (16) of Lemma 2 also holds. Therefore by (17) of Lemma 2, for $n > c_7$,

$$(\log p) h_t < (1 + \varepsilon) \ell \left(\frac{\log t}{t} + \frac{\log \left(\frac{e}{\varepsilon_1} \right)}{t} + \frac{2 \log n \cdot \max(1, \log \log t)}{(t - 1)^2} \right) \tag{48}$$

for $t = 3, \ldots, u$. Since the expression on the right hand side of (48) is a decreasing function of t for $t > e$, we see that

$$(\log p)(h_4 + \cdots + h_u)$$

$$< (1 + \varepsilon) \ell \int_3^{c\ell^{\frac{2}{3}}} \frac{\log t}{t} + \frac{\log \left(\frac{e}{\varepsilon_1} \right)}{t} + \frac{2 \log n \max(1, \log \log t)}{(t - 1)^2} \, dt$$

and so

$$(\log p)(h_4 + \cdots + h_u) < (1 + \varepsilon) \frac{2}{9} \ell (\log \ell)^2 + c_8 n \log n. \tag{49}$$

Since

$$(h_1 + h_2 + h_3) \log p < 3n \log n, \tag{50}$$

(42) now follows from (47), (49) and (50). □

Lemma 5. *Let ε be a positive real number. There exists a positive number c, which depends on ε, such that if p is a prime number, n an integer with $n \ge 2$ and I a subinterval of $[1, n]$ of length ℓ then*

$$(\log p) \operatorname{ord}_p \left(\prod_{i \in I} (i! + 1) \right) < \frac{2}{9} \ell (\log n)^2 + \varepsilon n (\log n)^2 + cn \log n. \tag{51}$$

PROOF. Let $c_1, c_2, \ldots$ denote positive numbers which depend on ε. By (42) of Lemma 4, if ℓ exceeds εn and n exceeds c_1,

$$(\log p)\, \mathrm{ord}_p \left(\prod_{i \in I}(i! + 1) \right) < \frac{2}{9}\ell(\log n)^2 + \frac{2}{9}\varepsilon n(\log n)^2 + c_2 n \log n. \quad (52)$$

On the other hand if $2 \leq \ell \leq \varepsilon n$ then by (41) of Lemma 4,

$$(\log p)\, \mathrm{ord}_p \left(\prod_{i \in I}(i! + 1) \right) < \frac{2}{3}\varepsilon n(\log n)^2 + c_3 n \log n; \quad (53)$$

plainly (53) holds if $\ell \leq 2$. Therefore from (52) and (53), we obtain (51) with c_4 in place of c for $n > c_1$, hence (51) holds for $n \geq 2$ and our result follows. □

3. Proof of Theorem 1

Let δ be a positive real number with $\delta < \frac{1}{100}$. Put $\delta' = 10\delta$,

$$\lambda = \frac{\sqrt{145} - 1}{8} + \delta', \quad (54)$$

and $\lambda_1 = \lambda + \delta'$. Note that $\lambda < \frac{3}{2}$. Let $c_1, c_2, \ldots$ denote positive numbers which depend on δ. We may suppose that there exist only finitely many odd positive integers n for which

$$p(n! + 1) \leq \lambda_1 n,$$

since otherwise the theorem holds. Thus there exists a positive integer c_1 such that for each odd integer n with $n > c_1$,

$$p(n! + 1) > \lambda_1 n. \quad (55)$$

We shall show that this leads to a contradiction and the theorem then follows.

Since (55) holds, we also have

$$P(n! + 1) < \frac{\lambda}{\lambda - 1}n, \quad (56)$$

for all odd integers n with $n > c_2$. To see this, observe that if $q = P(n!+1)$ with $n > 1$ and

$$q \geq \frac{\lambda}{\lambda - 1} n, \tag{57}$$

then q is odd and, by (4),

$$q \mid (q - n - 1)! + 1.$$

But then

$$p((q - n - 1)! + 1) \leq q = \frac{1}{1 - \frac{n+1}{q}}(q - n - 1).$$

We have $q > \lambda$ and, by (57), $\frac{n}{q} \leq \frac{\lambda - 1}{\lambda}$ hence

$$\frac{1}{1 - \frac{n+1}{q}} \leq \frac{1}{1 - \frac{1}{q} - \left(\frac{\lambda - 1}{\lambda}\right)} = \frac{q\lambda}{q - \lambda}.$$

But $\frac{q\lambda}{q-\lambda} < \lambda_1$ for $n > c_3$ since $q > n$. Thus

$$p((q - n - 1)! + 1) \leq \lambda_1(q - n - 1).$$

Furthermore, $q - n - 1 > c_1$ for $n > c_4$ by (57) and this contradicts (55). Therefore (56) holds.

The proof now proceeds by a comparison of estimates for

$$Z = \prod_{\substack{n=1 \\ n \text{ odd, } n > c_2}}^{N} (n! + 1).$$

Put $R = \{n \in \mathbb{Z} \mid n \text{ odd, } c_2 < n \leq N\}$. Observe that if $p \mid n! + 1$ with n in R then, by (55), $p > \lambda_1 n$ and, by (56), $p < \frac{\lambda}{\lambda - 1} n$. In particular,

$$\frac{\lambda - 1}{\lambda} p < n < \frac{1}{\lambda_1} p.$$

Put

$$I_p = \left(\frac{\lambda - 1}{\lambda} p, \min\left(N, \frac{1}{\lambda_1} p\right)\right).$$

Since $n! \geq \left(\frac{n}{e}\right)^n$,

$$Z > \exp\left(\sum_{n \in R} (n \log n - n)\right)$$

so

$$\log Z > (1 - \delta)\frac{N^2}{4} \log N, \tag{58}$$

provided that N exceeds c_5.

On the other hand

$$Z = \prod_p |Z|_p^{-1} \leq \prod_p \left| \prod_{n \in I_p \cap R} (n! + 1) \right|_p^{-1}.$$

Put

$$A(p) = (\log p) \operatorname{ord}_p \left(\prod_{n \in I_p \cap R} (n! + 1) \right).$$

Then

$$Z \leq \exp\left(\sum_{p < \frac{\lambda}{\lambda - 1} N} A(p) \right).$$

Thus by (51) of Lemma 5, with $\varepsilon = \delta$,

$$\log Z \leq \frac{2}{9}(\log N)^2 \sum_{p < \frac{\lambda}{\lambda - 1} N} \ell(p)$$

$$+ (\delta N (\log N)^2 + c_6 N \log N)\pi \left(\frac{\lambda}{\lambda - 1} N \right), \tag{59}$$

where $\ell(p)$, the length of I_p, is given by

$$\ell(p) = \left(\frac{1}{\lambda_1} - \frac{\lambda - 1}{\lambda} \right) p \quad \text{when } p \leq \lambda_1 N$$

and by

$$\ell(p) = N - \left(\frac{\lambda - 1}{\lambda} \right) p \quad \text{when } p \geq \lambda_1 N.$$

By (54), (59) and the prime number theorem,

$$\log Z \leq \frac{2}{9}(\log N)^2 \sum_{p < \frac{\lambda}{\lambda - 1} N} \ell(p) + 4\delta N^2 \log N + c_7 N^2. \tag{60}$$

Further

$$\sum_{p < \frac{\lambda}{\lambda-1}N} \ell(p) = \sum_{p \leq \lambda_1 N} \left(\frac{1}{\lambda_1} - \frac{\lambda-1}{\lambda} \right) p + \sum_{\lambda_1 N < p < \frac{\lambda}{\lambda-1}N} \left(N - \frac{\lambda-1}{\lambda} p \right)$$

$$= \frac{1}{\lambda_1} \sum_{p \leq \lambda_1 N} p + N \left(\sum_{\lambda_1 N < p < \frac{\lambda}{\lambda-1}N} 1 \right) - \frac{\lambda-1}{\lambda} \sum_{p < \frac{\lambda}{\lambda-1}N} p.$$

Thus by the prime number theorem and Abel summation, for $N > c_8$,

$$\sum_{p < \frac{\lambda}{\lambda-1}N} \ell(p) < (1+\delta) \left(\frac{\lambda_1}{2} + \left(\frac{\lambda}{\lambda-1} - \lambda_1 \right) - \frac{\lambda}{2(\lambda-1)} \right) \frac{N^2}{\log N}$$

$$< (1+\delta) \left(\frac{\lambda}{2(\lambda-1)} - \frac{\lambda_1}{2} \right) \frac{N^2}{\log N},$$

and so by (60), and the fact that λ_1 exceeds λ,

$$\log Z < \frac{(1+\delta)}{9} \left(\frac{\lambda}{\lambda-1} - \lambda \right) N^2 \log N + 4\delta N^2 \log N + c_7 N^2. \tag{61}$$

Comparing (58) and (61) we find that for $N > c_9$,

$$\frac{1-\delta}{4} < \frac{(1+\delta)}{9} \frac{\lambda(2-\lambda)}{(\lambda-1)} + 4\delta + \frac{c_8}{\log N}.$$

By (54) we obtain a contradiction for N sufficiently large and the result now follows.

4. Proof of Theorem 2

We may suppose that $0 < \varepsilon < \frac{1}{4}$. Put $\gamma = \frac{11}{2} - 18\varepsilon$ and let $B(\gamma)$ be the set of positive integers for which

$$P(n! + 1) \geq \gamma n. \tag{62}$$

We shall show that for n sufficiently large the set $B(\gamma) \cap \{1, \ldots, n\}$ has cardinality at least $\frac{\varepsilon}{3}n$ and hence the result follows. Accordingly suppose that N is a positive integer for which

$$|B(\gamma) \cap \{1, \ldots, N\}| \leq \frac{\varepsilon}{3}N. \tag{63}$$

Let $c_1, c_2, \ldots$ denote positive numbers which depend on ε. Our proof proceeds by a comparison of estimates for

$$Z = \prod_{\substack{n=1 \\ n \notin B(\gamma)}}^{N} (n! + 1).$$

Since $n! \geq \left(\frac{n}{e}\right)^n$,

$$Z > \exp\left(\sum_{n=2}^{N} (n \log n - n) - |B(\gamma) \cap \{1, \ldots, N\}| N \log N\right)$$

whence, by (63),

$$\log Z > (1 - \varepsilon) \frac{N^2 \log N}{2}, \tag{64}$$

provided that $N > c_1$.

Notice that if $p \mid n! + 1$ and $n \notin B(\gamma)$ then $n < p < \gamma n$ hence $\frac{1}{\gamma} p < n < p$. Put

$$I_p = \left(\frac{1}{\gamma} p, \min\{N, p\}\right),$$

and

$$A(p) = (\log p) \operatorname{ord}_p \left(\prod_{n \in I_p} (n! + 1)\right).$$

Then

$$Z \leq \exp\left(\sum_{p < \gamma N} A(p)\right).$$

Thus, by (51) of Lemma 5 with $\frac{\varepsilon}{6}$ in place of ε,

$$\log Z < \frac{2}{9} (\log N)^2 \sum_{p < \gamma N} \ell(p) + \left(\frac{\varepsilon}{6} N (\log N)^2 + c_2 N \log N\right) \pi(\gamma N), \tag{65}$$

where $\ell(p)$, the length of I_p, is given by

$$\ell(p) = \left(1 - \frac{1}{\gamma}\right) p \quad \text{for } p \leq N$$

and

$$\ell(p) = N - \frac{1}{\gamma} p \quad \text{for } p > N.$$

Since $\gamma < \frac{11}{2}$ it follows from (65) and the prime number theorem that

$$\log Z < \frac{2}{9}(\log N)^2 \sum_{p < \gamma N} \ell(p) + \varepsilon N^2 \log N + c_3 N^2. \qquad (66)$$

Further

$$\sum_{p < \gamma N} \ell(p) = \sum_{p \leq N} \left(1 - \frac{1}{\gamma}\right) p + \sum_{N < p < \gamma N} \left(N - \frac{1}{\gamma}p\right)$$

$$= \sum_{p \leq N} p + \sum_{N < p < \gamma N} N - \frac{1}{\gamma} \sum_{p < \gamma N} p.$$

Thus, by the prime number theorem and Abel summation, for $N > c_4$,

$$\sum_{p < \gamma N} \ell(p) < (1 + \varepsilon)\left(\frac{N^2}{2 \log N} + \frac{(\gamma - 1)N^2}{\log N} - \frac{\gamma}{2} \frac{N^2}{\log N}\right)$$

and so by (66),

$$\log Z < \frac{(1 + \varepsilon)}{9}(\gamma - 1)N^2 \log N + \varepsilon N^2 \log N + c_3 N^2. \qquad (67)$$

Comparing (64) and (67) we find that for $N > c_5$,

$$\frac{(1 - \varepsilon)}{2} < (1 + \varepsilon)\frac{(\gamma - 1)}{9} + \varepsilon + \frac{c_3}{\log N}.$$

But $\gamma < \frac{11}{2}$ and so for N sufficiently large we obtain a contradiction. Thus (63) does not hold for N sufficiently large and the result now follows.

References

[1] B. BRINDZA and P. ERDŐS, On some Diophantine problems involving powers and factorials, *J. Austral. Math. Soc. Ser. A* **51** (1991), 1–7.

[2] P. ERDŐS and R. GRAHAM, Old and new problems and results in combinatorial number theory, Monographie 28 de L'Enseignement Mathématique, *Genève*, 1980.

[3] P. ERDŐS and C. L. STEWART, On the greatest and least prime factors of $n! + 1$, *J. London Math. Soc.* **13**, no. 2 (1976), 513–519.

[4] D. R. HEATH-BROWN, The difference between consecutive primes, III, *J. London Math. Soc.* **20**, no. 2 (1979), 177–178.

[5] M. Le, On the Diophantine equation $x^{p-1} + (p-1)! = p^n$, *Publ. Math. Debrecen* **48** (1996), 145–149.

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

[7] F. Luca and I. E. Shparlinski, Prime divisors of shifted factorials (*to appear*).

[8] M. R. Murty and S. Wong, The *ABC* conjecture and prime divisors of the Lucas and Lehmer sequences, *Number Theory for the Millenium, III*, (Urbana, IL, 2000), A. K. Peters, Natick, MA, 2002, 43–54.

[9] C. L. Stewart, On divisor properties of arithmetical sequences, Ph.D. thesis, *University of Cambridge*, 1976.

[10] K. Yu and D. Liu, A complete resolution of a problem of Erdős and Graham, *Rocky Mountain J. Math.* **26** (1996), 1235–1244.

CAMERON L. STEWART
DEPARTMENT OF PURE MATHEMATICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO, N2L 3G1
CANADA

E-mail: cstewart@uwaterloo.ca
URL: http://www.math.uwaterloo.ca/PM_Dept/Homepages/Stewart/stewart.html