ACTA ARITHMETICA
LVII (1991)

On prime divisors of Mersenne numbers

by

PAUL ERDŐS (Budapest), PÉTER KISS* (Eger) and
CARL POMERANCE** (Athens, Ga.)

1. Introduction and results. The divisors of Mersenne numbers, i.e., the divisors of numbers of the form 2^n-1, have been investigated by several authors. Recently C. Pomerance [8] has obtained results on the magnitude of the reciprocal sum of the primitive divisors of Mersenne numbers proving and disproving some conjectures of P. Erdős [3].

In this note we consider only the prime divisors of Mersenne numbers. Put

$$f(n) = \sum_{p|2^n-1} \frac{1}{p}, \qquad n > 1;$$

that is, $f(n)$ is the reciprocal sum of the distinct prime divisors of the nth Mersenne number. P. Erdős [3] showed that there is a positive constant c such that

$$f(n) < \log\log\log n + c$$

for all large n. (Throughout the paper, we use c as a generic absolute constant, not necessarily the same at each appearance.) It can be easily seen that, apart from the precise value of c, this result is best possible: if $n = m!$, then $p|2^n-1$ for all odd primes $p \leqslant m$ and so

$$f(n) \geqslant \sum_{2<p\leqslant m} \frac{1}{p} > \log\log m + c > \log\log\log n + c.$$

On the other hand the reciprocal sum of the prime divisors can be arbitrarily small. For example, by a superficial argument, $f(n) < c/\log n$ follows if n is prime, since in this case every prime divisor of 2^n-1 is greater than n and the number of distinct prime divisors is less than $cn/\log n$. Furthermore from a result of P. Kiss and B. M. Phong [6], obtained for Lucas numbers,

* Research partially supported by Hungarian National Foundation for Scientific Research grant No. 273.

** Research partially supported by an NSF grant.

it follows that the average order of $f(n)$ is less than an absolute constant in the interval $(x, x+\log\log x)$ if x is sufficiently large.

In this paper we show that $f(n)$ can be "large", but not "too large", for arbitrarily many consecutive integers and we give an asymptotic formula for the average order of $f(n)$ which holds in "short" intervals.

THEOREM 1. *For any positive number C and integer s there exist consecutive integers $n, n+1, n+2, \ldots, n+s$ such that*

$$f(n+i) > C \quad \text{for } i = 0, 1, \ldots, s.$$

In fact we are able to prove the following stronger form of Theorem 1. Let $\log_k n$ denote the k-fold iterated natural logarithm.

THEOREM 2. *For each integer $k \geqslant 2$, there are infinitely many n with*

$$\min\{f(n), f(n+1), \ldots, f(n+k-1)\} \geqslant \log_{k+2} n + c\log_{k+3} n,$$

where c is an absolute constant.

One might wonder how close Theorem 2 is to the truth. It might seem that saying (in the case $k = 2$)

$$\min\{f(n), f(n+1)\} \geqslant \log_4 n + c\log_5 n$$

for infinitely many n is a quite weak result and that one might expect

$$\min\{f(n), f(n+1)\} = \Omega(\log_3 n).$$

In fact this is false. We show

THEOREM 3. *There is an absolute constant c such that*

$$\min\{f(n), f(n+1)\} \leqslant c(\log_3 n)^{2/3}(\log_4 n)^{1/3} \quad \text{for all large } n.$$

There is still a huge gap between Theorem 2 in the case $k = 2$ and Theorem 3. Almost certainly Theorem 2 is closer to the truth and in fact we can show this conditionally on the Extended Riemann Hypothesis.

THEOREM 4. *Assume the Extended Riemann Hypothesis holds for the Dedekind zeta functions for the fields K_p for each prime p, where K_p is the Galois closure of $Q(2^{1/p})$ and Q is the field of rational numbers. Then for every integer $k \geqslant 2$ we have*

$$\min\{f(n), f(n+1), \ldots, f(n+k-1)\} \leqslant 3\log_{k+2} n + ck,$$

where c is an absolute constant, for all sufficiently large n.

In fact, with a bit more work and assuming a stronger form of the ERH (namely, that it holds for each K_d where d is squarefree), we can replace the coefficient "3" in Theorem 4 with "1". However, we do not present this proof here. Note that this does not contradict Theorem 2 since the coefficient c in that theorem turns out to be negative.

Consider the function

$$g(n) = \sum_p \frac{(p-1, n)}{p(p-1)}$$

where the sum is over all primes p. In some ways $g(n)$ models the function $f(n)$, since we can view $g(n)$ as taking $1/p$ with "weight" $(p-1, n)/(p-1)$, while the "probability" that $p|2^n-1$ is $(p-1, n)/(p-1)$, since $p|2^n-1$ if and only if 2 is a $(p-1)/(p-1, n)$ power mod p. (This heuristic is not completely accurate since it ignores the special nature of the quadratic character of 2 mod p.) We can prove unconditionally that the maximal order of

$$\min\{g(n), g(n+1), \ldots, g(n+k-1)\}$$

is $\log_{k+2} n + O(\log_{k+3} n)$ for every $k \geqslant 1$, but we do not give the proof here.

It is easy to see that there is a constant $c_0 > 0$ such that

$$\sum_{n=1}^{x} f(n) = c_0 x + o(x)$$

for any integer x. We show this average result continues to hold for quite short intervals.

THEOREM 5. *If $z = z(x)$ is an integer valued function for which*

$$\frac{z}{\log\log\log x} \to \infty \quad \textit{as } x \to \infty,$$

then for any natural number x,

$$\sum_{n=x}^{x+z} f(n) = c_0 z + o(z).$$

Throughout the paper the letters p, q will always denote primes.

2. Notes and problems. From Theorem 5 it follows that

$$\frac{1}{x} \cdot \sum_{n=1}^{x} f(n) \to c_0 \quad \text{as } x \to \infty.$$

Now put

$$f_T(n) = \sum_{\substack{p|2^n-1 \\ p<T}} \frac{1}{p}.$$

It is easy to prove that for any $T>0$

$$\frac{1}{x}\cdot\sum_{n=1}^{x} f_T(n)\to c_T \qquad (n\to\infty)$$

and $c_T\to c_0$ as $T\to\infty$. Further we can prove that if $y\to\infty$ as slowly as we please then

$$\frac{1}{y}\cdot\sum_{x<n\leqslant x+y} f_T(n)\to c_T.$$

Note that Theorem 5 is best possible; i.e., it fails if $z\ll \log\log\log x$. We only have to remark that, as we have seen above, $f(n)\gg \log\log\log n$ is possible.

It can also be proved by more or less standard methods that the density of integers n for which $f(n)\leqslant C$ exists and is a continuous function of C. The same distribution holds for any interval $x<n\leqslant x+g(x)\log\log\log x$, where $g(x)\to\infty$ as slowly as we please. We suppress the proofs.

Perhaps the following problem is of some interest and not unattackable. Is it true that

$$\sum_{n=x}^{x+z}{}' f(n)=c_0 z+o(z)$$

whenever $z\geqslant \log\log\log x$, where the dash indicates that the largest term in the sum is deleted? In the spirit of Theorems 3 and 4, perhaps this is true under the assumption

$$z/((\log_3 x)^{2/3}(\log_4 x)^{1/3})\to\infty$$

or even, assuming the ERH,

$$z/\log_4 x\to\infty.$$

As we see from Theorem 2, we cannot hope to do better than $z/\log_4 x\to\infty$.

Generalizing, it is possible that for each fixed k,

$$\sum_{n=x}^{x+z}{}^{(k)} f(n)=c_0 z+o(z)$$

when $z/\log_{k+3} x\to\infty$, where $\sum^{(k)}$ indicates that the k largest terms are omitted from the sum.

In the introduction we remarked that $f(p)\ll 1/\log p$ is fairly trivial. In fact using the fact that primes $q\,|\,2^p-1$ satisfy $q\equiv 1(\bmod p)$ and the Brun–Titchmarsh inequality, we can prove

$$f(p)\ll (\log\log p)/p.$$

We conjecture that $pf(p)$ is unbounded, but this is probably a very hard problem. Note that there is a "large" infinite set S of primes p (large in the sense that the sum of the reciprocals of the members of S up to x is asymptotically loglog x) such that $f(p) = o(1/p)$ for $p \in S$—this is shown in [8]. Further, it is shown there that if the Extended Riemann Hypothesis holds, then $\sum_p f(p)$ converges.

We close this section with the solution of another problem from P. Erdős [3]. In (28) of this paper it is suggested that

$$f(n) \leqslant \max_{m \leqslant n} \sum_{\substack{p|m \\ p>2}} \frac{1}{p} + o(1) \quad \text{as } n \to \infty.$$

To see that this is untrue, let x be large and let n be the least common multiple of the integers up to x. Note that $2^n - 1$ is of course divisible by every odd prime p with $p-1 | n$. Every odd prime $p \leqslant x$ satisfies this condition. But from a result of P. Erdős [1], there are absolute constants $c > 0$, $\alpha > 0$ such that for all x large and all t with $t \leqslant x^{1+c}$, there are at least $\alpha\pi(t)$ primes $p \leqslant t$ with $p-1 | n$. Thus

$$\sum_{\substack{x < p < x^{1+c} \\ p|2^n-1}} \frac{1}{p} \gg 1.$$

Hence

$$f(n) - \sum_{2 < p \leqslant x} \frac{1}{p} \geqslant \sum_{\substack{x < p < x^{1+c} \\ p|2^n-1}} \frac{1}{p} \gg 1.$$

But by the prime number theorem,

$$\sum_{2 < p \leqslant x} \frac{1}{p} = \max_{m \leqslant n} \sum_{\substack{p|m \\ p>2}} \frac{1}{p} + o(1),$$

thus completing our disproof of (28) in [3].

Perhaps the following is true:

$$f(n) \leqslant \max_{m \leqslant n} \sum_{\substack{p-1|m \\ p>2}} \frac{1}{p} + o(1).$$

3. Proofs of Theorems 1–4. First we introduce a notation and recall some elementary properties of the sequence of Mersenne numbers.

For any odd positive integer m there are terms in the sequence $2^n - 1$, $n = 1, 2, \ldots$, divisible by m. Denote by $r(m)$ the rank of apparition of m in the sequence; i.e., $r(m)$ is a positive integer for which $m | 2^{r(m)} - 1$ but $m \nmid 2^n - 1$ if $0 < n < r(m)$. It is known that $m | 2^n - 1$ if and only if $r(m) | n$; furthermore $r(p) | p-1$ for any odd prime p and $r(m_1 m_2) = [r(m_1), r(m_2)]$ for any odd relatively prime integers m_1, m_2 ([,] denotes the least common multiple of numbers).

For the proof of Theorem 1 we need an auxiliary result.

LEMMA 1. *For any positive real number C and any positive integer m, there is an integer n such that $(m, n) = 1$ and $f(n) > C$.*

Proof. Let $C > 0$ be a real number and let m be an integer. We can choose primes $p_1, p_2, \dots, p_t$ of the form $8km-1$ such that

$$\sum_{i=1}^{t} \frac{1}{p_i} > C.$$

For these primes

$$2^{(p_i-1)/2} \equiv 1 \pmod{p_i}, \qquad i = 1, 2, \dots, t,$$

since 2 is quadratic residue modulo p_i, so that

$$r(p_i) \,\Big|\, \frac{p_i - 1}{2}.$$

Thus for the number $p_1 p_2 \dots p_t$, the rank of apparition

$$n := r(p_1 p_2 \dots p_t) = [r(p_1), r(p_2), \dots, r(p_t)]$$

satisfies $(n, m) = 1$ and $p_i | (2^n - 1)$ for $i = 1, 2, \dots, t$. Thus

$$f(n) \geqslant \sum_{i=1}^{t} \frac{1}{p_i} > C$$

follows and the lemma is proved.

From this lemma, Theorem 1 follows.

Proof of Theorem 1. By Lemma 1 we can construct integers n_0, $n_1, \dots, n_s$ such that $(n_i, n_j) = 1$ for any $i \neq j$ and $f(n_i) > C$ for $i = 0, 1, \dots, s$. By the Chinese remainder theorem there are integers $n, n+1, \dots, n+s$ such that $n_i | n+i$ for any i with $0 \leqslant i \leqslant s$ and by the properties of the sequence $2^n - 1$, mentioned above, we have

$$f(n+i) \geqslant f(n_i) > C, \qquad i = 0, 1, \dots, s,$$

which proves the theorem.

Proof of Theorem 2. Let $k \geqslant 2$. Let

$$\alpha_j(n) = \exp\big((\log_j n)/(\log_{j+1} n)^2\big)$$

and let $A_j(n)$ be the least common multiple of the integers up to $\alpha_j(n)$. Let $B_0(n) = A_{k+1}(n)$ and let $B_j(n)$ be the largest divisor of $A_{k+1-j}(n)$ coprime to $A_{k+2-j}(n)$ for $j = 1, \dots, k-1$. Then

(i) $B_0(n), \dots, B_{k-1}(n)$ are pairwise coprime,

(ii) $B_0(n) \dots B_{k-1}(n) \leqslant A_2(n) = n^{o(1)}$,

the last following from the prime number theorem. Thus by the Chinese remainder theorem, there are infinitely many integers n with

(1) $$B_j(n) | n+j \quad \text{for } j = 0, 1, \dots, k-1.$$

Suppose (1) holds for n. Then $B_0(n)|n$, so that $p-1|n$ for every prime $p \leqslant \alpha_{k+1}(n)$. Thus

$$f(n) \geqslant \sum_{2<p\leqslant \alpha_{k+1}(n)} \frac{1}{p} = \log\log \alpha_{k+1}(n) + O(1) \tag{2}$$
$$= \log_{k+2} n - 2\log_{k+3} n + O(1).$$

Suppose (1) holds for n and $1 \leqslant j \leqslant k-1$. Let S_j be the set of primes p such that

(i) $p \leqslant \alpha_{k+1-j}(n)$,
(ii) $p \equiv 7 \pmod 8$,
(iii) $\big((p-1)/2,\ A_{k+2-j}(n)\big) = 1$.

Note that if a prime $q|A_{k+2-j}(n)$, then

$$q \leqslant \alpha_{k+2-j}(n) \leqslant \big(\log \alpha_{k+1-j}(n)\big)^{o(1)}.$$

Since S_j is the set of primes p satisfying (i), (ii) such that $(p-1)/2$ is sifted out by the primes up to $\alpha_{k+2-j}(n)$, it follows from A. Selberg's sieve (see H. Halberstam and H.-E. Richert [4], Theorem 7.1) and a moderately strong form of the prime number theorem for arithmetic progressions that

$$\sum_{\substack{p\in S_j \\ p\leqslant t}} 1 \sim \tfrac{1}{4}\pi(t) \prod_{2<q\leqslant \alpha_{k+2-j}(n)} \left(1-\frac{1}{q-1}\right)$$

uniformly for

$$\exp\big((\log \alpha_{k+1-j}(n))^{\varepsilon}\big) \leqslant t \leqslant \alpha_{k+1-j}(n)$$

for every $\varepsilon > 0$. Thus

$$\sum_{p\in S_j} \frac{1}{p} \sim \frac{1}{4}\log\log \alpha_{k+1-j}(n) \prod_{2<q\leqslant \alpha_{k+2-j}(n)} \left(1-\frac{1}{q-1}\right) \tag{3}$$
$$\gg \frac{\log\log \alpha_{k+1-j}(n)}{\log \alpha_{k+2-j}(n)} \gg \frac{\log_{k+2-j} n}{(\log_{k+2-j} n)/(\log_{k+3-j} n)^2} = (\log_{k+3-j} n)^2.$$

But if $p \in S_j$ and (1) holds for n, then $p|2^{n+j}-1$. Thus from (3),

$$f(n+j) \gg (\log_{k+3-j} n)^2 \geqslant (\log_{k+2} n)^2$$

for $j = 1, \ldots, k-1$. Together with (2), this proves the theorem.

To prove Theorem 3 we first prove the following key lemma.

LEMMA 2. *Uniformly for all $x \geqslant 3$ and all natural numbers n,*

$$\sum_{\substack{x<p\leqslant x^e \\ p|2^n-1}} \frac{1}{p} \ll \exp\left(-\sum_{\substack{\log\log x<q<\log x \\ q\nmid n}} \frac{1}{q}\right).$$

Proof. Let

$$m = \prod_{\substack{\log\log x < q < (\log x)^{1/7} \\ q \nmid n}} q.$$

Then

$$\sum_{\substack{x < p \leqslant x^{e} \\ p \mid 2^n - 1}} \frac{1}{p} \leqslant \sum_{\substack{x < p \leqslant x^{e} \\ (p-1, m) = 1}} \frac{1}{p} + \sum_{q \mid m} \sum_{\substack{x < p \leqslant x^{e} \\ p \equiv 1 \pmod q \\ p \mid 2^n - 1}} \frac{1}{p}. \tag{4}$$

By the sieve we have

$$\sum_{\substack{x < p \leqslant x^{e} \\ (p-1, m) = 1}} \frac{1}{p} \ll \exp\left(-\sum_{q \mid m} \frac{1}{q}\right) \ll \exp\left(-\sum_{\substack{\log\log x < q < \log x \\ q \nmid n}} \frac{1}{q}\right). \tag{5}$$

Suppose $q \mid m$, $p \equiv 1 \pmod q$ and $p \mid 2^n - 1$. Since $q \nmid n$, it follows that 2 is a qth power mod p. Since $q \leqslant (\log x)^{1/7}$, it follows from Theorems 1.3 and 1.4 of J. C. Lagarias and A. M. Odlyzko [7], that

$$\sum_{\substack{x < p \leqslant x^{e} \\ p \equiv 1 \pmod q \\ p \mid 2^n - 1}} \frac{1}{p} \leqslant \sum_{\substack{x < p \leqslant x^{e} \\ p \equiv 1 \pmod q \\ 2 \text{ is } q^{\text{th}} \text{ power mod } p}} \frac{1}{p} \sim \frac{1}{q(q-1)} \tag{6}$$

uniformly. Since

$$\sum_{q \mid m} \frac{1}{q(q-1)} \ll \frac{1}{\log\log x} \ll \exp\left(-\sum_{q < \log x} \frac{1}{q}\right),$$

the lemma follows from (4) and (5).

Proof of Theorem 3. Let

$$a = \log_3 n, \qquad b = \exp\left((\log_3 n)^{2/3} (\log_4 n)^{1/3}\right)$$

and let

$$A_j = \sum_{\substack{a < q < b/e \\ q \nmid n + j}} \frac{1}{q} \quad \text{for } j = 0, 1.$$

No prime q can divide both n and $n+1$, so that

$$A_0 + A_1 \geqslant \sum_{a < q < b/e} \frac{1}{q} = \log\log b - \log\log a + o(1).$$

Thus

$$\max\{A_0, A_1\} \geqslant \tfrac{1}{2}(\log\log b - \log\log a) - 1$$

for all large n. We shall now prove that if $A_j = \max\{A_0, A_1\}$, then

$$f(n+j) \ll (\log_3 n)^{2/3} (\log_4 n)^{1/3}.$$

Without loss of generality, assume $j = 0$, that is, that

$$A_0 \geqslant \tfrac{1}{2}(\log\log b - \log\log a) - 1. \tag{7}$$

We have

$$f(n) = \sum_{\substack{p \leqslant e^b \\ p|2^n-1}} \frac{1}{p} + \sum_{\substack{e^b < p < \log n \\ p|2^n-1}} \frac{1}{p} + \sum_{\substack{p \geqslant \log n \\ p|2^n-1}} \frac{1}{p} = B_0 + B_1 + B_2, \quad \text{say.} \tag{8}$$

We trivially have

$$B_0 \leqslant \log b + O(1) = (\log_3 n)^{2/3}(\log_4 n)^{1/3} + O(1)$$

and from the proof of the main theorem in P. Erdős [3] it follows that

$$B_2 = \sum_{\substack{p \geqslant \log n \\ p|2^n-1}} \frac{1}{p} = O(1) \tag{9}$$

(without using the assumption (7)).

It remains to estimate B_1. We have by Lemma 2 that

$$B_1 \leqslant \sum_{[\log b] \leqslant i < a} \sum_{\substack{e^{e^i} < p \leqslant e^{e^{i+1}} \\ p|2^n-1}} \frac{1}{p} \ll \sum_{[\log b] \leqslant i < a} \exp\Big(-\sum_{\substack{i < q < e^i \\ q \nmid n}} \frac{1}{q}\Big)$$

$$\leqslant \sum_{[\log b] \leqslant i < a} \exp\Big(-\sum_{\substack{a < q < b/e \\ q \nmid n}} \frac{1}{q}\Big) \leqslant a \cdot \exp(-A_0).$$

By (7),

$$a \cdot \exp(-A_0) \ll a\left(\frac{\log a}{\log b}\right)^{1/2} = (\log_3 n)^{2/3}(\log_4 n)^{1/3}$$

and so the theorem follows from (8) and the above estimates for B_0, B_1, B_2.

Before we prove Theorem 4, we need the following stronger, but conditional analog to Lemma 2.

LEMMA 3. *Suppose the Extended Riemann Hypothesis holds for the Dedekind zeta-functions for the fields K_p for every prime p, where K_p is the Galois closure of $Q(2^{1/p})$. Then uniformly for all $x > 1$ and all natural numbers n we have*

$$\sum_{\substack{x < p \leqslant x^e \\ p|2^n-1}} \frac{1}{p} \ll \exp\Big(-\sum_{\substack{\log x < q < x \\ q \nmid n}} \frac{1}{q}\Big).$$

Proof. As in the proof of Lemma 2, we have (4) for any integer m. Let now

$$m = \prod_{\substack{\log x < q < x^{1/3} \\ q \nmid n}} q.$$

Then, as in (5), we have

$$\sum_{\substack{x < p \leqslant x^e \\ (p-1,m)=1}} \frac{1}{p} \ll \exp\left(-\sum_{q|m} \frac{1}{q}\right) \ll \exp\left(-\sum_{\substack{\log x < q < x \\ q \nmid n}} \frac{1}{q}\right). \tag{10}$$

It further follows from the hypothesis of the lemma and (115) on p. 56 of C. Hooley [5] that for each prime $q|m$, we have (6) uniformly. Thus

$$\sum_{q|m} \sum_{\substack{x < p \leqslant x^e \\ p \equiv 1 (\mathrm{mod}\, q) \\ p|2^n-1}} \frac{1}{p} \ll \sum_{q|m} \frac{1}{q(q-1)} \ll \frac{1}{\log x} \ll \exp\left(-\sum_{q<x} \frac{1}{q}\right)$$

and the lemma follows from this estimate, (4) and (10).

Proof of Theorem 4. Let $k \geqslant 2$ and let

$$\beta_j = \exp((\log_j n)^3) \quad \text{for } j = 2, 3, \ldots$$

For $m \in \{n, n+1, \ldots, n+k-1\}$, let

$$A_j(m) = \sum_{\substack{q \nmid m \\ \log \beta_{k-j} < q \leqslant \beta_{k+1-j}^{1/e}}} \frac{1}{q}, \quad \text{for } j = 0, 1, \ldots, k-2.$$

Note that we trivially have for $j = 0, 1, \ldots, k-2$,

$$\begin{aligned} A_j(m) &\leqslant \log\log \beta_{k+1-j} - \log\log\log \beta_{k-j} - 1 + o(1) \\ &= 2\log_{k+2-j} n - 1 - \log 3 + o(1). \end{aligned} \tag{11}$$

Further, if n is large and $q > \log \beta_k$ divides one of $n, n+1, \ldots, n+k-1$, it does not divide any other of those k numbers. Thus if $S \subset \{n, n+1, \ldots, n+k-1\}$, it follows that

$$\begin{aligned} \sum_{m \in S} A_j(m) &\geqslant (|S|-1)(\log\log \beta_{k+1-j} - \log\log\log \beta_{k-j} - 1 + o(1)) \\ &= (|S|-1)(2\log_{k+2-j} n - 1 - \log 3 + o(1)). \end{aligned} \tag{12}$$

Let $S_k = \{n, n+1, \ldots, n+k-1\}$. We claim that if n is large, then for all $m \in S_k$, but for at most one exception, we have

$$A_0(m) \geqslant \log_{k+2} n - 2. \tag{13}$$

For if there were two or more exceptions to (13), then from (11),

$$\sum_{m \in S_k} A_j(m) \leqslant (k-2)(2\log_{k+2} n - 1 - \log 3) + 2\log_{k+2} n - 4 + o(1),$$

contradicting (12) for n large. Let m_k be the exception to (13) if there is an exception and otherwise let $m_k = n+k-1$. Let $S_{k-1} = S_k \backslash \{m_k\}$. We similarly get that

$$A_1(m) \geqslant \log_{k+1} n - 2 \tag{14}$$

for all $m \in S_{k-1}$ but for at most one exception, so that we can construct $S_{k-2} \subset S_{k-1}$ of cardinality $k-2$ and where both (13) and (14) hold.

Continuing, we create a sequence (for large n)

$$S_k \supset S_{k-1} \supset \dots \supset S_1$$

where S_j has cardinality j and if m is the single element of S_1 we have

$$A_j(m) \geqslant \log_{k+2-j} n - 2 \qquad \text{for } j = 0, 1, \dots, k-2. \tag{15}$$

We now show that if (15) holds for $m \in \{n, n+1, \dots, n+k-1\}$, we have

$$f(m) \leqslant 3\log_{k+2} n + O(k)$$

which will establish the theorem. Without loss of generality, we will assume that (15) holds for $m = n$.

We have

$$f(n) = \sum_{p|2^n-1} \frac{1}{p} = \sum_{j=0}^{k} B_j, \tag{16}$$

where

$$B_0 = \sum_{\substack{p|2^n-1 \\ p \leqslant \beta_{k+1}}} \frac{1}{p}, \quad B_k = \sum_{\substack{p|2^n-1 \\ p > \beta_2}} \frac{1}{p}, \quad B_j = \sum_{\substack{p|2^n-1 \\ \beta_{k+2-j} < p \leqslant \beta_{k+1-j}}} \frac{1}{p}, \quad \text{for } j = 1, \dots, k-1.$$

We trivially have

$$B_0 \leqslant \sum_{p \leqslant \beta_{k+1}} \frac{1}{p} = \log\log \beta_{k+1} + O(1) = 3\log_{k+2} n + O(1)$$

and from (9) we have

$$B_k = O(1).$$

We now estimate each B_j for $j = 1, \dots, k-1$. We have by Lemma 3 and (15),

$$\begin{aligned} B_j &\leqslant \sum_{\substack{i \\ e^{e^i} < \beta_{k+1-j} \\ e^{e^{i+1}} > \beta_{k+2-j}}} \sum_{\substack{p|2^n-1 \\ e^{e^i} < p \leqslant e^{e^{i+1}}}} \frac{1}{p} \ll \sum_{\substack{i \\ e^{e^i} < \beta_{k+1-j} \\ e^{e^{i+1}} > \beta_{k+2-j}}} \exp\Big(-\sum_{\substack{q \nmid n \\ e^i < q < e^{e^i}}} \frac{1}{q}\Big) \\ &\leqslant \sum_{i < \log\log\beta_{k+1-j}} \exp\Big(-\sum_{\substack{q \nmid n \\ \log\beta_{k+1-j} < q < \beta_{k+2-j}^{1/e}}} \frac{1}{q}\Big) \\ &\leqslant (\log\log \beta_{k+1-j}) \exp\big(-A_{j-1}(n)\big) \ll (\log_{k+2-j} n)(\log_{k+2-j} n)^{-1} = 1. \end{aligned}$$

Thus by (16),

$$f(n) \leqslant 3\log_{k+2} n + O(k),$$

which was to be proved.

4. The proof of Theorem 5. In the proof of Theorem 5 we shall use two more lemmas.

LEMMA 4. *For any* $y > 3$ *we have*

$$\sum_{\substack{p\,\text{prime}\\ r(p)\leqslant y}} 1/p = \log\log y + O(1).$$

Proof. Since $r(p) \leqslant p-1$ for any odd prime, we obtain a trivial lower estimation

$$\sum_{r(p)\leqslant y} \frac{1}{p} \geqslant \sum_{p\leqslant y} \frac{1}{p} + O(1) = \log\log y + O(1). \tag{17}$$

On the other hand $2^n - 1$ has at most n distinct prime factors so in the sum there are at most y^2 primes and by the prime number theorem we get, for y large,

$$\sum_{r(p)\leqslant y} \frac{1}{p} \leqslant \sum_{p<y^3} \frac{1}{p} = \log\log y + O(1). \tag{18}$$

From (17) and (18) the lemma follows.

LEMMA 5. *The sum*

$$\sum_{\substack{p\,\text{prime}\\ p>2}} \frac{1}{p\cdot r(p)}$$

converges.

Proof. This follows from the papers of P. Erdős [2] and N. P. Romanoff [9] where it is shown the larger sum

$$\sum_{d\,\text{odd}} \frac{1}{d\cdot r(d)}$$

converges. However, Lemma 5 is completely trivial since $2^{r(p)} - 1 \geqslant p$ implies $r(p) \gg \log p$. It remains to note that

$$\sum \frac{1}{p \log p}$$

converges.

Proof of Theorem 5. Let x and z be sufficiently large positive integers with $z < x$. (For $z \geqslant x$, the theorem follows easily from the case $z < x$.) By the definitions of $f(n)$ and $r(p)$ we can write

$$\sum_{n=x}^{x+z} f(n) = A(x) + B(x), \tag{19}$$

where

$$A(x) = \sum_{n=x}^{x+z} \sum_{\substack{d|n \\ d \leqslant z}} \sum_{r(p)=d} \frac{1}{p} \quad \text{and} \quad B(x) = \sum_{n=x}^{x+z} \sum_{\substack{d|n \\ d > z}} \sum_{r(p)=d} \frac{1}{p}.$$

First we deal with $A(x)$. Since $p|2^n-1$ if and only if $r(p)|n$, by Lemmas 4 and 5 we have

$$\begin{aligned} A(x) &= \sum_{d \leqslant z} \left(\frac{z}{d}+O(1)\right) \sum_{r(p)=d} \frac{1}{p} \\ &= z \sum_{r(p) \leqslant z} \frac{1}{p \cdot r(p)} + O\left(\sum_{r(p) \leqslant z} \frac{1}{p}\right) = c_0 z + o(z), \end{aligned} \tag{20}$$

where c_0 is the infinite sum in Lemma 5.

In order to give an estimation for the sum $B(x)$ we cut it into three parts. Let

$$B_1(x) = \sum_{n=x}^{x+z} \sum_{\substack{d|n \\ z<d \leqslant (\log x)^4}} \sum_{r(p)=d} \frac{1}{p}.$$

Since every d with $d > z$ occurs at most once in the sum, by Lemma 4 we get

$$B_1(x) \leqslant \sum_{d \leqslant (\log x)^4} \sum_{r(p)=d} \frac{1}{p} = \sum_{r(p) \leqslant (\log x)^4} \frac{1}{p} = \log\log\log x + O(1). \tag{21}$$

For the sum

$$B_2(x) = \sum_{n=x}^{x+z} \sum_{\substack{d|n \\ d>z \\ d>(\log x)^4}} \sum_{\substack{r(p)=d \\ p \geqslant d^3}} \frac{1}{p},$$

note that there are at most d distinct primes with $r(p) = d$, so that

$$B_2(x) \leqslant \sum_{d>(\log x)^4} d \cdot \frac{1}{d^3} < \sum_{d=1}^{\infty} \frac{1}{d^2} = O(1). \tag{22}$$

The most difficult part of this proof is to give an estimation for

$$B_3(x) = \sum_{n=x}^{x+z} \sum_{\substack{d|n \\ d>z \\ d>(\log x)^4}} \sum_{\substack{r(p)=d \\ p<d^3}} \frac{1}{p}. \tag{23}$$

Since $p > r(p)$, we have

$$B_3(x) \leqslant \sum_{i \geqslant 2} \sum\nolimits_i \frac{1}{p} \tag{24}$$

where the summation in $\sum_i$ is the same as in $B_3(x)$ but we only take primes p for which

$$(\log x)^{2^i} < p \leqslant (\log x)^{2^{i+1}}.$$

Let Q denote the integer $x(x+1)\dots(x+z)$. Fix some $i \geqslant 2$. If p is counted in $\sum_i$, then

$$(p-1, Q) \geqslant r(p) > p^{1/3} > (\log x)^{2^i/3} \tag{25}$$

Let y be a real number such that

$$y_1 := (\log x)^{2^i} < y \leqslant (\log x)^{2^{i+1}} =: y_2$$

and let $S(y)$ be the set of primes $p \leqslant y$ for which (25) holds. By (25) it is clear that

$$\prod_{p \leqslant y} (p-1, Q) \geqslant \prod_{p \in S(y)} (p-1, Q) \geqslant (\log x)^{2^i |S(y)|/3}. \tag{26}$$

We now proceed in a manner analogous to that in P. Erdős [3]. Note that (where Λ is von Mangoldt's function and $\pi(y, d, 1)$ is the number of primes $p \leqslant y$ with $p \equiv 1 \pmod d$)

$$\log \prod_{p \leqslant y} (p-1, Q) = \sum_{p \leqslant y} \log(p-1, Q) = \sum_{p \leqslant y} \sum_{d|(p-1, Q)} \Lambda(d) \tag{27}$$
$$= \sum_{d|Q} \Lambda(d)\pi(y, d, 1) = S_1 + S_2,$$

say, where in S_1 we have $d \leqslant y^{2/3}$ and in S_2 we have $d > y^{2/3}$.

For S_1 we use the Brun–Titchmarsh inequality to get

$$S_1 \ll \sum_{d|Q} \frac{\Lambda(d)}{\varphi(d)} \frac{y}{\log y} \ll \frac{y \log\log Q}{\log y}.$$

For S_2 we estimate $\pi(y, d, 1)$ trivially as $\leqslant y/d$ and use the fact that Q has at most $O(\log Q)$ prime power divisors to get

$$S_2 \leqslant \sum_{\substack{d|Q \\ y^{2/3} < d < y}} \frac{\Lambda(d)}{d} y \ll \frac{y \log y \log Q}{y^{2/3}} \ll \frac{y \log Q}{y^{1/2} \log y} < \frac{y \log Q}{(\log x)^2 \log y}.$$

Putting these estimates in (26) and (27) we get

$$|S(y)| \leqslant \frac{3}{2^i \log\log x}(S_1 + S_2) \ll \frac{y}{2^i \log y}\left(\frac{\log\log Q}{\log\log x} + \frac{\log Q}{(\log x)^2}\right).$$

But $z < x$ implies

$$\log Q \ll z \log x, \quad \log\log Q \ll \log z + \log\log x,$$

so that

$$|S(y)| \ll \frac{y}{2^i \log y}\left(1 + \frac{\log z}{\log\log x} + \frac{z}{\log x}\right). \tag{28}$$

By partial summation, we have

$$\sum_i \frac{1}{p} \leqslant \sum_{p \in S(y_2)} \frac{1}{p} = y_2^{-1} |S(y_2)| + \int_{y_1}^{y_2} \frac{1}{y^2} |S(y)| \, dy$$
$$\ll 2^{-i}\left(1 + \frac{\log z}{\log\log x} + \frac{z}{\log x}\right),$$

where we use (28). Thus from (24) we have

$$B_3(x) \ll 1+\frac{\log z}{\log\log x}+\frac{z}{\log x} = o(z). \tag{29}$$

Since

$$B(x) = B_1(x)+B_2(x)+B_3(x),$$

by (21), (22) and (29)

$$B(x) \leqslant \log\log\log x + o(z),$$

so that by (19) and (20) we get

$$\sum_{n=x}^{x+z} f(n) = c_0 z + O(\log\log\log x) + o(z) = c_0 z + o(z).$$

References

[1] P. Erdős, *On the normal number of prime factors of $p-1$ and some related questions concerning Euler's ϕ-function*, Quart. J. Math. Oxford Ser. 6 (1935), 205–213.

[2] – *On some problems of Bellman and a theorem of Romanoff*, J. Chinese Math. Soc. (N.S.) 1 (1951), 409–421.

[3] – *On the sum $\sum_{d|2^n-1} d^{-1}$*, Israel J. Math. 9 (1971), 43–48.

[4] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London 1974.

[5] C. Hooley, *Applications of Sieve Methods to the Theory of Numbers*, Cambridge University Press, Cambridge 1976.

[6] P. Kiss and B. M. Phong, *Weakly composite Lucas numbers*, Ann. Univ. Sci. Budapest Eötvős, sect. Math. 31 (1988), 179–182.

[7] J. C. Lagarias and A. M. Odlyzko, *Effective versions of the Chebotarev density theorem*, in A. Fröhlich, ed., *Algebraic Number Fields* (Proceedings of the 1975 Durham Symposium), Academic Press, London 1977, 409–464.

[8] C. Pomerance, *On primitive divisors of Mersenne numbers*, Acta Arith. 46 (1986), 355–367.

[9] N. P. Romanoff, *Über einige Sätze der additiven Zahlentheorie*, Math. Ann. 109 (1934), 668–678.

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
Budapest, Hungary

TEACHER'S TRAINING COLLEGE
DEPARTMENT OF MATHEMATICS
Eger, Hungary

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GEORGIA
Athens, Georgia 30602, U.S.A.

Received on 12.10.1989 (1978)