ACTA ARITHMETICA
XLVI (1986)

On primitive divisors of Mersenne numbers

by

CARL POMERANCE* (Athens, Ga.)

An interesting source of problems in number theory is to study the restriction of a familiar function to a special set of integers. For example, it is common in the literature to see papers which study the various divisor functions at polynomials evaluated at natural numbers or at primes. In his interesting paper [2], Erdös considered the divisor function

$$\sigma_{-1}(m) = \sum_{d|m} 1/d$$

restricted to Mersenne numbers; that is, numbers m of the form $2^n - 1$.

It is well known that

$$\sigma_{-1}(m) = O(\log\log m).$$

Thus letting $m = 2^n - 1$, we trivially have

$$\sigma_{-1}(2^n - 1) = O(\log n).$$

What Erdös proved in [2] is the surprisingly difficult result that the $\log n$ can be replaced with $\log\log n$:

$$\sigma_{-1}(2^n - 1) = O(\log\log n).$$

It is not so hard to see that this result is best possible.

For d an odd natural number, let $l(d)$ denote the exponent to which 2 belongs modulo d. That is,

$$2^{l(d)} \equiv 1 \bmod d \quad \text{and} \quad 2^l \not\equiv 1 \bmod d \quad \text{for all } l,\ 1 \leqslant l < l(d).$$

Let

$$E(n) = \sum_{l(d)=n} 1/d.$$

---

* Research supported in part by an NSF grant.

The function $E(n)$ is connected to Erdös's result via the identity

$$\sigma_{-1}(2^n-1)=\sum_{m|n} E(m).$$

In another paper [1], Erdös showed that

$$\sum_{n\leqslant x} E(n)\leqslant c_1 \log x, \qquad c_2:=\sum_{d \text{ odd}} 1/dl(d)=\sum_n E(n)/n<\infty.$$

(The latter result had been shown earlier by Romanoff [9] by a complicated argument; Erdös's argument in [1] gives this as an easy corollary of the former result, which itself is not so hard.) The proof of the former result actually gives

$$\sum_{n\leqslant x} E(n)\leqslant \left(e^\gamma+o(1)\right)\log x, \tag{1}$$

where γ is Euler's constant. These results are of interest in that they immediately imply that

$$\sum_{n\leqslant x}\sigma_{-1}(2^n-1)\sim c_2 x \quad \text{as} \quad x\to\infty.$$

Thus the maximal and average orders of $\sigma_{-1}(m)$ for m a Mersenne number have been satisfactorily handled. It is to be remarked, however, that the situations for $d(m)$, $\omega(m)$ are much harder and are far from resolution. Here d, ω respectively count the number of natural divisors and the number of distinct prime divisors.

The function $E(n)$, which proved a useful tool in Erdös's papers [1], [2], seems interesting in its own right. In fact, in these papers, Erdös asks for the average order, normal order, minimal order, and maximal order of $E(n)$.

Concerning the average order, we trivially have

$$\sum_{n\leqslant x} E(n)\geqslant \sum_{d\leqslant x, d \text{ odd}} 1/d>\tfrac{1}{2}\log x. \tag{2}$$

In [2] Erdös conjectures that there is a c_3 with

$$\sum_{n\leqslant x} E(n)=\left(c_3+o(1)\right)\log x.$$

I am unable to prove this conjecture, but I show below that the constant e^γ in (1) can be replaced with a smaller constant. Probably the "correct" value of c_3 is 1/2.

In view of (1) and (2) it seems natural to measure $E(n)$ by its ratio with $1/n$. Thus in [2] Erdös suggests that $nE(n)$ has a distribution function and that

$$\limsup nE(n)=\infty, \qquad \liminf nE(n)=0. \tag{3}$$

The first limit was also conjectured in [1]. Erdös states he can prove $E(n) = o(1)$ and even that

(4) $$A(n) := \sum_{d \mid 2^n - 1, d > n} 1/d = o(1)$$

but he gives no explicit function tending to 0 as $n \to \infty$. He suggests that for every $\varepsilon > 0$,

(5) $$A(n) = O(n^{-1+\varepsilon}).$$

Below I establish the limits in (3), but show that $nE(n)$ does not have a distribution function. In fact I show that there is a set S of natural numbers of logarithmic density 1 such that

(6) $$\lim_{n \in S, n \to \infty} nE(n) = 0.$$

Recall that the logarithmic density, should it exist, of a set of natural numbers A is

$$\lim_{x \to \infty} \frac{1}{\log x} \sum_{n \in A, n \leqslant x} \frac{1}{n}.$$

Probably (6) holds for a set S of natural density 1, but I have been unable to show this.

I establish a more quantitative form for the limits in (3) by showing that

$$nE(n) \geqslant \exp\{(1+o(1))\sqrt{\log\log n}\}, \qquad nE(n) \leqslant c_4 (\log n)^{-17/24}$$

each hold for infinitely many n.

The relative logarithmic density, should it exist, of a set of primes B is

$$\lim_{x \to \infty} \frac{1}{\log\log x} \sum_{p \in B, p \leqslant x} \frac{1}{p}.$$

I show that (6) holds for a set S of primes of relative logarithmic density 1. This is accomplished as a corollary to a result of independent interest: the set of primes p with $l(p)$ prime has relative natural density 0.

I strengthen (4) by showing

$$A(n) \leqslant \exp(-\log n \log\log\log n / 2 \log\log n)$$

for all large n. Probably this result is near to best possible. A heuristic argument which implies that

$$E(n) \geqslant \exp\{-(1+o(1)) \log n \log\log\log n / \log\log n\}$$

for infinitely many n is presented. If this conjecture is to be believed, then (5) fails for every $\varepsilon < 1$.

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

THEOREM 1. *There is a set of natural numbers S of logarithmic density* 1 *such that*

$$\lim_{n\in S,\, n\to\infty} nE(n) = 0.$$

Proof. We will show the equivalent assertion that for each $\varepsilon > 0$ the set $T(\varepsilon)$ of n with $nE(n) > \varepsilon$ has logarithmic density 0. For each prime q, let

$$\boldsymbol{P}(q) = \{p \text{ prime: } p \equiv 1 \bmod q,\ 2 \text{ is not a } q\text{th power} \bmod p\}.$$

Note that if $p \in \boldsymbol{P}(q)$, then $q \mid l(p)$. By the Prime Ideal Theorem of Landau [6], the number of members of $\boldsymbol{P}(q)$ not exceeding x is

$$\frac{1}{q}\frac{x}{\log x} + O_q\left(\frac{x}{\log^2 x}\right).$$

Thus

$$\prod_{p\leqslant x,\, q\nmid l(p)} \frac{p}{p-1} \leqslant \prod_{p\leqslant x,\, p\notin \boldsymbol{P}(q)} \frac{p}{p-1} = O_q\big((\log x)^{1-1/q}\big). \tag{7}$$

For each $\varepsilon > 0$ and prime q, let $T(\varepsilon, q)$ denote the set of n with $nE(n) > \varepsilon$ and $q \nmid n$. Then

$$\begin{aligned}\sum_{n\in T(\varepsilon,q),\, n\leqslant x} \frac{1}{n} &< \frac{1}{\varepsilon} \sum_{n\in T(\varepsilon,q),\, n\leqslant x} E(n) \leqslant \frac{1}{\varepsilon} \sum_{n\leqslant x,\, q\nmid n} E(n)\\ &= \frac{1}{\varepsilon} \sum_{l(d)\leqslant x,\, q\nmid l(d)} \frac{1}{d} < \frac{1}{\varepsilon} \sum_{p\mid d \Rightarrow l(p)\leqslant x,\, q\nmid l(p)} \frac{1}{d}\\ &= \frac{1}{\varepsilon} \prod_{l(p)\leqslant x,\, q\nmid l(p)} \frac{p}{p-1}.\end{aligned}$$

We now use the trivial fact that the number of primes p with $l(p) = d$ is less than d. Thus the number of primes p with $l(p) \leqslant x$ is less than x^2 which is less than the number of primes $p \leqslant x^3$ with $q \nmid l(p)$ provided $x \geqslant x_0(q)$. (To see this last assertion, note that for $q \geqslant 3$, if $p \not\equiv 1 \bmod q$, then $q \nmid l(p)$, while for $q = 2$, if $p \equiv 7 \bmod 8$, then $2 \nmid l(p)$.) Thus for $x \geqslant x_0(q)$,

$$\sum_{n\in T(\varepsilon,q),\, n\leqslant x} \frac{1}{n} < \frac{1}{\varepsilon} \prod_{p\leqslant x^3,\, q\nmid l(p)} \frac{p}{p-1} = O_q\left(\frac{1}{\varepsilon}(\log x)^{1-1/q}\right) \tag{8}$$

by (7). That is, the logarithmic density of $T(\varepsilon, q)$ is 0 for any choice of ε, q. But

$$T(\varepsilon) \subset T(\varepsilon, q) \cup q\boldsymbol{N}, \tag{9}$$

so that the upper logarithmic density of $T(\varepsilon)$ is at most $1/q$ for any ε, q. It thus follows that the logarithmic density of $T(\varepsilon)$ is 0 for any ε, which was to be proved.

COROLLARY 1. $\limsup nE(n) = \infty$.

Proof. If not, then there is a number M such that $nE(n) \leqslant M$ for all n. Then from (2)

$$\text{(10)} \qquad \tfrac{1}{2}\log x \leqslant \sum_{n\leqslant x} E(n) = \sum_{n\leqslant x,\, nE(n)\leqslant 1/4} E(n) + \sum_{n\leqslant x,\, nE(n)>1/4} E(n)$$

$$\leqslant \sum_{n\leqslant x} \frac{1}{4n} + \sum_{n\leqslant x,\, n\in T(1/4)} \frac{M}{n} = \left(\tfrac{1}{4}+o(1)\right)\log x,$$

a contradiction.

COROLLARY 2. *There is an infinite sequence of integers n on which $n(\log n)^{17/24}E(n)$ is bounded.*

Proof. From Odoni [7],

$$\sum_{p\leqslant x,\, 2\nmid l(p)} 1 = \frac{7}{24}\frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right).$$

Using this in (8) with $q = 2$, we have for $x \geqslant 2$

$$\sum_{n\in T(\varepsilon,2),\, n\leqslant x} \frac{1}{n} \ll \frac{1}{\varepsilon}(\log x)^{7/24}$$

uniformly for any $\varepsilon > 0$. Thus if c_4 is a sufficiently large constant and $\varepsilon = c_4(\log x)^{-17/24}$, then for $x \geqslant 2$

$$\sum_{n\in T(\varepsilon,2),\, n\leqslant x} 1/n < \tfrac{1}{3}\log x.$$

Thus for $x \geqslant x_0$, there are at least $x^{1/4}$ values of odd $n \leqslant x$ such that $n \notin T(\varepsilon, 2)$, that is,

$$nE(n) \leqslant c_4(\log x)^{-17/24} \leqslant c_4(\log n)^{-17/24}.$$

COROLLARY 3. *Let $\delta(n)\downarrow 0$ arbitrarily slowly. Then the set of n with*

$$nE(n) < (\log n)^{-\delta(n)}$$

has logarithmic density 1.

Proof. This result follows immediately from (8) and (9).

Our last corollary of Theorem 1 requires a little more work. It is a quantified version of Corollary 1.

COROLLARY 4. *There is an infinite sequence of integers n for which*

$$nE(n) \geqslant \exp\left\{(1+o(1))\sqrt{\log\log n}\right\}.$$

Proof. Let $t \geqslant 2$ be arbitrary. We apply (7), (8), (9) for each prime $q \leqslant t$, so that for any ε, $0 < \varepsilon < 1$,

$$\text{(11)} \qquad \sum_{n\in T(\varepsilon),\, n\leqslant x} \frac{1}{n} < \frac{1}{\varepsilon}\sum_{q\leqslant t} \prod_{\substack{p\leqslant x^3\\ p\notin P(q)}} \frac{p}{p-1} + \Big(\prod_{q\leqslant t} q\Big)^{-1}\log x.$$

From Lagarias and Odlyzko [5], we have for each $\delta > 0$ an $x_0(\delta)$ such that for all $x \geqslant x_0(\delta)$ and primes $q < (\log x)^{1/6-\delta}$,

$$\sum_{p \leqslant x,\, p \notin P(q)} 1 = \frac{q-1}{q} \frac{x}{\log x} + O_\delta\left(\frac{x}{\log^2 x}\right).$$

Thus uniformly for $q < (\log x)^{1/7}$,

$$\sum_{p \leqslant x,\, p \notin P(q)} \frac{1}{p} = \frac{1}{x} \sum_{p \leqslant x,\, p \notin P(q)} 1 + \int_2^x \frac{1}{y^2} \sum_{p \leqslant y,\, p \notin P(q)} 1\, dy$$

$$= \int_2^x \frac{1}{y^2} \frac{q-1}{q} \frac{y}{\log y} dy + O\left(\int_2^x \frac{1}{y^2} \frac{y}{\log^2 y} dy\right) + O\left(\int_2^{e^{q^7}} \frac{1}{y^2} \frac{y}{\log y} dy\right)$$

$$= \frac{q-1}{q} \log\log x + O(\log q).$$

It thus follows that uniformly for each $q < (\log x)^{1/7}$,

$$\prod_{p \leqslant x,\, p \notin P(q)} \frac{p}{p-1} = (\log x)^{(q-1)/q} q^{O(1)}. \tag{12}$$

Let $t = \sqrt{\log\log x}$. Then

$$\prod_{q \leqslant t} q = \exp\{(1+o(1))\sqrt{\log\log x}\}$$

and from (12),

$$\sum_{q \leqslant t} \prod_{\substack{p \leqslant x^3 \\ p \notin P(q)}} \frac{p}{p-1} = \log x \exp\{-(1+o(1))\sqrt{\log\log x}\}.$$

Thus from (11),

$$\sum_{n \leqslant x,\, n \in T(\varepsilon)} \frac{1}{n} \leqslant \frac{1}{\varepsilon} \log x \exp\{-(1+o(1))\sqrt{\log\log x}\}.$$

We apply this inequality with $\varepsilon = 1/4$. If $M(x)$ is the maximum value of $nE(n)$ for $n \leqslant x$, then from (10) we have

$$\tfrac{1}{2}\log x \leqslant \left(\tfrac{1}{4} + M(x)\exp\{-(1+o(1))\sqrt{\log\log x}\}\right)\log x,$$

so that

$$M(x) \geqslant \exp\{(1+o(1))\sqrt{\log\log x}\}.$$

Thus our result is established.

THEOREM 2. *For all large* n,

$$\sum_{d|2^n-1,\, d>n} 1/d < \exp(-\log n \log\log\log n/2\log\log n).$$

Proof. Recall that we have denoted the sum in the theorem $A(n)$. For each $m|n$, let

$$A_m(n) = \sum_{l(d)=m,\, d>n} 1/d$$

so that

$$A(n) = \sum_{m|n} A_m(n).$$

Theorem 1 in [8] asserts that there is an x_0 such that for all $x \geqslant x_0$ and any m,

$$\#\{d \leqslant x\colon\ l(d) = m\} \leqslant x\exp\left(-\log x\frac{3+\log\log\log x}{2\log\log x}\right).$$

Thus by partial summation,

$$A_m(n) = \int_n^\infty \frac{1}{x^2} \sum_{n<d\leqslant x,\, l(d)=m} 1\,dx$$

$$\leqslant \int_n^\infty \frac{1}{x}\exp\left(-\log x\frac{3+\log\log\log x}{2\log\log x}\right)dx$$

$$\leqslant (\log n)^2\exp\left(-\log n\frac{3+\log\log\log n}{2\log\log n}\right)\int_n^\infty \frac{1}{x\log^2 x}dx$$

$$= \log n\exp\left(-\log n\frac{3+\log\log\log n}{2\log\log n}\right)$$

for all $n \geqslant x_1 \geqslant x_0$. Therefore

$$A(n) = \sum_{m|n} A_m(n) \leqslant d(n)\log n\exp\left(-\log n\frac{3+\log\log\log n}{2\log\log n}\right)$$

for $n \geqslant x_1$, where $d(n)$ is the number of divisors of n. Using the well-known fact that

$$d(n) \leqslant 2^{(1+o(1))\log n/\log\log n},$$

we have

$$A(n) \leqslant \exp(-\log n\log\log\log n/2\log\log n)$$

for all large n, which is what we wanted to show.

In [8] and in several papers mentioned there, a heuristic argument is presented that implies for each $x \geqslant 3$ there is an $n = n(x) \leqslant x$ such that

$$\# \{d \leqslant x\colon\ l(d) = n\} > x \exp\{-(1+o(1)) \log x \log\log\log x / \log\log x\}.$$

It therefore would follow that

$$E(n) \geqslant \exp\{-(1+o(1)) \log x \log\log\log x / \log\log x\}.$$

It does not seem unreasonable to ask also that for an unbounded set of x we have $n(x) \geqslant x^{1-o(1)}$. It thus would follow that for infinitely many n,

$$E(n) \geqslant \exp\{-(1+o(1)) \log n \log\log\log n / \log\log n\}.$$

I conjecture that this is the true maximal order of $E(n)$ and also that Theorem 2 can be strengthened to

$$A(n) \leqslant \exp\{-(1+o(1)) \log n \log\log\log n / \log\log n\}$$

for all n.

THEOREM 3. *There is a positive constant c_5 such that*

$$\sum_{n \leqslant x} E(n) \leqslant (e^{\gamma} - c_5 + o(1)) \log x.$$

Proof. Let $\varepsilon > 0$ be arbitrarily small, but fixed. Let

$$A(\varepsilon) = \{p\colon\ l(p) \leqslant p^{1/(2+2\varepsilon)}\}.$$

Thus the number of $p \in A(\varepsilon)$ with $p \leqslant x$ is $O(x^{1/(1+\varepsilon)})$. Write

$$\sum_{n \leqslant x} E(n) = \sum_{l(d) \leqslant x} 1/d = S_1 + S_2, \tag{13}$$

where in S_1 each d is divisible by some prime $p \in A(\varepsilon)$ with $p > x$ and in S_2 the remaining d's are considered. If d is counted by S_1, then d is of the form pd_1 where $p \in A(\varepsilon)$, $p > x$, and $l(d_1) \leqslant x$. Thus by (1)

$$S_1 \leqslant \Big(\sum_{p \in A(\varepsilon),\, p > x} 1/p\Big)\Big(\sum_{l(d) \leqslant x} 1/d\Big) \ll x^{-\varepsilon} \log x = o(1). \tag{14}$$

We now consider S_2. Note that for every d counted by S_2, if $p \mid d$, then $l(p) \leqslant x$ and $l(p) > p^{1/(2+2\varepsilon)}$, so that $p < x^{2(1+\varepsilon)}$.

For each prime q, let

$$\omega_q(n) = \sum_{p \mid n,\, l(p) \equiv 0 \bmod q} 1.$$

We have

$$S_2 = S_{2,0} + S_{2,1} + S_{2,2} \tag{15}$$

where

$$\text{in } S_{2,0}, \quad \text{each } d \text{ has} \quad \sum_{q>\sqrt{x}} \omega_q(d) = 0,$$

$$\text{in } S_{2,1}, \quad \text{each } d \text{ has} \quad \sum_{q>\sqrt{x}} \omega_q(d) = 1,$$

$$\text{in } S_{2,2}, \quad \text{each } d \text{ has} \quad \sum_{q>\sqrt{x}} \omega_q(d) \geqslant 2.$$

Let $P(m)$ denote the largest prime factor of m and let

$$c_6(x) = \prod_{p \leqslant x^{2(1+\varepsilon)},\, P(l(p)) > \sqrt{x}} p/(p-1).$$

Then

$$c_6 = \liminf_{x\to\infty} c_6(x) > 1. \tag{16}$$

To see (16), note that it is a consequence of the Bombieri–Vinogradov theorem and the Brun–Titchmarsh inequality (see Goldfeld [3]) that there are positive constants c_7, c_8 with

$$\sum_{p \leqslant z,\, P(p-1) > z^{1/2+c_7}} 1 \geqslant c_8 \frac{z}{\log z}$$

for all large z. (In fact, from recent work of Fouvry, we can take $c_7 > 1/6$.) But

$$\sum_{p \leqslant z,\, l(p) < z^{1/2-c_7}} 1 = O(z^{1-2c_7}),$$

so that

$$\sum_{p \leqslant z,\, P(l(p)) > z^{1/2+c_7}} 1 \geqslant \sum_{p \leqslant z,\, P(p-1) > z^{1/2+c_7}} 1 - \sum_{p \leqslant z,\, l(p) < z^{1/2-c_7}} 1 \geqslant (c_8 + o(1)) \frac{z}{\log z}.$$

We apply this estimate for $x^{1-c_9} < z \leqslant x$ where $c_9 = 2c_7/(1+2c_7)$. For z in this interval, $z^{1/2+c_7} > \sqrt{x}$, so that

$$\sum_{p \leqslant x,\, P(l(p)) > \sqrt{x}} \frac{1}{p} \geqslant \int_{x^{1-c_9}}^{x} \frac{1}{z^2} \sum_{\substack{x^{1-c_9} < p \leqslant z \\ P(l(p)) > \sqrt{x}}} \frac{1}{p}\, dz \geqslant -c_8 \log(1-c_9) + o(1),$$

which implies (16) with $c_6 \geqslant (1-c_9)^{-c_8}$.

We have

$$S_{2,0} \leqslant \sum_{\substack{p|d \Rightarrow p \leqslant x^{2(1+\varepsilon)} \\ P(l(p)) < \sqrt{x}}} \frac{1}{d} = \prod_{\substack{2 < p \leqslant x^{2(1+\varepsilon)} \\ P(l(p)) < \sqrt{x}}} \frac{p}{p-1} = e^{\gamma} c_6(x)^{-1} (1+\varepsilon+o(1)) \log x, \tag{17}$$

by Mertens' theorem.

Let

$$c_{10}(x) = \sum_{\substack{p \leqslant x^{2(1+\varepsilon)} \\ P(l(p)) > \sqrt{x}}} 1/p.$$

Thus $e^{c_{10}(x)} \sim c_6(x)$ as $x \to \infty$. If d is counted by $S_{2,1}$, then d is of the form pd_1 where $p \leqslant x^{2(1+\varepsilon)}$, $P(l(p)) > \sqrt{x}$, and d_1 is counted by $S_{2,0}$. Thus by (17),

$$S_{2,1} \leqslant c_{10}(x) S_{2,0} \leqslant \frac{c_{10}(x) e^{\gamma}}{c_6(x)} (1+\varepsilon+o(1)) \log x \tag{18}$$

$$= \frac{e^{\gamma} \log c_6(x)}{c_6(x)} (1+\varepsilon+o(1)) \log x.$$

Suppose d is counted by $S_{2,2}$. Then d has two prime factors p_1, p_2 with $P(l(p_i)) > \sqrt{x}$. Since $l(d) \leqslant x$, we have $P(l(p_1)) = P(l(p_2))$. Therefore

$$S_{2,2} \leqslant \Bigg(\sum_{q > \sqrt{x}} \sum_{\substack{p_1, p_2 \leqslant x^{2(1+\varepsilon)} \\ P(l(p_1)) = P(l(p_2)) = q}} \frac{1}{p_1 p_2} \Bigg) \Bigg(\sum_{l(d) \leqslant x} \frac{1}{d} \Bigg)$$

$$\ll \Bigg(\sum_{q > \sqrt{x}} \Bigg(\sum_{\substack{p \leqslant x^{2(1+\varepsilon)} \\ P(l(p)) = q}} \frac{1}{p} \Bigg)^2 \Bigg) \log x \leqslant \Bigg(\sum_{q > \sqrt{x}} \Bigg(\sum_{\substack{1 < m \leqslant x^{2(1+\varepsilon)} \\ m \equiv 1 \bmod q}} \frac{1}{m} \Bigg)^2 \Bigg) \log x$$

$$\ll \Bigg(\sum_{q > \sqrt{x}} \left(\frac{\log x}{q} \right)^2 \Bigg) \log x \ll x^{-1/2} \log^3 x = o(1).$$

Putting this estimate in with (15), (16), (17), and (18), we have

$$S_2 \leqslant (1+\log c_6(x)) \frac{e^{\gamma}}{c_6(x)} (1+\varepsilon+o(1)) \log x \tag{19}$$

$$\leqslant (1+\log c_6) \frac{e^{\gamma}}{c_6} (1+\varepsilon+o(1)) \log x$$

$$= (e^{\gamma} - c_5)(1+\varepsilon+o(1)) \log x$$

where $c_5 = e^{\gamma}(c_6 - 1 - \log c_6)/c_6 > 0$. Since $\varepsilon > 0$ is arbitrary, our theorem follows from (13), (14), and (19).

THEOREM 4.

$$\# \{ p \leqslant x \colon\ l(p) \ \text{is prime} \} = O\left(\frac{x \log\log\log x}{\log x \log\log x} \right).$$

Proof. From Brun's method, the number of primes $p \leqslant x$ such that $p-1$ is divisible by no prime from the interval $I := [\log\log x, (\log x)^{1/7}]$ is

$$O\left(\frac{x \log\log\log x}{\log x \log\log x}\right).$$

Thus we need only consider primes $p \leqslant x$ such that $p-1$ is divisible by at least one $q \in I$. For each $q \in I$, let N_q denote the number of $p \leqslant x$ such that $l(p)$ is prime and $p \equiv 1 \bmod q$. Thus if p is counted by N_q, then either $l(p) = q$ or 2 is a qth power $\bmod p$. As in the proof of Corollary 4 of Theorem 1, we may use the results of Lagarias and Odlyzko [5] to show that

$$N_q \leqslant \frac{1}{q(q-1)} \frac{x}{\log x} + q + O\left(\frac{x}{q^2 \log^2 x}\right) \tag{20}$$

uniformly for each q in I. Thus

$$\sum_{q \in I} N_q = O\left(\frac{x}{\log x \log\log x}\right),$$

which implies our theorem.

Remark. Assuming the Extended Riemann Hypothesis (ERH) it is possible to show that

$$\#\{p \leqslant x\colon\ l(p) \text{ is prime}\} = O\left(\frac{x \log\log x}{\log^2 x}\right).$$

This result is accomplished by changing I in the above proof to $[\log x, x^{1/5}]$ and using the fact that the ERH implies (20) holds uniformly for $q \leqslant x^{1/5}$ (see Hooley [4], Ch. 3). Presumably, the "true" order of magnitude of $\#\{p \leqslant x\colon\ l(p) \text{ is prime}\}$ is $x/\log^2 x$.

COROLLARY. *There is a set* $\boldsymbol{P}$ *of primes of relative logarithmic density* 1 *such that*

$$\lim_{p \in \boldsymbol{P},\, p \to \infty} pE(p) = 0.$$

Proof. This result will follow if we can show

$$\sum_{p \leqslant x} E(p) = o(\log\log x). \tag{21}$$

For if (21) holds and $\varepsilon > 0$ is arbitrary, we have, using the notation of Theorem 1,

$$\sum_{p \in T(\varepsilon),\, p \leqslant x} \frac{1}{p} < \frac{1}{\varepsilon} \sum_{p \in T(\varepsilon),\, p \leqslant x} E(p) = o(\log\log x).$$

Thus all we need do is show (21).

We first note that

$$E(p) = \sum_{l(d)=p} \frac{1}{d} \leqslant \sum_{q|d \Rightarrow l(q)=p} \frac{1}{d} = \prod_{l(q)=p} \frac{q}{q-1} - 1 \tag{22}$$

$$= \exp\left\{ \sum_{l(q)=p} \frac{1}{q-1} + O\left(\sum_{l(q)=p} \left(\frac{1}{q-1} \right)^2 \right) \right\} - 1.$$

Next note that since $2^p - 1$ has fewer than p prime factors and they are all $1 \bmod p$, we have

$$\sum_{l(q)=p} \frac{1}{q-1} < \sum_{i<p} \frac{1}{ip} = O\left(\frac{\log p}{p} \right).$$

Thus from (22)

$$E(p) = \sum_{l(q)=p} \frac{1}{q} + O\left(\frac{\log^2 p}{p^2} \right).$$

It follows that

$$\sum_{p \leqslant x} E(p) = \sum_{p \leqslant x} \left(\sum_{l(q)=p} \frac{1}{q} + O\left(\frac{\log^2 p}{p^2} \right) \right) = \sum_{\substack{l(q) \leqslant x \\ l(q)\text{prime}}} \frac{1}{q} + O(1)$$

$$\leqslant \sum_{\substack{q \leqslant x^3 \\ l(q)\text{prime}}} \frac{1}{q} + O(1) = O\left((\log\log\log x)^2 \right)$$

where we use Theorem 4 for the last estimate. This shows (21) and thus the theorem.

References

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

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

[3] M. Goldfeld, *On the number of primes p for which p+a has a large prime factor*, Mathematika 16 (1969), pp. 23–27.

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

[5] J. C. Lagarias and A. M. Odlyzko, *Effective versions of the Chebotarev density theorem*, in *Algebraic Number Fields*, A. Frohlich, ed., Academic Press, London–New York 1977, pp. 409–464.

[6] E. Landau, *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale*, Chelsea Reprint, 1949.

[7] R. W. K. Odoni, *A conjecture of Krishnamurthy on decimal periods and some allied problems*, J. Number Theory 13 (1981), pp. 303–319.

[8] C. Pomerance, *On the distribution of pseudoprimes*, Math. Comp. 37 (1981), pp. 587–593.
[9] N. P. Romanoff, *Über einige Sätze der additiven Zahlentheorie*, Math. Ann. 109 (1934), pp. 668–678.

BELL COMMUNICATIONS RESEARCH, INC.
MORRISTOWN, NEW JERSEY 07960
U.S.A.

Current Address:
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602
U.S.A.

Received on 5.11.1984
and in revised form on 21.1.1985 (1471)