

THE RECIPROCAL SUM OF DIVISORS OF MERSENNE NUMBERS

ZEBEDIAH ENGBERG AND PAUL POLLACK

ABSTRACT. We investigate various questions concerning the reciprocal sum of divisors, or prime divisors, of the Mersenne numbers $2^n - 1$. Conditional on the Elliott–Halberstam Conjecture and the Generalized Riemann Hypothesis, we determine $\max_{n \leq x} \sum_{p|2^n-1} 1/p$ to within $o(1)$ and $\max_{n \leq x} \sum_{d|2^n-1} 1/d$ to within a factor of $1 + o(1)$, as $x \rightarrow \infty$. This refines, conditionally, earlier estimates of Erdős and Erdős–Kiss–Pomerance. Conditionally (only) on GRH, we also determine $\sum 1/d$ to within a factor of $1 + o(1)$ where d runs over all numbers dividing $2^n - 1$ for some $n \leq x$. This conditionally confirms a conjecture of Pomerance and answers a question of Murty–Rosen–Silverman. Finally, we show that both $\sum_{p|2^n-1} 1/p$ and $\sum_{d|2^n-1} 1/d$ admit continuous distribution functions in the sense of probabilistic number theory.

1. INTRODUCTION

Let $f(n) = \sum_{p|2^n-1} 1/p$ and $F(n) = \sum_{d|2^n-1} 1/d$, so that $f(n)$ is the reciprocal sum of the primes dividing the n th Mersenne number, and $F(n)$ is the corresponding sum over all divisors. In this note we are concerned with the statistical distribution of $f(n)$ and $F(n)$ as n varies, continuing earlier investigations of Erdős [6, 7], Pomerance [24], and Erdős–Kiss–Pomerance [8].

It will be convenient in what follows to write $\ell(d)$ for the multiplicative order of 2 modulo the odd number d .

Our starting point is the following observation. Let $z \geq 2$, and let n be the least common multiple of the natural numbers not exceeding z . Then

$$f(n) = \sum_{p|2^n-1} \frac{1}{p} = \sum_{\ell(p)|n} \frac{1}{p} \geq \sum_{2 < p \leq z} \frac{1}{p},$$

the last inequality holding on account of having $\ell(p) < p \leq z$ for each odd prime $p \leq z$. As is well-known [12, Theorem 427, p. 466], there is a constant $C_1 = 0.26149\dots$ (the Meissel–Mertens constant) with the property that $\sum_{p \leq z} \frac{1}{p} = \log \log z + C_1 + o(1)$ as $z \rightarrow \infty$. Hence,

$$f(n) \geq \log \log z + C_1 - \frac{1}{2} + o(1).$$

Now taking $z = \frac{1}{2} \log x$, we deduce from the prime number theorem that

$$(1) \quad \max_{n \leq x} f(n) \geq \log \log \log x + C_1 - \frac{1}{2} + o(1).$$

Perhaps surprisingly, given the ease with which (1) was established, this lower bound on $\max_{n \leq x} f(n)$ is close to best possible. Indeed, in 1971 Erdős proved the complementary upper bound

$$(2) \quad \max_{n \leq x} f(n) \leq \log \log \log x + O(1).$$

In the same paper [7] he asks whether equality holds in (1).

Erdős’s question seems to have lain dormant for 20 years. In 1991, the study of $f(n)$ was resurrected by Erdős, Kiss, and Pomerance [8]. The main thought there is to study the local

2010 *Mathematics Subject Classification*. Primary 11N37; Secondary 11A25, 11B37.

distribution of $f(n)$. However, on p. 271 of their paper they briefly digress to give a negative answer to Erdős’s question, by showing that

$$\max_{n \leq x} f(n) \geq \log \log \log x + C$$

for some $C > C_1 - \frac{1}{2}$ and all large x . In the PhD thesis of the first author [5], it is shown using lower bounds for counts of smooth shifted primes (cf. the discussion in §2.1 below) that C can be taken as $C_1 - \frac{1}{2} + 0.2069$.

Our first theorem determines the “correct” value of C , conditional on the Generalized Riemann Hypothesis (GRH) and (a weak form of) the Elliott–Halberstam Conjecture (EHC).¹

Theorem 1.1 (conditional on GRH and weak EHC). *As $x \rightarrow \infty$,*

$$\max_{n \leq x} f(n) = \log \log \log x + C_1 - \frac{1}{2} + C_2 + o(1).$$

Here C_1 is the Meissel–Mertens constant and $C_2 = \int_1^\infty \rho(u)u^{-1} du$, where $\rho(u)$ is Dickman’s function (see §2.1).

One can compute that $C_2 = 0.521908\dots$ and that $C_1 - \frac{1}{2} + C_2 = 0.283405\dots$

Since $\sum_{d|m} \frac{1}{d} \asymp \exp(\sum_{p|m} \frac{1}{p})$ for all positive integers m , the estimate (2) is equivalent to the bound

$$\max_{n \leq x} F(n) \ll \log \log x,$$

and this is how Erdős advertises the result in [7]. Our proof of Theorem 1.1 is easily adapted to determine, under the same hypotheses as Theorem 1.1, $\max_{n \leq x} F(n)$ to within a factor of $1 + o(1)$.

Theorem 1.2 (conditional on GRH and weak EHC). *As $x \rightarrow \infty$,*

$$\max_{n \leq x} F(n) = \exp(\gamma - \log 2 + C_2 + o(1)) \log \log x.$$

Here γ is the familiar Euler–Mascheroni constant.

Observe that $F(n) = \sum_{m|n} E(m)$, where $E(m) = \sum_{\ell(d)=m} 1/d$. Erdős showed in [6] (see that paper’s eq. (8)) that $\sum_{m \leq x} E(m) \ll \log x$, as a part of a simple (re)proof of Romanov’s theorem asserting that $\sum_{d \text{ odd}} 1/d\ell(d) < \infty$. Several questions on the average and extremal orders of $E(m)$ are left open in [6]; many of these were later taken up by Pomerance in [24]. Our concern here is with the average order. Clearly,

$$\sum_{m \leq x} E(m) = \sum_{\ell(d) \leq x} \frac{1}{d} \geq \sum_{\substack{d \text{ odd} \\ d \leq x}} \frac{1}{d} > \frac{1}{2} \log x,$$

and it was conjectured in [24] that this is asymptotically sharp. Our next theorem shows that this conjecture follows from GRH.

Theorem 1.3 (conditional on GRH). *As $x \rightarrow \infty$,*

$$\sum_{m \leq x} E(m) = \left(\frac{1}{2} + o(1) \right) \log x.$$

¹By GRH, we always mean the Riemann Hypothesis for Dedekind zeta functions. The relevant version of the Elliott–Halberstam Conjecture is stated precisely in §2.1.

As we remark after its proof, Theorem 1.3 allows us to (conditionally) answer a question of Murty, Rosen, and Silverman concerning the behavior of $\sum_{d \text{ odd}} 1/d\ell(d)^\epsilon$ as $\epsilon \downarrow 0$.

Our final result belongs to the domain of probabilistic number theory. Recall that a function $D: \mathbb{R} \rightarrow \mathbb{R}$ is a distribution function if D is nondecreasing, right-continuous, and obeys the boundary conditions

$$\lim_{c \rightarrow -\infty} D(c) = 0 \quad \text{and} \quad \lim_{c \rightarrow \infty} D(c) = 1.$$

If $g(n)$ is a real-valued arithmetic function, we say that $g(n)$ has a distribution function (or possesses a limit law) if there is a distribution function D with the following property: For every real number c at which D is continuous, the set of natural numbers n with $g(n) \leq c$ has asymptotic density $D(c)$. It is easy to show that if $g(n)$ has a distribution function D , then D is uniquely determined. If this D is continuous everywhere, we say that g has a continuous distribution function.

On p. 270 of [8], Erdős, Kiss, and Pomerance claim that it can be proved by “more or less standard methods” that $f(n)$ possesses a continuous distribution function. That $f(n)$ and $F(n)$ have distribution functions follows immediately from (later) general results of Luca and Shparlinski [18], who used the method of moments. (See the proof of Theorem 3(2) in [17] for an alternative approach to these results.) However, we are not aware of any proof in the literature — before [8] or since — that the distribution function of f is continuous.² As the details of the proof seem to us both interesting and nontrivial, we have made our second goal with this note to fill this (much-unneeded) gap in the literature.

Theorem 1.4. *The distribution function of $f(n)$ is continuous. The same holds for $F(n)$.*

The reader interested in these problems may also wish to refer to [21], which describes another application of the GRH to the study of prime divisors of Mersenne numbers.

2. THE SHARP MAXIMAL ORDER OF $f(n)$

2.1. Preliminaries on smooth shifted primes. Our proofs rely heavily on the notion of “smoothness”. Let $P(n)$ denote the largest prime factor of n , with the convention that $P(1) = 1$. The natural number n is called Y -smooth (or Y -friable) if $P(n) \leq Y$. For $X, Y \geq 1$, we let

$$\Psi(X, Y) = \#\{n \leq X : P(n) \leq Y\}.$$

After work of de Bruijn, Hildebrand, Tenenbaum, and others, we have precise estimates for $\Psi(X, Y)$ in a wide range of X and Y ; a useful survey is [10]. A fundamental role in this area is played by the Dickman function $\rho(u)$, defined as the solution to the delay differential equation

$$u\rho'(u) + \rho(u-1) = 0$$

with the initial condition $\rho(u) = 1$ for $0 \leq u \leq 1$. Hildebrand, extending earlier work of de Bruijn, showed that with $u := \frac{\log X}{\log Y}$,

$$\Psi(X, Y) \sim X\rho(u), \quad \text{as } X \rightarrow \infty \text{ with } X \geq Y > \exp((\log \log X)^{1.7}).$$

(In fact, 1.7 may be replaced with any constant exceeding $5/3$.) Given the close connection between $\Psi(x, y)$ and $\rho(u)$ it is important to have accurate estimates of $\rho(u)$. While much finer results are known, the crude estimate

$$\rho(u) = 1/u^{u+o(u)}, \quad \text{as } u \rightarrow \infty,$$

will suffice for all of our purposes.

²Related distribution functions are claimed to be continuous in [17, Theorem 3(2)]. However, the offered proofs only show existence of the distribution functions, leaving continuity unaddressed.

In the arguments below, it is important to understand the frequency with which shifted primes $p - 1$ are smooth. For $X, Y \geq 1$, set

$$\Pi(X, Y) = \#\{p \leq X : P(p - 1) \leq Y\}.$$

It is natural to expect that shifted primes are smooth with the same relative frequency as generic numbers of the same size. This heuristic, coupled with the prime number theorem, suggests that

$$(3) \quad \Pi(X, Y) \sim \frac{\Psi(X, Y)}{\log X}$$

in a wide range of X and Y . An explicit conjecture of this kind appears in [23], where Pomerance proposes that (3) holds whenever $X, Y \rightarrow \infty$ with $X \geq Y$.

We require two results which represent partial progress towards Pomerance's conjecture. The first is conditional on the following weak form of the Elliott–Halberstam Conjecture.

Conjecture 2.1 (weak Elliott–Halberstam Conjecture). *For any fixed $\epsilon > 0$, we have*

$$\sum_{n \leq x^{1-\epsilon}} \max_{\gcd(a, n)=1} \left| \pi(x; n, a) - \frac{\pi(x)}{\phi(n)} \right| = o(\pi(x)), \quad \text{as } x \rightarrow \infty.$$

Proposition 2.2 (conditional on weak EHC). *Fix a real number $U > 1$. As $X \rightarrow \infty$,*

$$\Pi(X, Y) \sim \rho(u) \frac{X}{\log X},$$

uniformly in the range $1 \leq u \leq U$.

Proposition 2.2, which is due to Granville (unpublished), conditionally confirms (3) when $X \rightarrow \infty$ in the range $X \geq Y \geq X^{1/U}$. A generalization of Proposition 2.2 appears (with a full proof) as [16, Theorem 1]. See also Lemma 4.1 in [30]. We note that the proofs of Proposition 2.2 use only the even weaker form of EHC where $\pi(x; n, a)$ is replaced with $\pi(x; n, 1)$ and the max on a is removed, and that EHC enters the proofs of Theorems 1.1 and 1.2 only via Proposition 2.2.

The next result, due to Pomerance and Shparlinski (see [26, Theorem 1]), is unconditional and applies in a much wider range of X and Y . But it is only an upper bound, not an asymptotic formula, and the bound itself is slightly weaker than the prediction (3).

Proposition 2.3. *If X is sufficiently large, and $X \geq Y \geq \exp(\sqrt{\log X \log \log X})$, then*

$$\Pi(X, Y) \ll u \rho(u) \frac{X}{\log X}.$$

2.2. Proof of the lower bound in Theorem 1.1.

Lemma 2.4. *As $z \rightarrow \infty$,*

$$\sum_{\substack{p > z \\ P(p-1) \leq z}} \frac{1}{p} \rightarrow C_2,$$

where $C_2 = \int_1^\infty \rho(u) u^{-1} du$ (as in Theorem 1.1).

Proof. Fix a real number $U > 1$. Using \sum for the sum appearing in the lemma statement, we rewrite $\sum = \sum_1 + \sum_2 + \sum_3$, partitioning the range of summation as follows;

$$\sum_1 : z < p \leq z^U, \quad \sum_2 : z^U < p \leq \exp((\log z)^{3/2}), \quad \sum_3 : p > \exp((\log z)^{3/2}).$$

We estimate \sum_1 using Proposition 2.2. That result implies that, as $z \rightarrow \infty$,

$$\begin{aligned}\sum_1 &= \int_z^{z^U} \frac{d\Pi(t, z)}{t} = \int_z^{z^U} \frac{\Pi(t, z)}{t^2} dt + o(1) \\ &= (1 + o(1)) \int_z^{z^U} \frac{1}{t \log t} \rho\left(\frac{\log t}{\log z}\right) dt + o(1) = \int_1^U \rho(u) u^{-1} du + o(1),\end{aligned}$$

where in the last step we made the change of variables $t = z^u$. Hence,

$$\lim_{z \rightarrow \infty} \sum_1 = \int_1^U \rho(u) u^{-1} du.$$

A similar calculation, using Proposition 2.3 in place of Proposition 2.2, reveals that

$$\limsup_{z \rightarrow \infty} \sum_2 \ll \int_U^{(\log z)^{1/2}} \rho(u) du \leq \int_U^\infty \rho(u) du,$$

where the constant implied by “ \ll ” is absolute (independent of U).

We handle \sum_3 using the trivial bound $\Pi(X, Y) \leq \Psi(X, Y)$, along with the uniform estimate $\Psi(X, Y) \ll X \exp(-u/2)$, which holds for all $X \geq Y \geq 2$ (see [29, Theorem 5.1, p. 512]). We find that for large z ,

$$\begin{aligned}\sum_3 &\leq \int_{\exp((\log z)^{3/2})}^\infty \frac{\Pi(t, z)}{t^2} dt \leq \int_{\exp((\log z)^{3/2})}^\infty \frac{\Psi(t, z)}{t^2} dt \ll \int_{\exp((\log z)^{3/2})}^\infty \frac{1}{t} \exp\left(-\frac{1}{2} \frac{\log t}{\log z}\right) dt \\ &= \log z \int_{(\log z)^{1/2}}^\infty \exp\left(-\frac{1}{2} u\right) du \ll \log z \cdot \exp\left(-\frac{1}{2} (\log z)^{1/2}\right).\end{aligned}$$

Thus,

$$\lim_{z \rightarrow \infty} \sum_3 = 0.$$

Since $\sum = \sum_1 + \sum_2 + \sum_3$, we deduce that

$$(4) \quad \limsup_{z \rightarrow \infty} \sum_{\substack{p > z \\ P(p-1) \leq z}} \frac{1}{p} \leq \int_1^U \rho(u) u^{-1} du + O\left(\int_U^\infty \rho(u) du\right),$$

and, since $\sum \geq \sum_1$, that

$$(5) \quad \liminf_{z \rightarrow \infty} \sum_{\substack{p > z \\ P(p-1) \leq z}} \frac{1}{p} \geq \int_1^U \rho(u) u^{-1} du.$$

The lemma follows from (4) and (5) upon letting U tend to infinity. \square

Proof of the lower bound in Theorem 1.1. Let x be large, let $z = \frac{1}{2} \log x$, and let n be the least common multiple of the natural numbers not exceeding z . By the prime number theorem, $\log n = (\frac{1}{2} + o(1)) \log x$ as $x \rightarrow \infty$, and so in particular $n \leq x$ for large x . Thus, for the lower bound in Theorem 1.1, it is enough to prove that as $x \rightarrow \infty$,

$$f(n) \geq \log \log \log x + C_1 - \frac{1}{2} + C_2 + o(1).$$

Observe that

$$\begin{aligned} f(n) &= \sum_{\ell(p)|n} \frac{1}{p} \geq \sum_{\substack{p>2 \\ p-1|n}} \frac{1}{p} = \sum_{2<p\leq z} \frac{1}{p} + \sum_{\substack{p>z \\ p-1|n}} \frac{1}{p} \\ &= \sum_{2<p\leq z} \frac{1}{p} + \sum_{\substack{p>z \\ P(p-1)\leq z}} \frac{1}{p} - E, \end{aligned}$$

where

$$E := \sum_{\substack{p>z \\ P(p-1)\leq z \\ p-1\nmid n}} \frac{1}{p}.$$

We will show shortly that $E = o(1)$ as $x \rightarrow \infty$. Assuming this for now, Mertens' theorem and Lemma 2.4 yield

$$\begin{aligned} f(n) &\geq (\log \log z + C_1 - 1/2) + C_2 + o(1) \\ &= \log \log \log x + C_1 - \frac{1}{2} + C_2 + o(1), \end{aligned}$$

as desired.

We turn now to proving that $E = o(1)$, as $z \rightarrow \infty$. In view of the upper bounds for \sum_2 and \sum_3 established in the proof of Lemma 2.4, it will suffice to show that

$$\sum_{\substack{z<p\leq z^A \\ P(p-1)\leq z \\ p-1\nmid n}} \frac{1}{p} = o(1),$$

for each fixed $A > 1$. So suppose that $P(p-1) \leq z$ and $p-1 \nmid n$. Choose a prime power ℓ^e dividing $p-1$ with $\ell \leq z$ and $\ell^e > z$. Then $e > 1$, and ℓ^e is a squarefull divisor of $p-1$ exceeding z . It follows that

$$\begin{aligned} \sum_{\substack{z<p\leq z^A \\ P(p-1)\leq z \\ p-1\nmid n}} \frac{1}{p} &\leq \sum_{\substack{m \text{ squarefull} \\ m>z}} \sum_{\substack{p\leq z^A \\ m|p-1}} \frac{1}{p} \leq \sum_{\substack{m \text{ squarefull} \\ m>z}} \sum_{\substack{p\leq z^A \\ m|p-1}} \frac{1}{p-1} \\ &\leq \sum_{\substack{m \text{ squarefull} \\ m>z}} \sum_{k\leq z^A} \frac{1}{mk} \leq \log(z^A + 1) \sum_{\substack{m \text{ squarefull} \\ m>z}} \frac{1}{m} \ll \log(z^A + 1) z^{-1/2}. \end{aligned}$$

The final expression tends to 0 as z tends to infinity. \square

2.3. Proof of the upper bound in Theorem 1.1. The Generalized Riemann Hypothesis enters into our proof by way of the following lemma of Kurlberg and Pomerance [14, Theorem 23].

Lemma 2.5 (conditional on GRH). *For all real X, Y with $1 \leq Y \leq \log X$,*

$$\#\left\{p \leq X : \ell(p) \leq \frac{p}{Y}\right\} \ll \frac{X}{Y \log X} + \frac{X \log \log X}{(\log X)^2}.$$

Call the odd prime p *normal* if $\ell(p) > p/(\log \log p)^2$ and *abnormal* otherwise. By Lemma 2.5, the number of abnormal $p \in [X, 2X]$ is $\ll \frac{X}{\log X \cdot (\log \log X)^2}$ for all large X . Breaking $[3, X]$ into dyadic blocks, we deduce that the same bound holds for the number of abnormal $p \leq X$. Hence,

$$\sum_{p \text{ abnormal}} \frac{1}{p} < \infty,$$

which we will make use of momentarily.

Proof of the upper bound in Theorem 1.1. Let x be large, let $n \leq x$, and let L be the least common multiple of the natural numbers not exceeding $(\log \log x)^2$. If p is a normal odd prime and $\ell(p)$ divides n , then

$$p - 1 = \ell(p) \cdot \frac{p-1}{\ell(p)} \mid n \cdot L.$$

We also have trivially that $p - 1 \mid nL$ for all $p \leq (\log \log x)^2$. Thus,

$$\sum_{\substack{\ell(p) \mid n \\ p-1 \mid nL}} \frac{1}{p} \leq \sum_{\substack{p > 2 \\ p-1 \mid nL}} \frac{1}{p} + \sum_{\substack{p \text{ abnormal} \\ p > (\log \log x)^2}} \frac{1}{p} \leq \sum_{\substack{p > 2 \\ p-1 \mid nL}} \frac{1}{p} + o(1).$$

(In the last step, we use that $\sum_{p \text{ abnormal}} \frac{1}{p} < \infty$.) Now let

$$z = \log x \cdot \log \log x.$$

Clearly,

$$(6) \quad \sum_{\substack{p > 2 \\ p-1 \mid nL}} \frac{1}{p} \leq \sum_{\substack{p > 2 \\ P(p-1) \leq z}} \frac{1}{p} + \sum_{\substack{p-1 \mid nL \\ P(p-1) > z}} \frac{1}{p}.$$

The first right-hand sum satisfies

$$\begin{aligned} \sum_{\substack{p > 2 \\ P(p-1) \leq z}} \frac{1}{p} &= \sum_{2 < p \leq z} \frac{1}{p} + \sum_{\substack{p > z \\ P(p-1) \leq z}} \frac{1}{p} \\ &= \log \log z + C_1 - \frac{1}{2} + C_2 + o(1) \\ &= \log \log \log x + C_1 - \frac{1}{2} + C_2 + o(1), \end{aligned}$$

and so to finish the proof it is enough to show that the second sum on the right of (6) is $o(1)$.

Suppose that $P(p-1) > z$ and $p-1 \mid nL$. Then there is a prime $q > z$ dividing both n and $p-1$. Hence,

$$\sum_{\substack{p-1 \mid nL \\ P(p-1) > z}} \frac{1}{p} \leq \sum_{\substack{q \mid n \\ q > z}} \sum_{\substack{d \mid nL \\ q \mid d}} \frac{1}{d} \leq \sum_{\substack{q \mid n \\ q > z}} \frac{1}{q} \sum_{\substack{e \mid nL \\ e \mid n}} \frac{1}{e} = \frac{\sigma(nL)}{nL} \sum_{\substack{q \mid n \\ q > z}} \frac{1}{q} \leq \frac{\sigma(nL)}{nL} \cdot \frac{\omega(n)}{z}.$$

We finish the proof by appealing to the known maximal orders of σ and ω . By Theorem 323 on p. 350 of [12], $\frac{\sigma(nL)}{nL} \ll \log \log(nL)$. By the prime number theorem, $L < \exp(2(\log \log x)^2)$, so that $nL < x \exp(2(\log \log x)^2)$ and $\log \log(nL) \ll \log \log x$. Also (see p. 471 in [12]), $\omega(n) \ll \log x / \log \log x$. Thus,

$$\frac{\sigma(nL)}{nL} \cdot \frac{\omega(n)}{z} \ll \log \log x \cdot \frac{\log x / \log \log x}{\log x \cdot \log \log x} = \frac{1}{\log \log x},$$

which indeed tends to 0 as x tends to infinity. \square

3. THE SHARP MAXIMAL ORDER OF $F(n)$: PROOF OF THEOREM 1.2

The proof of Theorem 1.2 is based on the following variant of Theorem 1.1, where $1/p$ is replaced by $\log \frac{p}{p-1}$. The required changes in the proof are straightforward, bearing in mind that $\log \frac{p}{p-1} = 1/p + O(1/p^2)$.

Proposition 3.1 (conditional on GRH and weak EHC). *As $x \rightarrow \infty$, we have*

$$(7) \quad \sum_{\ell(p)|n} \log \frac{p}{p-1} \leq \sum_{2 < p \leq \frac{1}{2} \log x} \log \frac{p}{p-1} + C_2 + o(1)$$

uniformly for all $n \leq x$. Moreover, equality holds in (7) when n is taken as the least common multiple of the natural numbers not exceeding $\frac{1}{2} \log x$.

Proof of Theorem 1.2. If $\ell(d) \mid n$, then $\ell(p) \mid n$ for all primes p dividing d . Hence, $F(n) = \sum_{\ell(d)|n} 1/d \leq \prod_{\ell(p)|n} (1 + 1/p + 1/p^2 + \dots)$, so that by Proposition 3.1,

$$\begin{aligned} \log F(n) &\leq \sum_{\ell(p)|n} \log \frac{p}{p-1} \\ &\leq \log \left(\prod_{2 < p \leq \frac{1}{2} \log x} (1 - 1/p)^{-1} \right) + C_2 + o(1) \\ &= \log \log x + \gamma - \log 2 + C_2 + o(1), \end{aligned}$$

using Mertens' product theorem in the final step. Exponentiating gives the upper inequality in Theorem 1.2.

We now let n be the least common multiple of the integers up to $\frac{1}{2} \log x$ and prove that the lower inequality holds. Suppose that d has the form $p_1^{e_1} \cdots p_k^{e_k} q_1 \cdots q_l$, where the p_i are distinct primes in $(2, \log \log x]$, each $p_i^{e_i} \leq \frac{1}{2} \log x$, and the q_i are distinct primes exceeding $\log \log x$ where each $\ell(q_i) \mid n$. As $\ell(p_i^{e_i}) \mid (p_i - 1)p_i^{e_i - 1} \mid n$ for each i , we see that

$$\ell(d) = \text{lcm}[\ell(p_1^{e_1}), \dots, \ell(p_k^{e_k}), \ell(q_1), \dots, \ell(q_l)] \mid n.$$

Hence, with p and q running over primes and $e_p := \lfloor \frac{\log(\frac{1}{2} \log x)}{\log p} \rfloor$,

$$\begin{aligned} F(n) &= \sum_{\ell(d)|n} \frac{1}{d} \geq \prod_{2 < p \leq \log \log x} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots + \frac{1}{p^{e_p}} \right) \prod_{\substack{q > \log \log x \\ \ell(q)|n}} \left(1 + \frac{1}{q} \right) \\ &= \prod_{2 < p \leq \log \log x} \frac{p}{p-1} \left(1 - \frac{1}{p^{e_p+1}} \right) \prod_{\substack{q > \log \log x \\ \ell(q)|n}} \frac{q}{q-1} \left(1 - \frac{1}{q^2} \right) \\ &\geq \prod_{\ell(p)|n} \frac{p}{p-1} \prod_{2 < p \leq \log \log x} \left(1 - \frac{1}{p^{e_p+1}} \right) \prod_{q > \log \log x} \left(1 - \frac{1}{q^2} \right). \end{aligned}$$

Noting that $p^{e_p+1} > \frac{1}{2} \log x$,

$$\prod_{2 < p \leq \log \log x} \left(1 - \frac{1}{p^{e_p+1}} \right) \geq 1 - \sum_{2 < p \leq \log \log x} \frac{1}{p^{e_p+1}} \geq 1 - \frac{1}{\frac{1}{2} \log x} \sum_{2 < p \leq \log \log x} 1 = 1 + o(1).$$

The final product on q appearing above is also $1 + o(1)$. Hence,

$$F(n) \geq (1 + o(1)) \prod_{\ell(p)|n} \frac{p}{p-1} = \exp \left(\sum_{\ell(p)|n} \log \frac{p}{p-1} + o(1) \right).$$

Using that equality holds in (7) for our choice of n and applying Mertens' product theorem in the same manner as before, the desired lower bound follows. \square

4. THE RECIPROCAL SUM OF PRIMITIVE DIVISORS: PROOF OF THEOREM 1.3

Let $\lambda(n)$ denote the universal exponent for the multiplicative group modulo n . The next proposition, due to Friedlander, Pomerance, and Shparlinski [9, Theorem 5], shows that $\lambda(n)$ is usually not much smaller than n .

Proposition 4.1. *For all large X , and all $\Delta \geq (\log \log X)^3$, the inequality $\lambda(n) > n \exp(-\Delta)$ holds for all positive integers $n \leq X$ with at most $X \exp(-0.69(\Delta \log \Delta)^{1/3})$ exceptions.*

The following lower bound for $\ell(n)$, in terms of $\lambda(n)$ and the numbers $\ell(p)$ for p dividing n , is due to Kurlberg and Rudnick [15, see §5.1]. See also Lemma 5 of [14].

Proposition 4.2. *For all odd numbers n ,*

$$\ell(n) \geq \frac{\lambda(n)}{n} \prod_{p|n} \ell(p).$$

In [14], it is proved that $\ell(n)$ is usually a bit larger than $n^{1/2}$, for almost all n . Specifically, if $\epsilon(x)$ is any function tending to 0 as $x \rightarrow \infty$, then $\ell(n) > n^{1/2+\epsilon(n)}$ for all n on a set of asymptotic density 1. We need a variant of this result where the exceptional set is suitably small. Such a result is available if we are willing to slightly weaken the required lower bound on $\ell(n)$.

Lemma 4.3. *Fix $\delta \in (0, \frac{1}{2})$. For all large X , the inequality $\ell(n) > n^{\frac{1}{2}-\delta}$ holds for all odd $n \leq X$ with at most $O(X/(\log X)^2)$ exceptions.*

Proof. Clearly, we can restrict attention to

$$n > X/(\log X)^2.$$

Proposition 4.1 (with $\Delta = \frac{1}{4}\delta \log x$) licenses us to assume that

$$\lambda(n) \geq n \cdot X^{-\frac{1}{4}\delta}.$$

Using $\text{rad}(n)$ for the largest squarefree divisor of n , we can also assume that

$$\frac{n}{\text{rad}(n)} \leq (\log n)^4.$$

Indeed, if this inequality fails, let $u := n/\text{rad}(n)$. Then $u \cdot \text{rad}(u)$ is a squarefull divisor of n with $u \cdot \text{rad}(u) > 2(\log n)^4 > (\log X)^4$. But the number of $n \leq X$ with a squarefull divisor exceeding $(\log X)^4$ is $O(X/(\log X)^2)$, which is acceptable for us.

To continue, call the odd prime p *good* if $\ell(p) > p^{\frac{1}{2}-\frac{1}{2}\delta}$ and *bad* otherwise. Note that if p is bad, and $p \leq T$, then p divides $2^m - 1$ for some $m \leq T^{\frac{1}{2}-\frac{1}{2}\delta}$. Since each Mersenne number $2^m - 1$ has fewer than m distinct prime factors, the number of bad primes $p \leq T$ is at most

$$\sum_{m \leq T^{\frac{1}{2}-\frac{1}{2}\delta}} m \leq T^{1-\delta},$$

for every real number $T \geq 1$.

Next, we show that the product of the bad primes dividing n is usually small. Observe that

$$\begin{aligned} \sum_{n \leq X} \left(\sum_{\substack{p|n \\ p \text{ bad}}} \log p \right)^2 &\leq \sum_{p \leq X \text{ bad}} (\log p)^2 \sum_{\substack{n \leq X \\ p|n}} 1 + \sum_{\substack{p, q \leq X \text{ bad} \\ p \neq q}} \log p \log q \sum_{\substack{n \leq X \\ pq|n}} 1 \\ &\leq X \left(\sum_{p \text{ bad}} \frac{(\log p)^2}{p} + \left(\sum_{p \text{ bad}} \frac{\log p}{p} \right)^2 \right) \ll X. \end{aligned}$$

We use in the final step that the sums over bad primes are convergent, which follows by partial summation from the results of the last paragraph. Hence, by Markov's inequality,

$$(8) \quad \sum_{\substack{p|n \\ p \text{ bad}}} \log p < \frac{1}{2} \delta \log X$$

for all but $O(X/(\log X)^2)$ integers $n \leq X$. (The implied constant depends on δ .)

Now suppose that n is an odd number not exceeding X and that n is not one of the $O(X/(\log X)^2)$ exceptional integers above. Write n_{bad} for the largest divisor of n composed entirely of bad primes, and define n_{good} analogously. Then

$$\ell(n) \geq \frac{\lambda(n)}{n} \prod_{\substack{p \text{ good} \\ p|n}} \ell(p) \geq X^{-\frac{1}{4}\delta} \prod_{\substack{p \text{ good} \\ p|n}} p^{\frac{1}{2}-\frac{1}{2}\delta} = X^{-\frac{1}{4}\delta} \cdot \text{rad}(n_{\text{good}})^{\frac{1}{2}-\frac{1}{2}\delta}.$$

Since

$$\frac{n_{\text{good}}}{\text{rad}(n_{\text{good}})}, \frac{n_{\text{bad}}}{\text{rad}(n_{\text{bad}})} \leq \frac{n}{\text{rad}(n)} \leq (\log X)^4,$$

(8) yields

$$n_{\text{bad}} \leq (\log X)^4 \cdot \text{rad}(n_{\text{bad}}) \leq (\log X)^4 X^{\frac{1}{2}\delta},$$

so that

$$\text{rad}(n_{\text{good}}) \geq \frac{n_{\text{good}}}{(\log X)^4} = \frac{n/n_{\text{bad}}}{(\log X)^4} \geq nX^{-\frac{1}{2}\delta} (\log X)^{-8} \geq X^{1-\frac{1}{2}\delta} (\log X)^{-10}.$$

Here the final inequality is justified by our assumption that $n > X/(\log X)^2$. Putting this back above, we find that

$$\begin{aligned} \ell(n) &\geq X^{-\frac{1}{4}\delta} \cdot X^{(1-\frac{1}{2}\delta)(\frac{1}{2}-\frac{1}{2}\delta)} (\log X)^{-5} \\ &\geq X^{\frac{1}{2}-\delta} \geq n^{\frac{1}{2}-\delta}, \end{aligned}$$

thereby completing the proof. \square

The final piece of preparation needed for the proof of Theorem 1.3 is a theorem of Kurlberg (see [13, Theorem 1]).

Proposition 4.4 (conditional on GRH). *Fix $\epsilon \in (0, 1)$. Then $\ell(n) > n^{1-\epsilon}$ for all odd n except those belonging to a set of asymptotic density 0.*

Proof of Theorem 1.3. In view of the representation $\sum_{m \leq x} E(m) = \sum_{\ell(d) \leq x} 1/d$, it is enough to prove that for each $\epsilon > 0$ and all large x ,

$$(9) \quad \sum_{\substack{d > x \\ \ell(d) \leq x}} \frac{1}{d} < \epsilon \log x.$$

Certainly

$$\sum_{\substack{x < d < x^{1+\frac{1}{2}\epsilon} \\ \ell(d) \leq x}} \frac{1}{d} \leq \sum_{x < d < x^{1+\frac{1}{2}\epsilon}} \frac{1}{d} < \frac{2}{3} \epsilon \log x$$

for all large x . To handle the d with $x^{1+\frac{1}{2}\epsilon} < d \leq x^3$, we appeal to Proposition 4.4. If d is in this range, and $\ell(d) \leq x$, then $\ell(d) < d^\delta$ with $\delta = (1 + \frac{1}{2}\epsilon)^{-1}$. By Proposition 4.4, the count of $d \leq X$

with $\ell(d) < d^\delta$ is $o(X)$, as $X \rightarrow \infty$, and now partial summation yields

$$\sum_{\substack{x^{1+\frac{1}{2}\epsilon} < d \leq x^3 \\ \ell(d) \leq x}} \frac{1}{d} \leq \sum_{\substack{x^{1+\frac{1}{2}\epsilon} < d \leq x^3 \\ \ell(d) < d^\delta}} \frac{1}{d} = o(\log x),$$

as $x \rightarrow \infty$. Finally, if $d > x^3$ and $\ell(d) \leq x$, then $\ell(d) < d^{1/3}$. But the d with $\ell(d) < d^{1/3}$ comprise a set with finite reciprocal sum, by Lemma 4.3 and partial summation. Hence,

$$\sum_{\substack{d > x^3 \\ \ell(d) \leq x}} \frac{1}{d} \leq \sum_{\substack{d > x^3 \\ \ell(d) < d^{1/3}}} \frac{1}{d} = o(1),$$

as $x \rightarrow \infty$. Combining what was shown in the last few displays yields (9). \square

Remark. Fix an integer $a \geq 2$, and let $\ell_a(d)$ denote the order of a modulo d , where d is assumed coprime to a . On p. 376 of [22], Murty, Rosen, and Silverman observe that

$$\sum_{\substack{d \geq 1 \\ \gcd(a,d)=1}} \frac{1}{d \cdot \ell_a(d)^\epsilon} \geq \sum_{\substack{d \geq 1 \\ \gcd(a,d)=1}} \frac{1}{d^{1+\epsilon}} = \zeta(1+\epsilon) \cdot \prod_{p|a} \left(1 - \frac{1}{p^{1+\epsilon}}\right),$$

and thus $\liminf_{\epsilon \downarrow 0} \epsilon \sum 1/d\ell_a(d)^\epsilon \geq \phi(a)/a$. They remark that “it would be interesting to compute the exact value [of the limit] . . . if the limit exists”. Under GRH, we can show that this limit is exactly $\phi(a)/a$. To see the connection with our work, let $S_a(x) = \sum_{\ell_a(d) \leq x} 1/d$, and observe that

$$\sum_{\substack{d \geq 1 \\ \gcd(a,d)=1}} \frac{1}{d \cdot \ell_a(d)^\epsilon} = \int_{1^-}^{\infty} t^{-\epsilon} dS_a(t).$$

That the limit in question is $\phi(a)/a$ now follows in a straightforward way from the asymptotic relation $S_a(x) \sim \frac{\phi(a)}{a} \log x$, as $x \rightarrow \infty$.³ When $a = 2$, that relation is exactly the assertion of Theorem 1.3, while the proof for general a is entirely analogous.

5. CONTINUITY OF DISTRIBUTION FUNCTIONS: PROOF OF THEOREM 1.4

The following result in arithmetic combinatorics is the principal theorem of Pomerance and Sarközy’s paper [25].

Proposition 5.1. *There are absolute constants c_1 and N_1 such that, if N is a positive integer with $N \geq N_1$, \mathcal{P} is a set of prime numbers not exceeding N with*

$$\sum_{p \in \mathcal{P}} \frac{1}{p} > c_1,$$

$\mathcal{A} \subset \{1, 2, \dots, N\}$, and

$$\sum_{a \in \mathcal{A}} \frac{1}{a} > 10(\log N) \left(\sum_{p \in \mathcal{P}} \frac{1}{p} \right)^{-1/2},$$

then there exist integers $a < a'$ in \mathcal{A} for which $a \mid a'$ and

$$\frac{a'}{a} \mid \prod_{p \in \mathcal{P}} p.$$

³Conversely, if $\epsilon \sum 1/d\ell_a(d)^\epsilon \rightarrow K$ as $\epsilon \downarrow 0$, then $S_a(x) = (K + o(1)) \log x$, as $x \rightarrow \infty$. This follows from a Tauberian theorem of Hardy and Littlewood for Dirichlet series with nonnegative coefficients [11, Theorem 16].

We distill what we need from Proposition 5.1 into our next lemma (cf. [25, Corollary 2]).

Lemma 5.2. *Let $\delta > 0$, and let \mathcal{A} be a set of integers with upper logarithmic density exceeding δ . Then there are numbers $a < a'$ in \mathcal{A} for which $a \mid a'$ and $a'/a \ll_{\delta} 1$.*

Proof. Let \mathcal{P} consist of the first k primes, where $k = k(\delta)$ is chosen minimally to make $\sum_{p \in \mathcal{P}} \frac{1}{p} > c_1$ and $10(\sum_{p \in \mathcal{P}} \frac{1}{p})^{-1/2} < \delta$. Since \mathcal{A} has upper logarithmic density $> \delta$, there are arbitrarily large N for which

$$\sum_{\substack{a \in \mathcal{A} \\ a \leq N}} \frac{1}{a} > 10(\log N) \left(\sum_{p \in \mathcal{P}} \frac{1}{p} \right)^{-1/2}.$$

Proposition 5.1 now yields the existence of $a < a'$ in \mathcal{A} where $a \mid a'$ and $\frac{a'}{a} \leq \prod_{p \in \mathcal{P}} p \ll_{\delta} 1$. \square

We also need a classical result of Bang on primitive prime divisors of Mersenne numbers ([1]; see also [28]).

Lemma 5.3. *For each integer $m > 1$, $m \neq 6$, there is a prime p for which $\ell(p) = m$.*

We can now prove that the distribution function of $f(n)$ is continuous.

Proof of Theorem 1.4, for $f(n)$. Let $D(u)$ be the distribution function of f , and suppose for a contradiction that D is not continuous. Let α be a discontinuity of D (necessarily a jump discontinuity) and let $\delta = D(\alpha+) - D(\alpha-)$. Then $\delta > 0$ and for every $\epsilon > 0$,

$$D(\alpha + \epsilon) - D(\alpha - \epsilon) \geq \delta.$$

Except possibly on a countable set of ϵ (specifically, those ϵ where one of $\alpha \pm \epsilon$ is a discontinuity of D), the left-hand difference can be interpreted as the density of a with $f(a) \in (\alpha - \epsilon, \alpha + \epsilon)$. In particular, there are arbitrarily small positive ϵ where that density exists and is at least δ . For these ϵ , the corresponding logarithmic density also exists and is at least δ .

The set of a with squarefull part exceeding K has upper logarithmic density bounded by $\sum_{\text{squarefull } m \geq K} \frac{1}{m}$, which is smaller than $\frac{1}{2}\delta$ if K is fixed sufficiently large in terms of δ . Putting this together with the conclusion of the last paragraph, we conclude that there are arbitrarily small positive ϵ for which

$$\mathcal{A} = \{a : a \text{ has squarefull part at most } K, \text{ and } f(a) \in (\alpha - \epsilon, \alpha + \epsilon)\}$$

has lower logarithmic density exceeding $\frac{1}{2}\delta$. By Lemma 5.2, we may select $a < a'$ in \mathcal{A} for which $a \mid a'$ and $\frac{a'}{a} \ll_{\delta} 1$. Take q to be a prime appearing in a' to a higher power than in a , and say that $q^e \parallel a'$. If $e > 1$, then $q^e \leq K$, while if $e = 1$, then $q^e = q \mid a'/a$. Either way,

$$q^e \ll_{\delta} 1.$$

Now notice that

$$f(a') - f(a) \geq \sum_{\ell(p)=q^e} \frac{1}{p}.$$

By Bang's theorem, there is at least one prime p with $\ell(p) = q^e$, so the right-hand sum is nonempty. Moreover, if p is any such prime, $p < 2^{q^e} \ll_{\delta} 1$, and so

$$f(a') - f(a) \gg_{\delta} 1.$$

But $f(a), f(a') \in (\alpha - \epsilon, \alpha + \epsilon)$, so that

$$f(a') - f(a) < 2\epsilon.$$

The last two displays are incompatible for sufficiently small values of $\epsilon > 0$. \square

A bit of staring reveals that we have actually shown a more general result.

Theorem 5.4. *Suppose $a(n), b(n)$ are nonnegative-valued arithmetic functions connected by the relation*

$$(10) \quad a(n) = \sum_{d|n} b(d) \quad \text{for all positive integers } n.$$

Suppose that $b(q^e) > 0$ for every prime power q^e (including the case $e = 1$). Then for each real α and every $\delta > 0$, there is an $\epsilon > 0$ such that the upper logarithmic density of n with $a(n) \in (\alpha - \epsilon, \alpha + \epsilon)$ is smaller than δ . In particular, if $a(n)$ has a distribution function, that distribution function is continuous.

That the distribution function of $f(n)$ is continuous follows from choosing $a(n) = f(n)$ and $b(n) = \sum_{\ell(p)=n} 1/p$ in this general result. The second assertion of Theorem 5.4, that the distribution function of $F(n)$ is continuous, follows from choosing $a(n) = F(n)$ and $b(n) = E(n)$. (Note that Bang's theorem implies that $E(q^e) > 0$ for all prime powers q^e .)

Several variants of Theorem 5.4 are possible. For instance, instead of requiring that $b(q^e) > 0$ for all prime powers q^e , it is enough to impose this condition for the prime powers q^e with $q > q_0$, for any fixed q_0 . To see this, observe that the proof of Lemma 5.2 can be carried out including only primes larger than q_0 into \mathcal{P} . This allows one to impose the additional constraint in the conclusion of that lemma that the quotient a'/a have no prime factors up to q_0 . In the proof of Theorem 5.4, q will be chosen as a divisor of a'/a , and so the prime power q^e will have $q > q_0$.

An example and a non-example.

- (a) Let $a(n) = \sum_{p|F_n} 1/p$, where F_n is the n th Fibonacci number. It follows from the general results of [18] that $a(n)$ has a distribution function. To show continuity, we apply the variant of Theorem 5.4 just discussed. Let $z(p)$ denote the smallest natural number n for which $p | F_n$ (the rank of appearance of p). Then (10) holds with $b(n) = \sum_{z(p)|n} 1/p$. That $b(q^e) > 0$ for all odd prime powers q^e is a consequence of Carmichael's 1913 theorem [3] that $z(p) = m$ has a prime solution p for each $m \notin \{1, 2, 6, 12\}$. So our variant of Theorem 5.4 applies, and the distribution function of $a(n)$ is continuous.
- (b) In view of results such as Lemma 2.5, it is natural to compare $f(n)$ with $f^*(n) := \sum_{p-1|n} 1/p$. Theorem 5.4 does not apply in this example, and, unlike in (a), there is no obvious variant that does apply. There turns out to be a very good reason for these difficulties: $f^*(n)$ does not have a continuous distribution function. Indeed, is it clear that $f^*(n) = \frac{1}{2}$ precisely on the density $\frac{1}{2}$ set of odd numbers. With somewhat more effort, the ideas of [25] can be pushed to show that $f^*(n)$ has a discrete distribution function, with jumps precisely at the elements of $f^*(\mathbb{N})$.

6. CONCLUDING REMARKS ON GENERALIZATIONS

In all of the theorems appearing in the introduction, $2^n - 1$ can be replaced with $a^n - 1$ for any fixed integer $a > 1$. In the statement of Theorem 1.1, one should then replace $\frac{1}{2}$ with $\sum_{p|a} \frac{1}{p}$, and in Theorem 1.2, $\log 2$ should be replaced with $\log \frac{a}{\phi(a)}$. In Theorem 1.3, $\frac{1}{2}$ should be replaced with $\frac{\phi(a)}{a}$. Only superficial changes to the proofs are required. Theorem 1.4 remains true as stated, by more or less the same proof (we use that $\ell_a(p) = n$ has at least one solution p for each $n > 6$; see [28]).

A natural domain of generalization is provided by the nondegenerate Lucas sequences of the form $U_n(P, Q)$, where P and Q are coprime integers. (See Chapter 1 of [27] for a refresher on the basic properties of these sequences.) If a is an integer larger than 1, and we take $P = a + 1$ and $Q = a$, then $U_n(P, Q) = \frac{a^n - 1}{a - 1}$. For our purposes, $\frac{a^n - 1}{a - 1}$ is more or less the same as $a^n - 1$, and so this setting morally subsumes the one considered in the last paragraph.

For these sequences U_n , a positive integer m divides some term U_n precisely when m is coprime to Q . In that case, the set of corresponding n consists of all multiples of the rank of appearance $z(m)$, which is analogous to $\ell(m)$. If we set $\Delta = P^2 - 4Q$, then $z(p) \mid p - \left(\frac{\Delta}{p}\right)$ for each odd prime $p \nmid PQ\Delta$. Furthermore, for each prime p not dividing Q , and every positive integer k , we have $z(p^k) \mid z(p)p^{k-1}$. Also, $z(m) = \text{lcm}_{p^k \parallel m} z(p^k)$.

We expect U_n -analogues of all of our theorems. For Theorem 1.4, these analogues are fairly simple to show (and the case $P = 1, Q = -1$, corresponding to the Fibonacci numbers, was already discussed): Again [18] gives existence of the distribution function. It is known that for each $n > 30$, there is a prime p with $z(p) = n$ [2]. This allows us to deduce continuity from the variant of Theorem 5.4 discussed above.

The U_n versions of Theorems 1.1, 1.2, and 1.3 also appear attainable. Their statements should be modified as in the first paragraph of this remark, but with $|Q|$ replacing a . However, the proofs cannot be viewed as routine, as the literature does not seem to contain explicit analogues of Proposition 2.2, Lemma 2.5, Proposition 4.1 (see p. 13 of [27] for the U_n -analogue of λ), or Proposition 4.4. However, we believe all of these to be within reach. To give one example, it should be possible to prove the needed variant of Lemma 2.5 following Kurlberg and Pomerance, starting from Chen's arguments in [4].

One can also consider analogues of f and F associated to the "companion" V_n Lucas sequences. Extending Erdős's upper bound (2) to V_n requires slightly more arithmetic. Whereas all but finitely many primes divide some term in the U_n sequence, with finitely many exceptions a prime p divides some term in the V_n sequence if and only if $z(p)$ is defined and even. Using results of Moree [19, 20], the first author [5] showed that

$$\max_{n \leq x} \sum_{p \mid a^n + b^n} \frac{1}{p} = \frac{1}{3} \log \log \log x + O_{a,b}(1)$$

for a, b positive integers with a/b not a rational square. For an arbitrary nondegenerate Lucas sequence, the constant $1/3$ should be replaced by some density $\delta \in (0, 1)$ depending on a and b . We believe that Theorems 1.1, 1.2, and 1.3, can be extended to V_n Lucas sequences as well.

Establishing the analogue of Theorem 1.4 for V_n is easier. While V_n is not a divisibility sequence, it is still the case that $V_n \mid V_{mn}$ whenever m is odd, and this is enough to allow a modified version of the argument to run. (To get an idea of how this goes, suppose a, a' are positive integers with a'/a odd, and where q^e is an odd prime power dividing a' but not a . If $z(p) = 2^{r+1}q^e$, where $2^r \parallel a$, then $p \mid V_{a'}$ but $p \nmid V_a$. To make this observation useful, we use again that $z(p) = n$ is solvable for all large n , along with the fact that most integers are not divisible by a large power of 2.)

Acknowledgements. The authors thank Carl Pomerance and the anonymous referee for several helpful suggestions. The second author is supported by NSF Award DMS-2001581.

REFERENCES

- [1] A. S. Bang, *Taltheoretiske Undersøgelser*, Tidsskrift for Mathematik **4** (1886), 70–80, 130–137.
- [2] Y. Bilu, G. Hanrot, and P. M. Voutier (with an appendix by M. Mignotte), *Existence of primitive divisors of Lucas and Lehmer numbers*, J. Reine Angew. Math. **539** (2001), 75–122.
- [3] R. D. Carmichael, *On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$* , Ann. Math. **15** (1913), 30–70.
- [4] Y.-M. J. Chen, *On primitive roots of one-dimensional tori*, J. Number Theory **93** (2002), 23–33.
- [5] Z. Engberg, *The arithmetic of cyclic subgroups*, Ph.D. thesis, Dartmouth College, 2014.
- [6] P. Erdős, *On some problems of Bellman and a theorem of Romanoff*, J. Chinese Math. Soc. (N.S.) **1** (1951), 409–421.
- [7] ———, *On the sum $\sum_{d \mid 2^n - 1} d^{-1}$* , Israel J. Math. **9** (1971), 43–48.
- [8] P. Erdős, P. Kiss, and C. Pomerance, *On prime divisors of Mersenne numbers*, Acta Arith. **57** (1991), 267–281.
- [9] J. B. Friedlander, C. Pomerance, and I. E. Shparlinski, *Period of the power generator and small values of Carmichael's function*, Math. Comp. **70** (2001), 1591–1605, corrigendum in **71** (2002), 1803–1806.

- [10] A. Granville, *Smooth numbers: computational number theory and beyond*, Algorithmic number theory: lattices, number fields, curves and cryptography, Math. Sci. Res. Inst. Publ., vol. 44, Cambridge Univ. Press, Cambridge, 2008, pp. 267–323.
- [11] G. H. Hardy and J. E. Littlewood, *Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive*, Proc. London Math. Soc. **13** (1914), 174–191.
- [12] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, sixth ed., Oxford University Press, Oxford, 2008.
- [13] P. Kurlberg, *On the order of unimodular matrices modulo integers*, Acta Arith. **110** (2003), 141–151.
- [14] P. Kurlberg and C. Pomerance, *On the periods of the linear congruential and power generators*, Acta Arith. **119** (2005), 149–169.
- [15] P. Kurlberg and Z. Rudnick, *On quantum ergodicity for linear maps of the torus*, Comm. Math. Phys. **222** (2001), 201–227.
- [16] Y. Lamzouri, *Smooth values of the iterates of the Euler phi-function*, Canad. J. Math. **59** (2007), 127–147.
- [17] F. Luca, *Some mean values related to average multiplicative orders of elements in finite fields*, Ramanujan J. **9** (2005), 33–44.
- [18] F. Luca and I. E. Shparlinski, *Arithmetic functions with linear recurrence sequences*, J. Number Theory **125** (2007), 459–472.
- [19] P. Moree, *On the divisors of $a^k + b^k$* , Acta Arith. **80** (1997), 197–212.
- [20] ———, *Counting divisors of Lucas numbers*, Pacific J. Math. **186** (1998), 267–284.
- [21] L. Murata and C. Pomerance, *On the largest prime factor of a Mersenne number*, Number theory, CRM Proc. Lecture Notes, vol. 36, Amer. Math. Soc., Providence, RI, 2004, pp. 209–218.
- [22] M. R. Murty, M. Rosen, and J. H. Silverman, *Variations on a theme of Romanoff*, Internat. J. Math. **7** (1996), 373–391.
- [23] C. Pomerance, *Popular values of Euler's function*, Mathematika **27** (1980), 84–89.
- [24] ———, *On primitive divisors of Mersenne numbers*, Acta Arith. **46** (1986), 355–367.
- [25] C. Pomerance and A. Sárközy, *On homogeneous multiplicative hybrid problems in number theory*, Acta Arith. **49** (1988), 291–302.
- [26] C. Pomerance and I. E. Shparlinski, *Smooth orders and cryptographic applications*, Algorithmic number theory (Sydney, 2002), Lecture Notes in Comput. Sci., vol. 2369, Springer, Berlin, 2002, pp. 338–348.
- [27] P. Ribenboim, *My numbers, my friends*, Springer-Verlag, New York, 2000.
- [28] M. Roitman, *On Zsigmondy primes*, Proc. Amer. Math. Soc. **125** (1997), 1913–1919.
- [29] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, third ed., Graduate Studies in Mathematics, vol. 163, American Mathematical Society, Providence, RI, 2015.
- [30] Z. Wang, *Autour des plus grands facteurs premiers d'entiers consécutifs voisins d'un entier criblé*, Q. J. Math. **69** (2018), 995–1013.

WASATCH ACADEMY, MT. PLEASANT, UT 84647
 Email address: zeb.engberg@wasatchacademy.org

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GEORGIA, ATHENS, GA 30602
 Email address: pollack@uga.edu