MATHEMATICS OF COMPUTATION
VOLUME 46, NUMBER 173
JANUARY 1986, PAGES 259–279

On the Number of False Witnesses for a Composite Number

By Paul Erdös and Carl Pomerance*

Abstract. If a is not a multiple of n and $a^{n-1} \not\equiv 1 \bmod n$, then n must be composite and a is called a "witness" for n. Let $F(n)$ denote the number of "false witnesses" for n, that is, the number of $a \bmod n$ with $a^{n-1} \equiv 1 \bmod n$. Considered here is the normal and average size of $F(n)$ for n composite. Also considered is the situation for the more stringent Euler and strong pseudoprime tests.

1. Introduction. When presented with a large number n which one would like to test for primality, one usually begins with a modicum of trial division. If n is not revealed as composite, the next step is often to perform the simple and cheap test of computing $a^{n-1} \bmod n$ for some prechosen number $a > 1$ with $(a, n) = 1$. If this residue is not 1, then n is definitely composite (by Fermat's little theorem) and we say a is a *witness* for n. If the residue is 1, then n is probably prime, but there are exceptions. If we are in this exceptional case where

$$a^{n-1} \equiv 1 \bmod n \text{ and } n \text{ is composite,}$$

then we say a is a *false witness* for n, or equally, that n is a *pseudoprime* to the base a.

The problem of distinguishing between pseudoprimes and primes has been the subject of much recent work. For example, see [21].

Let

$$\mathbf{F}(n) = \{a \bmod n\colon a^{n-1} \equiv 1 \bmod n\}, \qquad F(n) = \#\mathbf{F}(n). \tag{1.1}$$

Thus, if n is composite, then $\mathbf{F}(n)$ is the set (in fact, group) of residues mod n that are false witnesses for n and $F(n)$ is the number of such residues. If n is prime, then $F(n) = n - 1$ and $\mathbf{F}(n)$ is the entire group of reduced residues mod n. For any n, Lagrange's theorem gives $F(n) | \phi(n)$, where ϕ is Euler's function.

There are composite numbers n for which $F(n) = \phi(n)$, such as $n = 561$. Such numbers are called Carmichael numbers and probably there are infinitely many of them, but this has never been proved. It is known that Carmichael numbers are much rarer than primes.

At the other extreme, there are infinitely many numbers n for which $F(n) = 1$. For example, any number of the form $2p$ will do, where p is prime. It is possible to show (in fact, we do so below) that while these numbers n with $F(n) = 1$ have asymptotic density 0, they are much more common than primes.

Received March 7, 1985; revised May 15, 1985.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11Y12; Secondary 11N56.
*Research supported in part by an NSF grant.

0025-5718/86 $1.00 + $.25 per page

So what is the normal and/or average behavior of the function $F(n)$? It is to these questions that this paper is addressed. We show (where $\sum'$ denotes a sum over composite numbers here and throughout the paper)

$$\frac{1}{x}\sum_{n\leqslant x}{}' F(n) > x^{15/23} \tag{1.2}$$

for x large and

$$\frac{1}{x}\sum_{n\leqslant x}{}' F(n) \leqslant x\exp\left\{\frac{-(1+o(1))\log x\log\log\log x}{\log\log x}\right\} \tag{1.3}$$

as $x\to\infty$. We conjecture that equality holds in (1.3).

We can compute the geometric mean value with more precision: there are positive constants c_1, c_2 such that

$$\left(\prod_{n\leqslant x} F(n)\right)^{1/x} = c_2(\log x)^{c_1} + o(1) \tag{1.4}$$

as $x\to\infty$. If the geometric mean is taken just for composite numbers, then the result is the same except that c_2 is replaced by c_2/e.

Concerning the normal value of $F(n)$, we show that $\log F(n)/\log\log n$ has a distribution function $D(u)$. That is, $D(u)$ is the asymptotic density of the integers n for which

$$F(n) \leqslant (\log n)^u.$$

The function $D(u)$ is continuous, strictly increasing, and singular on $[0,\infty)$. Moreover, $D(0) = 0$ and $D(+\infty) = 1$. Thus, for example, the set of n with $F(n) = 1$ has density 0.

The starting point for our results is the elegant and simple formula of Monier [17] and Baillie and Wagstaff [2]:

$$F(n) = \prod_{p\mid n}(p-1, n-1), \tag{1.5}$$

where here (and throughout the paper) p denotes a prime. For example, (1.5) immediately implies $F(2p) = 1$.

We are also able to prove analogous results for certain pseudoprime tests more stringent than the Fermat congruence, namely the Euler test and the strong pseudoprime test. It is to be expected that there will be similar results for all Fermat-type tests; for example, the Lucas tests. Such an undertaking might gain useful insights into the nature of these tests.

In the last section we address some further questions including the maximal order of $F(n)$ for n composite, the nature of the range of F, the normal number of prime factors of $F(n)$, and the universal exponent for the group $\mathbf{F}(n)$.

2. The Average Order of $F(n)$. In this section we shall establish the results (1.2) and (1.3). Recall that $\sum'$ denotes a sum over composite numbers.

THEOREM 2.1. *For all large x,*

$$\frac{1}{x}\sum_{n\leqslant x}{}' F(n) > x^{15/23}.$$

Proof. Let $M(x)$ denote the least common multiple of the integers up to x. For any y, let

$$\mathbf{P}(y, x) = \{ p \leqslant y: p - 1 | M(x) \}.$$

From Erdös [7] it follows that there is a value of $\alpha > 1$ with

$$\#\mathbf{P}(x^{\alpha'}, x) \gg x^{\alpha'}/\log x \quad \text{for all } 0 < \alpha' < \alpha, \tag{2.1}$$

while from Pomerance [20] it follows that there is a value of $\alpha > 9/4$ with (2.1) holding. (Note that the notation $f(x) \gg g(x)$ for positive functions f, g is equivalent to $g(x) = O(f(x))$.) We conjecture that (2.1) holds for every $\alpha > 0$. Let

$$\beta = \sup\{ \alpha: \#\mathbf{P}(x^{\alpha'}, x) = x^{\alpha' + o(1)} \text{ for all } 0 < \alpha' < \alpha \},$$

so that if (2.1) holds for α, then $\beta \geqslant \alpha$. From recent work of Balog [3] we have

$$\beta > 23/8. \tag{2.2}$$

Let L denote an upper bound for Linnik's constant, so that given positive integers a, m with $(a, m) = 1$ and $m > 1$, then there is a prime $p \equiv a \bmod m$ with $p < m^L$. Let α be such that $1 < \alpha < \beta$ and let $0 < \varepsilon < \alpha - 1$ be arbitrarily small, but fixed. Let

$$M = M(\log x/\log\log x),$$
$$\mathbf{P} = \mathbf{P}(\log^{\alpha} x, \log x/\log\log x) \setminus \{ p: p \leqslant \log^{\alpha-\varepsilon} x \}.$$

Note that

$$M = \exp\{(1 + o(1)) \log x/\log\log x\} = x^{o(1)}, \qquad \#\mathbf{P} = (\log x)^{\alpha + o(1)}. \tag{2.3}$$

Let $\mathbf{S}$ denote the set of integers composed of exactly

$$k = \left[\log(x/M^L)/\log(\log^{\alpha} x)\right]$$

distinct primes in $\mathbf{P}$. Thus, if $s \in \mathbf{S}$, then

$$x^{1-\varepsilon} < s \leqslant x/M^L \tag{2.4}$$

for all large x.

From (2.3) we have

$$\begin{aligned} \#\mathbf{S} = \binom{\#\mathbf{P}}{k} &\geqslant \left(\frac{\#\mathbf{P}}{k}\right)^k \geqslant \left(\frac{\#\mathbf{P}}{\log x}\right)^{(\alpha^{-1} + o(1)) \log x/\log\log x} \\ &= (\log x)^{(\alpha - 1 + o(1))(\alpha^{-1} + o(1)) \log x/\log\log x} = x^{(\alpha-1)\alpha^{-1} + o(1)}. \end{aligned} \tag{2.5}$$

If $s \in \mathbf{S}$, let $q = q(s)$ denote the least prime such that

$$sq \equiv 1 \mod M. \tag{2.6}$$

Thus $q \leqslant M^L$, so that $sq \leqslant x$ by (2.4). Let $\mathbf{S}'$ denote the set of such numbers sq. If $n \in \mathbf{S}'$, then n has at most $O(\log x)$ representations as sq for $s \in \mathbf{S}$, so that from (2.5), we have

$$\#\mathbf{S}' \geqslant x^{(\alpha-1)\alpha^{-1} + o(1)}. \tag{2.7}$$

If $n = sq \in \mathbf{S}'$, where $s \in \mathbf{S}$ and q is a prime satisfying (2.6), then

$$\begin{aligned} F(n) = \prod_{p|n} (p - 1, n - 1) &\geqslant \prod_{p|s} (p - 1, M) = \prod_{p|s} (p - 1) \\ &= \phi(s) \gg s/\log\log s \geqslant x^{1-\varepsilon}/\log\log x, \end{aligned} \tag{2.8}$$

by (1.5), Theorem 328 in [12], and (2.4). Thus by (2.7),

$$\frac{1}{x}\sum_{n\leqslant x}{}' F(n) \geqslant \frac{1}{x}\sum_{n\in \mathbf{S}'} F(n) \geqslant x^{-\varepsilon+o(1)}\cdot \#\mathbf{S}' \geqslant x^{(\alpha-1)\alpha^{-1}-\varepsilon+o(1)}.$$

Since $\varepsilon > 0$ is arbitrarily small and α is arbitrarily close to β, we have

$$\frac{1}{x}\sum_{n\leqslant x}{}' F(n) \geqslant x^{1-\beta^{-1}+o(1)}.$$

Our theorem now follows from (2.2).

Remark. For any choice of Θ with $\frac{1}{2} \leqslant \Theta < 1$, let $C(\Theta)$ be the least number such that for any choice of ε with $0 < \varepsilon \leqslant 1 - \Theta$, any integer a, and any $A > 0$, there is an $x_0(\varepsilon, a, A)$ such that for all $x \geqslant x_0(\varepsilon, a, A)$,

$$\sum_{\substack{p\leqslant x\\ p\equiv a \bmod q}} 1 \leqslant \frac{(C(\Theta)+\varepsilon)x}{\phi(q)\log x}$$

for all but at most $x^{\Theta}/(\log x)^A$ choices of $q \in (x^{\Theta}, 2x^{\Theta}]$. Let

$$C = \limsup_{\Theta\to\frac{1}{2}+} C(\Theta).$$

In [3], Balog shows that the number β defined in the proof of Theorem 2.1 above satisfies $\beta \geqslant 2e^{(C+1)^{-1}}$.

The estimate (2.2) is obtained using the result $C < 1.73$ which was recently obtained by Fouvry [10]. With continued improvements expected for estimates of C, we thus would have better estimates in Theorem 2.1. For example, if it can be shown that $C < 1.463$, then we could write $2/3$ in place of $15/23$ in Theorems 2.1 and 5.1.

Let

$$L(x) = \exp(\log x \log\log\log x/\log\log x).$$

THEOREM 2.2. *As* $x \to \infty$,

$$\frac{1}{x}\sum_{n\leqslant x}{}' F(n) \leqslant xL(x)^{-1+o(1)}.$$

Proof. For each integer $k \geqslant 1$, let $\mathbf{C}_k(x)$ denote the set of composite $n \leqslant x$ such that $F(n) = \phi(n)/k$, and let $C_k(x) = \#\mathbf{C}_k(x)$. We have

$$\begin{aligned}\sum_{n\leqslant x}{}' F(n) &= \sum_k \sum_{n\in\mathbf{C}_k(x)} F(n) \leqslant \sum_k \sum_{n\in\mathbf{C}_k(x)} \frac{n}{k}\\ &\leqslant \sum_{n\leqslant x}\frac{n}{L(x)} + \sum_{k\leqslant L(x)}\sum_{n\in\mathbf{C}_k(x)}\frac{n}{k} \qquad (2.9)\\ &\leqslant \frac{x^2}{L(x)} + x\sum_{k\leqslant L(x)}\frac{1}{k}C_k(x).\end{aligned}$$

It will thus be sufficient to obtain the uniform upper bound

$$C_k(x) \leqslant xL(x)^{-1+o(1)} \quad \text{for } k \leqslant L(x), \qquad (2.10)$$

for using (2.10) in (2.9) gives the theorem.

We begin by showing

$$F(n) = \phi(n)/k \quad \text{implies} \quad \lambda(n) | k(n-1), \tag{2.11}$$

where $\lambda(n)$ is Carmichael's universal exponent function. (Recall that $\lambda(n)$ is the smallest positive number such that $a^{\lambda(n)} \equiv 1 \bmod n$ for all a with $(a, n) = 1$.) It is not hard to show that there is actually some residue $a \bmod n$ that belongs to the exponent $\lambda(n)$. To see (2.11), suppose $F(n) = \phi(n)/k$, let $u = \lambda(n)/(\lambda(n), n-1)$ and let a belong to the exponent $\lambda(n)$ in the group $(Z/n)^*$ of all reduced residues $\bmod n$. Then the order of $a\mathbf{F}(n)$ in $(Z/n)^*/\mathbf{F}(n)$ (where $\mathbf{F}(n)$ is given by (1.1)) is exactly u. Thus, by Lagrange's theorem,

$$u | \phi(n)/F(n) = k,$$

so that $\lambda(n) | u(n-1) | k(n-1)$, which was to be shown.

To estimate $C_k(x)$ for $k \leqslant L(x)$ we consider three cases:

(i) $n < x/L(x)$,

(ii) n is divisible by some prime $p > kL(x)$,

(iii) $n \geqslant x/L(x)$ and every prime p in n is at most $kL(x)$.

If $n \in \mathbf{C}_k(x)$ and $p | n$, then by (2.11),

$$p - 1 | \lambda(n) | k(n-1),$$

so that

$$(p-1)/(k, p-1) | n - 1.$$

Since also $n > p$, the number of $n \in \mathbf{C}_k(x)$ divisible by p is at most

$$\left[\frac{x}{p(p-1)/(k, p-1)} \right] \leqslant \frac{xk}{p(p-1)}.$$

Thus the number of $n \in \mathbf{C}_k(x)$ in case (ii) is at most

$$\sum_{p > kL(x)} \frac{xk}{p(p-1)} < \frac{x}{L(x)}. \tag{2.12}$$

If $n \leqslant x$ is in case (iii), then n has a divisor d satisfying

$$\frac{x}{kL(x)^2} < d \leqslant \frac{x}{L(x)}. \tag{2.13}$$

If $n \in \mathbf{C}_k(x)$ and $d | n$, then from (2.11),

$$\lambda(d) | \lambda(n) | k(n-1).$$

Thus the number of $n \in \mathbf{C}_k(x)$ with $d | n$ is at most

$$\left[\frac{x}{d\lambda(d)/(k, \lambda(d))} \right] + 1.$$

If $\sum^*$ denotes a sum over d satisfying (2.13), we thus have the number of $n \in \mathbf{C}_k(x)$ in case (iii) is at most

$$\begin{aligned} \sum{}^* \left(\frac{x(k, \lambda(d))}{d\lambda(d)} + 1 \right) &\leqslant \frac{x}{L(x)} + x \sum{}^* \frac{(k, \lambda(d))}{d\lambda(d)} \\ &= \frac{x}{L(x)} + x \sum_{m \leqslant x} \frac{1}{m} \sum_{\lambda(d)/(k, \lambda(d)) = m}^* \frac{1}{d} \\ &\leqslant \frac{x}{L(x)} + x \sum_{m \leqslant x} \frac{1}{m} \sum_{u | k} \sum_{\lambda(d) = mu}^* \frac{1}{d}. \end{aligned} \tag{2.14}$$

To estimate the inner sum we use the fact that from (15) in [22], there is an x_0 such that for all integers s and $x \geqslant x_0$, we have

$$\#\{d \leqslant x: \lambda(d) = s\} \leqslant x/L(x).$$

Thus using partial summation,

$$\sum_{\lambda(d)=mu}^{*} \frac{1}{d} \leqslant \frac{L(x)}{x} \sum_{\lambda(d)=mu}^{*} 1 + \int_{x/kL(x)^2}^{x/L(x)} \frac{1}{t^2} \sum_{\substack{d<t \\ \lambda(d)=mu}} 1\, dt$$

$$\leqslant \frac{1}{L(x/L(x))} + \frac{\log x}{L\left(x/kL(x)^2\right)},$$

provided $x/kL(x)^2 \geqslant x_0$. Putting this estimate into (2.14), the number of $n \in \mathbf{C}_k(x)$ in case (iii) is

$$\ll \frac{x \log x}{L\left(x/kL(x)^2\right)} \sum_{m \leqslant x} \frac{1}{m} \sum_{u|k} 1$$

$$\leqslant \frac{x \log^2 x}{L\left(x/kL(x)^2\right)} 2^{(1+o(1))\log k/\log\log k}$$

$$= xL(x)^{-1+o(1)},$$

uniformly for $k \leqslant L(x)$. (We have used Theorem 317 in [12].)

Using this last estimate with (2.12) and the fact that evidently there are at most $x/L(x)$ choices of n in case (i), we have (2.10) and thus the theorem.

Remarks. We conjecture that Theorem 2.2 is sharp, i.e.,

$$\frac{1}{x} \sum_{n \leqslant x}{}' F(n) = xL(x)^{-1+o(1)}. \tag{2.15}$$

In [8], [24], and [22] a sequence of heuristic arguments is presented that culminate in the conjecture

$$C_1(x) = xL(x)^{-1+o(1)}.$$

(The upper bound in this assertion is proved in these papers and also follows from Theorem 2.2.) This conjecture together with Theorem 2.2 immediately gives the conjecture (2.15).

Let $P_a(x)$ denote the number of $n \leqslant x$ such that n is a pseudoprime to the base a. Thus $P_a(x)$ is the number of composite $n \leqslant x$ with $a \bmod n \in \mathbf{F}(n)$. For a fixed value of a, the sharpest results known on $P_a(x)$ are that

$$\exp\left\{(\log x)^{5/14}\right\} < P_a(x) < xL(x)^{-1/2} \tag{2.16}$$

for all $x \geqslant x_0(a)$; see [22], [23]. (Using (2.2), we may replace the "5/14" in the lower bound with 15/38.) We trivially have

$$\sum_{a \leqslant x} P_a(x) \geqslant \sum_{n \leqslant x}{}' F(n).$$

On the other hand,

$$\sum_{a \leqslant x} P_a(x) = \sum_{n \leqslant x}{}' \sum_{\substack{a \leqslant x \\ a^{n-1} \equiv 1 \bmod n}} 1 \leqslant \sum_{n \leqslant x}{}' F(n)\left(\frac{x}{n} + 1\right).$$

Thus, by using partial summation and Theorems 2.1 and 2.2 we can obtain a result that is, *on average*, much better than (2.16):

$$x^{15/23} < \frac{1}{x}\sum_{a\leqslant x} P_a(x) \leqslant xL(x)^{-1+o(1)}$$

for x large.

3. The Geometric Mean Value.

THEOREM 3.1. *There are positive constants c_1, c_2 such that*

$$\Big(\prod_{n\leqslant x} F(n)\Big)^{1/x} = c_2(\log x)^{c_1} + O\big((\log x)^{c_1-1}\big)$$

for $x \geqslant 2$.

Proof. Letting $\Lambda(n)$ denote von Mangoldt's function, we have

$$\sum_{n\leqslant x}\log F(n) = \sum_{n\leqslant x}\log\prod_{p\mid n}(p-1, n-1) = \sum_{n\leqslant x}\sum_{p\mid n}\sum_{d\mid(p-1,\,n-1)}\Lambda(d)$$

$$= \sum_{d\leqslant x}\Lambda(d)\sum_{\substack{p\leqslant x\\ d\mid p-1}}\sum_{\substack{n\leqslant x\\ d\mid n-1\\ p\mid n}} 1 = \sum_{d\leqslant x}\Lambda(d)\sum_{\substack{p\leqslant x\\ d\mid p-1}}\left(\left[\frac{x-p}{pd}\right]+1\right) \tag{3.1}$$

$$= \sum_{p\leqslant x}\sum_{d\mid p-1}\Lambda(d) + \sum_{d\leqslant x}\Lambda(d)\sum_{\substack{p\leqslant x\\ d\mid p-1}}\left[\frac{x-p}{pd}\right] = S_1 + S_2,$$

say. From the prime number theorem with error term, we have

$$S_1 = \sum_{p\leqslant x}\log(p-1) = x + O\big(x/e^{\sqrt{\log x}}\big). \tag{3.2}$$

To estimate S_2, we write $S_2 = S_{2,1} + S_{2,2}$, where in $S_{2,1}$ we have $d \leqslant \log^2 x$ and in $S_{2,2}$ we have $\log^2 x < d \leqslant x$. It is shown in Norton [18] and Pomerance [19], that uniformly for any $d \geqslant 2$ and $x \geqslant 3$, we have

$$\sum_{\substack{p\leqslant x\\ d\mid p-1}}\frac{1}{p} = \frac{\log\log x}{\phi(d)} + O\left(\frac{\log d}{\phi(d)}\right). \tag{3.3}$$

Thus,

$$S_{2,2} \leqslant \sum_{\log^2 x < d\leqslant x}\Lambda(d)\sum_{\substack{p\leqslant x\\ d\mid p-1}}\frac{x}{pd}$$

$$\ll x\sum_{\log^2 x < d\leqslant x}\left(\frac{\Lambda(d)\log\log x}{d\phi(d)} + \frac{\Lambda(d)\log d}{d\phi(d)}\right) \tag{3.4}$$

$$\ll \frac{x\log\log x}{\log^2 x}.$$

For any integer $d \geqslant 2$, let

$$C(d) = \lim_{x\to\infty}\left(-\frac{\log\log x}{\phi(d)} + \sum_{\substack{p\leqslant x\\ d\mid p-1}}\frac{1}{p}\right).$$

From the Siegel-Walfisz theorem it is known that $C(d)$ exists and that

$$\sum_{\substack{p \leqslant x \\ d \mid p-1}} \frac{1}{p} = \frac{\log\log x}{\phi(d)} + C(d) + O_A(e^{-\sqrt{\log x}}) \tag{3.5}$$

uniformly for $d \leqslant \log^A x$ for any $A > 0$. Moreover, from (3.3) it follows that for all $d \geqslant 2$

$$C(d) = O\left(\frac{\log d}{\phi(d)}\right). \tag{3.6}$$

Therefore from (3.5) and (3.6),

$$\begin{aligned} S_{2,1} &= \sum_{d \leqslant \log^2 x} \Lambda(d) \sum_{\substack{p \leqslant x \\ d \mid p-1}} \left[\frac{x-p}{pd}\right] \\ &= \sum_{d \leqslant \log^2 x} \Lambda(d) \sum_{\substack{p \leqslant x/d \\ d \mid p-1}} \frac{x}{pd} + O\left(\sum_{d \leqslant \log^2 x} \Lambda(d) \sum_{\substack{p \leqslant x/d \\ d \mid p-1}} 1\right) \\ &= \sum_{d \leqslant \log^2 x} \frac{x\Lambda(d)}{d}\left(\frac{\log\log(x/d)}{\phi(d)} + C(d) + O(e^{-\sqrt{\log x}})\right) \qquad (3.7) \\ &\quad + O\left(\sum_{d \leqslant \log^2 x} \frac{x\Lambda(d)}{d\phi(d)\log x}\right) \\ &= \sum_{d \leqslant \log^2 x} \frac{x\Lambda(d)\log\log(x/d)}{d\phi(d)} + \sum_{d \leqslant \log^2 x} \frac{x\Lambda(d)C(d)}{d} + O\left(\frac{x}{\log x}\right) \\ &= x\log\log x \sum_d \frac{\Lambda(d)}{d\phi(d)} + x\sum_d \frac{\Lambda(d)C(d)}{d} + O\left(\frac{x}{\log x}\right). \end{aligned}$$

Let

$$c_1 = \sum_d \frac{\Lambda(d)}{d\phi(d)}, \qquad c_2 = \exp\left(1 + \sum_d \frac{\Lambda(d)C(d)}{d}\right). \tag{3.8}$$

Combining (3.1), (3.2), (3.4), and (3.7), we have

$$\begin{aligned} \left(\prod_{n \leqslant x} F(n)\right)^{1/x} &= \exp\{c_1 \log\log x + \log c_2 + O(1/\log x)\} \\ &= c_2(\log x)^{c_1} + O\left((\log x)^{c_1 - 1}\right). \end{aligned}$$

Remarks. The exponent c_1 given by (3.8) can be numerically evaluated. Using the identity

$$\begin{aligned} c_1 + \frac{\zeta'}{\zeta}(2) + \frac{\zeta'}{\zeta}(3) + \frac{\zeta'}{\zeta}(4) + 2\frac{\zeta'}{\zeta}(5) \\ = \sum_p \frac{p^6 + 3p^5 + 2p^4 - 2p^3 - 2p^2 - p - 5}{(p^3-1)(p^4-1)(p^5-1)} \log p \end{aligned}$$

and values for these logarithmic derivatives of $\zeta(s)$ kindly supplied to us by Andrew M. Odlyzko, we have found that

$$c_1 = 0.898464899\cdots. \tag{3.9}$$

Since $c_1 < 1$, it follows that the error term $O((\log x)^{c_1-1})$ in the theorem is $o(1)$.

If we restrict the geometric mean to just composite integers, the result does not change very much. In fact,

$$\sum_{p\leqslant x} \log F(p) = \sum_{p\leqslant x} \log(p-1) = x + O\left(x/e^{\sqrt{\log x}}\right),$$

so that

$$\begin{aligned}\left(\prod_{\substack{n\leqslant x\\ n \text{ composite}}} F(n)\right)^{1/(x-\pi(x))} &= \exp\left\{c_1 \log\log x + \log c_2 - 1 + \frac{c_1 \log\log x}{\log x} + O\left(\frac{1}{\log x}\right)\right\}\\ &= \frac{c_2}{e}(\log x)^{c_1} + \frac{c_1 c_2}{e}(\log x)^{c_1-1}\log\log x + O\left((\log x)^{c_1-1}\right)\\ &= \frac{c_2}{e}(\log x)^{c_1} + o(1).\end{aligned}$$

4. The Normal Order of $F(n)$. The spirit of this section is to ignore sets of asymptotic density 0 and thus examine the behavior of $F(n)$ for typical values of n. Theorem 2.1 says that the arithmetic mean order of $F(n)$ for n composite is at least n^c where $c > 0$. On the other hand, Theorem 3.1 says that the geometric mean order of $F(n)$ is $c_2(\log n)^{c_1}$, where c_1, c_2 are given by (3.8). Of these two results, it is clear that the geometric mean order is more relevant for normal numbers. Indeed, an immediate corollary of Theorem 3.1 is that if $h(n) \to \infty$ arbitrarily slowly, then the asymptotic density of the n for which $F(n) < (\log n)^{h(n)}$ is 1.

Consider the additive function

$$G(n) = \sum_{d|n} \Lambda(d)/\phi(d), \tag{4.1}$$

where Λ is von Mangoldt's function. This function has essentially already appeared in the previous section, since the constant c_1 in (3.8) is the mean value of $G(n)$. Thus Theorem 3.1 says that the geometric mean order of $F(n)$ is about $\log n$ raised to the mean value of the function $G(n)$. In this section, we shall show that the normal order of $F(n)$ is about $(\log n)^{G(n-1)}$.

THEOREM 4.1. *There is a set S of natural numbers of asymptotic density* 1 *such that*

$$\lim_{n\in S,\, n\to\infty}\left(\frac{\log F(n)}{\log\log n} - G(n-1)\right) = 0,$$

where G is given by (4.1).

Proof. For $x \geqslant 3$, let $y = y(x)$ denote $\log\log x$. We have

$$\begin{aligned}\log F(n) &= \sum_{p|n} \log(p-1, n-1) = \sum_{p|n}\sum_{d|(p-1,\,n-1)} \Lambda(d)\\ &= \sum_{d|n-1} \Lambda(d) \sum_{\substack{p|n\\ d|p-1}} 1 = \sum_{d|n-1} \omega(n,d)\Lambda(d),\end{aligned}$$

where we write $\omega(n,d) = \sum_{p|n,\, d|p-1} 1$. Thus

$$\log F(n) = \sum_{\substack{d|n-1 \\ d\leqslant\sqrt{y}}} \omega(n,d)\Lambda(d) + \sum_{\substack{d|n-1 \\ d>\sqrt{y}}} \omega(n,d)\Lambda(d) \tag{4.2}$$
$$= S_1(n) + S_2(n), \quad \text{say.}$$

It is easy to show that $S_2(n)$ for $n \leqslant x$ is usually negligible. Indeed,

$$\sum_{n\leqslant x}{}' S_2(n) = \sum_{d>\sqrt{y}} \sum_{d|p-1} \sum_{\substack{n\leqslant x \\ d|n-1 \\ p|n}}{}' \Lambda(d) \leqslant \sum_{d>\sqrt{y}} \sum_{\substack{p\leqslant x \\ d|p-1}} \frac{x\Lambda(d)}{dp}$$

$$= \sum_{\sqrt{y}<d\leqslant x} \left(\frac{xy\Lambda(d)}{d\phi(d)} + O\left(\frac{x\Lambda(d)\log d}{d\phi(d)} \right) \right) = O(x\sqrt{y}),$$

where we have used (3.3). It therefore follows that the number of $n \leqslant x$ with $S_2(n) > y^{3/4}$ is at most $O(x/y^{1/4})$. Thus these values of n are negligible and we need only consider those $n \leqslant x$ with

$$S_2(n) \leqslant y^{3/4}. \tag{4.3}$$

To estimate $S_1(n)$ we shall use Theorem (1.13) in Norton [18] which implies that we can essentially take each $\omega(n,d) = y/\phi(d)$. Specifically, this result implies that the number of $n \leqslant x$ for which

$$\left| \omega(n,d) - \frac{y}{\phi(d)} \right| \leqslant \frac{y^{4/5}}{\phi(d)} \tag{4.4}$$

fails for some $d \leqslant \sqrt{y}$ is $O(x/e^{y^{1/11}})$. Thus we need only consider those values of $n \leqslant x$, for which (4.4) holds for every $d \leqslant \sqrt{y}$.

Thus, but for $O(x/y^{1/4})$ choices of $n \leqslant x$, we have

$$\log F(n) = S_1(n) + O(y^{3/4})$$
$$= \sum_{\substack{d|n-1 \\ d\leqslant\sqrt{y}}} \left(\frac{y\Lambda(d)}{\phi(d)} + O\left(\frac{y^{4/5}\Lambda(d)}{\phi(d)} \right) \right) + O(y^{3/4}) \tag{4.5}$$
$$= yG_{\sqrt{y}}(n-1) + O(y^{4/5}\log y),$$

where

$$G_T(m) := \sum_{\substack{d|m \\ d\leqslant T}} \Lambda(d)/\phi(d).$$

We now show that $G_{\sqrt{y}}(m)$ is usually a good approximation to $G(m)$. From Norton [18, Theorem (1.5)], the number of $n \leqslant x$ with more than $2\log y$ prime power factors below $\log x$ is $O(x/y^{1/3})$ while the number of $n \leqslant x$ with more than $2y$ prime power factors is $O(x/(\log x)^{1/3})$. Thus we need only consider these values of $n \leqslant x$ where $n-1$ has at most $2\log y$ prime power factors in the interval $(\sqrt{y}, \log x)$ and at most $2y$ prime power factors in the interval $[\log x, x]$. Thus, but for $O(x/y^{1/3})$ values of $n \leqslant x$,

$$G(n-1) = G_{\sqrt{y}}(n-1) + O(\log^2 y/\sqrt{y}) + O(y\log y/\log x).$$

Combined with (4.5), we have

$$\log F(n) = yG(n-1) + O\left(y^{4/5}\log y\right) \tag{4.6}$$

but for $O(x/y^{1/4})$ choices of $n \leqslant x$.

Now note that

$$\log\log n = y + O(1/\log x)$$

for $\sqrt{x} < n \leqslant x$. Also note that

$$G(n-1) \ll \log\log n.$$

Thus from (4.6), we have

$$\log F(n) = G(n-1)\log\log n + O\left((\log\log n)^{5/6}\right)$$

for all but $O(x/y^{1/4})$ choices of $n \leqslant x$. Dividing by $\log\log n$, we have the theorem.

COROLLARY. *For each $u \geqslant 0$, let $D_F(u)$ denote the asymptotic density of the set $\{n: F(n) \leqslant (\log n)^u\}$. Then $D_F(u)$ exists, is continuous, is strictly increasing, satisfies $D_F(0) = 0$, $D_F(+\infty) = 1$, and is singular.*

Proof. By the Erdös-Wintner theorem (see [6, Chapter 5]), the distribution function for $G(n)$ satisfies all of these properties. By Theorem 4.1, $D_F(u)$ is equal to the distribution function for G.

5. Euler Pseudoprimes and Strong Pseudoprimes. From the Euler criterion, we know that if n is an odd prime and $(a, n) = 1$, then

$$a^{(n-1)/2} \equiv (a/n) \bmod n, \tag{5.1}$$

where (a/n) is the Jacobi symbol. If an odd, composite number n satisfies (5.1), we say that n is an *Euler pseudoprime* to the base a, or equivalently, that a is a *false Euler witness* for n. For each odd n, let

$$\mathbf{E}(n) = \left\{a \bmod n: a^{(n-1)/2} \equiv (a/n) \bmod n \text{ and } (a, n) = 1\right\}$$

and let $E(n) = \#\mathbf{E}(n)$. Thus $\mathbf{E}(n)$ is a subgroup of $\mathbf{F}(n)$. It has been independently shown by Selfridge (unpublished), Lehmer [16], and Solovay and Strassen [26], that if n is odd and composite, then $\mathbf{E}(n)$ is always a proper subgroup of $(Z/n)^*$ and so that $E(n) \leqslant \phi(n)/2$.

If n is an odd prime, $2^k \| n-1$, and $(a, n) = 1$, then either

$$a^{(n-1)/2^k} \equiv 1 \bmod n \quad \text{or} \quad a^{(n-1)/2^i} \equiv -1 \bmod n \quad \text{for some } i = 1, 2, \ldots, k. \tag{5.2}$$

If n is odd, composite, and (5.2) holds, then n is said to be a *strong pseudoprime* to the base a, or equally, a is a *false strong witness* for n. Let $\mathbf{S}(n)$ denote the set of $a \bmod n$ for which (5.2) holds and let $S(n) = \#\mathbf{S}(n)$. Note that $\mathbf{S}(n)$ need not be a subgroup of $(Z/n)^*$. For example,

$$\mathbf{S}(65) = \{1, 8, 18, 47, 57, 64\}$$

and $8 \cdot 18 \equiv 14 \bmod 65$. It has been shown independently by Pomerance, Selfridge and Wagstaff [24], Monier [17], and Atkin and Larson [1] that $\mathbf{S}(n) \subset \mathbf{E}(n)$ for all odd n. It is also known (Monier [17], Rabin [25]) that $S(n) \leqslant \phi(n)/4$ for every odd composite n.

From the above comments, we have

$$S(n) \leqslant E(n) \leqslant F(n) \tag{5.3}$$

for every odd n.

In this section we shall prove the following results. They are analogous to Theorems 2.1, 3.1, and the Corollary to Theorem 4.1.

THEOREM 5.1. *For all large x,*

$$\frac{1}{x} \sum_{\substack{n \leqslant x \\ n \text{ odd}}}{}' S(n) > x^{15/23}.$$

THEOREM 5.2. *For c_1, c_2 given by* (3.8) *and c_3 defined in* (5.11) *below, we have*

$$\left(\prod_{\substack{n<x \\ n \text{ odd}}} E(n) \right)^{2/x} = ec_2(\log x)^{c_1} + O\big((\log x)^{c_1 - 1/2}\big),$$

$$\left(\prod_{\substack{1<n\leqslant x \\ n \text{ odd}}} S(n) \right)^{2/x} = c_3(\log x)^{c_1 - (2\log 2)/3} + o(1).$$

THEOREM 5.3. *Let $D_{F,even}(u)$, $D_E(u)$, $D_S(u)$, respectively, denote the asymptotic densities of the sets*

$$\{n \text{ even}: F(n) \leqslant (\log n)^u\}, \quad \{n \text{ odd}: E(n) \leqslant (\log n)^u\},$$
$$\{n \text{ odd}: S(n) \leqslant (\log n)^u\}.$$

Then $D_E = \frac{1}{2} D_F$ and $D_S = D_{F,even}$. In particular, D_E and D_S are continuous strictly increasing, and singular on $[0, \infty)$. They are 0 at 0 and $1/2$ at $+\infty$.

To show these theorems it is first necessary to find formulas for $E(n)$ and $S(n)$ that are analogous to (1.5). This in fact has been done by Monier in [17]. To state Monier's formulas, we need some notation. Let $v_p(n)$ denote the exponent on p in the prime factorization of n. For n odd, $n > 1$, let

$$n' = \text{largest odd divisor of } n-1, \qquad \nu(n) = \min_{p|n} \{v_2(p-1)\},$$

$$\omega(n) = \sum_{p|n} 1, \quad e(n) = \prod_{p|n} \left(\frac{n-1}{2}, p-1 \right), \quad s(n) = \prod_{p|n} (n', p'),$$

$$\delta(n) = \begin{cases} 2, & \text{if } \nu(n) = v_2(n-1), \\ 1/2, & \text{if } \exists\, p|n \text{ with } v_2(p-1) < v_2(n-1) \text{ and } v_p(n) \text{ odd}, \\ 1, & \text{otherwise.} \end{cases}$$

Then we have

$$E(n) = \delta(n)e(n), \qquad S(n) = \left(1 + \frac{2^{\nu(n)\omega(n)} - 1}{2^{\omega(n)} - 1}\right) s(n). \tag{5.4}$$

Proof of Theorem 5.1. We follow the proof of Theorem 2.1, but make a few alterations. Let $M_0(x)$ denote the least common multiple of the odd integers up to x. For any y, let

$$\mathbf{P}_0(y, x) = \{p \leqslant y: p - 1 | 2M_0(x)\}.$$

Let

$$\beta_0 = \sup\{\alpha\colon \#\mathbf{P}_0(x^{\alpha'}, x) = x^{\alpha'+o(1)} \text{ for all } 0 < \alpha' < \alpha\}.$$

Then from simple modifications of the Balog [3] result, we have

$$\beta_0 > 23/8 \tag{5.5}$$

in analogy to (2.2).

Let α, ε be arbitrary with $1 < \alpha < \beta_0$, $0 < \varepsilon < \alpha - 1$. Let

$$M_0 = M_0(\log x/\log\log x),$$
$$\mathbf{P}_0 = \mathbf{P}_0(\log^{\alpha} x, \log x/\log\log x)\setminus\{p\colon p \leqslant \log^{\alpha-\varepsilon} x\},$$

and define $\mathbf{S}_0$ in analogy with $\mathbf{S}$, but using M_0, $\mathbf{P}_0$ for M, $\mathbf{P}$. For large x, the numbers in $\mathbf{S}_0$ are all odd since they are free of prime factors $p \leqslant \log^{\alpha-\varepsilon} x$. For $s \in \mathbf{S}_0$, instead of (2.6) we choose $q_0 = q_0(s)$ the least prime such that

$$sq_0 \equiv 1 \bmod M_0, \qquad sq_0 \equiv 3 \bmod 4.$$

If n is such a number sq_0, then

$$\begin{aligned} S(n) > s(n) &= \prod_{p\mid n}(p', n') \geqslant \prod_{p\mid s}\left(\frac{p-1}{2}, \frac{n-1}{2}\right) \\ &= 2^{-\omega(s)}\prod_{p\mid s}(p-1, n-1) \geqslant 2^{-\omega(s)}\prod_{p\mid s}(p-1, 2M_0) \\ &= 2^{-\omega(s)}\phi(s) \gg 2^{-\omega(s)}x^{1-\varepsilon}/\log\log x = x^{1-\varepsilon+o(1)}, \end{aligned}$$

where we use (2.8) and the fact that $\omega(s) \leqslant (1 + o(1))\log s/\log\log s$.

Therefore, as in the final calculations in the proof of Theorem 2.1, we have

$$\frac{1}{x}\sum_{\substack{n\leqslant x\\ n \text{ odd}}}{}' S(n) \geqslant x^{1-\beta_0^{-1}+o(1)}.$$

The estimate (5.5) thus completes the proof.

Remarks. In view of (5.3), we have the same result as Theorem 5.1 with $E(n)$ replacing $S(n)$. Also, (5.3) implies we have Theorem 2.2 with either $E(n)$ or $S(n)$ replacing $F(n)$ (and n restricted to odd numbers).

Proof of Theorem 5.2. We begin by computing the geometric mean of $e(n)$ for n odd, $n \leqslant x$. As in (3.1), we have

$$\begin{aligned} \sum_{\substack{n\leqslant x\\ n \text{ odd}}} \log e(n) &= \sum_{d\leqslant x}\Lambda(d)\sum_{\substack{p\leqslant x\\ d\mid p-1}}\ \sum_{\substack{n\leqslant x\\ 2d\mid n-1\\ p\mid n}} 1 \\ &= \sum_{p\leqslant x}\sum_{2d\mid p-1}\Lambda(d)\left(\left[\frac{x-p}{2dp}\right]+1\right) \\ &\quad + \sum_{p\leqslant x}\sum_{\substack{d\mid p-1\\ 2d\nmid p-1\\ d\leqslant x/p-1}}\Lambda(d)\left(\left[\frac{x-p-dp}{2dp}\right]+1\right) \\ &= \sum_{p\leqslant x}\sum_{d\mid p-1}\Lambda(d)\left(\left[\frac{x-p}{2dp}\right]+1\right) + O\left(\frac{x}{\log x}\right), \end{aligned}$$

since if $d|p-1$ and $2d \nmid p-1$, then $\Lambda(d) > 0$ only when d is a power of 2. Using (3.2) and the arguments in (3.4) and (3.7) we thus have

$$\sum_{\substack{n \leqslant x \\ n \text{ odd}}} \log e(n) = x \log\log x \sum_d \frac{\Lambda(d)}{2d\phi(d)} + x\left(1 + \sum_d \frac{\Lambda(d)C(d)}{2d}\right) + O\left(\frac{x}{\log x}\right). \tag{5.6}$$

We now compute the geometric mean of $\delta(n)$ for $n \leqslant x$, n odd. If $n \equiv 3 \bmod 4$, then $\nu(n) = v_2(n-1) = 1$, so that $\delta(n) = 2$. If $n \equiv 1 \bmod 4$ then of course $v_2(n-1) \geqslant 2$. But, the number of $n \leqslant x$ not divisible by any prime $p \equiv 3 \bmod 4$ to an odd exponent is $O(x/\sqrt{\log x})$ (see, for example, Halberstam and Richert [11, Theorem 2.3]). Thus, but for $O(x/\sqrt{\log x})$ choices of $n \leqslant x$ with $n \equiv 1 \bmod 4$, we have $\delta(n) = 1/2$. Therefore,

$$\sum_{\substack{n \leqslant x \\ n \text{ odd}}} \log \delta(n) = O\left(\frac{x}{\sqrt{\log x}}\right).$$

Thus, with (5.6) and (3.8), we have

$$\frac{2}{x} \sum_{\substack{n \leqslant x \\ n \text{ odd}}} \log E(n) = c_1 \log\log x + 1 + \log c_2 + O\left(\frac{x}{\sqrt{\log x}}\right),$$

from which part of our theorem follows.

Again following (3.1), we have

$$\sum_{\substack{n \leqslant x \\ n \text{ odd}}} \log s(n) = \sum_{\substack{d \leqslant x \\ d \text{ odd}}} \Lambda(d) \sum_{\substack{p \leqslant x \\ d|p-1}} \sum_{\substack{n \leqslant x \\ n \text{ odd} \\ d|n-1,\, p|n}} 1$$

$$= \sum_{\substack{d \leqslant x \\ d \text{ odd}}} \Lambda(d) \sum_{\substack{p \leqslant x \\ d|p-1}} \left(\left[\frac{x-p}{2dp}\right] + 1\right).$$

Thus from the calculations in (3.2), (3.4), and (3.7), we have

$$\sum_{\substack{n \leqslant x \\ n \text{ odd}}} \log s(n) = x \log\log x \sum_{d \text{ odd}} \frac{\Lambda(d)}{2d\phi(d)} + x\left(1 + \sum_{d \text{ odd}} \frac{\Lambda(d)C(d)}{2d}\right) + O\left(\frac{x}{\log x}\right). \tag{5.7}$$

It thus remains to compute the geometric mean of the numbers

$$1 + (2^{\nu(n)\omega(n)} - 1)/(2^{\omega(n)} - 1) \quad \text{for } 1 < n \leqslant x,\ n \text{ odd}.$$

The main contribution comes from those n with $\nu(n) = 1$. We have (for notational convenience, define $\nu(n) = 0$ for $n = 1$, n even)

$$\sum_{\substack{n \leqslant x \\ \nu(n)=1}} \log\left(1 + \frac{2^{\nu(n)\omega(n)} - 1}{2^{\omega(n)} - 1}\right) = \sum_{\substack{n \leqslant x \\ \nu(n)=1}} \log 2 = x\frac{\log 2}{2} + O\left(\frac{x}{\sqrt{\log x}}\right). \tag{5.8}$$

For $\nu \geqslant 2$, we have uniformly

$$\sum_{\substack{n\leqslant x\\ \nu(n)=\nu}} \log\left(1+\frac{2^{\nu(n)\omega(n)}-1}{2^{\omega(n)}-1}\right) < \nu \sum_{\substack{n\leqslant x\\ \nu(n)=\nu}} \omega(n) \leqslant \nu \sum_{\substack{p,m\\ pm\leqslant x\\ \nu(pm)=\nu}} 1$$

(5.9)

$$\leqslant \nu \sum_{\substack{p\leqslant\sqrt{x}\\ 2^\nu\mid p-1}} \sum_{\substack{m\leqslant x/p\\ \nu(m)\geqslant 2}} 1 + \nu \sum_{\substack{m\leqslant\sqrt{x}\\ \nu(m)\geqslant 2}} \sum_{\substack{p\leqslant x/m\\ 2^\nu\mid p-1}} 1 = S_1(\nu)+S_2(\nu), \quad \text{say.}$$

From (3.3) we have

$$S_1(\nu) \ll \frac{\nu x}{\sqrt{\log x}} \sum_{\substack{p\leqslant\sqrt{x}\\ 2^\nu\mid p-1}} \frac{1}{p} \ll \frac{\nu x}{\sqrt{\log x}}\left(\frac{\log\log x}{2^\nu}+\frac{\nu}{2^\nu}\right). \tag{5.10}$$

To estimate $S_2(\nu)$, we distinguish two cases:

$$2 \leqslant \nu \leqslant \log x/\log\log x \quad \text{and} \quad \nu > \log x/\log\log x.$$

In the first case, the Brun-Titchmarsh inequality implies $S_2(\nu)$ is at most

$$\ll \frac{\nu x}{2^\nu \log x} \sum_{\substack{m\leqslant\sqrt{x}\\ \nu(m)\geqslant 2}} \frac{1}{m} \ll \frac{\nu x}{2^\nu\sqrt{\log x}}.$$

In the second case, a trivial estimate implies $S_2(\nu)$ is at most

$$\ll \frac{\nu x}{2^\nu} \sum_{\substack{m\leqslant\sqrt{x}\\ \nu(m)\geqslant 2}} \frac{1}{m} \ll \frac{\nu x\sqrt{\log x}}{2^\nu}.$$

Therefore, with these two estimates and (5.9), (5.10) we have

$$\sum_{\substack{n\leqslant x\\ \nu(n)\geqslant 2}} \log\left(1+\frac{2^{\nu(n)\omega(n)}-1}{2^{\omega(n)}-1}\right) \leqslant \sum_{\nu\geqslant 2}(S_1(\nu)+S_2(\nu)) \ll \frac{x\log\log x}{\sqrt{\log x}}.$$

Combining this with (5.7), (5.8), and (3.8) we have

$$\frac{2}{x}\sum_{\substack{1<n\leqslant x\\ n\text{ odd}}} \log S(n) = \left(c_1-\frac{2\log 2}{3}\right)\log\log x + \log c_2 + 1 + \log 2$$

$$-\log 2\sum_{i\geqslant 1}\frac{C(2^i)}{2^i} + O\left(\frac{\log\log x}{\sqrt{\log x}}\right).$$

We have thus proved the theorem with

$$c_3 = 2ec_2/2^{\Sigma C(2^i)/2^i}. \tag{5.11}$$

Remarks. It might seem paradoxical at first, when comparing Theorems 3.1 and 5.2, that the geometric mean order of $E(n)$ is larger than that of $F(n)$. However, it is more appropriate to compare Theorem 5.2 with the geometric mean order of $F(n)$ for odd numbers n. By the above methods, this is easily computed. We have

$$\left(\prod_{\substack{n\leqslant x\\ n\text{ odd}}} F(n)\right)^{2/x} = ec_2 2^{\Sigma C(2^i)/2^i}(\log x)^{c_1+(2\log 2)/3}\left(1+O\left(\frac{1}{\log x}\right)\right). \tag{5.12}$$

We also note that if we compute the geometric mean orders of $E(n)$, $S(n)$, $F(n)$ for odd, composite values of n, the expressions in Theorem 5.2 and (5.12) are divided by e^2 and in (5.12) there is an extra factor of $\log\log x$ in the error term.

Proof of Theorem 5.3. For all odd n except for a set of density 0,

$$1 + \frac{2^{\nu(n)\omega(n)} - 1}{2^{\omega(n)} - 1} = 2 = O(1).$$

Thus we may replace $E(n)$, $S(n)$ in the statement of the theorem with $e(n)$, $s(n)$, respectively. From the proof of Theorem 4.1, we have

$$\frac{\log e(n)}{\log\log n} - G\left(\frac{n-1}{2}\right) \to 0, \qquad \frac{\log s(n)}{\log\log n} - G(n') \to 0$$

on sets of odd numbers of asymptotic density $1/2$. Thus, by the Corollary to Theorem 4.1,

$$\begin{aligned} D_E(u) &= \lim_{x\to\infty} \frac{1}{x} \cdot \#\left\{m \leqslant x\colon m \text{ even}, G\left(\frac{m}{2}\right) \leqslant u\right\} \\ &= \lim_{x\to\infty} \frac{1}{x} \cdot \#\left\{k \leqslant \frac{x}{2}\colon G(k) \leqslant u\right\} = \frac{1}{2} D_F(u). \end{aligned}$$

Also, noting that the mapping $n \to n'$ gives a one-to-one correspondence from the odd numbers in $(x+1, 2x+1]$ to the odd numbers in $(0, x]$, we have

$$\begin{aligned} &\lim_{x\to\infty} \frac{1}{x} \cdot \#\left\{x < n \leqslant 2x\colon n \text{ odd}, \frac{\log s(n)}{\log\log n} \leqslant u\right\} \\ &= \lim_{x\to\infty} \frac{1}{x} \cdot \#\left\{x < n \leqslant 2x\colon n \text{ odd}, G(n') \leqslant u\right\} \\ &= \lim_{x\to\infty} \frac{1}{x} \cdot \#\left\{m \leqslant x\colon m \text{ odd}, G(m) \leqslant u\right\} \\ &= \lim_{x\to\infty} \frac{1}{x} \cdot \#\left\{k \leqslant x\colon k \text{ even}, \frac{\log F(k)}{\log\log k} \leqslant u\right\} = D_{F,\text{even}}(u). \end{aligned}$$

Therefore, $D_S = D_{F,\text{even}}$, which completes the proof of Theorem 5.3.

6. Further Problems. In this section we discuss some additional conjectures and results. Some of the proofs are omitted.

Maximal Order of $F(n)$ for n Composite. By definition, if n is a Carmichael number, then $F(n) = \phi(n)$. We conjecture that not only are there infinitely many Carmichael numbers, but that

$$\limsup_{n \text{ composite}} F(n)/n = 1.$$

From (2.8), we have $F(n)/n^{1-\varepsilon}$ unbounded on the composites for any $\varepsilon > 0$. However, we can do better.

THEOREM 6.1. $\limsup_{n\ composite} F(n)\log^2 n/n > 0$.

Proof. From Lemma 8 in Bateman, Pomerance and Vaughan [4], there is a positive constant c such that for infinitely many integers k, there are primes $p \equiv q \equiv 1 \bmod k$ with $p < q < ck\log k$. When such p, q exist for k, let $n_k = pq$.

Then

$$F(n_k) > k^2, \qquad n_k < c^2 k^2 \log^2 k,$$

so that $F(n_k) \log^2 n_k / n_k > 1/c^2$.

Small Values of $F(n)$.

THEOREM 6.2. *The number of* $n \leqslant x$ *with* $F(n) = 1$ *is*

$$\frac{(1 + o(1))x}{e^{\gamma} \log\log\log x}, \tag{6.1}$$

where γ *denotes Euler's constant.*

This follows from almost the same argument as in Erdös [9] that the number of $n \leqslant x$ with $(n, \phi(n)) = 1$ is given by (6.1).

In studying the equation $F(n) = k$ for a fixed k, there is a dramatic contrast between $k = 1$ and $k > 1$. Let $p(k)$ denote the least prime factor of k.

THEOREM 6.3. *For each* $k > 1$, *the number of* $n \leqslant x$ *with* $F(n) = k$ *is at most* $x/(\log x)^{(p(k)-1)^{-1}+o_k(1)}$.

To see this, note that if $F(n) = k$, then $n \equiv 1 \bmod p$, where $p = p(k)$. So, using the notation from Section 4,

$$p^{\omega(n,p)} | F(n) = k.$$

But the number of $n \leqslant x$ with $\omega(n, p)$ bounded is $x/(\log x)^{(p-1)^{-1}+o(1)}$.

Using a result of Halász (Theorem 21.5 in Elliott [6]), it is not so hard to prove the following theorem, which is uniform in k.

THEOREM 6.4. *Uniformly for* $x \geqslant 20$ *and* $k \geqslant 2$,

$$\#\{n \leqslant x: F(n) = k\} = O\left(\frac{x\,(\log\log\log x)^{2/3}}{(\log\log x)^{1/3}}\right).$$

Combining this theorem with Theorem 6.2, we see that for large x, 1 is the most popular value of $F(n)$ for $n \leqslant x$. Probably, this is true for all $x \geqslant 1$.

THEOREM 6.5. *For* $0 < \varepsilon \leqslant 1$, *the number of* $n \leqslant x$ *with* $F(n) < \exp\{(\log\log x)^{1-\varepsilon}\}$ *is* $O(x/(\varepsilon \log\log\log x))$.

From (4.4), we may assume that $\omega(n, q) > 2(\log\log x)^{1-\varepsilon}$ for every prime $q \leqslant (\log\log x)^{\varepsilon/2}$. We also may assume that $n - 1$ is divisible by some such prime q_0. Then

$$F(n) \geqslant q_0^{\omega(n,q_0)} > \exp\left\{(\log\log x)^{1-\varepsilon}\right\},$$

which proves the theorem.

The Range of $F(n)$.

THEOREM 6.6. *If* k *is odd or* $4|k$, *then there are infinitely many* n *with* $F(n) = k$. *If* $k \equiv 2 \bmod 4$, *then* $F(n) = k$ *has infinitely many solutions* n *or no solutions* n *depending on whether* $k = p - 1$ *for some prime* p. *In particular, the density of the range of* F *is* $3/4$.

Proof. Assume $k > 0$ is odd. Let p be any odd prime with

$$p \equiv \frac{-k+1}{2} \bmod \frac{k(k+1)}{2}.$$

Once p is chosen, let q be any prime that simultaneously satisfies

$$q \equiv 1 \bmod \frac{p-1}{2}, \qquad q \equiv k+1 \bmod (2p-1).$$

Then

$$(p-1, 2pq-1) = 1, \qquad (q-1, 2pq-1) = k,$$

so that $F(2pq) = k$.

Suppose $k \equiv 4 \bmod 8$. Let p denote any prime with

$$p \equiv \left(\frac{k}{2}-2\right)\left(\frac{k}{2}-1\right) - 1 \bmod \left(\left(\frac{k}{2}-2\right)\left(\frac{k}{2}-1\right)k\right)$$

and let q denote any prime with

$$q \equiv \frac{k}{2} - 1 \bmod (p-1).$$

Then

$$(q-1, pq^2-1) = 2, \qquad (p-1, pq^2-1) = k/2,$$

so that $F(pq^2) = k$.

Assume $8 \mid k$. Let p be any prime with

$$p \equiv -\frac{k}{2} - 1 \bmod \left(k\left(\frac{k}{4}+1\right)\right)$$

and let q be any prime with

$$q \equiv -1 \bmod kp(p-1).$$

Now choose r as any prime with

$$qr \equiv 2p-3 \bmod (p-1), \quad pr \equiv 2q-3 \bmod (q-1),$$
$$r \equiv \frac{k}{4} + 1 \bmod (pq-1)$$

all holding. We have

$$(p-1, pqr-1) = 2, \quad (q-1, pqr-1) = 2, \quad (r-1, pqr-1) = k/4,$$

so that $F(pqr) = k$.

Suppose finally that $k \equiv 2 \bmod 4$ and $F(n) = k$ has a solution n. Then n is odd, so that $2^{\omega(n)} \mid F(n)$, where $\omega(n)$ denotes the number of distinct prime factors of n. But $4 \nmid F(n)$, so that n is a prime power p^a. Note that $F(p^a) = p-1$ for every exponent $a \geqslant 1$. Thus $k = p-1$ for some prime p and $F(n) = k$ has infinitely many solutions.

The Normal Number of Prime Factors of $F(n)$. Let $\omega(n)$ denote the number of distinct prime factors of n and let $\Omega(n)$ denote the number of prime factors of n counted with multiplicity.

THEOREM 6.7. *The normal value of $\omega(F(n))$ is* $\log\log\log\log n$.

THEOREM 6.8. *Let* $h(n)$ *denote the additive function* $\sum_{p^a|n} 1/\phi(p^a)$. *Then the distribution function for* $h(n)$ *is identical to the distribution function of*

$$\Omega(F(n))/\log\log n.$$

COROLLARY. *There is a set of integers* S *with asymptotic density* 1 *such that* $\{F(n): n \in S\}$ *has asymptotic density* 0.

The Universal Exponent for the Group $\mathbf{F}(n)$. Let $L(n)$ denote the universal exponent for the group $\mathbf{F}(n)$. It is not so hard to see that

$$L(n) = \operatorname{lcm}\{(p-1, n-1): p|n\}$$

and that $L(n) = \lambda(n)$ (where $\lambda(n)$ is the universal exponent for the group of reduced residues mod n) if and only if $F(n) = \phi(n)$. The normal value of $L(n)$ is considered in the following two results.

THEOREM 6.9. *But for a set of* n *of density* 0, $L(n)$ *is the largest divisor of* $n-1$ *composed solely of primes below* $\log\log n$.

COROLLARY. *For every real* $u \geqslant 0$, *the asymptotic density of the* n *for which* $L(n) \leqslant (\log\log n)^u$ *is*

$$e^{-\gamma}\int_0^u \rho(t)\,dt,$$

where γ *is Euler's constant and* ρ *is the Dickman-deBruijn function.*

The function ρ is discussed in deBruijn [5]. We remark that it is known that $\int_0^\infty \rho(t)\,dt = e^\gamma$. This result is (9.1) in Knuth and Trabb Pardo [15].

Equal Values of $F(n)$ *for Neighboring Values of* n. It is easy to see that $F(n) = F(n+1)$ has only the single solution $n = 1$. This is because for $n > 1$, $F(n) \equiv n - 1 \bmod 2$. It is only slightly harder to show that $F(n) = F(n+2)$ has infinitely many solutions. Indeed, by the method of Theorem 6.2, the number of $n \leqslant x$ with $F(n) = F(n+2) = 1$ is $\gg x/(\log\log\log x)^2$. These values of n of course must be even. Probably there are infinitely many odd values of n with $F(n) = F(n+2)$, but this seems much harder.

We can prove that infinitely often there are five consecutive even numbers with F-value 1. To see this, consider numbers n with $n \equiv 16 \bmod 30$ such that n is not divisible by any prime $p \equiv 1 \bmod 3$ or $p \equiv 1 \bmod 5$ and such that $n-5$, $n-3$, $(n-1)/15$, $n+1$, and $n+3$ are divisible by no prime less than $\log n$. The number of such $n \leqslant x$ is $\gg x/(\log x)^{5/8}\,(\log\log x)^5$, and the number of such $n \leqslant x$ for which

$$F(n-4) = F(n-2) = F(n) = F(n+2) = F(n+4) = 1$$

fails is $O(x/\log x)$. This method runs into difficulties when six or more consecutive even numbers are considered, but perhaps the methods of C. Hooley [13] and K.-H. Indlekofer [14] would be of use. Thanks are due to Helmut Maier for the suggestion.

In any event, there do not exist 21 consecutive even numbers with F-value 1. Indeed, one of these numbers n is 7 mod 21, so $3|F(n)$. We conjecture that infinitely often there are 20 consecutive even numbers with F-value 1 and that for every k there are infinitely many n with $F(n) = F(n+2) = \cdots = F(n+2k)$.

Acknowledgment. The second author wishes to acknowledge the hospitality of the Discrete Mathematics Department at Bell Communications Research where much of the work for this paper was done.

Added in proof. E. Fouvry informs us that he and B. Rousselet have recently shown that the constant C discussed in the remark following the proof of Theorem 2.1 is less than 1.46. In addition, they have improved the theorem of Balog [3] which relates C to the constant β. In particular, they improve (2.2) to $\beta > 25/8$. Thus the exponent "15/23" appearing in the statements of Theorem 2.1 and Theorem 5.1 can now be replaced with "17/25". Also the exponents may be improved in the final remark of Section 2 and in the main results of the papers [20] and [23].

Mathematical Institute of the Hungarian
Academy of Sciences
Budapest, Hungary

Department of Mathematics
University of Georgia
Athens, Georgia 30602

1. A. O. L. ATKIN & R. G. LARSON, "On a primality test of Solovay and Strassen," *SIAM J. Comput.*, v. 11, 1982, pp. 789–791.

2. R. BAILLIE & S. S. WAGSTAFF, JR., "Lucas pseudoprimes," *Math. Comp.*, v. 35, 1980, pp. 1391–1417.

3. A. BALOG, "$p + a$ without large prime factors," *Séminaire de Théorie des Nombres de Bordeaux* (1983)–(1984), no. 31.

4. P. T. BATEMAN, C. POMERANCE & R. C. VAUGHAN, "On the size of the coefficients of the cyclotomic polynomial" (*Topics in Classical Number Theory*, Colloq. Math. Soc. János Bolyai, no. 34, 1981), North-Holland, Amsterdam, 1984, pp. 171–202.

5. N. G. DEBRUIJN, "On the number of positive integers $\leqslant x$ and free of prime factors $> y$," *Nederl. Akad. Wetensch. Proc. Ser. A*, v. 54, 1951, pp. 50–60.

6. P. D. T. A. ELLIOTT, *Probabilistic Number Theory*. I, II, Springer-Verlag, New York, 1980.

7. P. ERDÖS, "On the normal number of prime factors of $p - 1$ and some related questions concerning Euler's ϕ-function," *Quart. J. Math. Oxford Ser.*, v. 6, 1935, pp. 205–213.

8. P. ERDÖS, "On pseudoprimes and Carmichael numbers," *Publ. Math. Debrecen*, v. 4, 1955–1956, pp. 201–206.

9. P. ERDÖS, "Some asymptotic formulas in number theory," *J. Indian Math. Soc. (N.S.)*, v. 12, 1948, pp. 75–78.

10. E. FOUVRY, "Théorème de Brun-Titchmarsh; application au théorème de Fermat," *Invent. Math.*, v. 79, 1985, pp. 383–407.

11. H. HALBERSTAM & H. E. RICHERT, *Sieve Methods*, Academic Press, London, 1974.

12. G. H. HARDY & E. M. WRIGHT, *Introduction to the Theory of Numbers*, 4th ed., Oxford Univ. Press, London, 1965.

13. C. HOOLEY, "On the intervals between numbers that are sums of two squares. III," *J. Reine Angew. Math.*, v. 267, 1974, pp. 207–218.

14. K.-H. INDLEKOFER, "Scharfe untere Abschätzung für die Anzahlfunktion der B-Zwillinge," *Acta Arith.*, v. 26, 1974/75, pp. 207–212.

15. D. E. KNUTH & L. TRABB PARDO, "Analysis of a simple factorization algorithm," *Theoret. Comput. Sci.*, v. 3, 1976, pp. 321–348.

16. D. H. LEHMER, "Strong Carmichael numbers," *J. Austral. Math. Soc. Ser. A*, v. 21, 1976, pp. 508–510.

17. L. MONIER, "Evaluation and comparison of two efficient probabilistic primality testing algorithms," *Theoret. Comput. Sci.*, v. 12, 1980, pp. 97–108.

18. K. K. NORTON, "On the number of restricted prime factors of an integer. I," *Illinois J. Math.*, v. 20, 1976, pp. 681–705.

19. C. POMERANCE, "On the distribution of amicable numbers," *J. Reine Angew. Math.*, v. 293/294, 1977, pp. 217–222.

20. C. POMERANCE, "Popular values of Euler's function," *Mathematika*, v. 27, 1980, pp. 84–89.

21. C. POMERANCE, "Recent developments in primality testing," *Math. Intelligencer*, v. 3, 1981, pp. 97–105.

22. C. POMERANCE, "On the distribution of pseudoprimes," *Math. Comp.*, v. 37, 1981, pp. 587–593.

23. C. POMERANCE, "A new lower bound for the pseudoprime counting function," *Illinois J. Math.*, v. 26, 1982, pp. 4–9.

24. C. POMERANCE, J. L. SELFRIDGE & S. S. WAGSTAFF, JR., "The pseudoprimes to $25 \cdot 10^9$," *Math. Comp.*, v. 35, 1980, pp. 1003–1026.

25. M. O. RABIN, "Probabilistic algorithm for testing primality," *J. Number Theory*, v. 12, 1980, pp. 128–138.

26. R. SOLOVAY & V. STRASSEN, "A fast Monte-Carlo test for primality," *SIAM J. Comput.*, v. 6, 1977, pp. 84–85; erratum, *ibid.*, v. 7, 1978, p. 118.