ILLINOIS JOURNAL OF MATHEMATICS
Volume 26, Number 1, Spring 1982

A NEW LOWER BOUND FOR THE PSEUDOPRIME COUNTING FUNCTION

BY

CARL POMERANCE

1. Introduction

A composite natural number n is called a *pseudoprime* (to base 2) if

$$2^{n-1} \equiv 1 \pmod{n}.$$

The least pseudoprime is $341 = 11 \cdot 31$. Let $\mathscr{P}(x)$ denote the number of pseudoprimes not exceeding x. It is known that there are positive constants c_1, c_2 such that for all large x,

$$c_1 \log x \le \mathscr{P}(x) \le x \cdot \exp\{-c_2(\log x \cdot \log\log x)^{1/2}\}.$$

The lower bound is implicit in Lehmer [6] and the upper bound is due to Erdös [4]. Very recently in [9] we have obtained an improvement in the upper bound. There have been improvements on the lower bound, but they have only concerned the size of the constant c_1. For example, see Rotkiewicz [13].

In this paper we show that there is a positive constant α such that for all large x,

$$\mathscr{P}(x) \ge \exp\{(\log x^{\alpha})\}.$$

In particular, we may take $\alpha = 5/14$.

Erdös conjectures that $\mathscr{P}(x) = x^{1-\varepsilon(x)}$ where $\varepsilon(x) \to 0$ as $x \to \infty$. See Pomerance, Selfridge, Wagstaff [10] for more on this.

Our main result holds for pseudoprimes to any base and in fact for strong pseudoprimes to any base (see Section 2 for definitions). Moreover our result holds if we just count those pseudoprimes n with at least $(\log n)^{5/14}$ distinct prime factors.

On the negative side, if $\mathscr{P}'(x)$, $\mathscr{P}''(x)$, and $\mathscr{P}^k(x)$ denote respectively the counting functions for pseudoprimes that are square-free, not square-free, and have at most k distinct prime factors, then we cannot show any one of $\mathscr{P}'(x)/\log x$, $\mathscr{P}''(x)$, $\mathscr{P}^k(x)/\log x$ is unbounded.

We wish to thank H. W. Lenstra, Jr. and S. S. Wagstaff, Jr. for some helpful comments during early stages of this paper.

Received January 11, 1980.

2. Preliminaries

If b, n are natural numbers and $(b, n) = 1$, let $l_b(n)$ denote the exponent to which b belongs modulo n. Let $\lambda(n)$ denote the largest of all the $l_b(n)$ where b varies over a reduced residue system modulo n. We always have $l_b(n) \mid \lambda(n)$. From the theorem on the primitive root we have, for prime powers p^a,

$$\lambda(p^a) = \begin{cases} p^{a-1}(p-1) & \text{if } p > 2 \text{ or if } a \leq 2, \\ 2^{a-2} & \text{if } p = 2 \text{ and } a \geq 3. \end{cases}$$

For a general n we have $\lambda(n)$ equal to the least common multiple of the $\lambda(p^a)$ for the $p^a \| n$.

A composite natural number n is called a *pseudoprime to base* b if

$$b^{n-1} \equiv 1 \pmod{n}.$$

If n is an odd pseudoprime to base b and if there is an integer $k \geq 0$ such that $2^k \| l_b(p)$ for each prime factor p of n, then n is called a *strong pseudoprime to base* b. This slightly unorthodox definition is easily seen to be equivalent to the usual definition of strong pseudoprime (see [10], for example).

If $m \geq 1$, $b \geq 2$ are integers, we let $F_m(b)$ denote the mth cyclotomic polynomial evaluated at b. We have $F_m(b) \geq 1$. If $F_m(b)$ is divisible by a prime p with $l_b(p) \neq m$, then $m = p^k l_b(p)$ for some integer $k > 0$. In this case, p is called an *intrinsic* prime factor, and is evidently unique. The common case for prime factors q of $F_m(b)$ is for $l_b(q) = m$. Such prime factors q are called *non-intrinsic* or *primitive*. Moreover $F_m(b)$ has at least one primitive prime factor except in the cases $m = 1, b = 2$; $m = 2, b = 2^n - 1$ for some integer $n \geq 2$; $m = 6, b = 2$. This result is due to Bang [2] and many others. (Artin [1] is a more accessible reference on this topic.) Thus if $m = pc$ where p is prime and larger than the largest prime factor of c and if $c \neq l_b(p)$, then every prime factor of $F_{pc}(b)$ is primitive and $F_{pc}(b) > 1$.

If $\mathscr{S}$ is a set, by $\#\mathscr{S}$ we mean the cardinality of $\mathscr{S}$.

3. The constant E

If $n \geq 2$ is an integer, let $P(n)$ denote the largest prime factor of n. Let $\Pi(x, y)$ denote the number of primes $p \leq x$ such that $P(p-1) \leq y$. Let

$$E = \sup\{c\colon \Pi(x, x^{1-c}) \gg x/\log x\}.$$

Erdös [3] showed that $E > 0$. In [8] we showed that $E > 0.55092$. Furthermore we indicated that a new result of Iwaniec [5] and our method give $E > 0.55655$. Erdös [4] conjectured that $E = 1$. We remark that $E = 1$ follows from the method of [8] and the conjecture of Halberstam (see Montgomery [7], equation 15.10) that Bombieri's theorem holds for moduli up to $x^{1-\varepsilon}$ rather than just up to $x^{1/2-\varepsilon}$.

The interest in the constant E comes from the following result which is a variation on a theme of Erdös (see [3]).

THEOREM 1. *For every $\varepsilon > 0$, there is an $x_0(\varepsilon)$ such that for each $x \geq x_0(\varepsilon)$, if A is the least common multiple of the integers up to* $\log x/\log\log x$, *then*

$$\#\{a \leq x\colon \lambda(a) \mid A,\ a \text{ square-free}\} \geq x^{E-\varepsilon}.$$

Proof. We may assume $E > \varepsilon > 0$. Let $z = (\log x)^{(1-E+\varepsilon/2)^{-1}}$. Let

$$\mathscr{A} = \{p \leq z\colon p \text{ prime},\ p-1 \mid A\}.$$

From the definition of E, there is a $\delta > 0$ such that for all large x,

$$\Pi(z, \log x/\log\log x) \geq \delta z/\log z.$$

If p is a prime with the properties $p \leq z$, $P(p-1) \leq \log x/\log\log x$, and yet $p \notin \mathscr{A}$, then it must be that there is a prime power $q^c \mid p-1$ with $c \geq 2$ and $q^c > \log x/\log\log x$. Now the number of such primes p is at most

$$\sum [z/q^c] \ll z(\log\log x/\log x)^{1/2} = o(z/\log z).$$

Thus for all large x we have

$$\#\mathscr{A} \geq (\delta/2)z/\log z.$$

Now let $\mathscr{N}$ denote the set of square-free integers $a \leq x$ composed only of the primes in $\mathscr{A}$. Every member p of $\mathscr{A}$ satisfies $p \leq z$, so that $\mathscr{N}$ has at least as many elements as $\mathscr{A}$ has subsets of cardinality $[\log x/\log z]$. Thus, for large x,

$$\begin{aligned}\#\mathscr{N} &\geq \binom{\#\mathscr{A}}{[\log x/\log z]} \geq \left(\frac{\#\mathscr{A}}{[\log x/\log z]}\right)^{[\log x/\log z]} \\ &\geq \frac{1}{z}\left(\frac{(\delta/2)z/\log z}{\log x/\log z}\right)^{\log x/\log z} \\ &= \frac{1}{z}\left(\frac{\delta}{2}\right)^{\log x/\log z} \cdot x^{E-\varepsilon/2} \geq x^{E-\varepsilon}.\end{aligned}$$

But if $a \in \mathscr{N}$, then $a \leq x$, a is square-free, and $\lambda(a) \mid A$.

4. The main result

Let $\mathscr{P}_b(x)$ denote the number of pseudoprimes to base b that do not exceed x.

THEOREM 2. *For every $\varepsilon > 0$ and integer $b \geq 2$, there is an $x_0(\varepsilon, b)$ such that for all $x \geq x_0(\varepsilon, b)$, we have*

$$\mathscr{P}_b(x) \geq \exp\{(\log x)^{E/(E+1)-\varepsilon}\}.$$

Proof. Let $\varepsilon > 0$, $b \geq 2$ be given. Let x be large and let $y = (\log x)^{(E+1)^{-1}}$. Let A denote the least common multiple of the integers up to $\log y/\log\log y$. Let p denote the first prime that is congruent to 1 modulo 2A. By Linnik's theorem (see Prachar [11], Kapitel X, Satz 4.1) there is an absolute constant c with

$$(1) \qquad p \le A^c \le y^{2c/\log\log y}.$$

Let q be any fixed prime between $A + 1$ and $2A$. Let

$$\mathcal{N} = \{a \le y\colon \lambda(a) \mid A,\ a \text{ square-free},\ a \ne l_b(q),\ aq \ne l_b(p)\}.$$

The last two conditions delete at most 2 elements that otherwise would be in $\mathcal{N}$. By Theorem 1 and possibly deleting some elements of $\mathcal{N}$, we may assume $\#\mathcal{N} = [y^{E-\varepsilon}]$.

For each set $\mathcal{S} \subset \mathcal{N}$ with at least 2 elements, let

$$n(\mathcal{S}) = \prod_{a \in \mathcal{S}} F_{pqa}(b).$$

We claim that

(i) $n(\mathcal{S})$ is a pseudoprime to base b,
(ii) $n(\mathcal{S}) \le x$, and
(iii) if $\mathcal{S}' \subset \mathcal{N}$, $\#\mathcal{S}' \ge 2$, $\mathcal{S}' \ne \mathcal{S}$, then $n(\mathcal{S}') \ne n(\mathcal{S})$.

Our theorem then follows, for we have for large x

$$\begin{aligned}\mathcal{P}_b(x) &\ge 2^{\#\mathcal{N}} - \#\mathcal{N} - 1\\ &> 2^{y^{E-\varepsilon}-1} - y^{E-\varepsilon} - 1\\ &\ge \exp\{(\log x)^{E/(E+1)-\varepsilon}\}.\end{aligned}$$

We now show (i). Let m denote the least common multiple of the elements of $\mathcal{N}$. We claim that if $a \in \mathcal{N}$, then

$$(2) \qquad F_{pqa}(b) \equiv 1 \pmod{pqm}.$$

First, since every prime factor of $F_{pqa}(b)$ is primitive ($l_b(p) \ne qa$, $p > P(qa)$), we have

$$F_{pqa}(b) \equiv 1 \pmod{pq}.$$

Next, since every prime factor of $F_{qa}(b)$ is primitive ($l_b(q) \ne a$, $q > P(a)$), if r is such a prime factor, then $r \equiv 1 \pmod q$, so $r \nmid m$. Hence we have $(F_{qa}(b), m) = 1$. Thus

$$F_{pqa}(b) = \frac{F_{qa}(b^p)}{F_{qa}(b)} \equiv \frac{F_{qa}(b)}{F_{qa}(b)} = 1 \pmod m$$

since $\lambda(m) \mid A \mid (p-1)$ and m is square-free imply $b^p \equiv b \pmod m$. We thus have (2) and so $pqm \mid n(\mathcal{S}) - 1$. Thus

$$n(\mathcal{S}) \mid \prod_{d \mid pqm} F_d(b) = b^{pqm} - 1 \mid b^{n(\mathcal{S})-1} - 1.$$

Also, since $\mathcal{S}$ has at least 2 elements, $n(\mathcal{S})$ is composite. Thus $n(\mathcal{S})$ is a pseudoprime to base b.

For (ii), note that if x is large and using (1),

$$\begin{aligned} n(\mathcal{S}) < b^{pq\Sigma_{a\in\mathcal{S}}a} &\le \exp\left\{pq(\log b)\sum_{a\in\mathcal{N}} a\right\} \\ &\le \exp\{pq(\log b)y^{E-\varepsilon+1}\} \\ &\le \exp(y^{E+1}) \\ &= x. \end{aligned}$$

Now note that if r is a prime factor of $F_{pqa}(b)$, then $l_b(r) = pqa$. This immediately gives (iii).

Remarks. (1) We mentioned above that from [8] we have $E > 0.55655$. Thus

$$E/(E+1) > 0.35755 > 5/14.$$

(2) Some people like to insist in their definition of pseudoprime to base b that it be odd. Note that all of the pseudoprimes created in the proof of Theorem 2 are odd and in fact are relatively prime to every prime $r \le 2pq$. Also note that

$$2pq > \exp(\log\log x/\log\log\log x) \quad \text{for all large } x.$$

(3) In the proof of Theorem 1, if we insist in the definition of $\mathcal{A}$ that $p \ne 2$, we have the same theorem as before, but now every member of $\mathcal{N}$ is odd. Thus in the proof of Theorem 2, we conclude that if r is any prime factor of $n(\mathcal{S})$, then $l_b(r)$ is odd. Since also $n(\mathcal{S})$ is odd (Remark 2), we conclude that the pseudoprimes $n(\mathcal{S})$ are all strong pseudoprimes.

(4) We would still obtain our result if we restricted $\mathcal{S}$ to those subsets of $\mathcal{N}$ which have a majority of the elements of $\mathcal{N}$. The pseudoprimes so constructed have at least $(\log x)^{5/14}$ distinct prime factors.

(5) A slight modification of the above proof gives a lower bound for $\mathcal{P}_b(x)$ that has an explicit dependence on b:

$$\mathcal{P}_b(x) \ge \exp\{(\log x/\log b)^{E/(E+1)-\varepsilon}\}$$

for all $x \ge b^{x_0(\varepsilon)^2}$, where $x_0(\varepsilon)$ is the constant in Theorem 1. To see this, we change the definition of y in the proof of Theorem 2 to

$$y = (\log x/\log b)^{(E+1)^{-1}}.$$

Then if $x \ge b^{x_0(\varepsilon)^2}$, we have $y \ge x_0(\varepsilon)$, so that Theorem 1 can be used to estimate $\#\mathcal{N}$.

(6) Consolidating Remarks 1 and 5, we have an absolute constant C such that for all $b \ge 2$ and $x \ge b^C$,

$$\mathcal{P}_b(x) \ge \exp\{(\log x/\log b)^{5/14}\}.$$

5. Cyclotomic pseudoprimes

If $b \geq 2$ is an integer and if $1 \leq d_1 < d_2 < \cdots < d_k$ are integers, we shall call the number $\Pi F_{d_i}(b)$ a *cyclotomic number to base b*. A *cyclotomic pseudoprime to base b* is then a cyclotomic number to base b which is also a pseudoprime to base b. For example, $341 = F_5(2)F_{10}(2)$ is a cyclotomic pseudoprime to base 2. Let $\mathscr{C}_b(x)$, $\mathscr{P}\mathscr{C}_b(x)$ denote respectively the counting functions for the cyclotomic numbers to base b, the cyclotomic pseudoprimes to base b.

It is clear that Theorem 2 holds for $\mathscr{P}\mathscr{C}_b(x)$ in place of $\mathscr{P}_b(x)$. Our point is that Theorem 2 is near to best possible for cyclotomic pseudoprimes. Indeed $\mathscr{P}\mathscr{C}_b(x) \leq \mathscr{C}_b(x)$ and an argument which uses estimates for the partition function $p(n)$ (see Rademacher [12]) shows that

$$\mathscr{C}_b(x) = \exp \{(\log x)^{1/2 + o(1)}\}.$$

This is the same estimate we would have for $\mathscr{P}\mathscr{C}_b(x)$ if we knew, as Erdös has conjectured, that $E = 1$.

We conclude that if there is to be substantial further progress on lower bounds for $\mathscr{P}_b(x)$, one will have to consider pseudoprimes to base b that are not cyclotomic.

References

1. E. Artin, *The orders of the linear groups*, Comm. Pure Appl. Math., vol. 8 (1955), pp. 355–366.
2. A. S. Bang, *Taltheoretiske Undersøgelser*, Tidsskrift Math., vol. 5, IV (1886), pp. 70–80 and 130–137.
3. P. Erdös, *On the normal number of prime factors of p − 1 and some related problems concerning Euler's φ-function*, Quart. J. Math. (Oxford Ser.), vol. 6 (1935), pp. 205–213.
4. P. Erdös, *On pseudoprimes and Carmichael numbers*, Publ. Math. Debrecen, vol. 4 (1956), pp. 201–206.
5. H. Iwaniec, *On the Brun-Titchmarsh theorem*, to appear.
6. D. H. Lehmer, *On the converse of Fermat's theorem*, American Math. Monthly, vol. 43 (1936), pp. 347–354 (see the third footnote on p. 348).
7. H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math., vol. 227, Springer-Verlag, Berlin-Heidelberg-New York, 1971.
8. C. Pomerance, *Popular values of Euler's function*, Mathematika, vol. 27 (1980), pp. 84–89.
9. ———, *On the distribution of pseudoprimes*, Math. Comp., to appear.
10. C. Pomerance, J. L. Selfridge, and S. S. Wagstaff, Jr., *The pseudoprimes to* $25 \cdot 10^9$, Math. Comp., vol. 35 (1980), pp. 1003–1026.
11. K. Prachar, *Primzahlverteilung*, Springer-Verlag, Berlin, 1957.
12. H. Rademacher, *Topics in analytic number theory*, Springer-Verlag, New York, 1973 (see Section 121).
13. A. Rotkiewicz, *On the number of pseudoprimes* $\leq x$, Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz., No. 381–409 (1972), pp. 43–45.

University of Georgia
Athens, Georgia