The distribution of Lucas and elliptic pseudoprimes

Daniel M. Gordon* Carl Pomerance†

August 22, 2001

Abstract

Let $\mathcal{L}(x)$ denote the counting function for Lucas pseudoprimes, and $\mathcal{E}(x)$ denote the elliptic pseudoprime counting function. We prove that, for large x, $\mathcal{L}(x) \le x\ L(x)^{-1/2}$ and $\mathcal{E}(x) \le x\ L(x)^{-1/3}$, where

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

1 Introduction

A *pseudoprime* is a composite number n for which

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

The smallest pseudoprime is 341. Let $\mathcal{P}(x)$ be the number of pseudoprimes up to x. The second author, in [12] and [13], showed that for all large x

$$\exp\left\{(\log x)^{5/14}\right\} \le \mathcal{P}(x) \le xL(x)^{-1/2},$$

where $L(x) = \exp(\log x \log_3 x / \log_2 x)$ and $\log_k$ is the k-fold iteration of the natural logarithm. The exponent 5/14 has since been improved to 85/207, see [14].

*Supported in part by a grant from Sandia National Laboratories

†Supported in part by an NSF grant

[0]1980 *Mathematics subject classification number* Primary 11Y11 Secondary 11Y40, 11A51

Let D, P and Q be integers such that $D = P^2 - 4Q \neq 0$ and $P > 0$. Let $U_0 = 0$, $U_1 = 1$, and $U_k = PU_{k-1} - QU_{k-2}$ for $k \geq 2$. Then a composite number n is a *Lucas pseudoprime* if $(n, 2D) = 1$ and

$$U_{n-\epsilon(n)} \equiv 0 \pmod{n}, \tag{1}$$

where $\epsilon(n)$ denotes the Jacobi symbol $(D|n)$. Let $\mathcal{L}(x) = \mathcal{L}_{P,Q}(x)$ be the number of Lucas pseudoprimes up to x. The best known bounds for $\mathcal{L}(x)$ are:

$$\exp\{(\log x)^{c_1}\} \leq \mathcal{L}(x) \leq x \cdot \exp\left\{-c_2(\log x \log_2 x)^{1/2}\right\},$$

for some absolute positive constants c_1 and c_2. The upper bound is due to Baillie and Wagstaff [1], and the lower bound is due to Erdös, Kiss and Sárközy [5]. Of course, the counting function $\mathcal{L}(x)$ depends on the choice of P and Q. The above result is thus understood to hold for all $x \geq x_0(P, Q)$.

The first author introduced a similar test using elliptic curves. Let E be an elliptic curve over $\mathbf{Q}$ with complex multiplication by an order in $K = \mathbf{Q}(\sqrt{-r})$, for $r \in \mathbf{Z}^+$, and suppose E has a rational point $P = (x_0, y_0)$ of infinite order. Then if n is a prime which is inert in K and does not divide the discriminant of E,

$$(n+1)\ P \equiv \mathcal{O} \pmod{n}. \tag{2}$$

That is, when we view E as an elliptic curve over the finite field $\mathbf{Z}/n\mathbf{Z}$, the image of the point P has order dividing $n+1$. An *elliptic pseudoprime* is a composite number n for which $(-r|n) = -1$, n is coprime to the discriminant of E and n satisfies (2). (The concept of $(n+1)\ P \equiv \mathcal{O} \pmod{n}$ for composite n will be made precise in the next section.) Let $\mathcal{E}(x) = \mathcal{E}_{E,P}(x)$ be the number of elliptic pseudoprimes less than x. The best known upper bound for elliptic pseudoprimes was recently found by Balasubramanian and Murty, in [2]: for all sufficiently large x depending on the choice of curve E and point P, we have

$$\mathcal{E}(x) \leq x \cdot \exp\left\{-c_3(\log x \log_2 x)^{1/2}\right\}.$$

The number c_3 is positive and absolute. No good general lower bounds for elliptic pseudoprimes are known; the only result is from [6], that for certain curves and points,

$$\mathcal{E}(x) \geq \frac{\sqrt{\log x}}{\log_2 x}.$$

In this paper we improve the upper bounds for $\mathcal{E}(x)$ and $\mathcal{L}(x)$. The techniques used are similar to those of [12], with modifications to deal with elliptic curves similar to those of [2]. We show that $\mathcal{E}(x) \le x\ L(x)^{-1/3}$ and $\mathcal{L}(x) \le x\ L(x)^{-1/2}$ for large x.

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

2 Elliptic curve preliminaries

For a field k of characteristic > 3, an elliptic curve over k may be represented as

$$E(k) = \{(x, y) \in k^2 : y^2 = x^3 + ax + b\} \cup \mathcal{O},$$

where $a, b \in k$ and $\mathcal{O}$ is the point at infinity. E is nonsingular if the discriminant $\Delta = -16(4a^3 + 27b^2) \neq 0$. In this case, $E(k)$ can be naturally made into an additive group with $\mathcal{O}$ being the identity element.

Suppose E is a nonsingular elliptic curve defined over $\mathbf{Q}$. Let $End\ E$ denote the ring of endomorphisms of $E(\mathbf{Q})$. It is known that $End\ E$ is either equal to $\mathbf{Z}$ or an order in an imaginary quadratic field $K = \mathbf{Q}(\sqrt{-r})$. In the latter case, E is said to have complex multiplication by K. For instance, curves of the form $y^2 = x^3 - Dx$ have complex multiplication by $\mathbf{Q}(\sqrt{-1})$; the endomorphism corresponding to i sends a point (x, y) to $(-x, iy)$.

If E is defined over $\mathbf{Q}$ and has complex multiplication by K, then K must have class number one, so that $r \in \{1, 2, 3, 7, 11, 19, 43, 67, 163\}$. Conversely, for each such r there are elliptic curves with complex multiplication by O_K, the full ring of integers of K. In addition, the fields $\mathbf{Q}(\sqrt{-1})$, $\mathbf{Q}(\sqrt{-3})$, and $\mathbf{Q}(\sqrt{-7})$ have curves over $\mathbf{Q}$ with $End\ E = \mathbf{Z} + 2O_K$, and $\mathbf{Q}(\sqrt{-3})$ has curves with $End\ E = \mathbf{Z} + 3O_K$.

For a rational number x, let u/v be its representation in lowest terms. Then $\mathrm{Num}(x) = u$ will denote its numerator, $\mathrm{Den}(x) = v$ its denominator, and $\tilde{x} = uv$ their product.

Let $E(\mathbf{Q})$ be a nonsingular elliptic curve defined by the equation $y^2 = x^3+ax+b$, where the coefficients $a, b \in \mathbf{Q}$. If p is a prime with $(p, 6\tilde{\Delta}) = 1$, by an abuse of notation, we can use this same equation to define a nonsingular elliptic curve $E(\mathbf{F}_p)$ over $\mathbf{F}_p$, the field of p elements. In fact there is a natural homomorphic projection $E(\mathbf{Q}) \to E(\mathbf{F}_p)$ which takes $(x, y) \in E(\mathbf{Q})$ to $(x \bmod p, y \bmod p)$. If one of x, y has a factor p in the denominator, then (x, y) maps to $\mathcal{O}$ in $E(\mathbf{F}_p)$.

A celebrated theorem of Hasse is that for any nonsingular elliptic curve $E(\mathbf{F}_p)$, the number of points can be expressed as $p+1-a_p$, where $|a_p| \le 2\sqrt{p}$.

There is a polynomial time, deterministic algorithm, due to Schoof [15], for computing the number a_p. Nevertheless, for very large p, it is not an easy task to compute the order of $E(\mathbf{F}_p)$.

If E has complex multiplication by $K = \mathbf{Q}(\sqrt{-r})$, it is easier to compute $|E(\mathbf{F}_p)|$:

$$|E(\mathbf{F}_p)| = \begin{cases} p+1, & p \text{ inert in } K \\ p+1-2\beta, & p = (\beta+\gamma\sqrt{-r})(\beta-\gamma\sqrt{-r}) \end{cases} \tag{3}$$

where $2\beta, 2\gamma \in \mathbf{Z}$. Note that if p splits in K, formula (3) does not quite give $|E(\mathbf{F}_p)|$, since we don't know the sign of β (and if $K = \mathbf{Q}(\sqrt{-1})$ or $\mathbf{Q}(\sqrt{-3})$, there are extra units which add a few more possibilities). However, this is the only indeterminacy in (3), since primes p which split in K have a unique representation up to units as $\beta^2 + r\gamma^2$.

The representation of p as $\beta^2 + r\gamma^2$ can be found in random polynomial time by factoring the polynomial $x^2 + r$ in $\mathbf{F}_p$, using Berlekamp's algorithm [3]. Once a number c is found such that $c^2 + r \equiv 0 \pmod p$, one may use the method of Cornacchia [4] to determine β and γ.

Determining the sign of β in (3) can in principle be done using class field theory; it is worked out for $K = \mathbf{Q}(\sqrt{-1})$ and $\mathbf{Q}(\sqrt{-3})$ in [11].

For a nonsingular curve $E(\mathbf{Q})$ with coefficients $a, b \in \mathbf{Q}$, define the *division polynomial* $\psi_n(x, y)$ by

$$\psi_0 = 0,$$

$$\psi_1 = 1,$$

$$\psi_2 = 2y,$$

$$\psi_3 = 3x^4 + 6ax^2 + 12bx - a^2,$$

$$\psi_4 = 4y(x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - 8b^2 - a^3),$$

and the recursion

$$\psi_{m+n}\psi_{m-n} = \psi_{m-1}\psi_{m+1}\psi_n^2 - \psi_{n-1}\psi_{n+1}\psi_m^2.$$

Thus

$$\psi_{2n+1} = \psi_n^3\psi_{n+2} - \psi_{n+1}^3\psi_{n-1} \tag{4}$$

and

$$2y\psi_{2n} = \psi_n(\psi_{n+2}\psi_{n-1}^2 - \psi_{n-2}\psi_{n+1}^2). \tag{5}$$

The division polynomials characterize the division points of $E(\mathbf{Q})$. Namely, $P = (x_0, y_0) \in E(\mathbf{Q})$ is an m-division point (*i.e.*, $mP = \mathcal{O}$) if and only if $\psi_m(x_0, y_0) = 0$. This continues to make sense if we replace $\mathbf{Q}$ by some algebraic extension. However, we are primarily concerned here with the connection between the division polynomials and division points on $E(\mathbf{F}_p)$.

We now state three lemmas on division polynomials. See Chapter II of Lang [10] for many facts about these polynomials and, in particular, the following lemma.

Lemma 1 *Suppose $E(\mathbf{Q})$ is a nonsingular elliptic curve with coefficients $a, b \in \mathbf{Q}$ and let $P = (x_0, y_0)$ be a point of infinite order on $E(\mathbf{Q})$. For a prime p with $(p, 6\tilde{\Delta}) = 1$, let $\bar{P}$ be the image of P in $E(\mathbf{F}_p)$. Suppose $2\ \bar{P} \neq \mathcal{O}$ on $E(\mathbf{F}_p)$. Then for any integer $m > 2$ we have*

$$m\bar{P} = \mathcal{O} \text{ in } E(\mathbf{F}_p) \Longleftrightarrow \psi_m(x_0, y_0) \equiv 0 \pmod{p}.$$

Of course, we understand the rational number $\psi_m(x_0, y_0)$ to be $0 \pmod{p}$ if in reduced form, its numerator is $0 \pmod{p}$.

The second lemma involves the size of the values of the division polynomials:

Lemma 2 *Suppose E is a nonsingular elliptic curve, and $P = (x_0, y_0)$ is a point in $E(\mathbf{Q})$ of infinite order. Then for all natural numbers m,*

$$|\psi_m(x_0, y_0)| < c^{m^2-3}$$

for some constant c depending on the choice of curve E and point P.

Proof: Choose c such that $c^6 \geq \max\{2, y_0^{-2}\}$ and $|\psi_m(x_0, y_0)| < c^{m^2-3}$ for $m = 2, 3, 4$. It is easy to show by induction that $|\psi_m(x_0, y_0)| < c^{m^2-3}$ holds for all m, using (4) and (5). $\Box$

Corollary 1 *For E and P as in Lemmas 1 and 2, the number of primes p for which $mP = \mathcal{O}$ in $E(\mathbf{F}_p)$ is $O(m^2)$.*

Proof: By Lemma 1, all such primes p divide the numerator of $\psi_m(x_0, y_0)$, and by Lemma 2, $\psi_m(x_0, y_0) = O(c^{m^2})$. Therefore it suffices to show that the denominator of $\psi_m(x_0, y_0)$ is bounded by $c_2^{m^2}$.

Suppose we give a grading to the ring $\mathbf{Z}[a,b,x,y]$ by giving a weight 4, b weight 6, x weight 2 and y weight 3. Then $\psi_m(x,y)$ is homogeneous of weight m^2-1 with respect to this grading ([10], page 39). Therefore the denominator of $\psi_m(x_0,y_0)$ is less than

$$|\mathrm{Den}(y_0)^{m^2/3}\ \mathrm{Den}(x_0)^{m^2/2}\ \mathrm{Den}(a)^{m^2/4}\ \mathrm{Den}(b)^{m^2/6}| < c_2^{m^2}. \quad \square$$

Corollary 1 implies the case $r=1$ of Lemma 14 in Gupta and Murty [7]. They prove a more general result using a considerably more involved argument.

Suppose $E(\mathbf{Q})$, $P=(x_0,y_0)$ and p are as in Lemma 1, and $E(\mathbf{Q})$ has complex multiplication by $K=\mathbf{Q}(\sqrt{-r})$, where $(-r|p)=-1$. Suppose $2\ \bar{P}\neq\mathcal{O}$ on $E(\mathbf{F}_p)$. From (3), $(p+1)\bar{P}=\mathcal{O}$ in $E(\mathbf{F}_p)$, so that by Lemma 1,

$$\psi_{p+1}(x_0,y_0)\equiv 0 \pmod{p}.$$

The key observation is that even if we do not know for sure that p is prime, we can still check if the congruence $\psi_{p+1}(x_0,y_0)\equiv 0 \pmod{p}$ holds. We say a composite natural number n which satisfies $(n,6\tilde{\Delta})=1$ and $(-r|n)=-1$ is an *elliptic pseudoprime* (for the curve E and the point P) if

$$(\tilde{y_0},n)=1 \ \text{ and } \ \psi_{n+1}(x_0,y_0)\equiv 0 \pmod{n}. \tag{6}$$

This is what we mean by the congruence in (2) for n composite. Note that if n is prime, then the condition $(\tilde{y_0},n)=1$ assures that $2\ \bar{P}\neq\mathcal{O}$ on $E(\mathbf{F}_n)$.

For any natural number m with $(m,6\tilde{\Delta}\tilde{y_0})=1$, define $e_m=e_m(P)$ as the least positive number k for which $\psi_k(x_0,y_0)\equiv 0 \pmod{m}$. (If no such k exists, or if $(m,6\tilde{\Delta}\tilde{y_0})>1$, define $e_m=\infty$.) We will need the following lemma:

Lemma 3 *If m is a positive squarefree number with $(m,6\tilde{\Delta}\tilde{y_0})=1$, then*

$$e_m=\mathrm{lcm}\{e_q : q|m\}$$

and

$$\psi_k(x_0,y_0)\equiv 0 \pmod{m} \iff e_m|k.$$

Proof: The lemma is true for primes by Lemma 1, since e_p is the order of the point $\bar{P}$ in $E(\mathbf{F}_p)$. Suppose $m=q_1q_2\ldots q_s$, with the q_i's distinct primes.

Let $l = \operatorname{lcm}\{e_{q_1}, \ldots, e_{q_s}\}$. Then $\psi_l(x_0, y_0) \equiv 0 \pmod m$, so $e_m \le l$. But $\psi_{e_m}(x_0, y_0) \equiv 0 \pmod{q_i}$ for each q_i, so each $e_{q_i} | e_m$. Thus $e_m = l$. The second assertion in the lemma follows from similar considerations. $\Box$

A similar lemma was proved by Ward [16] for $a, b, x_0, y_0 \in \mathbf{Z}$, without the restriction that m be squarefree.

3 Elliptic pseudoprimes

Let $E(\mathbf{Q})$ be a nonsingular elliptic curve with coefficients $a, b \in \mathbf{Q}$ and complex multiplication by $\mathbf{Q}(\sqrt{-r})$, a complex quadratic field with class number one, and let $P = (x_0, y_0) \in E(\mathbf{Q})$ have infinite order.

Theorem 1 *There is a constant $X_0 = X_0(E, P)$ such that if n is a natural number and $x \ge X_0$ then*

$$\#\{m \le x : m \text{ is squarefree and } e_m = n\} \le x \cdot \exp\left(-\log x\,\frac{3 + \log_3 x}{3 \log_2 x}\right).$$

Proof: Unlike the exponent to which 2 belongs mod m studied with regular pseudoprimes, e_m may be greater than m. Thus n in the theorem may be greater than x. To determine an upper bound for n, if $m \le x$ is squarefree and $e_m = n$, note that

$$e_m \le \prod_{q|m} (q + 1 + 2\sqrt{q}) \le m \prod_{q|m} \left(1 + \frac{3}{\sqrt{q}}\right) \le x \prod_{q \le 2\log x} \left(1 + \frac{3}{\sqrt{q}}\right) \tag{7}$$

for x so large that $x \le \prod_{q \le 2\log x} q$. That such an inequality should eventually hold follows from the prime number theorem. Using partial summation and the prime number theorem, we have

$$\log \prod_{q \le 2\log x} \left(1 + \frac{3}{\sqrt{q}}\right) \ll \sum_{q \le 2\log x} \frac{1}{\sqrt{q}} \ll \frac{(\log x)^{1/2}}{\log_2 x},$$

and with (7) this implies that $e_m \le x^{1+\epsilon}$, for any $\epsilon > 0$ and $x \ge x_0(\epsilon)$. We shall take $\epsilon = 1/2$ and shall assume n in the theorem satisfies $n \le x^{3/2}$.

Let $c = 1 - (4 + \log_3 x)/(3 \log_2 x)$, and $c' = c - 1/(3 \log_2 x)$, with x large enough so that $c' \ge 7/8$. Then we need to estimate:

$$\sum_{\substack{m \le x \\ e_m = n}} 1 \le x^c \sum_{e_m = n} m^{-c} \le x^c \sum_{p|m \Rightarrow e_p|n} m^{-c} = x^c \prod_{e_p|n} (1 - p^{-c})^{-1} = x^c A,$$

say. To prove the theorem it is sufficient to show that

$$\log A = o(\log x/\log_2 x). \tag{8}$$

Since $c \geq 7/8$, we have

$$\log A = \sum_{e_p|n} p^{-c} + O(1) = \sum_{d|n} \sum_{e_p=d} p^{-c} + O(1).$$

There are only a finite number of primes p with $e_p = d$ for $d = 1$ or 2, since those primes divide either the numerator of y_0 (for $d = 2$) or the denominator of y_0 (for $d = 1$). Assume now that $d \geq 3$.

By Corollary 1, there are at most αd^2 primes p with $e_p = d$, where α is a constant depending only on E and P. Call them $q_1, q_2, \ldots, q_t$, where $0 \leq t \leq \alpha d^2$.

For each q_i, $E(\mathbf{F}_{q_i})$ has order kd where kd is a multiple of d satisfying

$$q_i + 1 - 2\sqrt{q_i} \leq kd \leq q_i + 1 + 2\sqrt{q_i}.$$

Therefore we have $q_i > kd/2$. If q_i is inert in K, then $kd = q_i + 1$. If q_i splits, say $q_i = (a + \sqrt{-r}b)(a - \sqrt{-r}b) = a^2 + rb^2$, then by (3)

$$kd = q_i + 1 - 2a = a^2 - 2a + 1 + rb^2 = (a-1)^2 + rb^2.$$

The number of representations of kd as $\beta^2 + r\gamma^2$ with $\beta,\ \gamma \geq 0$ is at most the number of divisors of kd: $\tau(kd)$ (see, for example Theorem 54 of [9]). In sum, the number of q_i with the order of $E(\mathbf{F}_{q_i})$ being kd is at most $2\tau(kd) + 1 < 3\tau(kd)$, and all of these q_i satisfy $q_i > kd/2$. From these facts, if $d \geq 3$,

$$\begin{aligned}
\sum_{e_p=d} p^{-c} &= \sum_{i=1}^{t} q_i^{-c} \leq 6 \sum_{k=1}^{t} \tau(kd)\ (kd)^{-c} \\
&\leq 6\ \tau(d)\ d^{-c} \sum_{k=1}^{[\alpha d^2]} \tau(k) k^{-c}.
\end{aligned}$$

Using partial summation, and $\sum_{k=1}^{N} \tau(k) = N \log N + O(N)$ (see [8]), this is

$$\begin{aligned}
&= 6\ \frac{\alpha^{1-c}}{1-c}\ \tau(d)\ d^{2-3c}(2\log d + \log \alpha)(1 + o(1)) \\
&\ll (1-c)^{-1}\ \tau(d)\ d^{2-3c} \log d.
\end{aligned} \tag{9}$$

To get rid of the $\log d$ factor, note that

$$\log d \ll \max\{d^{1/\log_2 x}, \log_2 x \log_3 x\} \leq d^{1/\log_2 x} \log_2 x \log_3 x.$$

Therefore,

$$d^{2-3c} \log d \ll d^{2-3c'} \log_2 x \log_3 x$$

so that (9) implies

$$\sum_{e_p = d} p^{-c} \ll (1-c)^{-1}\ \tau(d)\ d^{2-3c'} \log_2 x \log_3 x.$$

From the above computations, we have

$$\begin{aligned}
\log A &\ll (1-c)^{-1}\ \log_2 x \log_3 x \sum_{d|n} \tau(d) d^{2-3c'} \\
&< (1-c)^{-1}\ \log_2 x \log_3 x\ \prod_{p|n} (1 + 2p^{2-3c'} + 3(p^{2-3c'})^2 + \ldots) \\
&= (1-c)^{-1}\ \log_2 x \log_3 x\ \prod_{p|n} (1 - p^{2-3c'})^{-2} \qquad (10)
\end{aligned}$$

Since $2 - 3c' \leq -5/8$, we have

$$\log \prod_{p|n} (1 - p^{2-3c'})^{-2} = 2 \sum_{p|n} p^{2-3c'} + O(1) \leq 2 \sum_{p \leq 2\log x} p^{2-3c'} + O(1),$$

where x is large enough that $\prod_{p \leq 2\log x} p \geq x^{3/2}$. This implies

$$\log \prod_{p|n} (1 - p^{2-3c'})^{-2} \ll \frac{(\log x)^{3-3c'}}{(3-3c')\log_2 x} \ll \frac{\log_2 x}{\log_3 x}. \qquad (11)$$

Thus, if x is sufficiently large, we have

$$\prod_{p|n} (1 - p^{2-3c'})^{-2} \leq (\log x)^{1/2},$$

and with (10) we get

$$\log A \ll \frac{\log_2 x}{\log_3 x} \log_2 x \log_3 x\ (\log x)^{1/2}$$

which is $o(\log x / \log_2 x)$. $\square$

Theorem 2 *For all sufficiently large x, depending on the choice of E and P, the number of elliptic pseudoprimes for E, P up to x is at most*

$$x \cdot \exp\left(-\frac{\log x \log_3 x}{3 \log_2 x}\right).$$

Proof: As is now customary with proofs of upper bounds on pseudoprimes, we will divide the elliptic pseudoprimes $n \le x$ into several possibly overlapping classes:

(i) $n \le x\ L(x)^{-1}$,

(ii) there is a prime $p|n$ with $e_p \le L(x)^3, p > L(x)^{10}$,

(iii) there is a prime $p|n$ with $e_p > L(x)^3$ and $p \le 3x/L(x)$,

(iv) there is a prime $p|n$ inert in K with $e_p > L(x)^3$,

(v) there is a prime $p|n$ which splits in K with $L(x)^3 < e_p \le \sqrt{x}L(x)$ and $p > 3x/L(x)$,

(vi) there is a prime $p|n$ which splits in K with $e_p > \sqrt{x}L(x)$ and $p > 3x/L(x)$,

(vii) $n > x\ L(x)^{-1}$ and every prime $p|n$ is at most $L(x)^{10}$.

Clearly, the number of n in class (i) is at most $x\ L(x)^{-1}$.

From Corollary 1, the number of primes p with $e_p = m$ is $O(m^2)$. Thus the number of primes p with $e_p \le L(x)^3$ is

$$\sum_{m \le L(x)^3} \sum_{e_p = m} 1 \ll \sum_{m \le L(x)^3} m^2 < L(x)^9.$$

Therefore the number of elliptic pseudoprimes in class (ii) is at most

$$\sum_{\substack{p > L(x)^{10} \\ e_p \le L(x)^3}} x/p < x\ L(x)^{-10} \sum_{e_p \le L(x)^3} 1 \ll x\ L(x)^{-1}. \tag{12}$$

If p is a prime dividing an elliptic pseudoprime n, then from Lemma 3 (with $m = p$) we have

$$n \equiv 0 \pmod{p}, \quad n + 1 \equiv 0 \pmod{e_p}, \quad (p, e_p) = 1. \tag{13}$$

The number of $n \le x$ satisfying these conditions is at most

$$1 + \frac{x}{pe_p}. \tag{14}$$

Thus the number of elliptic pseudoprimes in class (iii) is at most

$$\sum_{\substack{p \le 3x/L(x) \\ e_p > L(x)^3}} \left(1 + \frac{x}{pe_p}\right) \le \sum_{p \le 3x/L(x)} 1 \; + \sum_{\substack{p \le 3x/L(x) \\ e_p > L(x)^3}} \frac{x}{pe_p}$$

The first sum on the right is at most $3x/L(x)$, and the final sum is at most of order $x \log_2 x / L(x)^3$. Thus the number of elliptic pseudoprimes in class (iii) is

$$\ll \frac{x}{L(x)}. \tag{15}$$

If p is inert in K, $e_p | (p+1)$, and so $n = p$ is a solution to (13). This solution is prime, so the number of elliptic pseudoprimes divisible by p is at most $x/(pe_p)$. Therefore the number of elliptic pseudoprimes in class (iv) is at most

$$\sum_{\substack{2 < p \le x \\ e_p > L(x)^3}} \frac{x}{pe_p} \ll \frac{x \log_2 x}{L(x)^3}. \tag{16}$$

For the special prime p dividing an elliptic pseudoprime n in class (v), let $k = n/p$, and $l = e_p$. Since p splits, we have $p = \beta^2 + r\gamma^2$ for some $|\beta|, |\gamma| < \sqrt{x}$, where $2\beta, 2\gamma \in \mathbf{Z}$. From (3), we have $p \equiv 2\beta - 1 \pmod{e_p}$, since $e_p | \; |E(\mathbf{F}_p)|$. Thus

$$n + 1 = kp + 1 \equiv k(2\beta - 1) + 1 \equiv 0 \pmod{l}, |\beta| < \sqrt{x}. \tag{17}$$

This means that possible integers 2β fall in a unique congruence class mod $l/(k,l)$. For a fixed k and l, the number of β satisfying (17) is at most

$$\frac{4\sqrt{x}}{l}(k,l) + O(1).$$

For each β and l, the number of solutions γ to

$$|E(\mathbf{F}_p)| = \beta^2 + r\gamma^2 + 1 - 2\beta \equiv 0 \pmod{l}$$

is bounded by $\tau(4l/(r,4l))(r,4l) \ll \tau(l)$, since $r \ll 1$. Since $|\gamma| < \sqrt{x}$, the number of γ's corresponding to any β and l is thus

$$\ll \left(\frac{\sqrt{x}}{l} + O(1) \right) \tau(l).$$

Summing over k and l, the number of elliptic pseudoprimes in class (v) is

$$\ll \sum_{\substack{k \le L(x) \\ L(x)^3 < l \le \sqrt{x} L(x)}} \left(\frac{\sqrt{x}}{l}(k,l) + O(1) \right) \left(\frac{\sqrt{x}}{l} + O(1) \right) \tau(l)$$

$$= x \sum_{k,l} \frac{(k,l)\tau(l)}{l^2} + O\left(\sqrt{x} \sum_{k,l} \frac{(k,l)\tau(l)}{l} \right) + O\left(\sum_{k,l} \tau(l) \right).$$

The final sum is easily seen to be $O(\sqrt{x} L(x)^2 \log x)$. The second sum is

$$\ll \sqrt{x} L(x) \sum_{k,l} \frac{\tau(l)}{l} \le \sqrt{x} L(x)^2 \sum_{l} \frac{\tau(l)}{l} \ll \sqrt{x} L(x)^2 \log^2 x.$$

Finally, the first sum is

$$\le x L(x) \sum_{k,l} \frac{\tau(l)}{l^2} \le x L(x)^2 \sum_{l} \frac{\tau(l)}{l^2} \le \frac{x}{L(x)} \sum_{l} \frac{\tau(l)}{l} \ll \frac{x \log^2 x}{L(x)}.$$

Combining these estimates, the number of elliptic pseudoprimes in class (v) is

$$\ll \frac{x \log^2 x}{L(x)}. \tag{18}$$

To estimate the size of class (vi), let $n = kp$ for some $k > 1$. We have $p \equiv -1 + a_p \pmod{e_p}$, since $e_p || E(\mathbf{F}_p)| = p + 1 - a_p$. Since $n + 1 \equiv 0 \pmod{e_p}$, we have

$$kp + 1 \equiv k(a_p - 1) + 1 \equiv 0 \pmod{e_p} \tag{19}$$

and so

$$|k(a_p - 1) + 1| \ge e_p > \sqrt{x} L(x).$$

Since $|a_p| \le 2\sqrt{p}$, this means that $k > L(x)/3$. But then $n = kp > x$, and so class (vi) is empty for x sufficiently large.

We will divide the pseudoprimes in class (vii) into two subclasses: those which have a squareful divisor s (*i.e.*, for each prime p dividing s, p^2 also divides s) with $s > L(x)^2$, and those which do not. The number of $n < x$ in the first subclass is at most

$$\sum_{\substack{s>L(x)^2 \\ s \text{ squareful}}} \frac{x}{s} \ll \frac{x}{L(x)}$$

using partial summation and the theorem that

$$\sum_{\substack{s\le t \\ s \text{ squareful}}} 1 \ll \sqrt{t}.$$

For the rest of class (vii), we have $x/L(x) < n \le x$, every prime $p|n$ satisfies $p \le L(x)^{10}$, and the squareful part of n does not exceed $L(x)^2$. Then n has a squarefree divisor d satisfying

$$x/L(x)^{13} < d \le x/L(x)^3. \tag{20}$$

(For let $m =$ the largest squarefree divisor of n. Then $x/L(x)^3 < m \le x$. We have some $d|m$ with $x/L(x)^{13} < d \le x/L(x)^3$. But d is squarefree and $d|n$.)

As in (13), we have from Lemma 3 that

$$n \equiv 0 \pmod{d}, \quad n+1 \equiv 0 \pmod{e_d}, \quad (d, e_d) = 1. \tag{21}$$

Therefore the number of such n is at most

$$\sum{}' \left(1 + \frac{x}{de_d}\right) \le x/L(x) + x\sum{}' \frac{1}{de_d} = x/L(x) + x\sum_{m\le x} \frac{1}{m} \sum_{e_d=m}{}' \frac{1}{d},$$

where $\sum'$ means the sum is over squarefree d in the range (20). By Theorem 1, and a partial summation argument, the inner sum is at most

$$\exp\left(-\log x \,\frac{2+\log_3 x}{3\log_2 x}\right)$$

uniformly in m, provided x is sufficiently large. Therefore, the number of n in class (vii) is at most

$$x \cdot \exp\left(-\log x \; \frac{1+\log_3 x}{3\log_2 x}\right) \tag{22}$$

for large x.

Summing the estimates for each of the classes gives the theorem. $\square$

4 Lucas pseudoprimes

The proof of the bound for $\mathcal{L}(x)$ will be similar to the proof for $\mathcal{E}(x)$. First we will need a few facts about Lucas pseudoprimes. See [1] for proofs.

Let ω_p denote the rank of apparition of p in the Lucas sequence U_k; *i.e.*, the least positive k for which $p|U_k$. Then if $(p, 2D) = 1$, we have

$$\omega_p|(p-\epsilon(p)),$$

where we recall that $\epsilon(p) = (D|p)$. Further, $\omega_{p^k}|p^{k-1}\omega_p$, and for any m with $(m, 2D) = 1$, we have $\omega_m = \mathrm{lcm}\{\omega_{p^k} : p^k \parallel m\}$. If $(m, 2D) = 1$ then $m|U_k$ if and only if $\omega_m|k$. Also, let α and β be the distinct roots of $x^2 - Px + Q = 0$. Then for $k \geq 0$,

$$U_k = \frac{\alpha^k - \beta^k}{\alpha - \beta}. \tag{23}$$

We are now ready to prove:

Theorem 3 *There is an $x_0 = x_0(P, Q)$ such that if n is a natural number and $x \geq x_0$ then*

$$\#\{m \leq x : \omega_m = n\} \leq x \cdot \exp\left(-\log x \; \frac{3+\log_3 x}{2\log_2 x}\right).$$

Proof: As in Theorem 1, we may assume that $n < x^{3/2}$. In fact, if the set in the theorem is not empty, it is possible to show that $n \ll x \log\log x$.

Let $c = 1-(4+\log_3 x)/(2\log_2 x)$, and let x be large enough that $c \geq 7/8$. Then

$$\sum_{\substack{m \leq x \\ \omega_m = n}} 1 \leq x^c \sum_{\omega_m = n} m^{-c} \leq x^c \sum_{p|m \Rightarrow \omega_p|n} m^{-c} = x^c \prod_{\omega_p|n} (1-p^{-c})^{-1} = x^c A,$$

say. As before, it suffices to show

$$\log A = o(\log x/\log_2 x). \tag{24}$$

Since $c \geq 7/8$, we have

$$\log A = \sum_{\omega_p | n} p^{-c} + O(1) = \sum_{d|n} \sum_{\omega_p = d} p^{-c} + O(1).$$

The primes p with $\omega_p = d$ are divisors of U_d, which is $O(\max\{|\alpha|, |\beta|\}^d)$ by (23), so there are at most $O(d)$ of them. Call them $q_1, q_2, \ldots, q_t$, where $0 \leq t \leq \delta d$, for some constant δ depending only on P and Q. Those p with $p|2D$ contribute at most $O(1)$ to $\log A$, so we may assume the primes q_i do not divide $2D$. Thus each $q_i \equiv \pm 1 \pmod{d}$, so

$$\sum_{\omega_p = d} p^{-c} = \sum_{i=1}^{t} q_i^{-c} \leq \sum_{k=1}^{t} 2(kd)^{-c} \leq 2d^{-c} \sum_{k=1}^{[\delta d]} k^{-c} \ll (1-c)^{-1} d^{1-2c}. \tag{25}$$

Thus,

$$\log A \ll (1-c)^{-1} \sum_{d|n} d^{1-2c} < (1-c)^{-1} \prod_{p|n} (1 - p^{1-2c})^{-1}. \tag{26}$$

Since $1 - 2c \leq -3/4$, we have

$$\log \prod_{p|n} (1 - p^{1-2c})^{-1} = \sum_{p|n} p^{1-2c} + O(1) \leq \sum_{p \leq 2 \log x} p^{1-2c} + O(1),$$

where x is large enough that $\prod_{p \leq 2 \log x} p \geq x^{3/2}$. This implies

$$\log \prod_{p|n} (1 - p^{1-2c})^{-1} \ll \frac{(\log x)^{2-2c}}{(2-2c)\log_2 x} \ll \frac{\log_2 x}{\log_3 x}. \tag{27}$$

Thus, if x is sufficiently large, we have

$$\prod_{p|n} (1 - p^{1-2c})^{-1} \leq (\log x)^{1/2},$$

and with (26) we get

$$\log A \ll \frac{\log_2 x}{\log_3 x} (\log x)^{1/2}$$

which is $o(\log x/\log_2 x)$. $\square$

Theorem 4 *For all sufficiently large x, depending on the choice of P, Q, the number of Lucas pseudoprimes up to x is at most $x\ L(x)^{-1/2}$.*

Proof: As in Theorem 2, we will divide the Lucas pseudoprimes $n \le x$ into several possibly overlapping classes:

(i) $n \le x\ L(x)^{-1}$,

(ii) there is a prime $p|n$ with $\omega_p \le L(x), p > L(x)^3$,

(iii) there is a prime $p|n$ with $\omega_p > L(x)$ and $\epsilon(p) = \epsilon(n)$,

(iv) there is a prime $p|n$ with $\omega_p > L(x)$ and $\epsilon(p) \neq \epsilon(n)$,

(v) $n > x\ L(x)^{-1}$ and every prime $p|n$ is at most $L(x)^3$.

Clearly, the number of n in class (i) is at most $x\ L(x)^{-1}$.
The number of primes p with $\omega_p \le L(x)$ is

$$\sum_{m \le L(x)} \sum_{\omega_p = m} 1 \ll \sum_{m \le L(x)} m < L(x)^2.$$

Therefore the number of Lucas pseudoprimes in class (ii) is at most

$$\sum_{\substack{p > L(x)^3 \\ \omega_p \le L(x)}} x/p < x\ L(x)^{-3} \sum_{\omega_p \le L(x)} 1 \ll x\ L(x)^{-1}. \tag{28}$$

If p is a prime dividing a Lucas pseudoprime n, we have

$$n \equiv 0 \pmod{p}, \quad n - \epsilon(n) \equiv 0 \pmod{\omega_p}, \quad (p, \omega_p) = 1. \tag{29}$$

For a fixed p, the numbers $n \le x$ that satisfy (29) can be split into two cases: those with $\epsilon(n) = \epsilon(p)$ and those with $\epsilon(n) = -\epsilon(p)$. In the first case, $n = p$ is a solution to (29), but is not a Lucas pseudoprime. Thus corresponding to a prime p in class (iii) there are at most $x/(p\omega_p)$ Lucas pseudoprimes $n \le x$. We conclude that the number of Lucas pseudoprimes in class (iii) is at most

$$\sum_{\substack{p \le x \\ \omega_p > L(x)}} \frac{x}{p\omega_p} \ll \frac{x \log_2 x}{L(x)}. \tag{30}$$

Suppose p, n are as in class (iv) and $n = kp$. From (29) we have

$$\epsilon(n) \equiv n = kp \equiv k\epsilon(p) \pmod{\omega_p},$$

so that $k \equiv -1 \pmod{\omega_p}$. The number of $k \le x/p$ with $k \equiv -1 \pmod{\omega_p}$ is exactly

$$\left[\frac{(x/p)+1}{\omega_p}\right],$$

so the number of Lucas pseudoprimes in class (iv) is at most

$$\sum_{\substack{p \le x \\ \omega_p > L(x)}} \left(\frac{x}{p\omega_p} + \frac{1}{\omega_p}\right) \ll \frac{x \log_2 x}{L(x)}. \tag{31}$$

Every n in class (v) has a divisor d with

$$x/L(x)^4 < d \le x/L(x). \tag{32}$$

As in (29), we have

$$n \equiv 0 \pmod{d}, \quad n - \epsilon(n) \equiv 0 \pmod{\omega_d}, \quad (d, \omega_d) = 1, \tag{33}$$

so that n is in one of two residue classes (mod $d\omega_d$), depending on whether $\epsilon(n) = 1$ or -1. Therefore the number of n in class (v) is at most

$$2\sum{}' \left(1 + \frac{x}{d\omega_d}\right) \le 2x/L(x) + x\sum{}' \frac{2}{d\omega_d} = 2x/L(x) + x\sum_{m \le x} \frac{2}{m} \sum_{\omega_d = m}{}' \frac{1}{d},$$

where $\sum'$ means the sum is over d in the range (32). By Theorem 3, and a partial summation argument, the inner sum is at most

$$\exp\left(-\log x\ \frac{2 + \log_3 x}{2\log_2 x}\right)$$

uniformly in m, provided x is sufficiently large. Therefore, the number of n in class (v) is at most

$$x \cdot \exp\left(-\log x\ \frac{1 + \log_3 x}{2\log_2 x}\right) \tag{34}$$

for large x.

Each of the classes has $o(x\ L(x)^{-1/2})$ Lucas pseudoprimes, which proves the theorem. □

References

[1] R. Baillie and S.S. Wagstaff, Jr., Lucas pseudoprimes, *Math. Comp.* **35** (1980), pp. 1391-1417.

[2] R. Balasubramanian and M. Ram Murty, Elliptic pseudoprimes, II, *Seminaire de Theorie des nombres, Paris 1988-89*, ed. C. Goldstein, Birkhaüser-Verlag, to appear.

[3] E.R. Berlekamp, Factoring polynomials over large finite fields, *Math. Comp.*, **24** (1970), pp. 713-735.

[4] G. Cornacchia, Su di un metodo per la risoluzione in numeri interi dell' equazione $\sum_{h=0}^{n} C_h x^{n-h} y^h = P$, *Giornale di Mathematiche di Battaglini*, **46** (1908), pp. 33-90.

[5] P. Erdös, P. Kiss and A. Sárközy, A lower bound for the counting function of Lucas pseudoprimes, *Math. Comp.* **41** (1988), pp. 315-323.

[6] D.M. Gordon, Pseudoprimes on elliptic curves, *Math. Comp.* **52** (1989), pp. 231-245

[7] R. Gupta and M. Ram Murty, Primitive points on elliptic curves, *Compositio Mathematica* **58** (1986), pp. 13-44.

[8] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, Fourth edition, Clarendon Press, Oxford, 1965.

[9] B.W. Jones, *The Arithmetic Theory of Quadratic Forms*, Mathematical Association of America, Baltimore, 1950.

[10] S. Lang, *Elliptic Curves: Diophantine Analysis*, Springer-Verlag, Heidelberg, 1978.

[11] H.W. Lenstra, Jr., Elliptic curves and number-theoretic algorithms, *Proc. Int. Congress Math. (Berkeley, 1986)*, American Mathematical Society, Providence, 1987, pp. 99-120.

[12] C. Pomerance, On the distribution of pseudoprimes, *Math. Comp.* **37** (1981), pp. 587-593.

[13] C. Pomerance, A new lower bound for the pseudoprime counting function, *Illinois J. Math.* **26** (1982), pp. 4-9.

[14] C. Pomerance, Two methods in elementary analytic number theory, *Number Theory and Applications*, ed. R.A. Mollin, Kluwer Academic Publishers, The Netherlands, 1989, pp. 135-161.

[15] R. Schoof, Elliptic curves over finite fields and the computation of square roots mod p, *Math. Comp.* **44** (1985), pp. 483-494.

[16] M. Ward, Memoir on elliptic divisibility sequences, *Amer. J. Math.*, 70 (1948), pp. 31-74.

Department of Computer Science
University of Georgia
Athens, GA 30602

and

Department of Mathematics
University of Georgia
Athens, GA 30602