

PSEUDOPRIMES IN CERTAIN LINEAR RECURRENCES

FLORIAN LUCA AND IGOR E. SHPARLINSKI

(Communicated by T. Shaska)

ABSTRACT. Let $b > 1$ be a fixed positive integer. We study the distribution of pseudoprimes to base b in certain linear recurrence sequences. We prove, in effective form, that most terms of these sequences are not pseudoprimes to base b .

1. INTRODUCTION

1.1. **Motivation.** Let $b \geq 2$ be an integer. Recall that a *pseudoprime to base b* is a composite positive integer m such that the congruence $b^m \equiv b \pmod{m}$ holds. The question of the distribution of pseudoprimes in certain sequences of positive integers has received some interest lately. For example, van der Poorten and Rotkiewicz show that any arithmetic progression $a \pmod{d}$ with a and d coprime contains infinitely many pseudoprimes to base b ; see [9] for details. Pseudoprime to base b values of the Fibonacci numbers, polynomials and the Euler function have been studied in [7], while pseudoprime Cullen and Woodall numbers are analyzed in [8]. In a recent paper, the authors jointly with Cojocaru, fixed an elliptic curve \mathbf{E} defined over \mathbb{Q} and studied the primes p such that the reductions of \mathbf{E} modulo p are base b pseudoprimes (see [2]).

Note that Fibonacci Cullen and Woodall numbers as well as polynomials, are all examples of linearly recurrence sequences. In this paper, we continue this program and look at the presence of pseudoprimes in linear recurrence sequences of certain general types. One application of our results is an upper bound on the number of pseudoprimes amongst the numbers of \mathbb{F}_{q^n} -rational points on a given elliptic curve over a finite field \mathbb{F}_q of q elements for $n \leq x$.

1.2. **The set up.** Let $u = (u_n)_{n \geq 0}$ be a linear recurrence sequence of integers satisfying a homogeneous linear recurrence relation

$$u_{n+k} = a_1 u_{n+k-1} + \cdots + a_{k-1} u_{n+1} + a_k u_n, \quad n = 1, 2, \dots,$$

with the characteristic polynomial

$$\psi(X) = X^k - a_1 X^{k-1} - \cdots - a_{k-1} X - a_k \in \mathbb{Z}[X].$$

Received by the editors June 13, 2007, and in revised form, July 14, 2007.

2000 *Mathematics Subject Classification.* 11N25, 11N37, 11A07.

Key words and phrases. pseudoprimes, linear recurrences.

We assume, without loss of generality, that $a_k \neq 0$. It is then well-known that

$$u_n = \sum_{i=1}^m A_i(n) \alpha_i^n,$$

where $\alpha_1, \dots, \alpha_m$ are the distinct roots of $\psi(X)$, of multiplicities $\sigma_1, \dots, \sigma_m$, respectively, and $A_i(X)$ are polynomials of degrees $\sigma_i - 1$ for $i = 1, \dots, m$, with coefficients in $\mathbb{K} = \mathbb{Q}[\alpha_1, \dots, \alpha_m]$.

We recall that $\alpha_1, \dots, \alpha_m$ are also called the *characteristic roots*. Further, assume that $(u_n)_{n \geq 0}$ is nondegenerate, namely that α_i/α_j is not a root of 1 for any $1 \leq i < j \leq m$. It is well-known that there exist only finitely many n such that $u_n = 0$ (see, for example, [10] for a bound on the number of such n). From now on, we may assume that $n > n_0$ is large so that $u_n \neq 0$.

We refer to [4] for these and other known facts about linear recurrence sequences.

In this paper, we study the number $N_{b,u}(x)$ of positive integers $n \leq x$ such that u_n is a base b pseudoprime where the sequence $u = (u_n)_{n \geq 0}$ satisfies one additional condition.

1.3. Divisibility sequences. Throughout the paper, we always assume that the sequence $(u_n)_{n \geq 0}$ is a *divisibility sequence*, that is, $u_m \mid u_n$ whenever $m \mid n$.

By the main result in [1] (see also [3] for a more general result), we know that $u_n \mid w_n$, where $(w_n)_{n \geq 0}$ is a recurrence whose general term has the shape

$$(1) \quad w_n = an^h \prod_{j=1}^s \frac{\beta_j^n - \gamma_j^n}{\beta_j - \gamma_j}$$

for some constants $a \in \mathbb{K}$, integer $h \geq 0$, and algebraic integers β_j, γ_j for $j = 1, \dots, s$ such that β_j/γ_j is not a root of unity for any $j = 1, \dots, s$. An immediate consequence of this representation is that

$$(2) \quad u_n = n^{h_0} v_n,$$

where $(v_n)_{n \geq 0}$ is a linear recurrence sequence having only simple roots and $h_0 \geq 0$ is some integer. It is also clear that the sequence $(v_n)_{n \geq 0}$ is of order at most k .

Note also that $h_0 = 0$ if and only if the characteristic polynomial $\Psi(X)$ of $(u_n)_{n \geq 0}$ has no multiple roots.

1.4. Examples. Let $(F_n)_{n \geq 0}$ be the Fibonacci sequence given by $F_0 = 0$, $F_1 = 1$ and $F_{n+2} = F_{n+1} + F_n$ for all $n \geq 0$. It is well-known that the sequence $(F_n)_{n \geq 0}$ is a divisibility sequence. In fact, $F_n = w_n$, where $(w_n)_{n \geq 0}$ is given by formula (1) with $s = 1$ and β_1, γ_1 are the golden section and its conjugate. In particular, it follows from our general results, that the set of n such that F_n is a base b pseudoprime is of asymptotic density zero. As we have mentioned, this is already proved in [7]. The same remarks apply to the Pell sequence $(P_n)_{n \geq 0}$ given by $P_0 = 0$, $P_1 = 1$ and $P_{n+2} = 2P_{n+1} + P_n$ for all $n \geq 0$. Our results show that even the product $F_n P_n$ is a base b pseudoprime only for a set of n of asymptotic density zero.

The Cullen and Woodall numbers, denoted C_n and D_n , respectively, are given by $C_n = n2^n + 1$ and $D_n = n2^n - 1$ for all $n \geq 1$. The sequences $(C_n)_{n \geq 0}$ and $(D_n)_{n \geq 0}$ are ternary recurrent of common characteristic polynomial $\psi(X) = (X-1)^2(X-2)$. However, none of them is a divisibility sequence so our general result does not apply to this sequence. However, using different arguments, it has been shown in [8] that

the set of n such that C_n or D_n is a base b pseudoprime is of asymptotic density zero.

The sequence of values of the Euler functions $(\varphi(n))_{n \geq 1}$ is a divisibility sequence because $\varphi(m) \mid \varphi(n)$ for all $m \mid n$, but it is not linearly recurrent. Nevertheless, it is shown in [7] that the set of n such that $\varphi(n)$ is a base b pseudoprime is of asymptotic density zero.

Let q be a prime power and let \mathbf{E} be an ordinary elliptic curve defined over a finite field of q elements \mathbb{F}_q . Let $m(n)$ be the number of points on \mathbf{E} defined over \mathbb{F}_{q^n} . Then both sequences $(m(n))_{n \geq 1}$ and $(m(n)/m(1))_{n \geq 1}$ are divisibility sequences. Indeed, $m(n) = (\tau^n - 1)(\bar{\tau}^n - 1)$, where τ and $\bar{\tau}$ are the two eigenvalues of the Frobenius. In the non-supersingular case, we know that $\tau/\bar{\tau}$ is not a root of 1 (see, for example, [6, Lemma 5]), therefore our results show that each one of the numbers $m(n)$ and $m(n)/m(1)$ is a base b pseudoprime only for a set of n of asymptotic density zero. This compliments the results of [2], where it is shown that for a fixed elliptic curve \mathbf{E} over \mathbb{Q} , under some natural assumptions, the set of primes p such that the reductions of \mathbf{E} modulo p are base b pseudoprimes forms a subset of primes of relative density zero (in the set of all primes).

1.5. Notation. Throughout this paper, for any positive real number x and any integer $\ell \geq 1$, we write $\log_\ell x$ for the function defined inductively by $\log_1 x = \max\{\ln x, 1\}$, where $\ln x$ is the natural logarithm of x , and $\log_\ell x = \log_1(\log_{\ell-1} x)$ for $\ell > 1$. When $\ell = 1$, we omit the subscript in order to simplify the notation; however, we continue to assume that $\log x \geq 1$ for any $x > 0$.

We use the Landau symbol O and the Vinogradov symbols \ll and \gg with their usual meanings, with the understanding that any implied constants depend on our data such as the sequence $(u_n)_{n \geq 0}$ and the number b . We recall that the notations $A \ll B$, $B \gg A$ and $A = O(B)$ are all equivalent to the fact that there exists a constant c such that the inequality $|A| \leq cB$ holds for all sufficiently large values of the input.

We always use the letters p and q to denote prime numbers, while m and n always denote positive integers.

1.6. Congruences with linear recurrence sequences. We make use of the following bound from [11] (see also [4, Theorem 5.11]).

Lemma 1. *Let $m \geq 2$ be an integer coprime to infinitely many elements of a nondegenerate linear recurrence sequence $(w_n)_{n \geq 0}$ of order k . Then for any integer $N \geq 1$ the number $R(N, m)$ of solutions of the congruence*

$$u(n) \equiv 0 \pmod{m}, \quad 0 \leq n \leq N - 1,$$

satisfies the bound

$$R(N, m) \leq C(k)(N/\log m + 1),$$

where $C(k)$ depends only on k .

2. MAIN RESULTS

2.1. Characteristic polynomial with multiple roots. Here we consider the case when $h_0 > 0$ in the representation (2).

Theorem 2. *Assume that a nondegenerate linear recurrence sequence $(u_n)_{n \geq 0}$ is a divisibility sequence. If $h_0 > 0$ in the representation (2), then*

$$N_{b,u}(x) \ll \frac{x \log_2 x \log_3 x}{\log x}.$$

Proof. We let x be large and set

$$(3) \quad w = (\log x)^2, \quad y = x^{1/(k+2)} \quad \text{and} \quad z = \exp(24 \log_2 x \log_3 x).$$

For a prime number p coprime to b we let $t(p)$ denote the multiplicative order of b modulo a prime p . We let \mathcal{Q} be the set of primes $p \in [z, y]$ with $t(p) \leq p^{1/3}$. It is shown in the proof of Theorem 1 in [8] that

$$(4) \quad \sum_{p \in \mathcal{Q}} \frac{1}{p} \ll \frac{1}{z^{1/3}} \ll 1.$$

For an integer m we write $P(m)$ for the largest prime divisor of m with the convention that $P(0) = P(\pm 1) = 1$. For a prime p we put $q_p = P(t(p))$.

We define \mathcal{R} as the set of primes $p \in [y, z] \setminus \mathcal{Q}$ with $q_p \leq w$. Clearly, each prime $p \in \mathcal{R}$ has the property that $p-1$ has a divisor $d \geq y^{1/3}$ with $P(d) \leq w$. Therefore,

$$\sum_{p \in \mathcal{R}} \frac{1}{p} \ll x \sum_{\substack{y^{1/3} \leq d \leq z \\ P(d) < w}} \sum_{\substack{p < d^3 \\ p \equiv 1 \pmod{d}}} \frac{1}{p}.$$

The arguments used in the proof of Theorem 2 in [8] lead easily to the bound

$$(5) \quad \sum_{p \in \mathcal{R}} \frac{1}{p} \ll \frac{(\log_2 x)^2}{z^{(\log w)/6}} = 1.$$

We let \mathcal{P} be the set of primes $p \in [z, y] \setminus (\mathcal{Q} \cup \mathcal{R})$. We let \mathcal{E} be the set of positive integers $n \leq x$ which do not have a divisor $p \in \mathcal{P}$.

By the Brun sieve inequality (see Theorem 2.2 in [5]), we have

$$\#\mathcal{E} \ll x \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

Using (4) and (5), we obtain

$$(6) \quad \#\mathcal{E} \ll x \prod_{p \in [z, y]} \left(1 - \frac{1}{p}\right).$$

By Mertens's formula (see [12] for a better error term) for a positive real number t we have

$$\prod_{p \leq t} \left(1 - \frac{1}{p}\right) = e^\gamma \log t \left(1 + O\left(\frac{1}{\log t}\right)\right).$$

Applying this with $t = y$ and $t = z$ and dividing the two relations obtained in this way we get, by estimate (6),

$$(7) \quad \#\mathcal{E} \ll x \frac{\log z}{\log y}.$$

We now let \mathcal{N} be the set of positive integers $n \leq x$ which are not in \mathcal{E} . Each positive integer $n \in \mathcal{N}$ has a prime factor $p \in \mathcal{P}$. To get an upper bound on $\#\mathcal{N}$,

it suffices, for every $p \in \mathcal{P}$, to count the number $M_{b,u}(x,p)$ of $m \leq x/p$ such that $n = mp \in \mathcal{N}$ and the congruence

$$b^{u_n} - b \equiv 0 \pmod{u_n}$$

holds. Since $h_0 > 0$, we see that

$$p \mid n \mid u_n \mid b(b^{u_n-1} - 1).$$

Clearly, if x is large enough, then

$$(8) \quad \gcd(q_p, a_k b D) = 1$$

for all $p \in \mathcal{P}$, where we recall that $a_k = \psi(0)$ is the constant term of the characteristic polynomial $\psi(X)$ of the sequence $(u_n)_{n \geq 0}$, and D is the product of the discriminants of the irreducible factors of ψ . Since $p \mid b^{u_n-1} - 1$, we get

$$q_p \mid t(p) \mid u_n - 1 = n^{h_0} v_n - 1.$$

Since $q_p \mid p - 1$ and $n = mp$, we derive that

$$(9) \quad m^{h_0} v_{pm} \equiv 1 \pmod{q_p}.$$

The classical theory of linear recurrence sequences (see [4]) implies that under the condition (8) the sequence $(v_{pm})_{m \geq 1}$ whose order is at most k is purely periodic modulo q_p with some period $T_p \leq q_p^k - 1$. Furthermore, we also have the divisibility

$$T_p \mid \prod_{\nu=1}^k (q_p^\nu - 1),$$

which in turn implies that

$$(10) \quad \gcd(T_p, q_p) = 1.$$

Thus, if we write $m = r + T_p s$ with some integers r and s such that $0 \leq r < T_p$ and $0 \leq s \leq x/pT_p$, then (9) implies that

$$(r + T_p s)^{h_0} v_{pr} \equiv 1 \pmod{q_p}.$$

Thanks to the condition (10), we see that the last congruence tells us that s belongs to at most h_0 arithmetic progressions modulo q_p . Namely, this is a polynomial congruence for s modulo q_p of degree exactly h_0 since p does not divide its leading term $T_p^{h_0} v_{pr}$. Thus, s may take at most $h_0(x/(pT_p q_p) + 1)$ possible values in the interval $[0, x/pT_p]$, leading to the bound

$$(11) \quad M_{b,u}(x,p) \leq T_p h_0 \left(\frac{x}{pT_p q_p} + 1 \right).$$

Notice that due to our choice of parameters,

$$pT_p q_p < pq_p^{k+1} < p^{k+2} \leq y^{k+2} = x.$$

Hence, the bound (11) simplifies to

$$M_{b,u}(x,p) \ll \frac{x}{pq_p}.$$

Using (7), we obtain

$$\begin{aligned} N_{b,u}(x) &\leq \#\mathcal{E} + \sum_{p \in \mathcal{P}} M_{b,u}(x, p) \ll x \frac{\log z}{\log y} + x \sum_{p \in \mathcal{P}} \frac{1}{pq_p} \\ &\ll x \frac{\log z}{\log y} + \frac{x}{w} \sum_{p \in \mathcal{P}} \frac{1}{p} \ll x \left(\frac{\log z}{\log y} + \frac{\log_2 y}{w} \right). \end{aligned}$$

Recalling the choice of w , y and z from (3), we obtain the desired result. \square

2.2. Characteristic polynomial without multiple roots. Here, we consider the case when $h_0 = 0$ in the representation (2) and get a slightly weaker result than in Theorem 2.

Theorem 3. *Assume that a nondegenerate linear recurrence sequence $(u_n)_{n \geq 0}$ is a divisibility sequence. If $h_0 = 0$ in the representation (2), then*

$$N_{b,u}(x) \ll x \frac{\log_3 x}{\sqrt{\log x}}.$$

Proof. We let again x be large and we now set

$$(12) \quad w = \exp\left(\sqrt{\log_2 x}\right), \quad y = x^{1/(k+2)}, \quad z = \exp\left(12\sqrt{\log_2 x} \log_3 x\right).$$

We redefine the sets of primes $\mathcal{P}, \mathcal{Q}, \mathcal{R}$ as well as the sets of integers \mathcal{E}, \mathcal{N} as in the proof of Theorem 2 but with the current choice of parameters w , y and z . Since we still have that

$$z \rightarrow \infty \quad \text{and} \quad z^{-1/6 \log w} (\log_2 x)^2 = 1,$$

as $x \rightarrow \infty$, the bound (7) holds for our new choice of parameters as well.

To estimate $M_{b,u}(x, p)$, we note that congruence (9) now becomes

$$(13) \quad u_{pm} \equiv 1 \pmod{q_p}$$

Note also that that $u_n = v_n$, for $n = 1, 2, \dots$, because $h_0 = 0$.

We now apply the bound of Lemma 1 to estimate the number of solutions of the congruence (13) with the linear recurrence sequence $(u_{np} - 1)_{n \geq 0}$, whose roots are now $\alpha_1^p, \dots, \alpha_k^p$ and 1. We note that if $\alpha_1^p, \dots, \alpha_k^p$ are not roots of unity, then Lemma 1 applies directly. Otherwise, if one of them is a root of unity ρ then all its conjugates must also be among $\alpha_1^p, \dots, \alpha_k^p$. However, since there are no roots of unity among α_i/α_j for $1 \leq i < j \leq m$, then $\rho = \pm 1$. Thus, $u_{np} = w_n + Ap^n$ where $(w_n)_{n \geq 0}$ is a linear recurrence sequence which has no roots of unity among its characteristic roots and their ratios. Thus, $(u_{2np} - 1)_{n \geq 0}$ and $(u_{(2n+1)p} - 1)_{n \geq 0}$ are nondegenerate linear recurrence sequences to which Lemma 1 applies.

Therefore,

$$M_{b,u}(x, p) \ll \frac{x}{p \log q_p} + 1 \ll \frac{x}{p \log q_p}.$$

Once again, recalling estimate (7), we obtain

$$\begin{aligned} N_{b,u}(x) &\leq \#\mathcal{E} + \sum_{p \in \mathcal{P}} M_{b,u}(x, p) \ll x \frac{\log z}{\log y} + x \sum_{p \in \mathcal{P}} \frac{1}{p \log q_p} \\ &\ll x \frac{\log z}{\log y} + \frac{x}{\log w} \sum_{p \in \mathcal{P}} \frac{1}{p} \ll x \left(\frac{\log z}{\log y} + \frac{\log_2 y}{\log w} \right). \end{aligned}$$

Recalling our choice of w , y and z from (12), we obtain the desired result. \square

ACKNOWLEDGEMENT

We thank the referee for suggestions which improved the quality of the manuscript. Research of F. L. was supported by grant SEP-CONACyT 46755 that of I. S. by ARC grant DP0556431.

REFERENCES

- [1] J.-P. Bézivin, A. Pethö and A. J. van der Poorten, ‘A full characterisation of divisibility sequences’, *Amer. J. Math.* **112** (1990), no. 6, 985–1001.
- [2] A. C. Cojocaru, F. Luca and I. E. Shparlinski, ‘Pseudoprime reductions of elliptic curves’, *Preprint*, 2007.
- [3] P. Corvaja and U. Zannier, ‘Finiteness of integral values for the ratio of two linear recurrences’, *Invent. Math.* **149** (2002), no. 2, 431–451.
- [4] G. Everest, A. van der Poorten, I. E. Shparlinski and T. Ward, *Recurrence sequences*, Mathematical Surveys and Monographs, **104**, American Mathematical Society, Providence, RI, 2003.
- [5] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, London, 1974.
- [6] F. Luca and I. E. Shparlinski, ‘On the exponent of the group of points on elliptic curves in extension fields’, *Int. Math. Res. Not.* **2005**, no. 23, 1391–1409.
- [7] F. Luca and I. E. Shparlinski, ‘Pseudoprime values of the Fibonacci sequence, polynomials and the Euler function’, *Indag. Math.* **17** (2006), 611–625.
- [8] F. Luca and I. E. Shparlinski, ‘Pseudoprime Cullen and Woodall numbers’, *Colloq. Math.* **107** (2007), 35–43.
- [9] A. J. van der Poorten and A. Rotkiewicz, ‘On strong pseudoprimes in arithmetic progressions’, *J. Austral. Math. Soc. Ser. A* **29** (1980), no. 3, 316–321.
- [10] H. P. Schlickewei and W. M. Schmidt, ‘The number of solutions of polynomial-exponential equations’, *Compositio Math.* **120** (2000), no. 2, 193–225.
- [11] I. E. Shparlinski, ‘The number of different prime divisors of recurrence sequences’, *Matem. Zametki* **42** (1987), 494–507 (in Russian).
- [12] A. I. Vinogradov, ‘On the remainder in Mertens’s formula,’ *Dokl. Akad. Nauk SSSR* **148** (1963), 262–263 (in Russian).

FLORIAN LUCA, INSTITUTO DE MATEMÁTICAS, UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO,
C.P. 58089, MORELIA, MICHOACÁN, MÉXICO
E-mail address: `fluca@matmor.unam.mx`

IGOR E. SHPARLINSKI, DEPT. OF COMPUTING, MACQUARIE UNIVERSITY, SYDNEY, NSW 2109,
AUSTRALIA
E-mail address: `igor@ics.mq.edu.au`