

On numbers n dividing the n th term of a linear recurrence

JUAN JOSÉ ALBA GONZÁLEZ

Instituto de Matemáticas,
Universidad Nacional Autónoma de México,
C.P. 58089, Morelia, Michoacán, México
jjalba@gmail.com

FLORIAN LUCA

Instituto de Matemáticas,
Universidad Nacional Autónoma de México,
C.P. 58089, Morelia, Michoacán, México
fluca@matmor.unam.mx

CARL POMERANCE

Mathematics Department, Dartmouth College,
Hanover, NH 03755, USA
carl.pomerance@dartmouth.edu

IGOR E. SHPARLINSKI

Dept. of Computing, Macquarie University
Sydney, NSW 2109, Australia
igor.shparlinski@mq.edu.au

Abstract

Here, we give upper and lower bounds on the count of positive integers $n \leq x$ dividing the n th term of a nondegenerate linearly recurrent sequence with simple roots.

1 Introduction

Let $\{u_n\}_{n \geq 0}$ be a linear recurrence sequence of integers satisfying a homogeneous linear recurrence relation

$$u_{n+k} = a_1 u_{n+k-1} + \cdots + a_{k-1} u_{n+1} + a_k u_n, \quad \text{for } n = 0, 1, \dots, \quad (1)$$

where a_1, \dots, a_k are integers with $a_k \neq 0$.

In this paper, we study the set of indices n which divide the corresponding term u_n ; that is, the set:

$$\mathcal{N}_u := \{n \geq 1 : n \mid u_n\}.$$

But first, some background on linear recurrence sequences.

To the recurrence (1), we associate its *characteristic polynomial*

$$f_u(X) := X^k - a_1 X^{k-1} - \cdots - a_{k-1} X - a_k = \prod_{i=1}^m (X - \alpha_i)^{\sigma_i} \in \mathbb{Z}[X],$$

where $\alpha_1, \dots, \alpha_m \in \mathbb{C}$ are the distinct roots of $f_u(X)$ with multiplicities $\sigma_1, \dots, \sigma_m$, respectively. It is then well-known that the general term of the recurrence can be expressed as

$$u_n = \sum_{i=1}^m A_i(n) \alpha_i^n, \quad \text{for } n = 0, 1, \dots, \quad (2)$$

where $A_i(X)$ are polynomials of degrees at most $\sigma_i - 1$ for $i = 1, \dots, m$, with coefficients in $K := \mathbb{Q}[\alpha_1, \dots, \alpha_m]$. We refer to [6] for this and other known facts about linear recurrence sequences.

For upper bounds on the distribution of \mathcal{N}_u , the case of a linear recurrence with multiple roots already can pose problems (but see below). For example, the sequence of general term $u_n = n2^n$ for all $n \geq 0$ having characteristic polynomial $f_u(X) = (X - 2)^2$ shows that \mathcal{N}_u may contain all the positive integers. So, we look at the case when $f_u(X)$ has only simple roots. In this case, the relation (2) becomes

$$u_n = \sum_{i=1}^k A_i \alpha_i^n, \quad \text{for } n = 0, 1, \dots \quad (3)$$

Here, A_1, \dots, A_k are constants in K . We may assume that none of them is zero, since otherwise, a little bit of Galois theory shows that the integer sequence $\{u_n\}_{n \geq 0}$ satisfies a linear recurrence of a smaller order.

We remark in passing that there is no real obstruction in reducing to the case of the simple roots. Indeed, let $D \in \mathbb{N}$ be a common denominator of all the coefficients of all the polynomials $A_i(X)$ for $i = 1, \dots, m$. That is, the coefficients of each DA_i are algebraic integers. Then

$$Du_n = \sum_{i=1}^m DA_i(0)\alpha_i^n + \sum_{i=1}^m D(A_i(n) - A_i(0))\alpha_i^n.$$

If $n \in \mathcal{N}_u$, then $n \mid Du_n$. Since certainly n divides¹ the algebraic integer

$$\sum_{i=1}^m D(A_i(n) - A_i(0))\alpha_i^n,$$

it follows that n divides $\sum_{i=1}^m DA_i(0)\alpha_i^n$. If this is identically zero (i.e., $A_i(0) = 0$ for all $i = 1, \dots, m$), then we are in an instance similar to the instance of the sequence of general term $u_n = n2^n$ for all $n \geq 0$ mentioned above. In this case, \mathcal{N}_u contains at least a positive proportion of all the positive integers (namely, all n coprime to D). Otherwise, we may put

$$w_n = \sum_{i=0}^m DA_i(0)\alpha_i^n \quad \text{for } n = 0, 1, \dots$$

A bit of Galois theory shows that w_n is an integer for all $n \geq 0$, and the sequence $\{w_n\}_{n \geq 0}$ satisfies a linear recurrence relation of order $\ell := \#\{1 \leq i \leq m : A_i(0) \neq 0\}$ with integer coefficients, which furthermore has only simple roots. Hence, $\mathcal{N}_u \subseteq \mathcal{N}_w$, and therefore there is indeed no loss of generality when proving upper bounds in dealing only with linear recurrent sequences with distinct roots.

We put

$$\Delta_u := \prod_{1 \leq i < j \leq k} (\alpha_i - \alpha_j)^2 = \text{disc}(f_u) \tag{4}$$

¹Here, for two algebraic integers α and β and a positive integer m we write $\alpha \equiv \beta \pmod{m}$ to mean that $(\alpha - \beta)/m$ is an algebraic integer. When $\beta = 0$ we say that m divides α .

for the (nonzero) discriminant of the sequence $\{u_n\}_{n \geq 0}$, or of the polynomial $f_u(X)$. It is known that Δ_u is an integer. We also assume that $(u_n)_{n \geq 0}$ is nondegenerate, which means that α_i/α_j is not a root of 1 for any $1 \leq i < j \leq m$. Throughout the remainder of this paper, all linear recurrences have only simple roots and are nondegenerate.

When $k = 2$, $u_0 = 0$, $u_1 = 1$ and $\gcd(a_1, a_2) = 1$, the sequence $\{u_n\}_{n \geq 0}$ is called a *Lucas sequence*. The formula (3) of its general term is

$$u_n = \frac{\alpha_1^n - \alpha_2^n}{\alpha_1 - \alpha_2}, \quad \text{for } n = 0, 1, \dots \quad (5)$$

That is, we can take $A_1 = 1/(\alpha_1 - \alpha_2)$ and $A_2 = -1/(\alpha_1 - \alpha_2)$ in the formula of the general term (3). In the case of a Lucas sequence $(u_n)_{n \geq 0}$, the fine structure of the set \mathcal{N}_u has been described in [18] (see the references therein and particularly [8]). We also note that divisibility of terms of a linear recurrence sequence by arithmetic functions of their index have been studied in [13] (see also [12] for the special case of Fibonacci numbers).

For a set \mathcal{A} and a positive real number x we put $\mathcal{A}(x) = \mathcal{A} \cap [1, x]$. Throughout the paper, we study upper and lower bounds for the number $\#\mathcal{N}_u(x)$. In particular, we prove that \mathcal{N}_u is of asymptotic density zero.

Observe first that if $k = 1$, then $u_n = Aa_1^n$ holds for all $n \geq 0$ with some integers $A \neq 0$ and $a_1 \notin \{0, \pm 1\}$. Its characteristic polynomial is $f_u(X) = X - a_1$. It is easy to see that in this case $\#\mathcal{N}_u(x) = O((\log x)^{\omega(|a_1|)})$, where for an integer $m \geq 2$ we use $\omega(m)$ for the number of distinct prime factors of m . So, from now on, we assume that $k \geq 2$.

Note next that for the sequence of general term $u_n = 2^n - 2$ for all $n \geq 0$ having characteristic polynomial $f_u(X) = (X - 1)(X - 2)$, Fermat's little theorem implies that every prime is in \mathcal{N}_u , so that the Prime Number Theorem and estimates for the distribution of pseudoprimes² show that it is possible for the estimate $\#\mathcal{N}_u(x) = (1 + o(1))x/\log x$ to hold as $x \rightarrow \infty$. However, we show that $\#\mathcal{N}_u(x)$ cannot have a larger order of magnitude.

Theorem 1. *For each $k \geq 2$, there is a positive constant $c_0(k)$ depending only on k such that if the characteristic polynomial of a nondegenerate linear recurrence sequence $\{u_n\}_{n \geq 0}$ of order k has only simple roots, then the*

²A pseudoprime is a composite number n which divides $2^n - 2$. The paper [14] shows that there are few odd pseudoprimes compared with primes, while [10] (unpublished) does the same for even pseudoprimes.

estimate

$$\#\mathcal{N}_u(x) \leq c_0(k) \frac{x}{\log x}$$

holds for x sufficiently large.

In case of a Lucas sequence, we have a better bound. To simplify notations, for a positive integer ℓ we define $\log_\ell x$ iteratively as $\log_1 x := \max\{\log x, 1\}$ and for $\ell > 1$ as $\log_\ell x := \log_1(\log_{\ell-1} x)$. When $\ell = 1$ we omit the index but understand that all logarithms are ≥ 1 . Let

$$L(x) := \exp\left(\sqrt{\log x \log_2 x}\right). \quad (6)$$

Theorem 2. *Assume that $\{u_n\}_{n \geq 0}$ is a Lucas sequence. Then the inequality*

$$\#\mathcal{N}_u(x) \leq \frac{x}{L(x)^{1+o(1)}} \quad (7)$$

holds as $x \rightarrow \infty$.

It follows from a result of Somer [19, Theorem 8] that \mathcal{N}_u is finite if and only if $\Delta_u = 1$, and in this case $\mathcal{N}_u = \{1\}$.

For Lucas sequences with $a_2 = \pm 1$, we also have a rather strong lower bound on $\#\mathcal{N}_u(x)$. Our result depends on the current knowledge of the distribution of y -smooth values of $p^2 - 1$ for primes p , that is values of $p^2 - 1$ which do not have prime divisors exceeding y . We use $\Pi(x, y)$ to denote the number of primes $p \leq x$ for which $p^2 - 1$ is y -smooth. Since the numbers $p^2 - 1$ with p prime are likely to behave as “random” integers from the point of view of the size of their prime factors, it seems reasonable to expect that behavior of $\Pi(x, y)$ resembles the behavior of the counting function for smooth integers. We record this in a very relaxed form of the assumption that for some fixed real $v \geq 1$ we have

$$\Pi(y^v, y) \geq y^{v+o(1)} \quad (8)$$

as $y \rightarrow \infty$. In fact, a general result from [3, Theorem 1.2] implies that (8) holds with any $v \in [1, 4/3)$.

Theorem 3. *There is a set of integers \mathcal{L} such that $\mathcal{L} \subset \mathcal{N}_u$ for any Lucas sequence u with $a_2 = \pm 1$, and such that if (8) holds with some $v > 1$, we have*

$$\#\mathcal{N}_u(x) \geq \#\mathcal{L}(x) \geq x^{\vartheta+o(1)}$$

as $x \rightarrow \infty$, where

$$\vartheta := 1 - 1/v.$$

In particular, since as we have already mentioned, any value of $v < 4/3$ is admissible, we can take

$$\vartheta = 1/4.$$

Furthermore, since (8) is expected to hold for *any* $v > 1$, it is very likely that the bound of Theorem 3 holds with $\vartheta = 1$.

Finally, we record a lower bound on $\#\mathcal{N}(x)$ when $a_2 \neq \pm 1$ but $\Delta_u \neq 1$.

Theorem 4. *Let $\{u_n\}_{n \geq 0}$ be any Lucas sequence with $\Delta_u \neq 1$. Then there exist positive constants c_1 and x_0 depending on the sequence such that for $x > x_0$ we have*

$$\#\mathcal{N}_u(x) > \exp(c_1(\log_2 x)^2).$$

Throughout the paper, we use x for a large positive real number. We use the Landau symbol O and the Vinogradov symbol \ll with the usual meaning in analytic number theory. The constants implied by them may depend on the sequence $\{u_n\}_{n \geq 0}$, or only on k . We use c_0, c_1, \dots for positive constants which may depend on $\{u_n\}_{n \geq 0}$. We label such constants increasingly as they appear in the paper.

2 Preliminary results

As in the proof of [6, Theorem 2.6], put

$$D_u(x_1, \dots, x_k) := \det(\alpha_i^{x_j})_{1 \leq i, j \leq k}.$$

For a prime number p not dividing a_k , let $T_u(p)$ be the maximal nonnegative integer T with the property that

$$p \nmid \prod_{0 \leq x_2, \dots, x_k \leq T} \max\{1, |N_{K/\mathbb{Q}}(D_u(0, x_2, \dots, x_k))|\}.$$

It is known that such T exists. In the above relation, x_2, \dots, x_k are integers in $[1, T]$, and for an element α of K we use $N_{K/\mathbb{Q}}(\alpha)$ for the norm of α over \mathbb{Q} . Since $\alpha_1, \dots, \alpha_k$ are algebraic integers in K , it follows that the numbers $N_{K/\mathbb{Q}}(D_u(0, x_2, \dots, x_k))$ are integers.

Observe that $T_u(p) = 0$ if and only if $k = 2$ and p is a divisor of $\Delta_u = (\alpha_1 - \alpha_2)^2$.

More can be said in the case when $\{u_n\}_{n \geq 0}$ is a Lucas sequence. In this case, we have

$$|N_{K/\mathbb{Q}}(D_u(0, x_2))| = |\alpha_2^{x_2} - \alpha_1^{x_2}|^2 = |\Delta_u|^2 |u_{x_2}|^2, \quad x_2 = 1, 2, \dots$$

Thus, if p does not divide the discriminant $\Delta_u = (\alpha_1 - \alpha_2)^2 = a_1^2 + 4a_2$ of the sequence $\{u_n\}_{n \geq 0}$, then $T_u(p) + 1$ is in fact the minimal positive integer ℓ such that $p \mid u_\ell$. This is sometimes called the *index of appearance* of p in $\{u_n\}_{n \geq 0}$ and is denoted by $z_u(p)$. The index of appearance $z_u(m)$ can be defined for composite integers m in the same way as above, namely as the minimal positive integer ℓ such that $m \mid u_\ell$. This exists for all positive integers m coprime to a_2 , and has the important property that $m \mid u_n$ if and only if $z_u(m) \mid n$. For any $\gamma \in (0, 1)$, let

$$\mathcal{P}_{u, \gamma} = \{p : T_u(p) < p^\gamma\}.$$

Lemma 1. *For $x^\gamma, y \geq 2$, the estimates*

$$\#\{p : T_u(p) \leq y\} \ll \frac{y^k}{\log y}, \quad \#\mathcal{P}_{u, \gamma}(x) \ll \frac{x^{k\gamma}}{\gamma \log x}$$

hold, where the implied constants depend only on the sequence $\{u_n\}_{n \geq 0}$.

Proof. It is clear that the second inequality follows immediately from the first with $y = x^\gamma$, so we prove only the first one. Suppose that $T_u(p) \leq y$. In particular, there exists a choice of integers x_2, \dots, x_k all in $[1, y + 1]$ such that

$$p \mid \max\{1, |N_{K/\mathbb{Q}}(D_u(0, x_2, \dots, x_k))|\}.$$

This argument shows that

$$\prod_{T_u(p) \leq y} p \mid \prod_{1 \leq x_2, \dots, x_k \leq y+1} \max\{1, |N_{K/\mathbb{Q}}(D_u(0, x_2, \dots, x_k))|\}. \quad (9)$$

There are at most $(y + 1)^{k-1} = O(y^{k-1})$ possibilities for the $(k - 1)$ -tuple (x_2, \dots, x_k) . For each one of these $(k - 1)$ -tuples, we have that

$$|N_{K/\mathbb{Q}}(D_u(0, x_2, \dots, x_k))| = \exp(O(y)).$$

Hence, the right hand side in (9) is of size $\exp(O(y^k))$. Taking logarithms in the inequality implied by (9), we get that

$$\sum_{T_u(p) \leq y} \log p = O(y^k).$$

If there are a total of n primes involved in this sum and if p_i denotes the i th prime, then

$$\sum_{i=1}^n \log p_i = O(y^k),$$

so that in the language of the prime number theorem, $\theta(p_n) \ll y^k$. It follows that $p_n \ll y^k$ and $n \ll y^k / (k \log y)$, which is what we wanted to prove. \square

The parameter $T_u(p)$ is useful to bound the number of solutions $n \in [1, x]$ of the congruence $u_n \equiv 0 \pmod{p}$ (see [17] and [6, Theorem 5.11]). The following result, whose proof is similar, relates $T_u(p)$ to the solutions to $u_{np} \equiv 0 \pmod{p}$.

Lemma 2. *There exists a constant $c_2(k)$ depending only on k with the following property. Suppose that $\{u_n\}_{n \geq 0}$ is a linearly recurrent sequence of order k satisfying the recurrence (1). Suppose that p is a prime coprime to $a_k \Delta_u$ and to the denominators of the numbers A_i in (3). Assume that there exists a positive integer s such that u_s is coprime to p . Then for any real $X \geq 1$ the number of solutions $R(X, p)$ of the congruence*

$$u_{np} \equiv 0 \pmod{p} \quad \text{with } 1 \leq n \leq X$$

satisfies the bound

$$R_u(X, p) \leq c_2(k) \left(\frac{X}{T_u(p)} + 1 \right).$$

Proof. By a result of Schlickewei [15] (see also [16]) there is a constant $C(k)$, depending only on k , such that for any $B_1, \dots, B_k \in K$, not all zero, the equation

$$\sum_{i=1}^k B_i \alpha_i^x = 0$$

has at most $C(k)$ solutions in positive integers x .

Let $w_n = u_{np}$, so that $\{w_n\}_{n \geq 0}$ is also a linearly recurrent sequence of order k that is clearly closely related to u . Note first that if $\alpha_1, \dots, \alpha_k$ are the characteristic roots of $\{u_n\}_{n \geq 0}$, then $\alpha_1^p, \dots, \alpha_k^p$ are the characteristic roots of $\{w_n\}_{n \geq 0}$. Hence,

$$f_w(X) = \prod_{i=1}^k (X - \alpha_i^p).$$

Observe that $T_u(p)$ exists because p does not divide a_k . Further, from the following calculation:

$$\begin{aligned} N_{K/\mathbb{Q}}(D_w(x_1, \dots, x_k)) &= N_{K/\mathbb{Q}}(\det(\alpha_i^{px_j})) \equiv N_{K/\mathbb{Q}}(\det(\alpha_i^{x_j})^p) \pmod{p} \\ &\equiv (N_{K/\mathbb{Q}}(D_u(x_1, \dots, x_k)))^p \pmod{p}, \end{aligned}$$

we conclude that if $0 < x_2 < \dots < x_k$ are any positive integers, then p divides $N_{K/\mathbb{Q}}(D_u(0, x_2, \dots, x_k))$ if and only if p divides $N_{K/\mathbb{Q}}(D_w(0, x_2, \dots, x_k))$.

Let \mathcal{I} be any interval of length $T_u(p)$ and $n_1 < \dots < n_\ell$ be all the integers $n \in \mathcal{I}$ such that $w_n \equiv 0 \pmod{p}$. Then we have

$$\sum_{i=1}^k A_i \alpha_i^{pn_j} \equiv 0 \pmod{p}, \quad j = 1, 2, \dots, \ell.$$

We re-write each congruence as

$$\sum_{i=1}^k (A_i \alpha_i^{pn_1}) \alpha_i^{p(n_j - n_1)} \equiv 0 \pmod{p}, \quad j = 1, 2, \dots, \ell. \quad (10)$$

Let π be any prime ideal dividing p in \mathcal{O}_K . We view the ‘‘unknowns’’ $A_i \alpha_i^{pn_1}$ in the residue ring $\mathcal{O}_K/\pi\mathcal{O}_K$. By assumption, no denominator of A_1, \dots, A_k is in π . Since u is not identically 0 \pmod{p} , not all A_i are in π , so the above solution $(A_i \alpha_i^{pn_1})$ is nonzero in $(\mathcal{O}_K/\pi\mathcal{O}_K)^k$.

Assume that

$$\ell \geq C(k) + k. \quad (11)$$

Put $x_1 = 0$ and out of the set $\mathcal{X}_2 = \{n_j - n_1 : j = 2, \dots, \ell\}$ choose $x_2 \in \mathcal{X}_2$ with

$$\det(\alpha_i^{x_j})_{1 \leq i, j \leq 2} \neq 0.$$

This is possible by the above result of Schlickewei [15] since $\#\mathcal{X}_2 = \ell - 1 \geq C(k) + k - 1 > C(k)$. For $k \geq 3$, put $\mathcal{X}_3 = \mathcal{X}_2 \setminus \{x_2\}$ and choose $x_3 \in \mathcal{X}_3$ with

$$\det(\alpha_i^{x_j})_{1 \leq i, j \leq 3} \neq 0,$$

which is still possible since $\#\mathcal{X}_3 = \ell - 2 \geq C(k) + k - 2 > C(k)$. and by the choice of x_2 this is a nontrivial exponential equation in x_3 . Continuing like this, after $k - 1$ steps, we obtain $x_2, \dots, x_k \in \mathcal{X}$ with

$$D_u(0, x_2, \dots, x_k) \neq 0, \quad (12)$$

However, by (10), we have $\pi \mid D_w(0, x_2, \dots, x_k)$, therefore

$$p \mid N_{K/\mathbb{Q}}(\pi) \mid N_{K/\mathbb{Q}}(D_w(0, x_2, \dots, x_k)),$$

which is impossible by the definition of $T_u(p)$ and the condition (12). Hence, the inequality (11) is false and the result follows. \square

When $\{u_n\}_{n \geq 0}$ is a Lucas sequence, we put

$$\mathcal{Q}_{u,\gamma} = \{p : z_u(p) \leq p^\gamma\}.$$

The remarks preceding Lemma 1 show that $\#\mathcal{Q}_{u,\gamma}(x) = \#\mathcal{P}_{u,\gamma}(x) + O(1)$. Hence, Lemma 1 implies the following result.

Lemma 3. *For $x > 1$, the estimate*

$$\#\mathcal{Q}_{u,\gamma}(x) \ll \frac{x^{2\gamma}}{\log x}$$

holds, where the implied constant depends only on the sequence $\{u_n\}_{n \geq 0}$.

As usual, we denote by $\Psi(x, y)$ the number of integers $n \leq x$ with $P(n) \leq y$. By [2, Corollary to Theorem 3.1], we have the following well-known result.

Lemma 4. *For $x \geq y > 1$, the estimate*

$$\Psi(x, y) = x \exp(-(1 + o(1))v \log v)$$

uniformly in the range $y > (\log x)^2$ as long as $v \rightarrow \infty$, where

$$v := (\log x)/(\log y).$$

3 The proof of Theorem 1

We assume that x is large. We split the set $\mathcal{N}_u(x)$ into several subsets. Let $P(n)$ be the largest prime factor of n and let $y := x^{1/\log \log x}$. Let

$$\begin{aligned} \mathcal{N}_1(x) &:= \{n \leq x : P(n) \leq y\}; \\ \mathcal{N}_2(x) &:= \{n \leq x : n \notin \mathcal{N}_1(x) \text{ and } P(n) \in \mathcal{P}_{u,1/(k+1)}\}; \\ \mathcal{N}_3(x) &:= \mathcal{N}(x) \setminus (\cup_{i=1}^2 \mathcal{N}_i(x)). \end{aligned}$$

We now bound the cardinalities of each one of the above sets.
For $\mathcal{N}_1(x)$, by Lemma 4, we obtain

$$\#\mathcal{N}_1(x) = \Psi(x, y) = x \exp(-(1 + o(1))v \log v) = o\left(\frac{x}{\log x}\right) \quad (13)$$

as $x \rightarrow \infty$, where

$$v = \frac{\log x}{\log y} = \log \log x.$$

Suppose now that $n \in \mathcal{N}_2(x)$. Then $n = pm$, where $p = P(n) \geq \max\{y, P(m)\}$. In particular, $p \leq x/m$ therefore $m \leq x/y$. Since we also have $p \in \mathcal{P}_{u, 1/(k+1)}(x/m)$, Lemma 1 implies that the number of such primes $p \leq x/m$ is $O\left((x/m)^{k/(k+1)}\right)$, where the implied constant depends on the sequence $\{u_n\}_{n \geq 0}$. Summing up the above inequality over all possible values of $m \leq x/y$, we get

$$\begin{aligned} \#\mathcal{N}_2(x) &\leq x^{k/(k+1)} \sum_{1 \leq m \leq x/y} \frac{1}{m^{k/(k+1)}} \ll x^{k/(k+1)} \int_1^{x/y} \frac{dt}{t^{k/(k+1)}} \\ &= ((k+1)x^{k/(k+1)}t^{1/(k+1)}) \Big|_1^{x/y} \ll \frac{x}{y^{1/(k+1)}}. \end{aligned} \quad (14)$$

Now let $n \in \mathcal{N}_3(x)$. As previously, we write $n = pm$, where $p = P(n) > y$. Thus, $m \leq x/p < x/y$. Since $n \in \mathcal{N}_u$, we have that $n \mid u_n$, therefore $p \mid u_n$. Furthermore, $T_u(p) \geq p^{1/(k+1)}$. We fix p and count the number of possibilities for m , using Lemma 2. We assume that x , hence y , hence p , is sufficiently large. Thus, $p \nmid u_s$ for some $s \geq 0$, since u is not identically 0. Further, we may assume the other hypotheses of Lemma 2 hold as well, using again that x is sufficiently large. Thus, by Lemma 2 the number of integers $m \leq x/p$ for which $u_{mp} \equiv 0 \pmod{p}$ satisfies

$$R_u(x/p, p) \leq c_2(k) \left(\frac{x}{pT_u(p)} + 1 \right). \quad (15)$$

Since $n \in \mathcal{N}_3(x)$, we have that $T_u(p) \geq p^{1/(k+1)}$. The inequality (15) now guarantees that the number of choices for m once p is fixed is

$$R_u(x/p, p) \leq c_2(k) \left(\frac{x}{p^{1+1/(k+1)}} + 1 \right).$$

To summarize, we have

$$\begin{aligned}
\mathcal{N}_3(x) &\leq \sum_{y \leq p \leq x} c_2(k) \left(\frac{x}{p^{1+1/(k+1)}} + 1 \right) \\
&\leq c_2(k) \left(\pi(x) + x \sum_{y \leq p} \frac{1}{p^{1+1/(k+1)}} \right) \\
&\leq c_2(k) \left(\pi(x) + x \int_y^\infty \frac{dt}{t^{1+1/(k+1)}} \right).
\end{aligned}$$

Therefore

$$\mathcal{N}_3(x) \leq c_2(k) \left(\pi(x) + O\left(\frac{(k+1)x}{y^{1/(k+1)}}\right) \right). \quad (16)$$

Comparing (13), (14) and (16), we get that

$$\#\mathcal{N}(x) \leq c_2(k)\pi(x) + \frac{x}{\exp((1+o(1))v \log v)} + O\left(\frac{x}{y^{1/(k+1)}}\right) \quad (17)$$

as $x \rightarrow \infty$, where the implied constant depends on the recurrence $\{u_n\}_{n \geq 0}$. By our choice of y as $x^{1/\log \log x}$, the second and third terms on the right side of (17) are both $o(\pi(x))$ as $x \rightarrow \infty$, so we have the theorem.

4 The proof of Theorem 2

We divide the numbers $n \in \mathcal{N}_u(x)$ into several classes:

- (i) $\mathcal{N}_1(x) := \{n \in \mathcal{N}_u(x) : P(n) \leq L(x)^{1/2}\}$;
- (ii) $\mathcal{N}_2(x) := \{n \in \mathcal{N}_u(x) : P(n) \geq L(x)^3\}$;
- (iii) $\mathcal{N}_3(x) := \mathcal{N}_u(x) \setminus (\mathcal{N}_1(x) \cup \mathcal{N}_2(x))$.

It follows from Lemma 4 that

$$\#\mathcal{N}_1(x) \leq \Psi(x, L(x)^{1/2}) \leq \frac{x}{L(x)^{1+o(1)}}$$

as $x \rightarrow \infty$.

For $n \in \mathcal{N}_u$ and $p \mid n$, we have $n \equiv 0 \pmod{p}$ and $n \equiv 0 \pmod{z_u(p)}$. For p not dividing the discriminant of the characteristic polynomial for u (and

so for p sufficiently large), we have $z_u(p) \mid p \pm 1$, so that $\gcd(p, z_u(p)) = 1$. Thus, the conditions $n \in \mathcal{N}_u$, $p \mid n$, and p sufficiently large jointly force $n \equiv 0 \pmod{pz_u(p)}$. Hence, if p is sufficiently large, the number of $n \in \mathcal{N}_u(x)$ with $P(n) = p$ is at most $\Psi(x/pz_u(p), p) \leq x/pz_u(p)$.

Thus, for large x ,

$$\#\mathcal{N}_2(x) \leq \sum_{p > L(x)^3} \frac{x}{pz_u(p)} = \sum_{\substack{p > L(x)^3 \\ z_u(p) \leq L(x)}} \frac{x}{pz_u(p)} + \sum_{\substack{p > L(x)^3 \\ z_u(p) > L(x)}} \frac{x}{pz_u(p)}.$$

The first sum on the right has, by Lemma 1, at most $L(x)^2$ terms for x large, each term being smaller than $x/L(x)^3$, so the sum is bounded by $x/L(x)$. The second sum on the right has terms smaller than $x/pL(x)$ and the sum of $1/p$ is of magnitude $\log \log x$, so the contribution here is $x/L(x)^{1+o(1)}$ as $x \rightarrow \infty$. Thus, $\#\mathcal{N}_2(x) \leq x/L(x)^{1+o(1)}$ as $x \rightarrow \infty$.

For any nonnegative integer j , let $I_j := [2^j, 2^{j+1})$. For \mathcal{N}_3 , we cover $I := [L(x)^{1/2}, L(x)^3)$ by these dyadic intervals, and we define a_j via $2^j = L(x)^{a_j}$. We shall assume the variable j runs over just those integers with I_j not disjoint from I . For any integer k , define $\mathcal{P}_{j,k}$ as the set of primes $p \in I_j$ with $z_u(p) \in I_k$. Note that, by Lemma 1, we have $\#\mathcal{P}_{j,k} \ll 4^k$. We have

$$\begin{aligned} \#\mathcal{N}_3(x) &\leq \sum_j \sum_k \sum_{\substack{p \in \mathcal{P}_{j,k} \\ n \in \mathcal{N}_u(x) \\ P(n)=p}} 1 \leq \sum_j \sum_k \sum_{p \in \mathcal{P}_{j,k}} \Psi\left(\frac{x}{pz_u(p)}, p\right) \\ &\leq \sum_j \sum_k \sum_{p \in \mathcal{P}_{j,k}} \frac{x}{pz_u(p)L(x)^{1/2a_j+o(1)}}, \end{aligned}$$

as $x \rightarrow \infty$, where we have used Lemma 4 for the last estimate. For $k > j/2$, we use the estimate

$$\sum_{p \in \mathcal{P}_{j,k}} \frac{1}{pz_u(p)} \leq 2^{-k} \sum_{p \in I_j} \frac{1}{p} \leq 2^{-k}$$

for x large. For $k \leq j/2$, we use the estimate

$$\sum_{p \in \mathcal{P}_{j,k}} \frac{1}{pz_u(p)} \ll \frac{4^k}{2^j 2^k} = 2^{k-j},$$

since there are at most order of magnitude 4^k such primes, as noted before. Thus,

$$\begin{aligned} \sum_k \sum_{p \in \mathcal{P}_{j,k}} \frac{1}{pz_u(p)} &= \sum_{k > j/2} \sum_{p \in \mathcal{P}_{j,k}} \frac{1}{pz_u(p)} + \sum_{k \leq j/2} \sum_{p \in \mathcal{P}_{j,k}} \frac{1}{pz_u(p)} \\ &\ll 2^{-j/2} = L(x)^{-a_j/2}. \end{aligned}$$

We conclude that

$$\#\mathcal{N}_3(x) \leq \sum_j \frac{x}{L(x)^{a_j/2+1/2a_j+o(1)}} \quad \text{as } x \rightarrow \infty.$$

Since the minimum value of $t/2 + 1/(2t)$ for $t > 0$ is 1 occurring at $t = 1$, we conclude that $\#\mathcal{N}_3(x) \leq x/L(x)^{1+o(1)}$ as $x \rightarrow \infty$. With our prior estimates for $\#\mathcal{N}_1(x)$ and $\#\mathcal{N}_2(x)$, this completes our proof.

It is possible that using the methods of [5] and [7] a stronger estimate can be made.

5 The proof of Theorem 3

Since $a_2 = \pm 1$, it is easy to see that the sequence u is purely periodic modulo any integer m . So, the index of appearance $z_u(m)$ defined in Section 2 exists for all positive integers m . Further, by examining the explicit formula (5) one can see that for any prime power $q = p^k$ we have

$$z_u(p^k) \mid z_u(p)p^{k-1}. \quad (18)$$

In fact this is known in much wider generality.

Now, for any real number $y \geq 1$ let

$$M_y := \text{lcm}[m : m \leq y].$$

We say that a positive integer n is *Lucas special* if it is of the form $n = 2sM_y$ for some $y \geq 3$ and for some squarefree positive integer s such that $\gcd(s, M_y) = 1$ and for every prime $p \mid s$ we have $p^2 - 1 \mid M_y$. Let \mathcal{L} denote the set of Lucas special numbers.

We now show that $\mathcal{L} \subset \mathcal{N}_u$ for any Lucas sequence u with $a_2 = \pm 1$. To see this it suffices to show for any $n = 2sM_y \in \mathcal{L}$ and for any prime power $q \mid n$, we have $z_u(q) \mid n$. This is easy for $q \mid s$, since then $q = p$ is prime and

either $z_u(p) = p$ (in the case $p \mid \Delta_u$) or $z_u(p) \mid p \pm 1$ (otherwise). And since $p^2 - 1 \mid M_y$, we have $z_u(p) \mid n$ in either case.

If $q \mid 2M_y$, we consider the cases of odd and even q separately.

- When q is odd, we have $q \mid M_y$ so $q \leq y$. Write $q = p^k$ with p prime, so that (18) implies $z_u(q) \mid (p-1)p^{k-1}, p^k$ or $(p+1)p^{k-1}$. We have $p^{k-1} \leq y$ and if $p+1 \leq y$, then $z(q) \mid M_y$. The only case not covered is $p+1 > y$ (so $p \in (y-1, y]$), $k=1$, $z_u(p) = p+1$. Write $p+1 = 2^j m$ where m is odd. Then $2^j \mid 2M_y$ and $m \mid 2M_y$, so $p+1 \mid 2M_y$. Thus, in all cases, $z_u(q) \mid 2M_y$ so $z_u(q) \mid n$.
- When $q = 2^k$ is a power of 2 with $q \mid 2M_y$, then since $z_u(2) \in \{2, 3\}$, we see from (18) that either $z_u(2^k) \mid 2^k$ or $z_u(2^k) \mid 3 \cdot 2^{k-1}$. Since $y \geq 3$, in either case we have $z_u(q) \mid 2M_y$.

We now use the method of Erdős [4] to show that the set \mathcal{L} is rather large. For this we take

$$y := \frac{\log x}{\log \log x} \quad \text{and} \quad z := y^v.$$

We say that q is a *proper prime power* if $q = \ell^k$ for a prime ℓ and an integer $k \geq 2$.

We define \mathcal{P} as the set of primes p such that:

- $p \in [y+1, z]$;
- $p^2 - 1$ is y -smooth;
- $p^2 - 1$ is not divisible by any proper prime power $q > y$.

Note that if q is a proper prime power and $q \mid p^2 - 1$, then $q \mid p \pm 1$, unless q is even, in which case $q/2 \mid p \pm 1$. Since trivially there are only $O(t^{1/2})$ proper prime powers $q \leq t$, there are only $O(zy^{-1/2})$ primes $p \leq z$ for which $p^2 - 1$ is divisible by a proper prime power $q > y$. Thus, recalling the assumption (8), we obtain

$$\#\mathcal{P} \geq \Pi(z, y) - y + O(zy^{-1/2}) = z^{1+o(1)},$$

provided that $x \rightarrow \infty$.

It is also obvious that for any squarefree positive integer s composed out of primes $p \in \mathcal{P}$, the integer $n = 2sM_y$ is Lucas special.

We now take the set $\mathcal{L}_v(x)$ of all such Lucas special integers $n = 2sM_y$, where s is composed out of

$$r := \left\lfloor \frac{\log x - 2y}{\log z} \right\rfloor$$

distinct primes $p \in \mathcal{P}$. Since by the prime number theorem the estimate $M_y = \exp((1 + o(1))y)$ holds as $x \rightarrow \infty$, we see that for sufficiently large x we have $n \leq x$ for every $n \in \mathcal{L}_v(x)$.

For the cardinality of $\mathcal{L}_v(x)$ we have

$$\#\mathcal{L}_v(x) \geq \binom{\#\mathcal{P}}{r} \geq \left(\frac{\#\mathcal{P}}{r}\right)^r.$$

Since

$$r = (v^{-1} + o(1)) \frac{\log x}{\log \log x} \quad \text{and} \quad \frac{\#\mathcal{P}}{r} = (\log x)^{v^{-1} + o(1)}$$

as $x \rightarrow \infty$, we obtain $\#\mathcal{L}_v(x) \geq x^{1-1/v+o(1)}$ as $x \rightarrow \infty$. Noting that $\mathcal{L}_v(x) \subset \mathcal{L}(x)$ concludes the proof.

6 The proof of Theorem 4

Since $\Delta_u \equiv 0, 1 \pmod{4}$ and $\Delta_u \neq 0, 1$, it follows that $|\Delta_u| > 1$. Let r be some prime factor of Δ_u . Then $r^k \in \mathcal{N}_u$ for all $k \geq 0$ (see [11, Pages 210 and 295]). We let k be a large positive integer and look at $u_{r^{k+4}}$. By Bilu, Hanrot and Voutier's primitive divisor theorem (see [1]), u_n has a primitive prime factor for all $n \geq 31$. Recall that a primitive prime factor of u_n is a prime factor p of u_n which does not divide $\Delta_u u_m$ for any positive integer $m < n$. Such a primitive prime factor p always satisfies $p \equiv \pm 1 \pmod{n}$. Since there are at most 5 values of $k \geq 0$ such that $r^k \leq 30$ for the same integer $r > 1$, and since $u_m \mid u_n$ if $m \mid n$, we conclude that $u_{r^{k+4}}$ has at least $\tau(r^{k+4}) - 5 = k$ distinct prime factors $p \neq r$, where $\tau(m)$ is the number of divisors of the positive integer m . Let them be $p_1 < \dots < p_k$. Assume that $|\alpha_1| \geq |\alpha_2|$. For large n , we have that $|\alpha_1|^{n/2} < |u_n| < 2|\alpha_1|^n$ (see [6, Theorem 2.3]). If β_1, \dots, β_k are nonnegative exponents such that

$$\beta_i \leq \frac{\log(x/r^{k+4})}{k \log p_i},$$

then $r^{k+4}p_1^{\beta_1} \cdots p_k^{\beta_k} \leq x$ is in \mathcal{N}_u (see [11, Page 210]), so it is counted by $\#\mathcal{N}_u(x)$. Hence,

$$\begin{aligned} \#\mathcal{N}_u(x) &\geq \prod_{i=1}^k \left(\left\lfloor \frac{\log(x/r^{k+4})}{k \log p_i} \right\rfloor + 1 \right) \geq \left(\frac{\log(x/r^{k+4})}{k} \right)^k \frac{1}{\prod_{i=1}^k \log p_i} \\ &\geq \left(\frac{\log(x/r^{k+4})}{2r^{k+4} \log |\alpha_1|} \right)^k, \end{aligned}$$

where the last inequality follows from the mean value inequality

$$\begin{aligned} \prod_{i=1}^k \log p_i &\leq \left(\frac{1}{k} \sum_{i=1}^k \log p_i \right)^k \leq \left(\frac{\log(|u_{r^{k+4}}|)}{k} \right)^k \\ &< \left(\frac{r^{k+4} \log |\alpha_1| + \log 2}{k} \right)^k \\ &< \left(\frac{2r^{k+4} \log |\alpha_1|}{k} \right)^k, \end{aligned}$$

for $k \geq 2$. In the above, we have also used the fact that $|u_n| < 2|\alpha_1|^n$ holds for all $n \geq 1$ with the choice $n := r^{k+4}$. Let $c_3 := 2 \log |\alpha_1|$. The above lower bound is

$$\begin{aligned} \#\mathcal{N}_u(x) &\geq \left(\frac{\log x}{r^{k+4}c_3} + O\left(\frac{k}{r^k}\right) \right)^k = \left(\frac{\log x}{r^{k+4}c_3} \right)^k \left(1 + O\left(\frac{k^2}{\log x}\right) \right) \\ &\gg \left(\frac{\log x}{r^{k+4}c_3} \right)^k \end{aligned}$$

provided that

$$k = o(\sqrt{\log x}), \tag{19}$$

as $x \rightarrow \infty$, which is now assumed. So, it suffices to look at

$$\left(\frac{\log x}{r^{k+4}c_3} \right)^k = \exp(k \log(\log x/c_3) - k(k+4) \log r).$$

Let $A := \log(\log x/c_3)$. In order to maximize the function $f(y) := yA - y(y+4) \log r$, we take its derivative and set it equal to zero to get $A - 2y \log r - 4 \log r = 0$, therefore $y = (A - 4 \log r)/(2 \log r) = A/(2 \log r) - 2$.

Thus, taking $k := \lfloor A/(2 \log r) - 2 \rfloor$ (so that (19) is satisfied), we get that $f(k) = f(y) + O(f'(y)) = A^2/(4 \log r) + O(A)$, hence

$$\begin{aligned} \#\mathcal{N}_u(x) &\geq \exp\left(\frac{(\log(\log x/c_3))^2}{4 \log r} + O(\log \log x)\right) \\ &= \exp\left(\frac{(\log \log x)^2}{4 \log r} + O(\log \log x)\right), \end{aligned}$$

which implies the desired conclusion with any constant $c_1 < 1/(4 \log r)$.

7 Remarks

We end with a result showing that it is quite possible for $\#\mathcal{N}_u(x)$ to be large under quite mild conditions. Observe that the sequence $u_n = 2^n - 2$ has the property that $u_1 = 0$. Here is a more general version of this fact.

Proposition 1. *Let $k \geq 2$ and $\{u_n\}_{n \geq 0}$ be a linearly recurrent sequence of order k satisfying recurrence (1). Assume that there exists a positive integer n_0 coprime to a_k such that $u_{n_0} = 0$. Then*

$$\#\mathcal{N}_u(x) \gg x/\log x,$$

where the implied constant depends on the sequence $\{u_n\}_{n \geq 0}$.

Proof. Since n_0 is coprime to a_k , it follows that $\{u_n\}_{n \geq 0}$ is purely periodic modulo n_0 . Let t_{n_0} be this period. Now, let \mathcal{R}_u be the set of primes $p \equiv 1 \pmod{t_{n_0}}$ such that $f_u(X)$ splits into linear factors modulo p . Alternatively, \mathcal{R}_u is the set of primes p such that the polynomial $f_u(X)(X^{t_{n_0}} - 1)$ splits into linear factors modulo p . The set of such primes has a positive density by the Chebotarev density theorem. We claim that

$$\mathcal{S}_u \subseteq \mathcal{N}_u, \tag{20}$$

where

$$\mathcal{S}_u := \{pn_0 : p \in \mathcal{R}_u \text{ and } p > n_0|\Delta_u|\}.$$

The above inclusion implies the desired bound since then

$$\#\mathcal{N}_u(x) \geq \#\mathcal{R}_u(x/n_0) + O(1) \gg x/\log x.$$

So, let us suppose that $p > n_0|\Delta_u|$ is in \mathcal{R}_u . Then $p \equiv 1 \pmod{t_{n_0}}$, therefore $p = 1 + \lambda t_{n_0}$ for some positive integer λ . Thus, $pn_0 = n_0 + \lambda n_0 t_{n_0}$ and since $\{u_n\}_{n \geq 1}$ is purely periodic with period t_{n_0} modulo n_0 , we get that

$$u_{pn_0} = u_{n_0 + \lambda n_0 t_{n_0}} \equiv u_{n_0} \equiv 0 \pmod{n_0}. \quad (21)$$

Next, observe that since the polynomial $f_u(X)$ factors in linear factors modulo p , we get that $\alpha_i^p \equiv \alpha_i \pmod{p}$ for all $i = 1, \dots, k$. In particular, $\alpha_i^{pn_0} \equiv \alpha_i^{n_0} \pmod{p}$ for all $i = 1, \dots, k$. Since the denominators of the coefficient A_i , $i = 1, \dots, k$, in (3) are divisors of Δ_u and $p > |\Delta_u|$, it follows that such denominators are invertible modulo p , therefore $A_i \alpha_i^{pn_0} \equiv A_i \alpha_i^{n_0} \pmod{p}$ for all $i = 1, \dots, k$. Summing up these congruences for $i = 1, \dots, k$, we get

$$u_{pn_0} = \sum_{i=1}^k A_i \alpha_i^{pn_0} \equiv \sum_{i=1}^k A_i \alpha_i^{n_0} \equiv u_{n_0} \equiv 0 \pmod{p}. \quad (22)$$

From the congruences (21) and (22), we get that both p and n_0 divide u_{pn_0} , and since p is coprime to n_0 , we get that $pn_0 \mid u_{pn_0}$. This completes the proof of the inclusion (20) and of the proposition. \square

The condition that n_0 is coprime to a_k is not always necessary. The conclusion of Proposition 1 may hold without this condition like in the example of the sequence of general term

$$u_n = 10^n - 7^n - 2 \cdot 5^n - 1 \quad \text{for all } n \geq 0,$$

for which we can take $n_0 = 2$. Observe that $k = 4$,

$$f_u(X) = (X - 10)(X - 7)(X - 5)(X - 1),$$

and n_0 is not coprime to $a_4 = -350$, yet one can check that the divisibility relation $2p \mid u_{2p}$ holds for all primes $p \geq 11$. We do not give further details.

Let $\mathcal{M}_u(x)$ be the set of integers $n \leq x$ with $n \mid u_n$ and n is *not* of the form pn_0 , where p is prime and $u_{n_0} = 0$. It may be that in the situation of Theorem 1, we can get a smaller upper bound for $\#\mathcal{M}_u(x)$ than for $\#\mathcal{N}_u(x)$. We can show this in a special case.

Proposition 2. *Assume that $\{u_n\}_{n \geq 0}$ is a linearly recurrent sequence of order k whose characteristic polynomial splits into distinct linear factors in $\mathbb{Z}[X]$. There is a positive constant $c_4(k)$ depending on k such that for all sufficiently large x (depending on the sequence u), we have $\#\mathcal{M}_u(x) \leq x/L(x)^{c_4(k)}$.*

Proof. Let $y = L(x)$. We partition $\mathcal{M}_u(x)$ into the following subsets:

$$\begin{aligned}\mathcal{M}_1(x) &:= \{n \in \mathcal{M}_u(x) : P(n) \leq y\}; \\ \mathcal{M}_2(x) &:= \{n \in \mathcal{M}_u(x) : \text{there is a prime } p \mid n, p > y, pT_u(p) \leq kx\}; \\ \mathcal{M}_3(x) &:= \mathcal{M}_u(x) \setminus (\mathcal{M}_1(x) \cup \mathcal{M}_2(x)).\end{aligned}$$

As in the proof of Theorem 2, we see that Lemma 4 implies that $\#\mathcal{M}_1(x) \leq x/L(x)^{1/2+o(1)}$ as $x \rightarrow \infty$.

As in the proof of Theorem 1,

$$\#\mathcal{M}_2(x) \ll \sum_{\substack{y < p \leq x \\ pT_u(p) \leq kx}} \left(\frac{x}{pT_u(p)} + 1 \right) \ll \sum_{y < p \leq x} \frac{x}{pT_u(p)}.$$

We break this last sum according as $p \in \mathcal{P}_{u,1/(k+1)}$ and $p \notin \mathcal{P}_{u,1/(k+1)}$, respectively. Lemma 1 shows that $\#\mathcal{P}_{u,1/(k+1)}(t) \ll t^{k/(k+1)}/\log t$. Thus,

$$\sum_{\substack{y < p \leq x \\ p \in \mathcal{P}_{u,1/(k+1)}}} \frac{x}{pT_u(p)} \leq \sum_{\substack{y < p \leq x \\ p \in \mathcal{P}_{u,1/(k+1)}}} \frac{x}{p} \ll \frac{x}{y^{1/(k+1)}}$$

and

$$\sum_{\substack{y < p \leq x \\ p \notin \mathcal{P}_{u,1/(k+1)}}} \frac{x}{pT_u(p)} \leq \sum_{y < p \leq x} \frac{x}{py^{1/(k+1)}} \ll \frac{x \log_2 x}{y^{1/(k+1)}}.$$

Hence,

$$\#\mathcal{M}_2(x) \ll \frac{x}{L(x)^{1/(k+1)+o(1)}} \quad \text{as } x \rightarrow \infty.$$

Suppose now that $n \in \mathcal{M}_3(x)$. Let $p \mid n$ with $pT_u(p) > kx$. Using as before the notation t_p for the period of u modulo p , as well as the fact that $T_u(p) \leq kt_p$ and $t_p \mid p-1$ (since f_u splits in linear factors over $\mathbb{Q}[X]$), we have

$$kx < pT_u(p) \leq kpt_p \leq kp^2,$$

so that $p > \sqrt{x}$. Thus, n can have at most one prime factor p with $pT_u(p) > kx$. So, if $n \in \mathcal{M}_3(x)$, we may assume that $n = mp$ where $p > \sqrt{x} > m$, and $P(m) \leq y$. Further, we may assume that $u_m \neq 0$. Since $p \mid u_{pm}$ and $t_p \mid p-1$, we have $p \mid u_m$. Now the number of prime factors of u_m is $O(m)$. Since the

number of $n \in \mathcal{M}_3(x)$ with such a prime $p \mid n$ is $O(x/(pT_u(p)) + 1) = O(1)$, we have

$$\#\mathcal{M}_3(x) \ll \sum_{\substack{m < \sqrt{x} \\ P(m) \leq y}} m \leq \sqrt{x} \Psi(\sqrt{x}, y) = \frac{x}{L(x)^{1/4+o(1)}} \quad \text{as } x \rightarrow \infty,$$

using Lemma 4.

We conclude that the result holds with $c_4(k) := \min\{1/5, 1/(k+2)\}$, say. \square

Finally, we note that for a given non-constant polynomial $g(X) \in \mathbb{Z}[X]$ one can consider the more general set

$$\mathcal{N}_{u,g} := \{n \geq 1 : g(n) \mid u_n\}.$$

We fix some real $y < x^{1/2}$ and note that by the Brun sieve (see [9, Theorem 2.3]), there are at most

$$N_1 \ll x \left(\frac{\log y}{\log x} \right) \quad (23)$$

values of $n \leq x$ such that $g(n)$ does not have a prime divisor in the interval $[y, x^{1/2}]$. We also note that for a prime p not dividing the content of g , the divisibility $p \mid g(n)$ puts n in at most $\deg g$ arithmetic progressions. Thus, using Lemma 2 as it was used in the proof of Theorem 1, the number of other $n \leq x$ with $g(n) \mid u_n$ can be estimated as

$$N_2 \leq \sum_{p \in [y, x^{1/2}]} \sum_{\substack{n \leq x \\ p \mid g(n) \\ p \mid u_n}} 1 \ll \sum_{p \in [y, x^{1/2}]} \left(\frac{x}{pT_u(p)} + 1 \right) \ll x \sum_{p \in [y, x^{1/2}]} \frac{1}{pT_u(p)} + O(x^{1/2}).$$

Using Lemma 1 for any $\gamma \in (0, 1)$ and the trivial estimate $T_u(p) \gg \log p$, we derive

$$\sum_{p \in [z, 2z]} \frac{1}{pT_u(p)} \leq \frac{1}{z} \sum_{p \in [z, 2z]} \frac{1}{T_u(p)} \ll \frac{1}{z} \left(\frac{z^{k\gamma}}{(\log z)^2} + \frac{z^{1-\gamma}}{\log z} \right).$$

Taking γ to satisfy

$$z^\gamma = (z \log z)^{1/(k+1)},$$

we obtain

$$\frac{1}{z} \sum_{p \in [z, 2z]} \frac{1}{pT_u(p)} \ll z^{-1/(k+1)} (\log z)^{-(k+2)/(k+1)}.$$

Summing over dyadic intervals, we now have

$$\sum_{p \in [y, x^{1/2}]} \frac{1}{pT_u(p)} \ll y^{-1/(k+1)} (\log y)^{-(k+2)/(k+1)}.$$

Therefore,

$$N_2 \ll xy^{-1/(k+1)} (\log y)^{-(k+2)/(k+1)} + x^{1/2}. \quad (24)$$

Taking, for example, $y := (\log x)^{k+1}$, we obtain from (23) and (24) the estimate

$$\#\mathcal{N}_{u,g}(x) \leq N_1 + N_2 \ll x \left(\frac{\log \log x}{\log x} \right). \quad (25)$$

It is certainly an interesting question to bring the bound (25) to the same shape as that of Theorem 1. However, the method of proof of Theorem 1 does not apply due to the possible existence of large prime divisors of $g(n)$.

Acknowledgements

The authors are grateful to Chris Smyth who attracted their attention to the problems considered in this paper and to Larry Somer for many valuable comments.

During the preparation of this paper, F. L. was supported in part by grants SEP-CONACyT 79685 and PAPIIT 100508, C. P. was supported in part by NSF grant DMS-1001180 and I. S. was supported in part by ARC grant DP1092835.

References

- [1] Yu. Bilu, G. Hanrot, P. M. Voutier, ‘Existence of primitive divisors of Lucas and Lehmer numbers. With an appendix by M. Mignotte’, *J. Reine Angew Math.* **539** (2001), 75–122.
- [2] E. R. Canfield, P. Erdős and C. Pomerance, ‘On a problem of Oppenheim concerning ‘Factorisatio Numerorum’, *J. Number Theory* **17** (1983), 1–28.

- [3] C. Dartyge, G. Martin and G. Tenenbaum, ‘Polynomial values free of large prime factors’, *Periodica Math. Hungar.*, **43** (2001), 111–119.
- [4] P. Erdős, ‘On the normal number of prime factors of $p - 1$ and some other related problems concerning Euler’s ϕ function’, *Quart. J. Math. (Oxford Ser.)* **6** (1935), 205–213.
- [5] P. Erdős, F. Luca, and C. Pomerance, ‘On the proportion of numbers coprime to a given integer’, Proceedings of the Anatomy of Integers Conference, Montréal, March 2006, J.-M. De Koninck, et al., eds., CRM Proceedings and Lecture Notes, vol. 46 (2008), 47–64.
- [6] 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.
- [7] D. Gordon and C. Pomerance, ‘The distribution of Lucas and elliptic pseudoprimes’, *Math. Comp.* **57** (1991), 825–838.
- [8] K. Györy and C. Smyth, ‘The divisibility of $a^n - b^n$ by powers of n ’, *Integers* **10** (2010), 319–334.
- [9] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.
- [10] S. Li, ‘On the distribution of even pseudoprimes’, *Unpublished Manuscript*, 1996.
- [11] E. Lucas, ‘Théorie des fonctions numériques simplement périodiques’, *Amer. J. Math.* **1** (1878), 184–240, 289–321.
- [12] F. Luca, ‘On positive integers n for which $\Omega(n)$ divides F_n ’, *Fibonacci Quart.*, **41** (2003), 365–371.
- [13] F. Luca and I. E. Shparlinski, ‘Some divisibilities amongst the terms of linear recurrences’, *Abh. Math. Sem. Univ. Hamburg*, **46** (2006), 143–156.
- [14] C. Pomerance, ‘On the distribution of pseudoprimes’, *Math. Comp.* **37** (1981), 587–593.

- [15] H. P. Schlickewei, 'Multiplicities of recurrence sequences' *Acta Math.*, **176** (1996), 171–243.
- [16] H. P. Schlickewei and W. M. Schmidt, 'The number of solutions of polynomial-exponential equations', *Compositio Math.* **120** (2000), 193–225.
- [17] I. E. Shparlinski, 'The number of different prime divisors of recurrent sequences', (Russian), *Mat. Zametki* **42** (1987), 494–507, 622.
- [18] C. Smyth, 'The terms in Lucas sequences divisible by their indices', *J. Integer Sequences* **13** (2010), Article 10.2.4.
- [19] L. Somer, 'Divisibility of terms in Lucas sequences by their subscripts', *Applications of Fibonacci numbers*, **5** (*St. Andrews, 1992*), Kluwer Acad. Publ., Dordrecht, 1993, 515–525.