Proceedings of the Edinburgh Mathematical Society (2012) **55**, 271–289
DOI:10.1017/S0013091510001355

ON NUMBERS n DIVIDING THE nTH TERM OF A LINEAR RECURRENCE

JUAN JOSÉ ALBA GONZÁLEZ[1], FLORIAN LUCA[1], CARL POMERANCE[2] AND IGOR E. SHPARLINSKI[3]

[1] *Instituto de Matemáticas, Universidad Nacional Autonoma de México, CP 58089, Morelia, Michoacán, México* (jjalba@gmail.com; fluca@matmor.unam.mx)
[2] *Mathematics Department, Dartmouth College, Hanover, NH 03755, USA* (carl.pomerance@dartmouth.edu)
[3] *Department of Computing, Macquarie University, Sydney, NSW 2109, Australia* (igor.shparlinski@mq.edu.au)

(Received 25 October 2010)

Abstract We give upper and lower bounds on the count of positive integers $n \leqslant x$ dividing the nth term of a non-degenerate linearly recurrent sequence with simple roots.

Keywords: linear recurrence; Lucas sequence; divisibility

2010 *Mathematics subject classification:* Primary 11B37
Secondary 11A07; 11N25

1. Introduction

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

$$u_{n+k} = a_1 u_{n+k-1} + \cdots + a_{k-1} u_{n+1} + a_k u_n \quad \text{for } n = 0, 1, \dots, \tag{1.1}$$

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

$$\mathcal{N}_u := \{n \geqslant 1 \colon n | u_n\}.$$

But first, some background on linear recurrence sequences.

To the recurrence (1.1) 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 $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$ are the distinct roots of $f_u(X)$ with multiplicities $\sigma_1, \ldots, \sigma_m$, respectively. It is then well known that the general term of the recurrence can be expressed as

$$u_n = \sum_{i=1}^{m} A_i(n)\alpha_i^n \quad \text{for } n = 0, 1, \ldots, \tag{1.2}$$

where $A_i(X)$ are polynomials of degrees at most $\sigma_i - 1$ for $i = 1, \ldots, m$, with coefficients in $K := \mathbb{Q}[\alpha_1, \ldots, \alpha_m]$. We refer the reader 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 can pose problems (but see below). For example, the sequence of the general term $u_n = n2^n$ for all $n \geqslant 0$ 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 (1.2) becomes

$$u_n = \sum_{i=1}^{k} A_i\alpha_i^n \quad \text{for } n = 0, 1, \ldots. \tag{1.3}$$

Here, $A_1, \ldots, A_k$ 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 $\{u_n\}_{n\geqslant 0}$ satisfies a linear recurrence of a smaller order.

We remark in passing that there is no real obstruction in reducing to the case of simple roots. Indeed, let $D \in \mathbb{N}$ be a common denominator of all the coefficients of all the polynomials $A_i(X)$ for $i = 1, \ldots, m$. 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 $n|Du_n$. Since n certainly divides* 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 $i = 1, \ldots, m$), then we are in an instance similar to the instance of the sequence of general term $u_n = n2^n$ for all $n \geqslant 0$ 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 set

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

A bit of Galois theory shows that w_n is an integer for all $n \geqslant 0$, and the sequence $\{w_n\}_{n\geqslant 0}$ satisfies a linear recurrence relation of order $\ell := \#\{1 \leqslant i \leqslant m \colon A_i(0) \neq 0\}$

* Here, 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.

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 set

$$\Delta_u := \prod_{1 \leqslant i < j \leqslant k} (\alpha_i - \alpha_j)^2 = \operatorname{disc}(f_u) \tag{1.4}$$

for the (non-zero) discriminant of the sequence $\{u_n\}_{n \geqslant 0}$, or of the polynomial $f_u(X)$. It is known that Δ_u is an integer. We also assume that $\{u_n\}_{n \geqslant 0}$ is non-degenerate, which means that α_i/α_j is not a root of 1 for any $1 \leqslant i < j \leqslant m$. Henceforth, all linear recurrences have only simple roots and are non-degenerate.

When $k = 2$, $u_0 = 0$, $u_1 = 1$ and $\gcd(a_1, a_2) = 1$, the sequence $\{u_n\}_{n \geqslant 0}$ is called a *Lucas sequence*. The formula (1.3) for the general term is

$$u_n = \frac{\alpha_1^n - \alpha_2^n}{\alpha_1 - \alpha_2} \quad \text{for } n = 0, 1, \dots. \tag{1.5}$$

That is, we can take $A_1 = 1/(\alpha_1 - \alpha_2)$ and $A_2 = -1/(\alpha_1 - \alpha_2)$ in (1.3). In the case of a Lucas sequence $\{u_n\}_{n \geqslant 0}$, the fine structure of the set $\mathcal{N}_u$ has been described in [**8**,**17**] (see also the references therein). We also note that divisibility of terms of a linear recurrence sequence by arithmetic functions of their index have been studied in [**12**] (see also [**11**] for the special case of Fibonacci numbers).

For a set $\mathcal{A}$ and a positive real number x we set $\mathcal{A}(x) = \mathcal{A} \cap [1, x]$. 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 $n \geqslant 0$ 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 $m \geqslant 2$ we use $\omega(m)$ for the number of distinct prime factors of m. So, from now on, we assume that $k \geqslant 2$.

Note next that for the sequence of general term $u_n = 2^n - 2$ for all $n \geqslant 0$ 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 pseudoprimes* show that it is possible for the estimate $\#\mathcal{N}_u(x) = (1 + o(1))x/\log x$ to hold as $x \to \infty$. However, we show that $\#\mathcal{N}_u(x)$ cannot have a larger order of magnitude.

Theorem 1.1. *For each $k \geqslant 2$, there is a positive constant $c_0(k)$ depending only on k such that if the characteristic polynomial of a non-degenerate linear recurrence sequence $\{u_n\}_{n \geqslant 0}$ of order k has only simple roots, then the estimate*

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

holds for x sufficiently large.

* A 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**] does the same for even pseudoprimes.

In the case of a Lucas sequence, we have a better bound. Let

$$L(x) := \exp(\sqrt{\log x \log\log x}). \tag{1.6}$$

Theorem 1.2. *Assume that $\{u_n\}_{n\geqslant 0}$ is a Lucas sequence. Then the inequality*

$$\#\mathcal{N}_u(x) \leqslant \frac{x}{L(x)^{1+o(1)}} \tag{1.7}$$

holds as $x \to \infty$.

It follows from a result of Somer [**18**, 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 $p^2 - 1$ that do not have prime divisors exceeding y. We use $\Pi(x, y)$ to denote the number of primes $p \leqslant x$ 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 behaviour of $\Pi(x, y)$ resembles the behaviour of the counting function for smooth integers. We record this in a very relaxed form as the assumption that for some fixed real $v \geqslant 1$ we have

$$\Pi(y^v, y) \geqslant y^{v+o(1)} \tag{1.8}$$

as $y \to \infty$. In fact, a general result from [**3**, Theorem 1.2] implies that (1.8) holds with any $v \in [1, \frac{4}{3})$.

Theorem 1.3. *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 (1.8) holds with some $v > 1$, we have*

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

as $x \to \infty$, where

$$\vartheta := 1 - \frac{1}{v}.$$

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

$$\vartheta = \tfrac{1}{4}.$$

Furthermore, since (1.8) is expected to hold for *any* $v > 1$, it is very likely that the bound of Theorem 1.3 holds with $\vartheta = 1$.

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

Theorem 1.4. *Let $\{u_n\}_{n\geqslant 0}$ 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*

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

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

2. Preliminary results

As in the proof of [**6**, Theorem 2.6], set

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

For a prime number p not dividing a_k, let $T_u(p)$ be the maximal non-negative integer T with the property that p does not divide

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

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 $\{u_n\}_{n\geqslant 0}$ 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 $\{u_n\}_{n\geqslant 0}$, then $T_u(p)+1$ is in fact the minimal positive integer ℓ such that $p|u_\ell$. This is sometimes called the *index of appearance* of p in $\{u_n\}_{n\geqslant 0}$ 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 $m|u_\ell$. This exists for all positive integers m coprime to a_2, and has the important property that $m|u_n$ if and only if $z_u(m)|n$. For any $\gamma \in (0,1)$, let

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

Lemma 2.1. *For $x^\gamma, y \geqslant 2$, the estimates*

$$\#\{p \colon T_u(p) \leqslant y\} \ll \frac{y^k}{\log y}, \qquad \#\mathcal{P}_{u,\gamma}(x) \ll \frac{x^{k\gamma}}{\gamma \log x}$$

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

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 $T_u(p) \leqslant y$. In particular, there exists a choice of integers $x_2,\dots,x_k$ all in $[1, y+1]$ such that p divides

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

This argument shows that

$$\prod_{T_u(p)\leqslant y} p \Bigg| \prod_{1\leqslant x_2,\dots,x_k\leqslant y+1} \max\{1, |N_{K/\mathbb{Q}}(D_u(0,x_2,\dots,x_k))|\}. \tag{2.1}$$

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

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 (2.1) is of size $\exp(O(y^k))$. Taking logarithms in the inequality implied by (2.1), we get that

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

If there are a total of n primes involved in this sum and if p_i denotes the ith 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 $n \ll y^k/\log y$, which is what we wanted to prove. □

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 2.2. *There exists a constant $c_2(k)$ depending only on k with the following property. Suppose that $\{u_n\}_{n\geqslant 0}$ is a linearly recurrent sequence of order k satisfying recurrence (1.1). Suppose that p is a prime coprime to $a_k\Delta_u$ and to the denominators of the numbers A_i in (1.3). Assume that there exists a positive integer s such that u_s is coprime to p. Then, for any real $X \geqslant 1$, the number of solutions $R_u(X, p)$ of the congruence*

$$u_{pn} \equiv 0 \pmod p \quad \text{with } 1 \leqslant n \leqslant X$$

satisfies the bound

$$R_u(X, p) \leqslant c_2(k)\biggl(\frac{X}{T_u(p)} + 1\biggr).$$

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 $\{w_n\}_{n\geqslant 0}$ 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 $\{u_n\}_{n\geqslant 0}$, then $\alpha_1^p, \dots, \alpha_k^p$ are the characteristic roots of $\{w_n\}_{n\geqslant 0}$. Hence,

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

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

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

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

we conclude that if $0 < x_2 < \cdots < x_k$ 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 let $n_1 < \cdots < n_\ell$ 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 rewrite each congruence as

$$\sum_{i=1}^{k} (A_i\alpha_i^{pn_1})\alpha_i^{p(n_j-n_1)} \equiv 0 \pmod p, \quad j = 1,2,\dots,\ell. \tag{2.2}$$

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 non-zero in $(\mathcal{O}_K/\pi\mathcal{O}_K)^k$.

Assume that

$$\ell \geqslant C(k) + k. \tag{2.3}$$

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

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

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

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

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

$$D_u(0,x_2,\dots,x_k) \neq 0. \tag{2.4}$$

However, by (2.2), we have $\pi | D_w(0,x_2,\dots,x_k)$; therefore,

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

which is impossible by the definition of $T_u(p)$ and the condition (2.4). Hence, the inequality (2.3) is false and the result follows. □

When $\{u_n\}_{n\geqslant 0}$ is a Lucas sequence, we set

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

The remarks preceding Lemma 2.1 show that $\#\mathcal{Q}_{u,\gamma}(x) = \#\mathcal{P}_{u,\gamma}(x) + O(1)$. Hence, Lemma 2.1 implies the following result.

Lemma 2.3. *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 $\{u_n\}_{n\geqslant 0}$.

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

Lemma 2.4. *For $x \geqslant y > 1$, the estimate*

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

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

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

3. The proof of Theorem 1.1

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. Let $y := x^{1/\log\log x}$ and let

$$\begin{aligned}
\mathcal{N}_1(x) &:= \{n \leqslant x\colon P(n) \leqslant y\},\\
\mathcal{N}_2(x) &:= \{n \leqslant x\colon n \notin \mathcal{N}_1(x) \text{ and } P(n) \in \mathcal{P}_{u,1/(k+1)}\},\\
\mathcal{N}_3(x) &:= \mathcal{N}(x) \setminus (\mathcal{N}_1(x) \cup \mathcal{N}_2(x)).
\end{aligned}$$

We now bound the cardinalities of each one of the above sets.

For $\mathcal{N}_1(x)$, by Lemma 2.4, we obtain

$$\#\mathcal{N}_1(x) = \Psi(x,y) = x\exp(-(1+o(1))v\log v) \tag{3.1}$$

as $x \to \infty$, where

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

Suppose now that $n \in \mathcal{N}_2(x)$. Then $n = pm$, where $p = P(n) \geqslant \max\{y, P(m)\}$. In particular, $p \leqslant x/m$; therefore, $m \leqslant x/y$. Since we also have $p \in \mathcal{P}_{u,1/(k+1)}(x/m)$, Lemma 2.1 implies that the number of such primes $p \leqslant x/m$ is $O((x/m)^{k/(k+1)})$, where the implied constant depends on the sequence $\{u_n\}_{n\geqslant 0}$. Summing the above inequality over all possible values of $m \leqslant x/y$, we get

$$\begin{aligned}\#\mathcal{N}_2(x) &\leqslant x^{k/(k+1)} \sum_{1\leqslant m\leqslant x/y} \frac{1}{m^{k/(k+1)}} \\ &\ll x^{k/(k+1)} \int_1^{x/y} \frac{\mathrm{d}t}{t^{k/(k+1)}} \\ &= ((k+1)x^{k/(k+1)})t^{1/(k+1)}|_1^{x/y} \\ &\ll \frac{x}{y^{1/(k+1)}}. \end{aligned} \tag{3.2}$$

Now let $n \in \mathcal{N}_3(x)$. As previously, we write $n = pm$, where $p = P(n) > y$. Thus, $m \leqslant x/p < x/y$. We assume that x (hence, y) is sufficiently large. Since $n \in \mathcal{N}_u$, we have that $n|u_n$; therefore, $p|u_n$. Furthermore, $T_u(p) \geqslant p^{1/(k+1)}$. We fix p and count the number of possibilities for m. To this end, let $\{w_\ell\}_{\ell\geqslant 0}$ be the sequence defined as $w_\ell = u_{p\ell}$ for all $\ell \geqslant 0$. This is a linearly recurrent sequence of order k. We would like to apply Lemma 2.2 to it to bound the number of solutions to the congruence

$$w_m \equiv 0 \pmod p, \quad \text{where } 1 \leqslant m \leqslant x/p.$$

If the conditions of Lemma 2.2 are satisfied, then this number, denoted by $R_w(x/p, p)$, satisfies

$$R_w(x/p, p) \leqslant c_2(k)\biggl(\frac{x}{pT_w(p)} + 1\biggr).$$

Let us check the conditions of Lemma 2.2. Note first that if $\alpha_1, \dots, \alpha_k$ are the characteristic roots of $\{u_n\}_{n\geqslant 0}$, then $\alpha_1^p, \dots, \alpha_k^p$ are the characteristic roots of $\{w_\ell\}_{\ell\geqslant 1}$. Hence,

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

In particular, the term $a_{w,k}$ corresponding to the recurrence $\{w_\ell\}_{\ell\geqslant 1}$ satisfies $a_{w,k} = a_k^p$ assuming that $y > 2$. Thus, assuming further that $y > |a_k|$, we then have that p does not divide a_k; therefore, p does not divide $a_{w,k}$ either. Next, note that

$$\Delta_w = \prod_{1\leqslant i<j\leqslant k} (\alpha_i^p - \alpha_j^p)^2.$$

Modulo p, we have that

$$\Delta_w \equiv \biggl(\prod_{1\leqslant i<j\leqslant k} (\alpha_i - \alpha_j)^2\biggr)^p \equiv \Delta_u^p \pmod p.$$

From the above congruence, we easily get that $p|\Delta_w$ if and only if $p|\Delta_u$. Thus, assuming that x is sufficiently large such that $y > |\Delta_u|$, we then have that $p \nmid \Delta_u$, therefore $p \nmid \Delta_w$ either.

So far, we have checked that p does not divide $a_{w,k}\Delta_w$, which is the first assumption in the statement of Lemma 2.2.

Let us check the next assumption.

Note that, since $p\nmid\Delta_u$, the characteristic polynomial $f_u(X)$ of $\{u_\ell\}_{\ell\geqslant 0}$ has only simple roots modulo p. Since p does not divide the last coefficient a_k for the recurrence for $\{u_n\}_{n\geqslant 0}$ either, it follows that this sequence is purely periodic modulo p. Let t_p be its period modulo p. It is known that t_p is coprime to p. In fact, t_p is a divisor of the number

$$\operatorname{lcm}[p^i-1\colon i=1,2,\dots,k].$$

Choose some $n_0>0$ such that $u_{n_0}\neq 0$. Let x be so large that $y>|u_{n_0}|$. Since $p>y$, we have $p\nmid u_{n_0}$. And since $\gcd(p,t_p)=1$, there exists an integer s with $sp\equiv n_0 \pmod{t_p}$. Thus,

$$w_s=u_{sp}\equiv u_{n_0}\pmod p.$$

In particular, w_s is coprime to p. Hence, for x sufficiently large, the second assumption from Lemma 2.2 holds for the sequence $\{w_\ell\}_{\ell\geqslant 0}$.

Next we show that $T_u(p)=T_w(p)$. Observe that this number exists (for both the sequences $\{u_\ell\}_{\ell\geqslant 0}$ and $\{w_\ell\}_{\ell\geqslant 0}$) because p does not divide a_k. Indeed, the claimed equality follows easily from the following calculation:

$$\begin{aligned}D_w(x_1,\dots,x_k)&=\det(\alpha_i^{px_j})_{1\leqslant i,j\leqslant k}\\&\equiv(\det(\alpha_i^{x_j}))^p \pmod p\\&\equiv D_u(x_1,\dots,x_k)^p \pmod p.\end{aligned}$$

Since $n\in\mathcal{N}_3(x)$, we have that $T_w(p)=T_u(p)\geqslant p^{1/(k+1)}$.

Lemma 2.2 now guarantees that the number of choices for m once p is fixed is

$$R_w(x/p,p)\leqslant c_2(k)\biggl(\frac{x}{p^{1+1/(k+1)}}+1\biggr).$$

To summarize, we have

$$\begin{aligned}\mathcal{N}_3(x)&\leqslant\sum_{y\leqslant p\leqslant x}c_2(k)\biggl(\frac{x}{p^{1+1/(k+1)}}+1\biggr)\\&\leqslant c_2(k)\biggl(\pi(x)+x\sum_{y\leqslant p}\frac{1}{p^{1+1/(k+1)}}\biggr)\\&\leqslant c_2(k)\biggl(\pi(x)+x\int_y^\infty\frac{\mathrm{d}t}{t^{1+1/(k+1)}}\biggr).\end{aligned}$$

Therefore,

$$\mathcal{N}_3(x)\leqslant c_2(k)\biggl(\pi(x)+O\biggl(\frac{(k+1)x}{y^{1/(k+1)}}\biggr)\biggr). \tag{3.3}$$

Comparing (3.1)–(3.3), we get that

$$\#\mathcal{N}(x)\leqslant c_2(k)\pi(x)+\frac{x}{\exp((1+o(1))v\log v)}+O\biggl(\frac{x}{y^{1/(k+1)}}\biggr) \tag{3.4}$$

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

as $x \to \infty$, where the implied constant depends on the recurrence $\{u_n\}_{n \geqslant 0}$. By our choice of y as $x^{1/\log\log x}$, the second and third terms on the right-hand side of (3.4) are both $o(\pi(x))$ as $x \to \infty$, so we have the theorem.

4. The proof of Theorem 1.2

As in Theorem 1.1, we divide the numbers $n \in \mathcal{N}_u(x)$ into several classes:

(i) $\mathcal{N}_1(x) := \{n \in \mathcal{N}_u(x) \colon P(n) \leqslant L(x)^{1/2}\}$;

(ii) $\mathcal{N}_2(x) := \{n \in \mathcal{N}_u(x) \colon P(n) \geqslant L(x)^3\}$;

(iii) $\mathcal{N}_3(x) := \mathcal{N}_u(x) \setminus (\mathcal{N}_1(x) \cup \mathcal{N}_2(x))$.

It follows from Lemma 2.4 that

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

as $x \to \infty$.

For $n \in \mathcal{N}_u$ and $p|n$, 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 $z_u(p)|p \pm 1$, so that $\gcd(p, z_u(p)) = 1$. Thus, the conditions $n \in \mathcal{N}_u$, $p|n$ 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 $\Psi(x/pz_u(p), p) \leqslant x/pz_u(p)$.

Thus, for large x,

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

The first sum on the right has, by Lemma 2.1, 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 $x \to \infty$. Thus, $\#\mathcal{N}_2(x) \leqslant x/L(x)^{1+o(1)}$ as $x \to \infty$.

For any non-negative 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 b_j via $2^j = L(x)^{b_j}$. We shall assume that 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 2.1, we have $\#\mathcal{P}_{j,k} \ll 4^k$. We have

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

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

as $x \to \infty$, where we have used Lemma 2.4 for the last estimate. For $k > j/2$, we use the estimate

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

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

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

since there are at most $O(4^k)$ such primes, as noted before. Thus,

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

We conclude that

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

Since the minimum value of $t/2 + 1/(2t)$ for $t > 0$ is 1 occurring at $t = 1$, we conclude that $\#\mathcal{N}_3(x) \leqslant x/L(x)^{1+o(1)}$ as $x \to \infty$. 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**, **7**], a stronger estimate can be made.

5. The proof of Theorem 1.3

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 § 2 exists for all positive integers m. Further, by examining the explicit formula (1.5) one can see that for any prime power $q = p^k$ we have

$$z_u(p^k) | z_u(p) p^{k-1}. \tag{5.1}$$

In fact this is known in much wider generality.

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

$$M_y := \operatorname{lcm}[m \colon m \leqslant y].$$

We say that a positive integer n is *Lucas special* if it is of the form $n = 2sM_y$ for some $y \geqslant 3$ and for some square-free positive integer s such that $\gcd(s, M_y) = 1$ and for every prime $p|s$ we have $p^2 - 1 | M_y$. 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 $q|n$, we have $z_u(q)|n$. This is easy for $q|s$, since then $q = p$ is prime and either $z_u(p) = p$ (in the case $p|\Delta_u$) or $z_u(p)|p \pm 1$ (otherwise). And since $p^2 - 1 | M_y$, we have $z_u(p)|n$ in either case.

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

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

(i) When q is odd, we have $q|M_y$ so $q \leqslant y$. Write $q = p^k$ with p prime, so that (5.1) implies $z_u(q)|(p-1)p^{k-1}$, p^k or $(p+1)p^{k-1}$. We have $p^{k-1} \leqslant y$ and if $p+1 \leqslant y$, then $z(q)|M_y$. 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 $2^j|2M_y$ and $m|2M_y$, so $p+1|2M_y$. Thus, in all cases, $z_u(q)|2M_y$ so $z_u(q)|n$.

(ii) When $q = 2^k$ is a power of 2 with $q|2M_y$, since $z_u(2) \in \{2, 3\}$, we see from (5.1) that either $z_u(2^k)|2^k$ or $z_u(2^k)|3 \cdot 2^{k-1}$. Since $y \geqslant 3$, in either case we have $z_u(q)|2M_y$.

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

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

with v satisfying (1.8). We say that q is a *proper prime power* if $q = \ell^k$ for a prime ℓ and an integer $k \geqslant 2$.

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

(i) $p \in [y+1, z]$;

(ii) $p^2 - 1$ is y-smooth;

(iii) $p^2 - 1$ is not divisible by any proper prime power $q > y$.

Note that if q is a proper prime power and $q|p^2-1$, then $q|p \pm 1$, unless q is even, in which case $q/2|p \pm 1$. Since trivially there are only $O(t^{1/2})$ proper prime powers $q \leqslant t$, there are only $O(zy^{-1/2})$ primes $p \leqslant z$ for which $p^2 - 1$ is divisible by a proper prime power $q > y$. Thus, recalling the assumption (1.8), we obtain

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

provided that $x \to \infty$.

It is also obvious that for any square-free positive integer s composed 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 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 $x \to \infty$, we see that for sufficiently large x we have $n \leqslant x$ for every $n \in \mathcal{L}_v(x)$.

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

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

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

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

6. The proof of Theorem 1.4

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 $k \geqslant 0$ [**13**, pp. 210 and 295]. We let k be a large positive integer and look at $u_{r^{k+4}}$. By Bilu *et al.*'s primitive divisor theorem [**1**], u_n has a primitive prime factor for all $n \geqslant 31$. 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 five values of $k \geqslant 0$ such that $r^k \leqslant 30$ for the same integer $r > 1$, and since $u_m | u_n$ if $m | n$, 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 the first k be $p_1 < \cdots < p_k$. Assume that $|\alpha_1| \geqslant |\alpha_2|$. For large n, we have that $|\alpha_1|^{n/2} < |u_n| < 2|\alpha_1|^n$ [**6**, Theorem 2.3]. If $\beta_1, \ldots, \beta_k$ are non-negative integral exponents such that

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

then $r^{k+4} p_1^{\beta_1} \cdots p_k^{\beta_k} \leqslant x$ is in $\mathcal{N}_u$ [**13**, p. 210], so it is counted by $\#\mathcal{N}_u(x)$. Hence,

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

where the last inequality follows from the mean-value inequality

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

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

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

provided that

$$k = o(\sqrt{\log x}), \tag{6.1}$$

as $x \to \infty$, 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 $f(t) := tA - t(t+4)\log r$, we take its derivative and set it equal to zero to get $A - 2t\log r - 4\log r = 0$; therefore,

$$t = \frac{A - 4\log r}{2\log r} = \frac{A}{2\log r} - 2.$$

Thus, taking $k := \lfloor A/(2\log r) - 2\rfloor$ (so that (6.1) is satisfied), we get that

$$f(k) = f(t) + O(f'(t)) = \frac{A^2}{4\log r} + O(A).$$

Hence,

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

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 7.1. *Let $k \geqslant 2$ and $\{u_n\}_{n\geqslant 0}$ be a linearly recurrent sequence of order k satisfying recurrence (1.1). 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 $\{u_n\}_{n\geqslant 0}$.

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

$$\mathcal{S}_u \subseteq \mathcal{N}_u, \tag{7.1}$$

where

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

The above inclusion implies the desired bound since then

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

Since $\{u_n\}_{n\geqslant 0}$ modulo n_0 is purely periodic with period t_{n_0}, we get that

$$u_{pn_0} \equiv u_{n_0} \equiv 0 \pmod{n_0}. \tag{7.2}$$

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 (1.3) 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

$$u_{pn_0} = \sum_{i=1}^{k} A_i\alpha_i^{pn_0} \equiv \sum_{i=1}^{k} A_i\alpha_i^{n_0} \equiv u_{n_0} \equiv 0 \pmod p. \tag{7.3}$$

From the congruences (7.2) and (7.3), we get that both p and n_0 divide u_{pn_0}, and since p is coprime to n_0, we get that $pn_0 | u_{pn_0}$. This completes the proof of the inclusion (7.1) and of the proposition. □

The condition that n_0 is coprime to a_k is not always necessary. The conclusion of Proposition 7.1 may hold without this condition, as in the example of the sequence of general term

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

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 $2p|u_{2p}$ holds for all primes $p \geqslant 11$. We do not give further details.

Let $\mathcal{M}_u(x)$ be the set of integers $n \leqslant x$ with $n|u_n$ 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 1.1 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.

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

Proposition 7.2. *Assume that $\{u_n\}_{n\geqslant 0}$ 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 $\#\mathcal{M}_u(x) \leqslant x/L(x)^{c_4(k)}$.*

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

$$\mathcal{M}_1(x) := \{n \in \mathcal{M}_u(x) \colon P(n) \leqslant y\};$$
$$\mathcal{M}_2(x) := \{n \in \mathcal{M}_u(x) \colon \text{there is a prime } p|n,\ p > y,\ pT_u(p) \leqslant kx\};$$
$$\mathcal{M}_3(x) := \mathcal{M}_u(x) \setminus (\mathcal{M}_1(x) \cup \mathcal{M}_2(x)).$$

As in the proof of Theorem 1.2, we see that Lemma 2.4 implies that $\#\mathcal{M}_1(x) \leqslant x/L(x)^{1/2+o(1)}$ as $x \to \infty$.

As in the proof of Theorem 1.1,

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

We split this summation according to $p \in \mathcal{P}_{u,1/(k+1)}$ and $p \notin \mathcal{P}_{u,1/(k+1)}$, respectively. Lemma 2.1 shows that $\#\mathcal{P}_{u,1/(k+1)}(t) \ll t^{k/(k+1)}/\log t$. Thus,

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

and

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

Hence,

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

Suppose now that $n \in \mathcal{M}_3(x)$. Let $p|n$ 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 $T_u(p) \leqslant kt_p$ and $t_p|p-1$ (since f_u splits in linear factors over $\mathbb{Q}[X]$), we have

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

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 $P(m) \leqslant y$. Further, we may assume that $u_m \neq 0$. Since $p|u_{pm}$ and $t_p|p-1$, we have $p|u_m$. 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 $p|n$ is $O(x/(pT_u(p)) + 1) = O(1)$, we have

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

using Lemma 2.4.

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

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

$$\mathcal{N}_{u,g} := \{n \geqslant 1\colon g(n) | u_n\}.$$

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

$$N_1 \ll \frac{x \log y}{\log x} \tag{7.4}$$

values of $n \leqslant x$ 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 $p | g(n)$ puts n in at most $\deg g$ arithmetic progressions. Thus, using Lemma 2.2 as it was used in the proof of Theorem 1.1, the number of other $n \leqslant x$ with $g(n) | u_n$ can be estimated as

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

Using Lemma 2.1 for any $\gamma \in (0, 1)$ and the trivial estimate $T_u(p) \gg \log p$, we derive

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

Taking γ to satisfy

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

we obtain

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

Summing over dyadic intervals, we now have

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

Therefore,

$$N_2 \ll x y^{-1/(k+1)} (\log y)^{-(k+2)/(k+1)} + x^{1/2}. \tag{7.5}$$

Taking, for example, $y := (\log x)^{k+1}$, we obtain from (7.4) and (7.5) the estimate

$$\#\mathcal{N}_{u,g}(x) \leqslant N_1 + N_2 \ll \frac{x \log \log x}{\log x}. \tag{7.6}$$

This estimate is slightly worse than the estimate in Theorem 1.1 and it is certainly an interesting question if the gap can be closed. However, the method of proof of Theorem 1.1 does not apply due to the possible existence of large prime divisors of $g(n)$.

Acknowledgements. The authors are grateful to Chris Smyth, who brought the problems considered in this paper to our attention, 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.E.S. was supported in part by ARC Grant DP1092835.

https://doi.org/10.1017/S0013091510001355 Published online by Cambridge University Press

References

1. Yu. Bilu, G. Hanrot, P. M. Voutier, Existence of primitive divisors of Lucas and Lehmer numbers, *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, *Period. 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, *Q. J. Math.* **6** (1935), 205–213.
5. P. Erdős, F. Luca, and C. Pomerance, On the proportion of numbers coprime to a given integer, in *Anatomy of Integers* (ed. J.-M. De Koninck), CRM Proceedings and Lecture Notes, Volume 46, pp. 47–64 (2008).
6. G. Everest, A. van der Poorten, I. E. Shparlinski and T. Ward, *Recurrence sequences*, Mathematical Surveys and Monographs, Volume 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. Győry 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. F. Luca, On positive integers n for which $\Omega(n)$ divides F_n, *Fibonacci Q.* **41** (2003), 365–371.
12. F. Luca and I. E. Shparlinski, Some divisibilities amongst the terms of linear recurrences, *Abh. Math. Sem. Univ. Hamburg* **46** (2006), 143–156.
13. E. Lucas, Théorie des fonctions numériques simplement périodiques, *Am. J. Math.* **1** (1878), 184–240, 289–321.
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. C. Smyth, The terms in Lucas sequences divisible by their indices, *J. Integer Sequences* **13** (2010), 10.2.4.
18. L. Somer, Divisibility of terms in Lucas sequences by their subscripts, in *Applications of Fibonacci numbers*, Volume 5, pp. 515–525 (Kluwer Academic, Dordrecht, 1993).