

## Positive lower density for prime divisors of generic linear recurrences

BY OLLI JÄRVINIEMI

*Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland.*

*e-mail: [olli.a.jarviniemi@utu.fi](mailto:olli.a.jarviniemi@utu.fi)*

*(Received 20 September 2021; revised 13 March 2023; accepted 15 March 2023)*

### Abstract

Let  $d \geq 3$  be an integer and let  $P \in \mathbb{Z}[x]$  be a polynomial of degree  $d$  whose Galois group is  $S_d$ . Let  $(a_n)$  be a non-degenerate linearly recursive sequence of integers which has  $P$  as its characteristic polynomial. We prove, under the generalised Riemann hypothesis, that the lower density of the set of primes which divide at least one non-zero element of the sequence  $(a_n)$  is positive.

2020 Mathematics Subject Classification: 11B37, 11B50, 11R45 (Primary)

### 1. Introduction

Given a sequence of integers, it is natural to consider the set of primes which divide at least one of its values (the prime divisors of the sequence). Here we consider the prime divisors of linearly recursive sequences. We assume that the elements  $a_1, a_2, \dots$  of the sequence and the coefficients  $c_i$  in the defining recursion

$$a_{n+d} + c_{d-1}a_{n+d-1} + \dots + c_0a_n = 0$$

are integers. The minimal  $d$  for which such coefficients exist is called the order of  $(a_n)$ .

It is a well-known result often attributed to Pólya [17] that, excluding degenerate cases, a linearly recursive sequence has infinitely many prime divisors. It is natural to ask how dense the set of prime divisors of a linearly recursive sequence is with respect to the set of primes.

For second order recurrences this has been studied in a number of works assuming the generalised Riemann hypothesis (GRH), under which it is known that the density of prime divisors exists, is positive (unless there are only finitely many prime divisors) and can, at least in principle, be computed explicitly (see [1, 5, 7, 8, 10, 12–14, 22, 23] and [11, section 8.4]). Unconditional results are much more modest: in [15] it is proved that the number of primes  $p \leq x$  which are prime divisors of such a sequence is at least of magnitude  $\log x$ .

Higher order sequences have received considerably less attention. Roskam [19] states that “Essentially nothing is known for sequences of order larger than 2”, and it is mentioned that some very non-generic cases can be handled (see Ballot [1]). Roskam proceeds by proving that, under a certain generalisation of Artin’s conjecture on primitive roots, “generic” linear recurrences have a positive lower density of prime divisors. However, we note that a much stronger result under a significantly weaker assumption follows directly from the work of

Niederreiter [16, theorem 4.1]. Nevertheless, while the regular version of Artin's conjecture has been proven by Hooley [6] under GRH (see [11] for a survey), establishing this weaker assumption seems very difficult even under GRH. See Section 8 for more details.

Here we prove, under GRH, that the set of prime divisors of a “generic” linear recurrences has positive lower density. This seems to be the first such result which is applicable to “almost all” sequences and which assumes only standard conjectures.

**THEOREM 1.1.** *Assume GRH. Let  $d \geq 3$  be an integer and let  $P \in \mathbb{Z}[x]$  be a polynomial whose Galois group is the symmetric group  $S_d$  and such that the quotient of any two distinct roots of  $P$  is not a root of unity. Let  $(a_n)$  be a linearly recursive sequence of integers whose characteristic polynomial is  $P$ . The set of primes which divide some non-zero element of the sequence  $(a_n)$  has a lower density of at least  $1/(d-1)$ . In particular, this lower density is strictly positive.*

The assumption on the quotients of the roots of  $P$  (the *non-degeneracy*) follows from the assumption on the Galois group when  $d \geq 4$ . (Indeed, if  $(\alpha_i/\alpha_j)^m = 1$  for some distinct roots  $\alpha_i, \alpha_j$  of  $P$  and  $m \in \mathbb{Z}_+$ , then by considering suitable elements of the Galois group this holds for any distinct roots  $\alpha_i, \alpha_j$  of  $P$ , so  $P$  must be a binomial, and hence the Galois group is of size at most  $d(d-1)$ .)

It is well known that almost all polynomials of degree  $d$  have Galois group isomorphic to  $S_d$ , so the result applies to “100%” of linear recurrences. The proof actually works for a slightly larger class of recurrences, and proves that almost all primes  $p$  such that  $P$  has suitable factorisation modulo  $p$  are prime divisors of the sequence.

**THEOREM 1.2.** *Assume GRH. Let  $P \in \mathbb{Z}[x]$  be a polynomial which has the following properties:*

- (i)  $d := \deg(P) \geq 3$ ;
- (ii)  $P$  is irreducible;
- (iii) there are infinitely many primes  $p$  such that  $P$  factorises as the product of a linear polynomial and an irreducible polynomial of degree  $d-1$  modulo  $p$ , that is, the Galois group of  $P$  contains an element whose cycle type is  $(1, d-1)$ ;
- (iv) if  $|P(0)| > 1$ , the roots of  $P$  are multiplicatively independent, and if  $|P(0)| = 1$ , some (or, equivalently, any)  $d-1$  roots of  $P$  are multiplicatively independent.

*Let  $(a_n)$  be a linearly recursive sequence of integers whose characteristic polynomial is  $P$ . Then almost all primes  $p$  such that  $P$  factorises as the product of irreducibles of degree 1 and  $d-1$  modulo  $p$  divide some element of the sequence  $(a_n)$ . In particular, the set of primes which divide some element of the sequence  $(a_n)$  has a strictly positive lower density.*

Here and in what follows “almost all primes” means “for a set of primes of relative natural density 1 in the set of primes”.

We note that there exist non-trivial examples of polynomials  $P$  which do not satisfy condition (iv) above [2].

The proof can be adapted so that it also works for reducible characteristic polynomials in certain special cases, but as such cases are rare, we do not discuss them in detail.

The GRH we use in the proofs states that the non-trivial zeros of the Dedekind zeta-function of any number field lie on the line  $\operatorname{Re}(s) = 1/2$ . (For details, see Lemma 3.2 below and [9, theorem 3.1].)

We first provide a proof sketch, after which we give a detailed argument. We conclude by discussing challenges arising in the study of prime divisors of linear recurrences.

2. Overview of the method

For concreteness we consider the sequence  $a_n$  defined by

$$a_n = 5^n + (3 + \sqrt{2})^n + (3 - \sqrt{2})^n, n = 1, 2, \dots$$

The characteristic polynomial  $(x - 5)(x - (3 + \sqrt{2}))(x - (3 - \sqrt{2}))$  is reducible and thus not of the form of Theorem 1.1, but we only use this example to demonstrate the idea. (The proof of Theorem 1.1 can be adapted to this sequence, though.)

In the case when 2 is a quadratic residue modulo  $p$  the period of the sequence  $(a_n)$  modulo  $p$  divides  $p - 1$ . We are unable to say anything nontrivial about whether such primes are prime divisors of  $(a_n)$  or not.

The case when 2 is a quadratic nonresidue modulo  $p$ , however, turns out to be accessible. Write  $n = (p + 1)k + r, k, r \in \mathbb{Z}$ . We may view the numbers  $3 \pm \sqrt{2}$  as elements of  $\mathbb{F}_{p^2}$ , and by norms we have  $(3 \pm \sqrt{2})^{p+1} = 7$  in this finite field. Hence

$$a_{(p+1)k+r} \equiv 5^{2k+r} + 7^k \left( (3 + \sqrt{2})^r + (3 - \sqrt{2})^r \right) \pmod{p}. \tag{2.1}$$

The equation  $a_{(p+1)k+r} \equiv 0 \pmod{p}$  may thus be written, for  $p > 7$ , as

$$\left( \frac{5^2}{7} \right)^k \equiv - \left( \frac{3 + \sqrt{2}}{5} \right)^r - \left( \frac{3 - \sqrt{2}}{5} \right)^r \pmod{p}. \tag{2.2}$$

Artin’s primitive root conjecture states that a given rational number  $a$  is a primitive root modulo  $p$  for infinitely many primes  $p$  as long as  $a$  is not  $-1$  or a square. Under GRH one can prove this in a quantitative form: the set of such primes has a positive density (as long as its infinite) [6]. This density is often quite large. For example, for  $a = 2$  the density is roughly 37 percent.

It turns out that the order of a rational number  $a$  modulo primes is almost always almost maximal assuming  $a \notin \{-1, 0, 1\}$  (under GRH). More precisely, the density of primes  $p$  with  $\text{ord}_p(a) \geq (p - 1)/C$  goes to 1 as  $C \rightarrow \infty$ . (See [24, section 5] or Lemma 3.2 below.)

Note also that if  $\text{ord}_p(a) = (p - 1)/h$ , then the function sending integers  $x$  to  $a^x \pmod{p}$  attains all non-zero  $h$ th powers modulo  $p$  as its values.

In this light, to prove that (2.2) is solvable for almost any prime  $p$  it suffices to show that the equation

$$x^h \equiv - \left( \frac{3 + \sqrt{2}}{5} \right)^r - \left( \frac{3 - \sqrt{2}}{5} \right)^r \pmod{p} \tag{2.3}$$

has a solution  $(x, r)$  with  $x \neq 0$  for almost any prime  $p$ .

Note the right-hand side of (2.3) satisfies a linear recurrence  $(b_r)$ . We aim to prove that the sequence  $b_1, b_2, \dots$  almost always attains a  $h$ th power as its value modulo  $p$ .

By using results from Galois theory and the Chebotarev density theorem, it is not very hard to show that this is true, for example, if there is an infinite subsequence of  $(b_r)$  whose elements are distinct primes.

Of course, we cannot guarantee that a linear recurrence has infinitely many prime values. However, there are only a very few cases where such an idea does not work. To name one, if  $b_r$  is always three times a square, then if 3 is a quadratic nonresidue modulo  $p$  (which happens for a positive density of primes), the sequence  $b_r$  may avoid all squares modulo  $p$ .

In general, the only obstructions arise when the values of the linear recurrence are almost perfect powers. By applying Zannier's result on Pisot's  $d$ th root conjecture [25] we reduce our problem to determining whether or not a linearly recursive sequence arising in the proof is the power of another recurrence. From here on only elementary observations are needed.

As we already mentioned, the sequence  $a_n$  considered here does not satisfy the conditions of Theorem 1.1, and the general case is more complicated. There are two notable differences.

To perform the "norm-trick" and to arrive to an equation of the form (2.2) we need to control the norms of the roots of the characteristic polynomial in finite fields. To do so, in the situation of Theorem 1.1 we consider those primes  $p$  for which  $P$  factorises modulo  $p$  as the product of two irreducibles of degree 1, respectively  $d - 1$ .

From here we are able to reduce to an equation similar to (2.3), though this time the right hand side is not necessarily a linear recurrence of integers but of algebraic numbers. By taking norms we reduce to the integer case, the same idea can be implemented and we are able to show that the reduction of at least one term of the sequence of algebraic numbers to  $\mathbb{F}_p$  is almost always an  $h$ th power.

In Section 3 we state the GRH-conditional result mentioned earlier. In Section 4 we reduce the problem to a polynomial equation of type (2.3) in a similar manner as above. We note that the equation is solvable for a set of primes of density 1 if certain field extensions are linearly disjoint. We present the tool to handle such questions in Section 5. In order to apply it we have to check that no subsequence  $(x_{An+B})$  of a linear recurrence  $(x_n)$  appearing in our proof consists only of perfect powers, which we do in Section 6. We wrap up the proof in Section 7.

### 3. Orders of reductions of algebraic numbers

The following lemma is used when transforming our problem into a polynomial equation. This lemma is the only part of the proof that relies on GRH.

LEMMA 3.1. *Assume GRH. Let  $P \in \mathbb{Z}[x]$  be non-constant and irreducible. Assume  $P$  is not a cyclotomic polynomial, i.e. at least one root of  $P$  is not a root of unity, and that  $P$  is not the identity. Let  $S$  denote the set of primes such that the equation  $P(x) \equiv 0 \pmod{p}$  has at least one solution  $f(p)$  for  $p \in S$ . For  $C > 0$  let  $S_C$  denote the set of primes  $p \in S$  for which the order of  $f(p)$  modulo  $p$  is at least  $(p - 1)/C$ . The (lower) density of  $S_C$  with respect to  $S$  approaches 1 as  $C \rightarrow \infty$ .*

Note that we do not say anything about which root  $f(p)$  of  $P$  modulo  $p$  we choose if there are several of them. The result is true no matter how the choices are made.

This result is equivalent to the following algebraic number theoretic formulation.

LEMMA 3.2. *Assume GRH. Let  $\alpha$  be a non-zero algebraic number which is not a root of unity. Let  $K = \mathbb{Q}(\alpha)$  and let  $O_K$  denote the ring of integers of  $K$ . Let  $T$  denote the set of prime ideals of  $O_K$  whose norm is a prime. For  $C > 0$  let  $T_C$  denote the set of primes  $\mathfrak{p}$  of  $T$  such that the reduction of  $\alpha$  in  $O_K/\mathfrak{p} \cong \mathbb{F}_p$  has order at least  $(p - 1)/C$ , where  $p$  is the norm*

of  $\mathfrak{p}$ . The (lower) density of  $T_C$  with respect to  $T$  approaches 1 as  $C \rightarrow \infty$  (where ideals are ordered by norm).

*Proof.* (Cf. [24, section 5].) Note that almost all prime ideals of  $O_K$  belong to  $T$ . For  $k \in \mathbb{Z}_+$  let  $T'_k = T_k \setminus T_{k-1}$ . The results of Lenstra [9] imply that  $T'_k$  has a density for all  $k$ . This density is given by

$$d(T'_k) = \sum_{t=1}^{\infty} \frac{\mu(t)}{[K(\zeta_{kt}, \alpha^{1/kt}) : K]},$$

where the sum is absolutely convergent. Let  $f(n) = 1/[K(\zeta_n, \alpha^{1/n}) : K]$ . Now rearranging gives

$$\sum_{k=1}^{\infty} \sum_{t=1}^{\infty} \mu(t)f(kt) = \sum_{K=1}^{\infty} f(K) \sum_{d|K} \mu(d) = f(1) = 1,$$

from which the result follows.

#### 4. Reduction to a polynomial equation

We consider the setup of Theorem 1.1. The proof also works in the situation of Theorem 1.2.

Let  $P$  and  $(a_n)$  be as in Theorem 1.1. Assume  $a_n \neq 0$  for all  $n$ , as otherwise we are done. Let  $S$  denote the set of primes  $p$  such that  $P$  factorises as the product of polynomials of degree 1 and  $d - 1$  modulo  $p$ . By the assumption and the Chebotarev density theorem, the relative density of  $S$  is  $1/(d - 1)$ . We prove that the density of primes of  $S$  which are prime divisors of  $(a_n)$  is one (relative to  $S$ ).

We write

$$a_n = \gamma_1 \alpha_1^n + \dots + \gamma_d \alpha_d^n,$$

where  $\alpha_1, \dots, \alpha_n$  are the roots of  $P$  and  $\gamma_1, \dots, \gamma_d$  are constants. For further purposes we define additionally

$$b_n = b_{h,n} := \alpha_1^{hn} \left( -\frac{a_n}{\gamma_1 \alpha_1^n} + 1 \right)$$

and

$$c_n = c_{h,n} := N_{F/\mathbb{Q}}(b_n) = (-1)^d N^{hn} \prod_{i=1}^d \sum_{j \neq i} \frac{\gamma_j \alpha_j^n}{\gamma_i \alpha_i^n},$$

where  $F$  is the splitting field of  $P$ ,  $N := N_{F/\mathbb{Q}}(\alpha_1)$  and  $h \in \mathbb{Z}_+$  is a parameter fixed later. For an algebraic number field  $K$  we let  $O_K$  denote its ring of integers.

Note that  $\gamma_i \in \mathbb{Q}(\alpha_i) \setminus \{0\}$  (see e.g. [19, section 2]).

For each (unramified) prime  $p \in S$  there exists a homomorphism  $\varphi : O_F \rightarrow \mathbb{F}_{p^{d-1}}$  mapping one of the roots  $\alpha_i$ , say  $\alpha_1$ , to an element of  $\mathbb{F}_p$ , and the other roots to elements of  $\mathbb{F}_{p^{d-1}}$  whose degrees over  $\mathbb{F}_p$  are  $d - 1$ . Let  $K = \mathbb{Q}(\alpha_1)$ .

Note that for  $n = k(p^{d-1} - 1)/(p - 1) + r$  one has

$$\varphi(\alpha_1)^n = \varphi(\alpha_1)^{kd+r},$$

$$\varphi(\alpha_i)^n = N_{\mathbb{F}_{p^{d-1}}/\mathbb{F}_p}(\varphi(\alpha_i))^k \varphi(\alpha_i)^r, \quad 2 \leq i \leq d$$

and

$$N_{\mathbb{F}_{p^{d-1}}/\mathbb{F}_p}(\alpha_i) = \varphi(\alpha_2) \cdots \varphi(\alpha_d) = \frac{N}{\varphi(\alpha_1)}, \quad 2 \leq i \leq d.$$

(Clearly  $\varphi(\alpha_1) \neq 0$  for all but finitely many  $p$ .) Hence

$$\varphi(a_n) = \varphi(\gamma_1)\varphi(\alpha_1)^{kd+r} + \left(\frac{N}{\varphi(\alpha_1)}\right)^k (\varphi(\gamma_2)\varphi(\alpha_2)^r + \cdots + \varphi(\gamma_d)\varphi(\alpha_d)^r) \tag{4.1}$$

so  $\varphi(a_n) = 0$  if and only if

$$\varphi(\alpha_1^{d+1}/N)^k = -\frac{\varphi(\gamma_2)\varphi(\alpha_2)^r + \cdots + \varphi(\gamma_d)\varphi(\alpha_d)^r}{\varphi(\gamma_1)\varphi(\alpha_1)^r}.$$

We then note that  $\alpha_1^{d+1}/N$  is not a root of unity, for otherwise the conjugates  $\alpha_i^{d+1}/N$  are roots of unity as well. Hence the product

$$\prod_{1 \leq i \leq d} \alpha_i^{d+1}/N = N$$

is a root of unity. Hence  $N = \pm 1$ , and thus the numbers  $\alpha_i$  are roots of unity. This contradicts the non-degeneracy of  $P$ .

We may hence apply Lemma 3.1 to  $\alpha_1^{d+1}/N$ . It suffices to show that for any  $h \in \mathbb{Z}_+$  the equation

$$x^h = b_r$$

is solvable in  $\mathbb{F}_p \setminus \{0\}$  for almost all primes  $p \equiv 1 \pmod{h}$ . This may be reformulated as follows: almost all prime ideals of  $O_{K(\zeta_h)}$  split in at least one of the fields  $K_r := K(\zeta_h, b_r^{1/h})$ ,  $r = 1, 2, \dots$

### 5. A linear disjointness result

We use the following basic result from Galois theory [4].

LEMMA 5.1. *Let  $h$  be a positive integer, let  $K$  be a number field containing a  $h$ th root of unity, and let  $a_1, a_2, \dots, a_k$  be a sequence of integers. Assume that for any integers  $0 \leq e_i < h$ , not all zero, one has  $a_1^{e_1/h} \cdots a_k^{e_k/h} \notin K$ . Then*

$$\left[ K \left( a_1^{1/h}, \dots, a_k^{1/h} \right) : K \right] = h^k.$$

LEMMA 5.2. *Let  $(x_n)$  be a linearly recursive sequence of integers. Assume that there do not exist a number field  $K$  and integers  $A, B \geq 1$ ,  $D \geq 2$  such that  $x_{An+B}$  is always a  $D$ th power of an element of  $K$ . Then there exist a subsequence  $x_{n_1}, x_{n_2}, \dots$  of  $(x_n)$  and primes  $p_1, q_1, p_2, q_2, \dots$  such that:*

- (i)  $\gcd(v_{p_i}(x_{n_i}), v_{q_i}(x_{n_i})) = 1$  for all  $i$ ;
- (ii) the primes  $p_1, q_1, p_2, q_2, \dots$  are pairwise distinct.

*Proof.* Note that the condition implies that  $(x_n)$  has infinitely many prime divisors – otherwise choose  $A = 1, B = 0, D = 2$  and  $K$  to be  $\mathbb{Q}(\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_n})$ , where  $p_1, \dots, p_n$  are the prime divisors of  $(x_n)$ . Note also that for all except finitely many primes  $p$ , say for all  $p$  not belonging to  $T$ , the sequence  $(x_n)$  is periodic modulo  $p^k$  for all  $k \in \mathbb{Z}_+$ .

We will inductively choose the indices  $n_i$  and the primes  $p_i, q_i$ , additionally requiring that  $p_i, q_i \notin T$ . Assume we have already chosen some  $n_1, \dots, n_k, p_1, \dots, p_k, q_1, \dots, q_k$  satisfying the conditions. We now pick  $n_{k+1}, p_{k+1}, q_{k+1}$ . Let  $S = \{p_1, q_1, \dots, p_k, q_k\} \cup T$ .

Pick some prime divisor  $p_{k+1} \notin S$ , let  $v_{p_{k+1}}(x_{n_0}) = t > 0$  for some  $n_0$  such that  $x_{n_0} \neq 0$ . By periodicity modulo  $p_{k+1}^{t+1}$ , there exists an arithmetic progression  $An + B, n = 1, 2, \dots$  such that  $v_{p_{k+1}}(x_{An+B}) = t$  for all  $n$

If there exist some prime  $q_{k+1} \notin S \cup \{p_{k+1}\}$  and  $n \in \mathbb{Z}_+$  such that  $\gcd(v_{q_{k+1}}(x_{An+B}), t) = 1$ , then we are done. Assume this is not the case.

If for any prime  $q_{k+1} \notin S$  and any  $n$  we had  $\gcd(v_{q_{k+1}}(x_{An+B}), t) = t$ , then we could write

$$|x_{An+B}| = f(n)^t \prod_{s \in S} s^{f_s(n)},$$

for some functions  $f, f_s : \mathbb{Z}_+ \rightarrow \mathbb{Z}_{\geq 0}$ . Then  $x_{An+B}$  would always be a perfect  $t$ th power in a number field containing the  $t$ th roots of all primes of  $S$  and a  $2t$ th root of unity, contrary to the assumption.

Hence there exist a prime  $q_{k+1} \notin S$  and  $n \in \mathbb{Z}_+$  such that  $t' := \gcd(v_{q_{k+1}}(x_{An+B}), t) < t$ . Repeat the above argument with  $q_{k+1}$  in place of  $p_{k+1}$  and  $t'$  in place of  $t$ . The value of  $t$  decreases. It must happen that for some value of  $p_{k+1}$  and  $t$  we find a prime  $q_{k+1}$  with  $\gcd(v_{q_{k+1}}(x_{An+B}), t) = 1$ .

LEMMA 5.3. Let  $h \in \mathbb{Z}_+$  and let  $F$  be a number field containing a  $h$ th root of unity. Let  $(x_n)$  be a linearly recursive sequence of integers satisfying the assumption of Lemma 5.2. Then there exists a subsequence  $x_{n_1}, x_{n_2}, \dots$  of  $(x_n)$  such that the extensions

$$F(x_{n_1}^{1/h})/F, F(x_{n_2}^{1/h})/F, \dots$$

are linearly disjoint and of degree  $h$ .

*Proof.* Note first that there exists a constant  $c$  (depending on  $F$ ) such that if  $x$  is an integer with  $x^{1/h} \in F$ , then all prime divisors of  $x$  which are greater than  $c$  have multiplicity divisible by  $h$ . (Indeed: If  $p$  is a prime with for which  $h$  does not divide  $v_p(x)$ , then  $p$  is ramified in the  $\mathbb{Q}(x^{1/h})$ . If  $\mathbb{Q}(x^{1/h}) \subset F$ , then  $p$  is ramified in  $F$ , and only finitely many primes ramify in  $F$ .)

Let  $(x_{n_i})$  be a subsequence constructed in Lemma 5.2. We may assume that the corresponding primes  $p_i, q_i$  are larger than  $c$ . We prove that this subsequence works by showing that

$$\left[ F(x_{n_1}^{1/h}, x_{n_2}^{1/h}, \dots, x_{n_k}^{1/h}) : F \right] = h^k$$

for all  $k$ . Apply Lemma 5.1. Assume that

$$x_{n_1}^{e_1/h} \dots x_{n_k}^{e_k/h} \in F, 0 \leq e_i < h. \tag{5.1}$$

By the choice of  $c$ , (5.1) implies that the prime divisors of  $x_{n_1}^{e_1} \cdots x_{n_k}^{e_k}$  which are larger than  $c$  have multiplicity divisible by  $h$ . By the choice of the primes  $p_i, q_i$  this implies  $h \mid e_i$  for all  $i$ .

6. Linear recurrences and perfect powers

In this section we show that Lemma 5.3 may be applied to the sequence  $(c_n)$ . Assume not, so  $c_{An+B}$  is always a  $D$ th power of an element in a fixed number field.

By a result of Zannier [25], the only case when a linear recurrence is always a  $D$ th power is when it is the  $D$ th power of a linear recurrence. We may hence write

$$c_{An+B} = d_n^D,$$

where  $d$  is a linearly recursive sequence (whose elements are not necessarily integers). Write  $d$  as an exponential polynomial

$$d_n = \sum_{m=0}^t \left( n^t \sum_{k=1}^{u_m} e_{m,k} \sigma_{m,k}^n \right),$$

where the coefficients  $e_{m,k}$  are non-zero,  $\sigma_{m,1}, \sigma_{m,2}, \dots, \sigma_{m,u_m}$  are non-zero and pairwise distinct, and  $i_m > 0$  for all  $m \leq t$ .

We first show that  $t = 0$ , i.e. that the characteristic polynomial of  $(d_n)$  has no repeated roots. Assume not. Now one sees that  $d_n^D$  may be written as

$$n^{Dt} \left( \sum_{k=1}^{u_t} e_{t,k} \sigma_{t,k}^n \right)^D + f_n,$$

where  $f_n$  is an exponential polynomial whose polynomial terms have degree less than  $Dt$ . Since the representation of a linear recurrence as an exponential polynomial is unique, one sees that  $c_{An+B} = d_n^D$  cannot hold for all  $n \in \mathbb{Z}$ .

We may thus write

$$d_n = e_1 \sigma_1^n + \cdots + e_u \sigma_u^n.$$

Every  $\sigma_i$  can be written in the form  $\alpha_1^{x_{i,1}} \alpha_2^{x_{i,2}} \cdots \alpha_d^{x_{i,d}}$ , where the  $x_{i,j}$  are rational numbers [18]. Let  $M$  be a positive integer such that  $x_{i,j}M$  is an integer for all  $i, j$ . Write the equation  $c_{AMn+B} = d_{Mn}^D$  as

$$(-1)^d N^{h(AMn+B)} \prod_{i=1}^d \sum_{j \neq i} \frac{\gamma_j \alpha_j^B \alpha_j^{AMn}}{\gamma_i \alpha_i^B \alpha_i^{AMn}} = \left( \sum_{i=1}^u e_i \prod_{j=1}^d \alpha_j^{x_{i,j}Mn} \right)^D$$

Note then that if  $\prod_{i=1}^d \alpha_i^{f_i} \in \mathbb{Q}$  for some integers  $f_i$ , then  $f_i = f_j$  for all  $i, j$ . Indeed: since the Galois group of  $P$  is  $S_d$ , we have  $\prod_{i=1}^d \alpha_i^{f_i'} = \prod_{i=1}^d \alpha_i^{f_i}$  for any permutation  $f_i'$  of  $f_i$ . Hence  $\alpha_i^{f_i} \alpha_j^{f_j} = \alpha_i^{f_j} \alpha_j^{f_i}$ , so  $(\alpha_i / \alpha_j)^{f_i - f_j} = 1$ , which by non-degeneracy of  $P$  implies  $f_i = f_j$ .

Hence, if  $N \neq \pm 1$ , one has  $\prod_{i=1}^d \alpha_i^{f_i} = 1$  only if  $f_i = 0$  for all  $i$ . By basic results on linear recurrences, this implies that for any  $Q \in \mathbb{C}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]$  in  $d$  variables we have

$$Q(\alpha_1^n, \alpha_2^n, \dots, \alpha_d^n) = 0$$

for all integers  $n$  if and only if  $Q$  is identically zero. Hence

$$(-1)^d N^{hB} (X_1 \cdots X_d)^{hAM} \prod_{i=1}^d \sum_{j \neq i} \frac{\gamma_j \alpha_j^B X_j^{AM}}{\gamma_i \alpha_i^B X_i^{AM}} = \left( \sum_{i=1}^u e_i \prod_{j=1}^d X_j^{x_{ij}M} \right)^D$$

identically as elements of  $\mathbb{C}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]$ .

In particular, the left hand side is a perfect  $D$ th power in  $\mathbb{C}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]$ . Perform a suitable transformation of the form  $X_i \rightarrow c_i X_i$ , clear out constants and simplify. One obtains that

$$(X_1 \cdots X_d)^{(h-1)A'} \prod_{i=1}^d (X_1^{A'} + \cdots + X_d^{A'} - X_i^{A'}) \tag{6.1}$$

is a  $D$ th power of a polynomial, where  $A' := AM$ . This is clearly impossible if  $D$  does not divide  $(h-1)A'$ . Otherwise we may drop the term  $(X_1 \cdots X_d)^{(h-1)A'}$ . One sees that the polynomials

$$X_1^{A'} + \cdots + X_d^{A'} - X_i^{A'}$$

are pairwise coprime, and hence each of them must be a  $D$ th power. This is not the case, as can be seen, for example, by considering the partial derivative of  $X_2^{A'} + \cdots + X_d^{A'}$  with respect to  $X_2$  at  $(X_2, X_3, \dots, X_d) = (x_0, 1, \dots, 1)$ , where  $x_0^{A'} + (d-2) = 0$ .

The case  $N = \pm 1$  is handled similarly: For any polynomial  $Q \in \mathbb{C}[X_1^{\pm 1}, \dots, X_{d-1}^{\pm 1}]$  we have

$$Q(\alpha_1^n, \alpha_2^n, \dots, \alpha_{d-1}^n) = 0$$

for all integers  $n$  only if  $Q$  is zero. Proceeding as before, we have that

$$\prod_{i=1}^d (X_1^{A'} + \cdots + X_d^{A'} - X_i^{A'})$$

is a  $D$ th power of a polynomial in the variables  $X_1, \dots, X_{d-1}$  (with possibly negative exponents in the monomials), where  $X_d$  is shorthand for  $N/X_1 \cdots X_{d-1}$ . Note that the term

$$X_1^{A'} + \cdots + X_{d-1}^{A'}$$

is coprime with all the other terms of the product, and, as before, this is not a  $D$ th power of a polynomial.

### 7. Concluding the proof

We aim to prove that almost all primes of  $K(\zeta_h)$  split in at least one of the fields  $K(\zeta_h, b_n^{1/h})$ . By the Chebotarev density theorem it suffices to construct a subsequence  $b_{n_1}, b_{n_2}, \dots$  of  $(b_n)$  such that the fields

$$K(\zeta_h, b_{n_1}^{1/h}) / K(\zeta_h), K(\zeta_h, b_{n_2}^{1/h}) / K(\zeta_h), \dots$$

are linearly disjoint

In Section 6 we checked that we may apply Lemma 5.3 to the norm sequence  $(c_n)$ . Let  $c_{n_1}, c_{n_2}, \dots$  denote a subsequence given by the lemma with the base field  $F(\zeta_h)$ , so

$$F(\zeta_h, c_{n_i}^{1/h}) / F(\zeta_h), i = 1, 2, \dots$$

are of degree  $h$  and linearly disjoint. We claim that this implies that

$$F(\zeta_h, b_{n_i}^{1/h}) / F(\zeta_h), i = 1, 2, \dots$$

are of degree  $h$  and linearly disjoint, too.

By Lemma 5.1 it suffices to show that

$$b_{n_1}^{e_1/h} \dots b_{n_k}^{e_k/h} \in F(\zeta_h), 0 \leq e_i < h$$

implies  $e_i = 0$  for all  $i$ . But if  $b_{n_1}^{e_1} \dots b_{n_k}^{e_k}$  is an  $h$ th power in  $F(\zeta_h)$ , the norm  $c_{n_1}^{e_1} \dots c_{n_k}^{e_k}$  is a  $h$ th power in  $F(\zeta_h)$ , too. By the choice of  $c_{n_i}$  this happens only if  $e_i = 0$  for all  $i$ .

Finally, note that linear disjointness over  $F(\zeta_h)$  implies linear disjointness over  $K(\zeta_h)$ , so  $K(\zeta_h, b_{n_1}^{1/h})/K(\zeta_h), K(\zeta_h, b_{n_2}^{1/h})/K(\zeta_h), \dots$  are linearly disjoint, as desired.

### 8. Discussion

The presented proof considers the primes  $p$  such that  $P$  has factorisation type  $(1, d - 1)$  modulo  $p$ . Naturally one wonders whether other factorisation types could be handled as well. Unfortunately, we are not able to do this.

There are two main limitations. First, performing the “norm trick” as in Section 4 requires that  $P$  has just two factors modulo  $p$ . Second, to reduce the exponential equation to a polynomial equation one needs a result similar to Lemma 3.1. The analogue of Lemma 3.1 is, however, not known when one considers the roots of  $P$  in  $\mathbb{F}_{p^k}$  for  $k \geq 2$ .

By an approach also based on reduction to a polynomial equation, Roskam has settled the case where  $P$  remains irreducible modulo  $p$ , assuming an analogue of Lemma 3.1 holds for roots of  $P$  of degree  $d$  over  $\mathbb{F}_p$  (namely that the multiplicative orders of the roots are of magnitude  $p^d$  almost always). Note that given a linear recurrence  $(a_n)$  with a squarefree characteristic polynomial  $P$ , the period of  $(a_n)$  modulo  $p$  is, for large enough primes  $p$ , equal to the least common multiple of the multiplicative orders of the roots of  $P$  in extensions of  $\mathbb{F}_p$ . Hence Roskam’s assumption is equivalent to the period of  $(a_n)$  modulo  $p$  being of order  $p^d$  for almost all primes  $p$  for which  $P$  is irreducible modulo  $p$ .

A theorem of Niederreiter [16, theorem 4.1] gives a stronger result under a weaker assumption: as long as the period of the sequence modulo  $p$  is at least of magnitude  $p^{d/2+1}$ , the sequence does not only attain the value  $0 \pmod{p}$ , but the sequence is approximately equidistributed modulo  $p$ , see [21, theorem 3] for a similar result. While we have managed to avoid the consideration of the period of  $a_n$  modulo  $p$ , our approach does not yield equidistribution results or even non-trivial lower bounds for the number of values attained by  $a_n$  modulo  $p$ .

We hence see that analogies of Lemma 3.1 to roots of  $P$  in extensions of  $\mathbb{F}_p$  are central to understanding the behavior of linear recurrences modulo primes. While it seems likely that such variants of Lemma 3.1 hold (one can present a similar heuristic as for Artin’s conjecture), our understanding is very limited. Unconditionally, we only know that given an integer  $a$  with  $|a| > 1$ , the order of  $a$  modulo  $p$  is almost always  $> p^{1/2}$  [3]. Under GRH one has Lemma 3.1, and an involved variant of Hooley’s classical (conditional) solution

of Artin's conjecture yields an analogue of Lemma 3.1 in the case where  $P$  is of degree 2 and remains irreducible modulo  $p$ , as shown by Roskam [20]. (Roskam only considers the case where the order is equal to exactly  $p^2 - 1$  in [20], but the argument may be modified for order  $(p^2 - 1)/C$  to give Lemma 3.1.) It seems that all other cases are open, and as explained in [20], Hooley's argument does not adapt to higher degrees without new ideas.

We conclude by remarking that the case where  $P$  splits into  $d$  linear factors modulo  $p$  seems to be the most difficult to analyse. In these cases the period of the linear recurrence modulo  $p$  divides  $p - 1$ , and heuristically there is a positive density of split primes which are not prime divisors of the sequence (see [19]). For example, we are not able to say essentially anything about the prime divisors of  $3^n + 2^n + 3$  other than that there are infinitely many of them. In contrast, heuristics suggest that if  $P$  is irreducible and  $\deg(P) \geq 3$ , then almost all non-split primes are prime divisors of the corresponding sequence (excluding degenerate cases). This suggests that the lower bound  $1/(d - 1)$  in Theorem 1.1 could be replaced with  $1 - 1/d!$ .

## REFERENCES

- [1] C. BALLOT. Density of prime divisors of linear recurrences, vol. 551. (*Amer. Math. Soc.*, 1995).
- [2] M. DRMOTA and M. SKABA. Relations between polynomial roots. *Acta Arith.* **71**(1) (1995), 64–77.
- [3] P. ERDŐS and M. R. MURTY. *On the order of a (mod p)*. In CRM Proc. Lecture Notes, vol. 19, (1999), 87–97.
- [4] P. GARRETT. Linear independence of roots (2011). Unpublished note, online at [https://www-users.cse.umn.edu/~garrett/m/v/linear\\_indep\\_roots.pdf](https://www-users.cse.umn.edu/~garrett/m/v/linear_indep_roots.pdf)
- [5] H. HASSE. Über die Dichte der Primzahlen  $p$ , für die eine vorgegebene ganzrationale Zahl  $a \neq 0$  von gerader bzw. ungerader Ordnung mod.  $p$  ist. *Math. Ann.* **166** (1966), 19–23.
- [6] C. HOOLEY. On Artin's conjecture. *J. Reine Angew. Math.* **225** (1967), 209–220.
- [7] J. C. LAGARIAS. The set of primes dividing the Lucas numbers has density  $2/3$ . *Pacific J. Math.* **118**(2) (1985), 449–461.
- [8] J. C. LAGARIAS. Errata to: The set of primes dividing the Lucas numbers has density  $2/3$ . *Pacific J. Math.* **162**(2) (1994), 393–396.
- [9] H. W. LENSTRA, Jr. On Artin's conjecture and Euclid's algorithm in global fields. *Invent. Math.* **42** (1977), 201–224.
- [10] P. MOREE. On the prime density of Lucas sequences. *J. Théor. Nombres Bordeaux* **8**(2) (1996), 449–459.
- [11] P. MOREE. Artin's primitive root conjecture -a survey. *Integers* **12** (2005), 01.
- [12] P. MOREE and P. STEVENHAGEN. Prime divisors of Lucas sequences. *Acta Arith.* **82**(4) (1997), 403–410.
- [13] P. MOREE and P. STEVENHAGEN. A two-variable Artin conjecture. *J. Number Theory* **85**(2) (2000), 291–304.
- [14] P. MOREE and P. STEVENHAGEN. Prime divisors of the Lagarias sequence. *J. Théor. Nombres Bordeaux* **13**(1) (2001), 241–251.
- [15] M. R. MURTY, F. SÉGUIN, and C. L. STEWART. A lower bound for the two-variable Artin conjecture and prime divisors of recurrence sequences. *J. Number Theory*, **194** (2019), 8–29.
- [16] H. NIEDERREITER. On the cycle structure of linear recurring sequences. *Math. Scand.* **38**(1) (1976), 53–77.
- [17] G. PÓLYA. Arithmetische Eigenschaften der Reihenentwicklungen rationaler Funktionen. *J. Reine Angew. Math.* **151** (1921), 1–31.
- [18] J. F. RITT. A factorisation theory for functions  $\sum_{i=1}^n a_i e^{a_i x}$ . *Trans. Amer. Math. Soc.* **29**(3) (1927), 584–596.
- [19] H. ROSKAM. Prime divisors of linear recurrences and Artin's primitive root conjecture for number fields. *J. Théor. Nombres Bordeaux* **13**(1) (2001), 303–314.

- [20] H. ROSKAM. Artin's primitive root conjecture for quadratic fields. *J. Théor. Nombres Bordeaux* **14**(1) (2002), 287–324.
- [21] I. E. SHPARLINSKI. Distribution of nonresidues and primitive roots in recurrent sequences. *Mat. Zametki* **24**(5) (1978), 603–613.
- [22] P.J. STEPHENS. Prime divisors of second-order linear recurrences. i. *J. Number Theory* **8**(3) (1976), 313–332.
- [23] P. STEVENHAGEN. Prime densities for second order torsion sequences (2007).
- [24] S. WAGSTAFF. Pseudoprimes and a generalisation of Artin's conjecture. *Acta Arith.* **41**(2) (1982), 141–150.
- [25] U. ZANNIER. A proof of Pisot's dth root conjecture. *Ann. Math.* **151**(1) (2000), 375–383.