New York Journal of Mathematics

New York J. Math. **12** (2006) 39–45.

Prime factors of $a^{f(n)} - 1$ with an irreducible polynomial $f(x)$

Christian Ballot and **Florian Luca**

Abstract. In this note, we show that if a is an integer not 0 or ± 1 and $f(X) \in \mathbb{Q}[X]$ is an integer valued irreducible polynomial of degree $d \geq 2$, then the set of primes p dividing $a^{f(n)} - 1$ for some positive integer n is of (relative) asymptotic density zero.

Contents

1. Introduction 39
2. The proof 40
3. An application 43
References 44

1. Introduction

Let a be an integer not 0 or ± 1. For any integer n prime to a we write $\ell_a(n)$ for the order of a as an element of the group $(\mathbb{Z}/n\mathbb{Z})^*$.

In recent years, partly motivated by cryptographic demands, several authors have investigated the arithmetic structure of the numbers $\ell_a(n)$, when n ranges either over all the positive integers, or only over the set of prime numbers. For example, the prime numbers p such that $\ell_a(p)$ is square-free have been investigated by Pappalardi in [6], while the set of primes p such that $\ell_a(p)$ is smooth have been investigated by Pomerance and Shparlinski in [7]. Recall here that a positive integer n is called smooth if its largest prime factor q is small; i.e., if the ratio $\log q / \log n$ is small.

In this paper, we fix an integer a different from 0 and ± 1 and an irreducible polynomial $f(X) \in \mathbb{Q}[X]$ of degree $d \geq 2$ which is integer valued and of positive leading term, and we study the set of prime factors of the numbers $u_n = a^{f(n)} - 1$ as

Received October 20, 2005.

Mathematics Subject Classification. 11N37, 11B37.

Key words and phrases. Prime factors, linear recurrences, Chebotarev Density Theorem.

This paper was written during a very enjoyable visit by the second author to the Laboratoire Nicolas Oresme of the University of Caen; he wishes to express his thanks to that institution for its hospitality and support. He was also partly supported by grants SEP-CONACYT 46755, PAPIIT IN104505 and a Guggenheim Fellowship.

ISSN 1076-9803/06

n ranges over the set of positive integers. Note that such primes p are precisely the ones for which the congruence $f(n) \equiv 0 \pmod{\ell_a(p)}$ admits one (hence, infinitely many) positive integer solutions n. Our main result shows, perhaps not unexpectedly, that most prime numbers do not divide u_n for any value of the positive integer n.

For a positive real number x we write

$$\mathcal{U}_f(x) = \left\{ p \le x : p \mid a^{f(n)} - 1 \text{ for some positive integer } n \right\}.$$

Theorem 1. *Let $f(X) \in \mathbb{Q}[X]$ be an integer valued irreducible of degree $d \ge 2$. There is a constant $r_f > 0$ such that for every $\varepsilon > 0$ there exists $x_\varepsilon > e^e$ such that*

$$\sharp\mathcal{U}_f(x) < \frac{x}{\log x (\log\log\log x)^{r_f - \varepsilon}} \qquad \text{for } x > x_\varepsilon.$$

In particular, $\sharp\mathcal{U}_f(x) = o(\pi(x))$ as $x \to \infty$.

The constant r_f is the asymptotic density of primes p that divide some $f(n)$, where n is a natural number. Bounds for r_f are given in terms of the degree d in Lemma 3.

We do not address the problem of determining a lower bound for $\sharp\mathcal{U}_f(x)$. But at least note that $\sharp\mathcal{U}_f(x) \to \infty$ as $x \to \infty$, since by the Primitive Divisor Theorem, for n large enough there is always a prime factor p of $a^n - 1$ which does not divide $a^m - 1$ for any $m < n$. Taking this assertion through the set of values of $f(n)$ proves our statement.

Under the Generalized Riemann Hypothesis, the $\log\log\log x$ appearing on the right-hand side of the inequality from Theorem 1 can be replaced by $\log\log x$. However, since the constant r_f is less than 1, the resulting bound on $\mathcal{U}_f(x)$ is not even strong enough to lead to the conclusion that the sum of the reciprocals of the primes in $\mathcal{U}_f$ is convergent. We give no details in this direction.

Throughout this paper, we use the Vinogradov symbols $\gg$ and $\ll$ and the Landau symbols O and o with their regular meanings. The constants implied by them may depend on the given integer a and polynomial $f(X)$. For a set $\mathcal{A}$ of positive integers we put $\mathcal{A}(x) = \mathcal{A} \cap [1, x]$. Some remarks having to do with primes dividing two consecutive terms of cubic integral linear recurring sequences whose associated polynomial has integer roots are presented in Section 3 as an application of Theorem 1.

Acknowledgements. The first author is thankful to Michael Filaseta for sharing some insight on primes that divide two consecutive terms of the sequence $a_n = 2 \cdot 3^n - 3 \cdot 2^n + 1$ at the West Coast Number Theory Conference of 1993. The authors are grateful to the referee for a fast appreciation of the paper.

2. The proof

The following lemma plays a crucial rôle in our proof and is a special instance of Theorem 1.3 in [6], where we have replaced the function $Li(x)$ by $\pi(x)$. This is justified since we know that

$$Li(x) = \pi(x) \left(1 + O\left(\exp(-c\sqrt{\log x})\right)\right),$$

with some constant $c > 0$, and certainly

$$\exp(-c\sqrt{\log x}) = o\left((\log x)^{-1/8}\right) \qquad \text{as } x \to \infty.$$

Lemma 2. *Assume that $a > 1$ is not the k-th power of an integer for any $k \geq 2$. Let m be an odd positive integer and x be a positive integer. Consider the set*

$$\mathcal{A}(x, m) = \{p \leq x : m \mid \ell_a(p)\}.$$

Then, for every $\varepsilon > 0$, the estimate

$$\sharp\mathcal{A}(x, m) = \kappa_m \left(1 + O\left(\frac{m^{1-2\varepsilon}}{(\log x)^{1/8-\varepsilon}}\right)\right) \pi(x) \tag{1}$$

holds uniformly in m and x where

$$\kappa_m = \frac{1}{m} \prod_{l|m} \frac{l^2}{l^2 - 1}, \qquad (l \text{ prime}). \tag{2}$$

We point out that Theorem 1.3 in [6] is more general since it covers the cases when m is even or a is a power of a positive integer. Note also that particular instances of Lemma 2 have appeared previously in works of Odoni [8], Ballot [1], p. 32, and Wiertelak [10].

We will also need the following well-known consequence of Chebotarev's Density Theorem. To be precise, it only requires the Kronecker and the Frobenius Density Theorems [5], [4], both of which being corollaries of the later, and now famous, Chebotarev Density Theorem, originally conjectured by Frobenius. The theorem we attribute to Kronecker appears in [5] as a consequence of an actual theorem, *provided* the existence of the densities δ_i is assumed. This existence was first shown by Frobenius.

Let $f(X) \in \mathbb{Q}[X]$ be irreducible of degree $d \geq 2$. Let

$$\mathcal{R}_f = \{p : f(n) \equiv 0 \ (\text{mod } p) \text{ does not admit an integer solution } n\}.$$

Lemma 3. *The set $\mathcal{R}_f$ has a positive (relative) asymptotic density r_f. Furthermore, r_f is a rational number in the interval $[(d-1)/d!, 1 - 1/d]$.*

Proof. By the Frobenius Density Theorem the set of primes p for which the factorization of $f(X)$ $(\text{mod } p)$ contains exactly i linear factors has a Dirichlet density δ_i. Therefore, $\sum_{i=0}^d \delta_i = 1$. By the Kronecker Density Theorem, we also have $\sum_{i=0}^d i\delta_i = 1$. Hence,

$$r_f = \delta_0 = \sum_{i=1}^{d} (i-1)\delta_i \geq (d-1)\delta_d \geq \frac{d-1}{\sharp G} \geq \frac{d-1}{d!},$$

where we wrote G for the Galois group of the splitting field of $f(X)$. But if H is the subgroup of G that fixes some root of $f(X)$, then, by the Frobenius Density Theorem, primes p that divide $f(n)$ for some n have the Dirichlet density

$$\frac{\sharp\left(\cup_{x\in G} H^x\right)}{\sharp G},$$

where for $x \in G$ we wrote $H^x = xHx^{-1}$. Thus,

$$r_f = \frac{\sharp G - \sharp\left(\cup_{x\in G} H^x\right)}{\sharp G} \leq \frac{\sharp G - \sharp H}{\sharp G} = 1 - \frac{1}{d}.$$

But sets of primes thus arising from applying the Frobenius Density Theorem possess a (relative) asymptotic density equal to their Dirichlet density. □

We are now ready to prove Theorem 1.

Proof of Theorem 1. We may assume that a is not a k-th power, $k \ge 2$. Indeed if $a = b^k$, then $u_n = b^{g(n)} - 1$, where $g(X)$ is the integer valued polynomial $kf(X)$. Furthermore, by replacing $f(X)$ by $2f(X)$ if needed, we may assume that a is positive.

Let x be large. We put

$$y = \frac{1}{20}\log\log x.$$

Let $q_1 < \cdots < q_s$ be all odd primes in $\mathcal{R}_f(y)$. It is clear that if $p \in \mathcal{U}_f(x)$, then $\ell_a(p) \mid f(n)$ for some positive integer n, therefore $\ell_a(p)$ cannot be divisible by any of the primes q_i for $i = 1, \ldots, s$. Thus,

$$\mathcal{U}_f(x) \subset \{p \le x\} \backslash \left(\bigcup_{i=1}^{s} \mathcal{A}(x, q_i) \right),$$

which shows, via the Principle of Inclusion and Exclusion, that

$$\sharp\mathcal{U}_f(x) \le \pi(x) - \sum_{t=1}^{s} (-1)^{t-1} \sum_{i_1 < \cdots < i_t} \sharp\mathcal{A}(x, q_{i_1} \ldots q_{i_t}).$$

Using estimate (1) and (2) with $\varepsilon = 1/40$, we get that

$$\sharp\mathcal{U}_f(x) \le \pi(x) \left(1 + \sum_{t=1}^{s} (-1)^t \sum_{i_1 < \cdots < i_t} \kappa_{q_{i_1} \ldots q_{i_t}} + 2^s O\left(\frac{q_1 \ldots q_s}{(\log x)^{1/10}} \right) \right).$$

Having in mind the estimate $\sum_{p \le y} \log p = y(1 + o(1))$, we note that

$$\begin{aligned}
2^s q_1 \ldots q_s &= \exp\left(s \log 2 + \sum_{q \in \mathcal{R}(y)} \log q \right) \\
&\le \exp\left(\pi(y) \log 2 + \int_{2^-}^{y} \log t \, d\sharp\mathcal{R}_f(t) \right) \\
&= \exp\left(\pi(y) \log 2 + r_f \int_{2^-}^{y} (1 + o(1)) \log t \, d\pi(t) \right) \\
&= \exp\left(\pi(y) \log 2 + r_f(1 + o(1)) \int_{2^-}^{y} \log t \, d\pi(t) + o(y) \right) \\
&= \exp\Big(o(y) + r_f y(1 + o(1)) \Big) \\
&< \exp(y) = (\log x)^{1/20},
\end{aligned}$$

for large x since $r_f<1$. Furthermore, we have

$$1+\sum_{t=1}^{s}(-1)^t\sum_{i_1<\cdots<i_t}\kappa_{q_{i_1}\cdots q_{i_t}}=1+\sum_{t=1}^{s}(-1)^t\sum_{i_1<\cdots<i_t}\prod_{j=1}^{t}\frac{q_{i_j}}{q_{i_j}^2-1}$$
$$=\prod_{i=1}^{s}\left(1-\frac{q_i}{q_i^2-1}\right)$$
$$=\alpha_f\prod_{q\in\mathcal{R}_f(y)}\left(1-\frac{1}{q}\right),$$

where

$$\alpha_f=\prod_{q\in\mathcal{R}_f(y)}\left(1-\frac{1}{q}\right)^{-1}\left(1-\frac{q}{q^2-1}\right)$$
$$=\prod_{q\in\mathcal{R}_f(y)}\frac{q(q^2-q-1)}{(q-1)(q^2-1)}$$
$$=\prod_{q\in\mathcal{R}_f(y)}\frac{q^3-q^2-q}{q^3-q^2-q+1}<1.$$

Hence,

$$1+\sum_{t=1}^{s}(-1)^t\sum_{i_1<\cdots<i_t}k_{q_{i_1}\cdots q_{i_t}}<\prod_{q\in\mathcal{R}_f(y)}\left(1-\frac{1}{q}\right).$$

Since

$$\prod_{q\in\mathcal{R}_f(y)}\left(1-\frac{1}{q}\right)<\exp\left(-\sum_{q\in\mathcal{R}_f(y)}\frac{1}{q}\right)$$
$$=\exp\left(-\int_{2^-}^{y}\frac{1}{t}\,d\sharp\mathcal{R}_f(t)\right)$$
$$=\exp\left(-r_f(1+o(1))\int_{2^-}^{y}\frac{1}{t}\,d\pi(t)\right)$$
$$\le\exp\left(-r_f(1+o(1))\log\log y\right)=(\log y)^{-r_f+o(1)}$$
$$=(\log\log\log x)^{-r_f+o(1)},$$

we get that

$$\sharp\mathcal{U}_f(x)\le\pi(x)\left(\frac{1}{(\log\log\log x)^{r_f+o(1)}}+O\left(\frac{1}{(\log x)^{1/20}}\right)\right),$$

which obviously implies the conclusion of the theorem. □

3. An application

Let $(a_n)_{n\ge0}$ be an integral linear recurring sequence whose minimal characteristic polynomial g is a monic polynomial in $\mathbb{Z}[X]$ of degree 3 with integral roots that are distinct in absolute value. A prime p is a *maximal* divisor of $(a_n)_{n\ge0}$ if it divides two consecutive terms a_n,a_{n+1} for some n. To such a g and such sequences $(a_n)_{n\ge0}$

one associates a group structure of infinite rank and finite torsion in a natural way (see, for example, [3], p. 4 and 106-7). Based on experimental evidence, it seems that sequences not in the torsion subgroup have few maximal prime divisors (see [1], p. 44). Since there is a method to compute the positive Dirichlet density of such primes for 'torsion' sequences (see [1]), it would be interesting to assess the density for 'nontorsion' sequences. Is it always 0? We merely observe that for some nontorsion sequences such as the one of general term $a_n = 2\cdot 3^n - 3\cdot 2^n + 1$, Theorem 1 yields the answer. Indeed, assume a prime $p > 3$ is a prime dividing both a_n and a_{n+1} for some n. Then p divides $a_{n+1} - a_n = 4 \cdot 3^n - 3 \cdot 2^n$. Thus, $3^{n-1} \equiv 2^{n-2} \pmod p$. Similarly $a_{n+1} - 3a_n = 3\cdot 2^n - 2$ so that $3\cdot 2^{n-1} \equiv 1 \pmod p$. But raising this latter congruence to the power $n-1$ and using $3^{n-1} \equiv 2^{n-2} \pmod p$, we get $2^{n^2-n-1} \equiv 1 \pmod p$. The polynomial $X^2 - X - 1 \in \mathbb{Q}[X]$ being irreducible, the prime density of maximal divisors of $(a_n)_{n\geq 0}$ is 0. Note that if a prime $p > 3$ divides both b_n and b_{n+1}, where $b_n = 3^{n+1} - 2^{n+2} + 6$, then p divides both $b_{n+1} - 2b_n$ and $b_{n+1} - 3b_n$, implying that

$$(3) \qquad 3^n \equiv 2 \pmod p \qquad \text{and} \qquad 2^n \equiv 3 \pmod p, \quad \text{for some } n \geq 0.$$

Here, we have $2^{n^2-1} \equiv 1 \pmod p$ and $X^2 - 1$ is reducible so our result does not apply. However, Skałba [9] showed recently that primes satisfying (3) also have (relative) asymptotic density 0. Finally, note that the sequence of general term $c_n = 3^n - 2^{n+1} - 1$ is torsion of order 2 in the aforementioned group and that its associated density is 65/224 (see [1], Ch. 4). The problem of the relative density of maximal prime divisors of such ternary recurrent sequences is investigated in more detail in [2].

References

[1] Ballot, Christian. Density of prime divisors of linear recurrences. *Mem. Amer. Math. Soc.* **115** (1995), no. 551. MR1257079 (95i:11110), Zbl 0827.11006.

[2] Ballot, Christian; Luca, Florian. Common prime factors of $a^n - b$ and $c^n - d$. Preprint, 2006.

[3] Everest, Graham; van der Poorten, Alf; Shparlinski, Igor; Ward, Thomas. Recurrence sequences. Mathematical Surveys and Monographs, 104, *Mathematical Society, Providence, RI*, 2003. MR1990179 (2004c:11015), Zbl 1033.11006.

[4] Frobenius, G. Über Beziehungen zwischen den Primidealen eines algebraischen Körpers und den Substitutionen seiner Gruppe. *S'ber. Akad. Wiss. Berlin* (1896), 689–703. JFM 27.0091.04.

[5] Kronecker, L. Über die Irreductibilität von Gleichungen. *Monatsb. König. Preuss. Akad. Wiss. Berlin* (1880), 155–162. JFM 12.0065.02.

[6] Pappalardi, Francesco. Square free values of the order function. *New York J. Math.* **9** (2003), 331–344. MR2028173 (2004i:11116), Zbl 1066.11044.

[7] Pomerance, Carl; Shparlinski, Igor E. Smooth orders and cryptographic applications. *Algorithmic number theory* (Sydney, 2002), 338–348, Lect. Notes Comput. Sci. 2369, *Springer, Berlin*, 2002. MR2041095 (2005e:11117), Zbl 1058.11059.

[8] Odoni, R. W. K. A conjecture of Krishnamurthy on decimal periods and some allied problems. *J. Number Theory* **13** (1981), 303–319. MR0634201, Zbl 0471.10031.

[9] Skałba, M. Primes dividing both $2^n - 3$ and $3^n - 2$ are rare. *Arch. Math.* (Basel) **84** (2005), 485–495. MR2148488 (2006b:11114).

[10] Wiertelak, Kazimierz. On the density of some sets of primes p for which $n \mid \mathrm{ord}_p a$. *Funct. Approx. Comment. Math.* **28** (2000), 237–241. MR1824009 (2003a:11120), Zbl 1009.11056.

Laboratoire de Mathématiques Nicolas Oresme, Université de Caen, BP 5186, 14032 Caen Cedex, France
Christian.Ballot@math.unicaen.fr

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

This paper is available via http://nyjm.albany.edu/j/2006/12-3.html.