Prime divisors of Lucas sequences and a conjecture of Skałba

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

Pantelimon Stănică
Auburn University Montgomery
Department of Mathematics
Montgomery, AL 36124-4023, USA
pstanica@mail.aum.edu

October 3, 2005

Abstract

In this paper, we give some heuristics suggesting that if $(u_n)_{n\geq 0}$ is the Lucas sequence given by $u_n = (a^n - 1)/(a - 1)$, where $a > 1$ is an integer, then $\omega(u_n) \geq (1 + o(1)) \log n \log\log n$ holds for almost all positive integers n.

1 Introduction

If n is a positive integer, we write $\omega(n)$ and $\Omega(n)$ for the number of distinct prime factors of n and total prime factors of n; i.e, including multiplicities,

[1]2000 *Mathematics Subject Classification*: 11B39,11N37,11N60
[2]*Keywords*: Lucas sequences, prime numbers, divisibility.

in the latter case. In what follows, for a real number $x > 1$ we write $\log x$ for the natural logarithm of x.

Hasse proved in [5] that the set of primes p dividing $2^n + 1$ for some positive integer n has relative density 17/24. Inspired by Hasse's result, Skałba proved the following results in [7].

Theorem 1. *Let p be a prime. If $\mathrm{ord}_p(2) \geq p^{0.8}$, then p divides a number of the form $2^a + 2^b + 1$ for some positive integers a and b. Here, $\mathrm{ord}_p(2)$ stands for the multiplicative order of 2 modulo p.*

Theorem 2. *If $\Omega(2^m - 1) < \log m / \log 3$, then there exists a prime divisor q of $2^m - 1$ such that q is not a divisor of $2^a + 2^b + 1$ for any positive integers a and b.*

We point out that the inequality in Theorem 2 in Skałba's paper [7] should be strict, since otherwise the assertion is not true. Moreover, he proposed two conjectures:

Conjecture 1.1. (*i*) *The number of primes $p \leq x$ that are divisors of some number of the form $2^a + 2^b + 1$ is $(1 + o(1))x/\log x$ as $x \to \infty$.*

(*ii*) *There are infinitely many primes q such that q does not divide any number of the form $2^a + 2^b + 1$.*

Regarding (*i*) above, we point out that a result of Pappalardi (see Theorem 2.3 in [6]) implies that $\mathrm{ord}_p(2) > p^{0.8}$ holds for almost all primes p under the Generalized Riemann Hypothesis, which, via Theorem 1, supports (*i*) above.

We cannot comment on (*ii*) above, but in this paper we look at the condition $\Omega(2^m - 1) < \log m / \log 3$, which, via Theorem 2, would support (*ii*) above.

In [2], Bugeaud et al., proved that for the Fibonacci sequence $(F_n)_{n \geq 0}$ given by $F_0 = 0$, $F_1 = 1$ and $F_{n+2} = F_{n+1} + F_n$ for all $n \geq 0$, if $\omega(F_n) \leq 2$, then either $n = 1, 2, 4, 8, 12$, or $n = \ell, 2\ell, \ell^2$ for some odd prime ℓ. Clearly, these are only necessary conditions for F_n to have at most two distinct prime factors but not sufficient. They also showed that the inequality $\omega(F_n) \geq (\log n)^{\log 2 + o(1)}$ holds for almost all positive integers n, and offered an heuristic to support that the inequality $\omega(F_n) \gg \log n$ holds for all composite positive integers n. Here and in what follows, we use the Vinogradov symbols $\gg$, $\ll$ and $\asymp$ and the Landau symbols O and o with their usual meanings.

We recall that $A \ll B$, $B \gg A$ and $A = O(B)$ are all equivalent and mean that $|A| < cB$ holds with some constant c, while $A \asymp B$ means that both $A \ll B$ and $B \ll A$ hold. The constants implied by such symbols may depend on our data a, ε, etc. Throughout, a property holds for "almost all" natural numbers if it holds for a set of asymptotic density 1.

Acknowledgements. The authors are grateful to the referee whose suggestions have led to an improvement in the presentation of this paper. During the preparation of this paper, F. L. was supported in part by projects PAPIIT 104505, SEP-CONACyT 46755 and a Guggenheim Fellowship.

2 The Results

Let $a > 1$ be an integer. Put $u_n = \dfrac{a^n - 1}{a - 1}$ for $n = 0, 1, \ldots$. In this paper, we offer the following conjecture, which complements the heuristics made in [2] and, if true, suggests that the positive integers m which fulfill the hypothesis of Theorem 2 are not typical ones.

Conjecture 2.1. *The inequality*

$$\omega(u_n) \geq (1 + o(1)) \log n \log\log n$$

holds for almost all positive integers n.

The same conjecture can be made for the sequence $(u_n)_{n\geq 0}$ replaced by any nondegenerate Lucas sequence.

In what follows, we offer an heuristic in support of the above conjecture.

Let $\varepsilon > 0$ be fixed. Let n be a positive integer from a set of asymptotic density one. Let $p_1 > p_2 > \ldots > p_t$ be all the prime factors of n in the interval

$$\mathcal{I}_n = \left[\log n, \exp\left(\frac{\log n}{\log\log n}\right)\right]. \tag{1}$$

We shall assume that n fulfills various conditions such as:

(*i*) If $p > \log n$ is a prime factor of n, then $p \parallel n$.

(*ii*) There do not exist primes $q > p > \log n$ dividing n such that $q \equiv 1 \pmod{p}$.

In particular, $p_i \mid\mid n$ ($p_i \mid n$ and $p_i^2 \nmid n$) for all $i = 1,\ 2,\ \ldots, t$, and there do not exist $i < j$ such that $p_i \equiv 1 \pmod{p_j}$. For a positive integer m we write $P(m)$ for the largest prime factor of m. Let $d(n)$ be the largest divisor of n such that $P(d(n)) < \log n$, and put $\overline{n} = n/d(n)$.

Define $m_0 = \overline{n}$ and $m_i = \overline{n}/(p_1 \ldots p_i)$ for $i = 1, \ldots, t$.

Consider the following finite sequence:

$$v_i \quad = \quad u_{m_{i-1}}/u_{m_i}, \qquad i = 1, 2, \ldots, t. \tag{2}$$

We observe that $v_i = (a_i^{p_i} - 1)/(a_i - 1)$, where $a_i = a^{\overline{n}/p_1 \cdots p_i}$. We also observe that v_i and v_j are coprime if $i \neq j$. Indeed, assume that $i < j$ and that there exists a prime q dividing v_i and v_j. Since $j > i$, we have that $v_j \mid u_{m_i}$, therefore

$$q \Big| \left(\frac{u_{m_{i-1}}}{u_{m_i}},\ u_{m_i} \right).$$

It is well-known that the above greatest common divisor divides m_i. Thus, $q \mid m_i \mid \overline{n}$. However, since $q \mid u_{m_{i-1}}$, it follows that there exists a unique minimal divisor d of n such that $q \mid u_d$. If $d = 1$, we then get $q \mid (a - 1)$, which is impossible if $\log n > a - 1$, because $q \mid \overline{n}$ and $\overline{n}$ is free of primes $\leq \log n$ (the case $\log n \leq a - 1$ need not be treated as there are only finitely many positive integers n satisfying that inequality). Thus, $d > 1$ is a divisor of m_i, and q is a primitive prime factor of u_d. It is then well-known that $q \equiv 1 \pmod{d}$ (see [3]). Since $d > 1$, there exists a prime factor p of d. Clearly, $p \mid \overline{n}$. Hence, $p \mid (q - 1)$, contradicting (ii).

It is known that the probability that a typical positive integer m is prime is $1/\log m$, and that a typical positive integer m has k distinct prime factors is

$$\frac{(\log \log m)^{k-1}}{(k-1)! \log m}.$$

We now make the following heuristic:

Heuristic 2.2. *With the above notations, we suppose that $\omega(v_i) = k_i$ happens with the probability*

$$\frac{(\log \log v_i)^{k_i - 1}}{(k_i - 1)! \log v_i}, \tag{3}$$

that this is uniform in the k_i's, and that these probabilities are independent for $i = 1, \ldots, t$ and uniformly in our range for t.

Under all these assumptions, the probability that u_n has at most K prime factors will be

$$\leq S(K) = \sum_{n\geq 1} \sum_{k_1+\ldots+k_t\leq K} \prod_{i=1}^{t} \frac{(\log\log v_i)^{k_i-1}}{(k_i-1)!\log v_i}. \tag{4}$$

We will also assume that

(*iii*) $t > (1-\varepsilon/2)\log\log n$.

(*iv*) $d(n) < n^{\varepsilon/4}$;

(*v*) $\prod_{i=1}^{t} p_i < n^{\varepsilon/4}$;

Under these assumptions, we shall show that:

Theorem 2.3. *If $K < (1-\varepsilon)t\log n$, then the series $S(K)$ converges.*

Theorem 2.3 has the following corollary.

Corollary 2.4. *Heuristic* 2.2 *implies that the inequality*

$$\omega(u_n) \geq (1-2\varepsilon)\log n \log\log n$$

holds for all n satisfying (*i*)–(*v*).

To end, we prove the following proposition.

Proposition 2.5. *Let $\varepsilon > 0$ be fixed. Then the set of positive integers n satisfying* (*i*)–(*v*) *has asymptotic density one.*

Clearly, letting ε to tend to zero, we get that Corollary 2.4 and Proposition 2.5 lead to the conclusion that Heuristic 2.2 implies Conjecture 2.1.

3 Proofs

Proof of Theorem 2.3. Fix $k_1, \ldots, k_t$, with $\sum_{i=1}^{t} k_i = k$. Obviously, $k_i \geq 1$. Since $a_i^{p_i-1} \leq v_i \leq a_i^{p_i}$, one sees immediately that

$$\log v_i \asymp p_i \log a_i \asymp m_{i-1} \asymp \frac{\overline{n}}{p_1\cdots p_{i-1}}. \tag{5}$$

Thus,

$$\prod_{i=1}^{t}\frac{(\log\log v_i)^{k_i-1}}{(k_i-1)!\log v_i} \ll \prod_{i=1}^{t}\frac{1}{(k_i-1)!}\frac{(\log(\overline{n}/(p_1\cdots p_{i-1})))^{k_i-1}}{\overline{n}/(p_1\cdots p_{i-1})}$$
$$= \frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\prod_{i=1}^{t}\frac{(\log m_{i-1})^{k_i-1}}{(k_i-1)!}.$$

Using the above inequality, we obtain

$$\sum_{k_1+\cdots+k_t=k}\prod_{i=1}^{t}\frac{(\log\log v_i)^{k_i-1}}{(k_i-1)!\log v_i} \ll \sum_{k_1+\cdots+k_t=k}\frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\prod_{i=1}^{t}\frac{(\log m_{i-1})^{k_i-1}}{(k_i-1)!}$$
$$\leq \frac{1}{(k-t)!}\frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\sum_{k_1+\cdots+k_t=k}\binom{k-t}{k_1-1,\ldots,k_t-1}\prod_{i=1}^{t}(\log m_{i-1})^{k_i-1}$$
$$= \frac{1}{(k-t)!}\frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\left(\sum_{i=1}^{t}\log m_{i-1}\right)^{k-t}$$
$$= \frac{1}{(k-t)!}\frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\left(\log\left(\prod_{i=1}^{t}m_{i-1}\right)\right)^{k-t}$$
$$= \frac{1}{(k-t)!}\frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\left(\log\left(\frac{\overline{n}^t}{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}\right)\right)^{k-t}.$$

Moreover, from Stirling's formula and the fact that $k_i \geq 1$, we obtain

$$s(n,K) = \sum_{k_1+\ldots+k_t\leq K}\prod_{i=1}^{t}\frac{(\log\log v_i)^{k_i-1}}{(k_i-1)!\log v_i}$$
$$\ll \sum_{k\leq K}\frac{1}{(k-t)!}\frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\left(\log\left(\frac{\overline{n}^t}{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}\right)\right)^{k-t} \quad (6)$$
$$\leq \frac{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}{\overline{n}^t}\sum_{0\leq j\leq K-t}\left(\frac{e\log\left(\frac{\overline{n}^t}{\prod_{i=1}^{t}(p_1\cdots p_{i-1})}\right)}{j}\right)^{j}.$$

As in [2], if y is fixed, then the function $x \longmapsto (ey/x)^x$ is increasing for $x < y$. Thus, if we assume that

$$K \le t + c_0 \log \left(\frac{\overline{n}^t}{\prod_{i=1}^t (p_1 \cdots p_{i-1})} \right), \tag{7}$$

where $c_0 = c_0(\varepsilon) < 1$ is a positive constant to be chosen later depending on ε, then estimate (6) leads to

$$s(n,K) \ll \frac{\prod_{i=1}^t (p_1 \cdots p_{i-1})}{\overline{n}^t} \log \left(\frac{\overline{n}^t}{\prod_{i=1}^t p_1 \cdots p_{i-1}} \right) \left(\frac{e}{c_0} \right)^{c_0 \log \left(\frac{\overline{n}^t}{\prod_{i=1}^t p_1 \cdots p_{i-1}} \right)}. \tag{8}$$

Furthermore, denoting

$$m(n) = \frac{\overline{n}^t}{\prod_{i=1}^t (p_1 \cdots p_{i-1})},$$

we get

$$S(K) = \sum_{n=1}^{\infty} s(n,k) \ll \sum_{n=1}^{\infty} \frac{\log m(n)}{m(n)} \left(\frac{e}{c_0} \right)^{c_0 \log m(n)} = \sum_{n=1}^{\infty} \frac{\log m(n)}{m(n)^{1 - c_0 \log(e/c_0)}}. \tag{9}$$

Now note that, by (iv) and (v),

$$m(n) \ge \left(\frac{\overline{n}}{p_1 \dots p_t} \right)^t \ge n^{(1-\varepsilon/2)t}.$$

It now follows easily that if

$$(1-\varepsilon/2)(1 - c_0 \log(e/c_0))t > 1, \tag{10}$$

then the series (9) converges, and by (iii) it is clear that for fixed ε and c_0, the above inequality (10) holds for all but finitely many n. Finally, to conclude, it remains to check that if K satisfies the inequality from the hypothesis of Theorem 2.3, it then satisfies inequality (7), as well. But clearly, the double inequality

$$t + c_0 \log(m(n)) > t(1-\varepsilon/2) c_0 \log n > t(1-\varepsilon) \log n$$

holds if we choose $c_0(\varepsilon)$ to be in the interval

$$\left(\frac{1-\varepsilon}{1-\varepsilon/2},\ 1\right).$$

□

Proof of Corollary 2.4. Theorem 2.3 together with (v) shows that if

$$K < (1-\varepsilon)t\log n < (1-\varepsilon)(1-\varepsilon/2)\log n\log\log n,$$

then the series $S(K)$ converges. Since

$$(1-\varepsilon)(1-\varepsilon/2) > 1-2\varepsilon,$$

it follows that the series $S(K)$ converges when $K < (1-2\varepsilon)\log n\log\log n$ as well. Heuristic 2.2 now completes the proof. □

Proof of Proposition 2.5. Let $\mathcal{A}$ be the set of positive integers satisfying (i)–(v). For a positive real number x, we let $\mathcal{A}(x) = \mathcal{A}\cap[1,x]$. It suffices to show that $\#\mathcal{A}(x) = (1+o(1))x$.

Let $\mathcal{B}_1(x) = \{n \le x : p^2 \mid n \text{ for some } p > \log x\}$. Let $n \in \mathcal{B}_1(x)$. There exists a prime $p > \log x$ such that $p^2 \mid n$. For fixed p, the number of such positive integers n is $\le x/p^2$. Hence,

$$\#\mathcal{B}_1(x) \le \sum_{p\ge\log x}\frac{x}{p^2} \ll \frac{x}{\log x} = o(x). \tag{11}$$

Let $\mathcal{B}_2(x) = \{n \le x : pq \mid n \text{ for some primes } q > p > \log x \text{ with } p \mid q-1\}$. Let $n \in \mathcal{B}_2(x)$. There exist primes $q > p > \log x$ such that $pq \mid n$ and $p \mid (q-1)$. For fixed p and q, the number of such positive integers n is $\le x/pq$. Hence,

$$\begin{aligned}
\#\mathcal{B}_2(x) &\le \sum_{p\ge\log x}\ \sum_{\substack{q<x\\ q\equiv 1 \pmod p}}\frac{x}{pq}\\
&\ll x\sum_{p>\log x}\frac{1}{p}\sum_{\substack{q<x\\ q\equiv 1 \pmod p}}\frac{1}{q}\\
&\ll x\log\log x\sum_{p>\log x}\frac{1}{p\phi(p)}\\
&\ll \frac{x\log\log x}{\log x} = o(x),
\end{aligned} \tag{12}$$

where in the above estimates we used the known fact that

$$\sum_{\substack{q<x\\ q\equiv 1 \pmod p}} \frac{1}{q} \ll \frac{\log\log x}{\phi(p)},$$

and that this estimate is uniform in $2 \le p \le x$ (see, for example, Lemma 1 in [1] or bound (3.1) in [4]).

Now put $\mathcal{B}_3(x) = \{n \le x/\log x\}$. Obviously,

$$\#\mathcal{B}_3(x) \le \frac{x}{\log x} = o(x). \tag{13}$$

From now on, we consider only those $n \le x$ not in $\cup_{i=1}^{3}\mathcal{B}_i(x)$. It is clear that such integers satisfy (i) and (ii). Put $y = \log x$. Let $f(s) = \exp(\log s/\log\log s)$ and put $z_1 = f(x/\log x)$ and $z_2 = f(x)$. The function $f(s)$ is increasing for $s > s_0 = e^e$. Thus, if $x > x_0$ is sufficiently large, then the inequalities

$$\log n \le y < z_1 < f(n) < z_2 \tag{14}$$

hold for all our n. Thus, by (14), we get

$$[y, z_1] \subset \mathcal{I}_n = [\log n, f(n)] \subset [1, z_2]. \tag{15}$$

For any $s > 1$ and positive integer m, we write $\omega_s(m)$ and $\Omega_s(m)$ for the number of distinct prime factors of m which are $\le s$, and the total number of prime factors of m which are $\le s$, respectively. By the well-known Turán-Kubilius estimates (see [9], for example), we have

$$\sum_{n\le x} |g_s(n) - \log\log s|^2 = O(x\log\log s), \tag{16}$$

where $g \in \{\Omega, \omega\}$. Further, the above estimates are uniform in $e^e < s \le x$.

We now put

$$\mathcal{B}_4(x) = \{n \le x : \Omega_y(n) \ge (\varepsilon/4)\log\log x\},$$

and

$$\mathcal{B}_5(x) = \{n \le x : \omega_{z_1}(n) \le (1-\varepsilon/4)\log\log x\}.$$

Using estimates (16) with $(g, s) = (\Omega, y)$ and (ω, z_1), together with the fact that $\log\log z_1 = (1 + o(1)) \log\log x$, we immediately get that

$$\#\mathcal{B}_4(x) \ll \frac{x \log\log\log x}{(\log\log x)^2} = o(x), \tag{17}$$

and

$$\#\mathcal{B}_5(x) \ll \frac{x}{\log\log x} = o(x). \tag{18}$$

Assume now that $n \notin \cup_{i=1}^{5} \mathcal{B}_i(x)$. Then,

$$t \geq \omega_{z_1}(n) - \Omega_y(n) > (1 - \varepsilon/2) \log\log x \geq (1 - \varepsilon/2) \log\log n,$$

so n satisfies (iii) (here, t is the number of distinct prime factors of n in $\mathcal{I}_n$, where $\mathcal{I}_n$ is given as in (1)). Furthermore, for large x we also have

$$d(n) \leq (\log n)^{\Omega_y(n)} \leq \exp\left((\varepsilon/4)(\log\log x)^2\right) < \left(\frac{x}{\log x}\right)^{\varepsilon} < n^{\varepsilon},$$

therefore n satisfies (iv) as well.

It remains to deal with condition (v). For this, we note that if n does not fulfill condition (v), then n has a divisor

$$\prod_{i=1}^{t} p_i \geq n^{\varepsilon/4} > \left(\frac{x}{\log x}\right)^{\varepsilon/4} > x^{\varepsilon/8}$$

whose largest prime factor is $\leq z_2$, by (1) and (15). Fix such a divisor d. Then the number of positive integers $n \leq x$ which are multiples of d is $\leq x/d$. Thus, writing $\mathcal{B}_6(x)$ for the set of $n \notin \cup_{i=1}^{5} \mathcal{B}_i(x)$ which do not fulfill (v), we get that

$$\#\mathcal{B}_6(x) \leq \sum_{\substack{x^{\varepsilon/8} < d < x \\ P(d) < z_2}} \frac{x}{d}. \tag{19}$$

Let

$$u = \log(x^{\varepsilon/8}) / \log z_2 = (\varepsilon/8) \log\log x.$$

It is known (see, for example, Chapter III of [8]), that

$$\sum_{\substack{x^{\varepsilon/8} < d < x \\ P(d) < z_2}} \frac{1}{d} \ll \rho(u) \log x, \tag{20}$$

where ρ is the Dickman function. Since $\rho(u) = u^{-(1+o(1))u}$ as $u \to \infty$, we get, by estimates (19) and (20), that

$$\begin{aligned} \#\mathcal{B}_6(x) &\ll x\rho(u)\log x \\ &= x \log x \exp\left(-(1+o(1))(\varepsilon/8)\log\log x \log\log\log x\right) \\ &= o(x). \end{aligned} \tag{21}$$

Thus, we conclude that the complement of $\cup_{i=1}^{6}\mathcal{B}_i(x)$ consists of positive integers $n \le x$ satisfying (i)–(v), and since by (11), (12), (13), (17), (18) and (21), we have that

$$\#\left(\cup_{i=1}^{6}\mathcal{B}_i(x)\right) \le \sum_{i=1}^{6} \#\mathcal{B}_i(x) = o(x),$$

the conclusion of the proposition follows. □

References

[1] N. L. Bassily, I. Kátai and M. Wijsmuller, *On the prime power divisors of the iterates of the Euler-ϕ function*, Publ. Math. Debrecen **55** (1999), 17–32.

[2] Y. Bugeaud, F. Luca, M. Mignotte and S. Siksek, *On Fibonacci numbers with few prime divisors*, Proc. Japan Acad. Sci. Ser. A **81** (2005), 17–20.

[3] R. D. Carmichael, *On the numerical factors of arithmetical forms $\alpha^n \pm \beta^n$*, Ann. Math. (2) **15** (1913), 30–70.

[4] P. Erdős, A. Granville, C. Pomerance and C. Spiro, *On the normal behavior of the iterates of some arithmetic functions*, *Analytic Number Theory*, Birkhäuser, Boston, 1990, 165–204.

[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. **168** (1966), 19-23.

[6] F. Pappalardi *On the Order of Finitely Generated Subgroups of $\mathbb{Q}^*$ (mod p) and Divisors of $p-1$*, J. Number Theory **57** (1996), 207–222.

[7] M. Skałba, *Two conjectures on primes dividing* $2^a + 2^b + 1$, Elem. Math. **59** (2004), 171-173.

[8] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge U. Press, 1995.

[9] P. Turán, *On a Theorem of Hardy and Ramanujan*, J. London Math. Soc. **9** (1934), 274–276.