THE DISTRIBUTION OF SELF-FIBONACCI DIVISORS

FLORIAN LUCA AND EMANUELE TRON

ABSTRACT. Consider the positive integers n such that n divides the n-th Fibonacci number, and their counting function A. We prove that

 $A(x) \le x^{1-(1/2+o(1))\log\log\log\log x/\log\log x}.$

1. INTRODUCTION

The Fibonacci numbers notoriously possess many arithmetical properties in relation to their indices. In this context, Fibonacci numbers divisibile by their index constitute a natural subject of study, yet there are relatively few substantial results concerning them in the literature.

Let $\mathcal{A} = \{a_n\}_{n \in \mathbb{N}}$ be the increasing sequence of natural numbers such that a_n divides F_{a_n} : this is OEIS A023172, and it starts

 $1, 5, 12, 24, 25, 36, 48, 60, 72, 96, 108, 120, 125, 144, 168, 180, \ldots$

(as they have no common name, we dub them *self-Fibonacci divisors*). Let moreover $A(x) := \#\{n \le x : n \in \mathcal{A}\}$ be its counting function.

This kind of sequences has already been considered by several authors; we limit ourselves to mentioning the current state-of-the-art result, due to Alba González–Luca–Pomerance–Shparlinski.

Proposition 1.1 ([1], Theorems 1.2 and 1.3).

$$\left(\frac{1}{4} + o(1)\right)\log x \le \log A(x) \le \log x - (1 + o(1))\sqrt{\log x \log \log x}.$$

We improve the upper bound above as follows.

Theorem 1.2.

$$\log A(x) \le \log x - \left(\frac{1}{2} + o(1)\right) \frac{\log x \log \log \log x}{\log \log x}.$$
(1.1)

The main element of the proof is a new classification of self-Fibonacci divisors. We now recall some basic facts about Fibonacci numbers. All statements in the next lemma are well-known and readily provable.

Lemma 1.3. Define z(n) to be the least positive integer such that n divides $F_{z(n)}$ (the Fibonacci entry point, or order of appearance, of n). Then the following properties hold.

- z(n) exists for all $n \in \mathbb{N}$. In fact, $z(n) \leq 2n$.
- $gcd(F_a, F_b) = F_{gcd(a,b)}$ for $a, b \in \mathbb{N}$.

²⁰¹⁰ Mathematics Subject Classification. 11B39.

- z(p) divides $p \left(\frac{p}{5}\right)$ for p prime, $\left(\frac{p}{5}\right)$ being the Legendre symbol.
- If a divides b, then z(a) divides z(b).
- $z(\operatorname{lcm}(a,b)) = \operatorname{lcm}(z(a), z(b))$ for $a, b \in \mathbb{N}$. In particular, $\operatorname{lcm}(z(a), z(b))$ divides z(ab).
- $z(p^n) = p^{\max(n-e(p),0)} z(p)$ for p prime, where $e(p) := v_p(F_{z(p)}) \ge 1$ and v_p is the usual p-adic valuation.

From now on, we shall use the above properties without citing them.

Next comes a useful result concerning the p-adic valuation of Fibonacci numbers.

Lemma 1.4 ([4], Theorem 1).

$$v_2(F_n) = \begin{cases} 0, & \text{if } n \equiv 1, 2 \pmod{3}; \\ 1, & \text{if } n \equiv 3 \pmod{6}; \\ 3, & \text{if } n \equiv 6 \pmod{12}; \\ v_2(n) + 2, & \text{if } n \equiv 0 \pmod{12}. \end{cases}$$

$$v_5(F_n) = v_5(n).$$

For $p \neq 2, 5$ prime,

$$v_p(F_n) = \begin{cases} v_p(n) + e(p), & \text{if } n \equiv 0 \pmod{z(p)}; \\ 0, & \text{if } n \not\equiv 0 \pmod{z(p)}. \end{cases}$$

To end the section, we point out an interesting feature of the upper bound in Theorem 1.2: it should be, up to a constant factor, best possible.

A squarefree integer n is a self-Fibonacci divisor if and only if z(p) divides n for every prime p that divides n. This is certainly true if $p - \left(\frac{p}{5}\right)$ divides n for every prime factor p of n. This is indeed strongly reminiscent of Korselt's criterion for Carmichael numbers: one should therefore expect heuristics for self-Fibonacci divisors similar to those for Carmichael numbers to be valid; in particular Pomerance's [7], which would predict

$$\log A(x) = \log x - (1 + o(1)) \frac{\log x \log \log \log x}{\log \log x}.$$

2. ARITHMETICAL CHARACTERISATION

In this section, we show how \mathcal{A} can be partitioned into subsequences that admit a simple description.

Note that n divides F_n if and only if z(n) divides n and set

$$\mathcal{A}_k := \{ n \in \mathbb{N} : n/z(n) = k \}.$$

Our next task is to prove the following characterisation of the \mathcal{A}_k 's. Let $c(k) := \min \mathcal{A}_k$ whenever \mathcal{A}_k is not empty.

Theorem 2.1. $\mathcal{A}_k = \emptyset$ if k is divisible by 8, 5 or $p^{e(p)+1}$ for an odd prime p. Otherwise, if $k = \prod_i p_i^{\alpha_i}$,

 $\mathbf{2}$

THE DISTRIBUTION OF SELF-FIBONACCI DIVISORS

Proof. We shall henceforth implicitly assume that the primes we deal with are distinct from 2 and 5, and all the proofs when some prime is 2 or 5 are easily adapted using the modified statement of Lemma 1.4.

Suppose that, for some n, n/z(n) = k, and p^d is the exact power of p that divides k. Upon writing $k = p^d k'$ and $n = p^d n'$, with k' coprime to p, this becomes $n'/k' = z(p^d n')$. In particular,

$$d + v_p(n') \le v_p(F_{z(p^d n')}) = v_p(F_{n'/k'}) \le v_p(n') + e(p),$$

which is absurd if $d \ge e(p) + 1$, so that $\mathcal{A}_k = \emptyset$ if $p^{e(p)+1}$ divides k.

Suppose on the other hand that k fulfills the conditions for \mathcal{A}_k to be nonempty. We want to know for which $m \in \mathbb{N}$, given $n \in \mathcal{A}_k$, mn is itself in \mathcal{A}_k : this will give the conclusion, once we know that all the numbers in the sequence are multiples of a smallest number c(k) which belongs itself to \mathcal{A}_k . The proof of this latter fact is deferred to Theorem 2.2 since it fits better within that setting.

Suppose we have $n \in \mathcal{A}_k$, and take $m = p_1^{a_1} \cdots p_w^{a_w}$ with $a_i > 0$ for each *i*; set $n = p_1^{\lambda_1} \cdots p_w^{\lambda_w} n'$ with n' coprime to *m* and $\lambda_i \ge 0$ for each *i*. Then one has

$$k = \frac{n}{z(n)} = \frac{p_1^{\lambda_1} \cdots p_w^{\lambda_w} n'}{z(p_1^{\lambda_1} \cdots p_w^{\lambda_w} n')}$$
$$= \frac{p_1^{\lambda_1} \cdots p_w^{\lambda_w} n'}{\operatorname{lcm}(p_1^{\max(\lambda_1 - e(p_1), 0)} z(p_1), \dots, p_w^{\max(\lambda_w - e(p_w), 0)} z(p_w), z(n'))}.$$

The p_i -adic valuation of this expression is $v_{p_i}(k)$, so in the denominator either $\max(\lambda_i - e(p_i), 0)$ is the greatest power of p_i , or some of $z(p_1), \ldots, z(p_w), z(n')$ has p-adic valuation $\lambda_i - v_{p_i}(k) \ge \lambda_i - e(p_i)$. Furthermore, one has $\lambda_i \ge v_{p_i}(k)$, as n has to be a multiple of k.

Now, the number

$$\frac{mn}{z(mn)} = \frac{p_1^{\lambda_1 + a_1} \cdots p_w^{\lambda_w + a_w} n'}{z(p_1^{\lambda_1 + a_1} \cdots p_w^{\lambda_w + a_w} n')}$$
$$= \frac{p_1^{\lambda_1 + a_1} \cdots p_w^{\lambda_w + a_w} n'}{\operatorname{lcm}(p_1^{\lambda_1 + a_1 - e(p_1)} z(p_1), \dots, p_w^{\lambda_w + a_w - e(p_w)} z(p_w), z(n'))}$$

is equal to k if and only if its p_i -adic valuation is $v_{p_i}(k)$ for each i, that is

$$v_{p_i}(k) = \lambda_i + a_i - \max(\lambda_i + a_i - e(p_i), \lambda_i - v_{p_i}(k)).$$

Suppose that the first term in the max is the greater, that is $a_i \ge e(p_i) - v_{p_i}(k)$. The above equality reduces to $v_{p_i}(k) = e(p_i)$: so in this case each value $\ge e(p_i) - v_{p_i}(k)$ for a_i , and each nonnegative value for λ_i is admissible, if $v_{p_i}(k) = e(p_i)$, and no value is admissible if $0 \le v_{p_i}(k) < e(p_i)$.

Suppose that the second term is the greater, that is $a_i < e(p_i) - v_{p_i}(k)$. The equality reduces to $a_i = 0$, which is impossible.

Starting from c(k) and building all the members of \mathcal{A}_k by progressively adding prime factors, we find exactly the statement of the theorem.

In the remainder of this section, we show that c(k) admits a more explicit description.

Theorem 2.2. $c(k) = k \operatorname{lcm} \{z^i(k)\}_{i=1}^{\infty}$.

Proof. To prove first that such an expression is well-defined, we show that the sequence of iterates of z eventually hits a fixed point.

First note that, for
$$k = \prod_{i} p_i^{\alpha_i}$$
, $z(k) = \operatorname{lcm} \left\{ p_i^{\max(\alpha_i - e(p_i), 0)} z(p_i) \right\}_i$: this is a

divisor of $\frac{k}{\operatorname{rad}(k)}$ lcm $\{z(p_i)\}_i$, where $\operatorname{rad}(k) = \prod_i p_i$ is the radical of k. Consider

now the largest prime factor P of k: if $P \ge {}^{i}7$, its exponent in the previous expression decreases by at least 1 at each step, since the largest prime factor of z(P) is strictly smaller than P. Consequently, after at most $v_P(k)$ steps, the exponent of P would have vanished. By iterating the argument concerning the largest prime factor at each step, after a finite number ℓ of steps, $z^{\ell}(k)$ will have only prime factors smaller than 7; set $z^{\ell}(k) = 2^{a}3^{b}5^{c}$.

Recall now Theorem 1.1 of [6]: the fixed points of z are exactly the numbers of the form 5^f and $12 \cdot 5^f$. By noting that $z(2^a) = 3 \cdot 2^{a-2}$, $z(3^b) = 4 \cdot 3^{b-1}$, $z(5^c) = 5^c$, we get that $z(2^a 3^b 5^c) = 2^{\max(a-2,2)} 3^{\max(b-1,1)} 5^c$. Since we can continue this until $a \leq 2$ and $b \leq 1$, we are left with a few cases to check to show that the sequence of iterates indeed reaches a fixed point.

As c(k) must be a multiple of k, call T := c(k)/k. Consider next the obvious equalities

$$T = z(kT),$$

$$z(T) = z^{2}(kT),$$

$$z^{2}(T) = z^{3}(kT),$$

$$\vdots$$

Write $x \stackrel{\text{div}}{\leftarrow} y$ for the statement "y divides x". Then we have that

$$T = z(kT)$$

$$\stackrel{\text{div}}{\leftarrow} z(\text{lcm}(k,T))$$

$$= \text{lcm}(z(k), z(T))$$

$$= \text{lcm}(z(k), z^2(kT))$$

4

$$\stackrel{\text{div}}{\leftarrow} \operatorname{lcm}(z(k), z(\operatorname{lcm}(z(k), z(T)))) \\ = \operatorname{lcm}(z(k), \operatorname{lcm}(z^2(k), z^2(T))) \\ = \operatorname{lcm}(z(k), z^2(k), z^3(kT)) \\ \vdots \\ = \operatorname{lcm}(z(k), z^2(k), z^3(k), \dots).$$

Note that we have not used yet that kT is the smallest member of \mathcal{A}_k ; this means the above reasoning works for any member of \mathcal{A}_k , so that any number in \mathcal{A}_k is a multiple of $k \operatorname{lcm}(z(k), z^2(k), z^3(k), \ldots)$. If we manage to prove that $k \operatorname{lcm}(z(k), z^2(k), z^3(k), \ldots)$ is indeed in the sequence, we will obtain the divisibility argument we needed in the proof of Theorem 2.1.

Thus, we want to prove that $T = \operatorname{lcm} \{z^i(k)\}_{i=1}^{\infty}$ works; it is enough to prove that the divisibilities we previously derived are equalities, or in other words that $z(kT) = z(\operatorname{lcm}(k,T))$ for T defined this way.

If
$$k = \prod_{i} p_i^{\alpha_i}$$
 with $\alpha_i \le e(p_i)$ for each i , then

$$T = \operatorname{lcm}\left(z\left(\prod_{i} p_i^{\alpha_i}\right), \ z^2\left(\prod_{i} p_i^{\alpha_i}\right), \ \ldots\right)$$

$$= \operatorname{lcm}\left(\operatorname{lcm}\left\{z\left(p_i^{\alpha_i}\right)\right\}_i, \ \operatorname{lcm}\left\{z^2\left(p_i^{\alpha_i}\right)\right\}_i, \ \ldots\right)$$

$$= \operatorname{lcm}\left(\left\{z\left(p_i^{\alpha_i}\right)\right\}_i, \ \left\{z^2\left(p_i^{\alpha_i}\right)\right\}_i, \ \ldots\right)$$

and

$$z(kT) = z\left(\left(\prod_{i} p_{i}^{\alpha_{i}}\right) \operatorname{lcm}\left(z\left(\prod_{i} p_{i}^{\alpha_{i}}\right), \ z^{2}\left(\prod_{i} p_{i}^{\alpha_{i}}\right), \ \ldots\right)\right).$$

We would like to bring the $\prod_{i} p_i^{\alpha_i}$ into the least common multiple, but some power of p_i could divide the iterated entry point of some other prime to a higher power. Define then $m(p_h)$ to be the largest exponent of a power of p_h that divides $z^i(p_j)$ as i and j vary; thus

$$\begin{split} & z\left(\left(\prod_{i} p_{i}^{\alpha_{i}}\right) \operatorname{lcm}\left(z\left(\prod_{i} p_{i}^{\alpha_{i}}\right), \ z^{2}\left(\prod_{i} p_{i}^{\alpha_{i}}\right), \ \ldots\right)\right) \\ &= & z\left(\operatorname{lcm}\left(\left\{p_{i}^{m(p_{i})+\alpha_{i}}\right\}_{i}, \ z\left(\prod_{i} p_{i}^{\alpha_{i}}\right), \ z^{2}\left(\prod_{i} p_{i}^{\alpha_{i}}\right), \ \ldots\right)\right) \\ &= & \operatorname{lcm}\left(\left\{z\left(p_{i}^{m(p_{i})+\alpha_{i}}\right)\right\}_{i}, \ \left\{z^{2}\left(p_{i}^{\alpha_{i}}\right)\right\}_{i}, \ \left\{z^{3}\left(p_{i}^{\alpha_{i}}\right)\right\}_{i}, \ \ldots\right). \end{split}$$

We need this to be equal to

lcm $(\{z(p_i^{\alpha_i})\}_i, \{z^2(p_i^{\alpha_i})\}_i, \{z^3(p_i^{\alpha_i})\}_i, \ldots) = \text{lcm}(z(k), z(T)) = z(\text{lcm}(k, T)).$ All that is left to do now is to remark that this is true if and only if their p_i -adic valuations are equal for each i, or in other words, as p_i is coprime to $z(p_i^{\alpha_i}) = z(p_i),$

$$\max(m(p_i) + \alpha_i - e(p_i)), m(p_i)) = m(p_i)$$

and this is evident.

3. The proof of Theorem 1.2

Let $x \ge 10$. One of our ingredients is the following result from [3].

Lemma 3.1 ([3], Theorem 3). As $x \to \infty$,

$$\# \{ n \le x : z(n) = m \} \le x^{1 - (1/2 + o(1)) \log \log \log x / \log \log x}$$

uniformly in m.

Let $n \in \mathcal{A}(x)$. By Theorem 2.1, every self-Fibonacci divisor is of the form c(k)m, where *m* is composed of primes that divide *k*. Thus, write n = c(k)m, where every prime factor of *m* divides *k*. Let $C(x) := x^{\log \log \log x/\log \log x}$. We distinguish two cases.

Case 1. $k \leq x/C(x)$.

Let $\mathcal{A}_1(x)$ be the subset of such $n \in \mathcal{A}(x)$. We fix k and count possible m's because c(k) is determined by k; we use an idea similar to the one of the proof of Theorem 4 in [2]. Clearly, m has at most $\omega(k)$ distinct prime factors. Define next $\Psi(x, y)$ to be the number of positive integers $\ell \leq x$ whose largest prime factor $P(\ell)$ satisfies the inequality $P(\ell) \leq y$, and let p_s be the s-th prime. If \mathcal{P}_k is the set of the prime divisors of k, the quantity of numbers $m \leq x$ all of whose prime factors are in \mathcal{P}_k is of course at most $\Psi(x, p_{\omega(k)}) \leq \Psi(x, 2\log x)$ for x large enough. Here we used the fact that $p_s < s(\log s + \log \log s)$ for all $s \geq 6$ (Theorem 3 of [8]) together with $\omega(k) < 2\log k/\log\log k$ for all $k \geq 3$. Classical estimates on $\Psi(x, y)$, such as the one of de Bruijn (see, for example, Theorem 2 on page 359 in [9]), show that if we put

$$Z := \frac{\log x}{\log y} \log \left(1 + \frac{y}{\log x} \right) + \frac{y}{\log y} \log \left(1 + \frac{\log x}{y} \right),$$

then the estimate

$$\log \Psi(x,y) = Z\left(1 + O\left(\frac{1}{\log y} + \frac{1}{\log\log(2x)}\right)\right)$$
(3.1)

holds uniformly in $x \ge y \ge 2$. The above estimates (3.1) with $y = 2 \log x$ imply that there are at most $C(x)^{(3\log 3 - 2\log 2 + o(1))/\log \log \log x} = C(x)^{o(1)}$ values of m for any fixed k. Summing up over k, we get that

$$#\mathcal{A}_1(x) \le C(x)^{o(1)} \sum_{k \le x/C(x)} 1 \le \frac{x}{C(x)^{1+o(1)}} \qquad (x \text{ large}).$$
(3.2)

Case 2. $x/C(x) < k \le x$.

Here, we have $kz(k) \leq c(k) \leq x$, whence $z(k) \leq C(x)$. Fix z(k) = z in [1, C(x)]. By Lemma 3.1, if we put $\mathcal{B}_z := \{n \in \mathbb{N} : z(n) = z\}$, then the inequality

$$#\mathcal{B}_z(t) \le t/C(t)^{1/2+o(1)} \quad \text{holds as} \quad t \to \infty.$$
(3.3)

6

We now let $k \in \mathcal{B}_z$. Then $n \leq x$ is a multiple of kz. The number of such n is $\lfloor x/kz \rfloor \leq x/kz$. Summing up the above inequality over $k \in \mathcal{B}_z$ and using partial summation and (3.3), we have

$$\begin{aligned} \frac{x}{z} \sum_{\substack{k \in \mathcal{B}_z \\ x/C(x) < k \le x}} \frac{1}{k} &= \frac{x}{z} \int_{x/C(x)}^x \frac{\mathrm{d}\#\mathcal{B}_z(t)}{t} \\ &= \frac{x}{z} \left(\frac{\#\mathcal{B}_z(t)}{t} \Big|_{t=x/C(x)}^{t=x} + \int_{x/C(x)}^x \frac{\#\mathcal{B}_z(t)}{t^2} \mathrm{d}t \right) \\ &\le \frac{x}{z} \left(\frac{\#\mathcal{B}_z(x)}{x} + \int_{x/C(x)}^x \frac{\mathrm{d}t}{tC(t)^{1/2+o(1)}} \right) \\ &= \frac{x}{z} \left(\frac{1}{C(x)^{1/2+o(1)}} + \frac{1}{C(x)^{1/2+o(1)}} \int_{x/C(x)}^x \frac{\mathrm{d}t}{t} \right) \\ &= \frac{(1+o(1))x \log C(x)}{zC(x)^{1/2+o(1)}} = \frac{x}{zC(x)^{1/2+o(1)}}, \end{aligned}$$

where in the above calculation we used the fact that

 $C(t)^{1/2+o(1)} = C(x)^{1/2+o(1)}$ uniformly in $t \in [x/C(x), x]$ as $x \to \infty$. We now sum over $z \in [1, C(x)]$, and obtain that

$$#\mathcal{A}_{2}(x) \leq \frac{x}{C(x)^{1/2+o(1)}} \sum_{1 \leq z \leq C(x)} \frac{1}{z}$$
$$= \frac{(1+o(1))x \log C(x)}{C(x)^{1/2+o(1)}} = \frac{x}{C(x)^{1/2+o(1)}} \qquad (x \to \infty). \quad (3.4)$$

The desired conclusion now follows from (3.2) and (3.4).

4. Comments

Of course, the methods we presented apply equally well to other Lucas sequences, where analogues of Theorems 2.1, 2.2 and 1.2 hold; we chose to display the Fibonacci case, when the classification takes a particularly simple form.

To conclude, we make some observations to promote future progress. The problem of finding lower bounds for A(x) requires completely different ideas; one can prove that

$$\log A(x) = \log \# \{ n \le x : c(n) \le x, n \text{ squarefree} \} + O\left(\frac{\log x \log \log \log x}{\log \log x}\right),$$

so that in order to prove $A(x) = x^{1+O(\log \log \log x/\log \log x)}$ unconditionally one would need to build many squarefree n with small c(n). The best we managed to prove is that $\log c(n) < 3P(n)$ (by double counting), and $\log c(n) < 7 \sum_{p|n} (\log p)^2$ (by

induction), but neither of these is sufficient. This hints at building numbers n for which their prime factors share most of their Pratt-Fibonacci trees (Pratt

trees built with the factors of z(p) as children of a node p^{δ} , taken with their exponents).

The set of numbers n with small c(n) is both small and large in a certain sense: it has asymptotic density 0 and exponential density 1, conjecturally.

It is indeed likely that c(n) is quite large for most n. Recall that putting

$$F(n) := \operatorname{rad}\left(\prod_{k \ge 1} \phi^k(n)\right),$$

then in [5] it is proved that the inequality

 $F(n) > n^{(1+o(1))\log\log n/\log\log\log n}$

holds for n tending to infinity through a set of asymptotic density 1. Since c(n) is quite similar to F(n), we conjecture that a similar result holds for c(n) as well.

Acknowledgements

We thank K. Ford, P. Leonetti, D. Marques, C. Pomerance, G. Tenenbaum for their help.

References

- J. J. Alba González, F. Luca, C. Pomerance, I. E. Shparlinski, On numbers n dividing the nth term of a linear recurrence, Proc. Edinburgh Math. Soc. 55.2 (2012), 271–289.
- [2] P. Erdős, F. Luca, C. Pomerance, On the proportion of numbers coprime to a given integer, Proceedings of the Anatomy of Integers Conference, Montréal, March 2006 (J.-M. De Koninck, A. Granville, F. Luca eds.), CRM Proceedings and Lecture Notes 46 (2008), 47–64.
- [3] D. M. Gordon, C. Pomerance, The distribution of Lucas and elliptic pseudoprimes, Math. Comp. 57 (1991), 825–838.
- [4] J. H. Halton, On the divisibility properties of Fibonacci numbers, Fibonacci Quart. 4.3 (1966), 217–240.
- [5] F. Luca, C. Pomerance, Irreducible radical extensions and Euler-function chains, Combinatorial number theory, 351361, de Gruyter, Berlin, 2007.
- [6] D. Marques, Fixed points of the order of appearance in the Fibonacci sequence, Fibonacci Quart. 50.4 (2012), 346–351.
- [7] C. Pomerance, On the Distribution of Pseudoprimes, Math. Comp. 37 (1981), 587–593.
- [8] J. Barkley Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, Illinois J. Math. 6.1 (1962), 64–94.
- [9] G. Tenenbaum, Introduction to analytic and probabilistic number theory, Cambridge U. Press, 1995.

School of Mathematics, University of the Witwatersrand, PO Box 2050, Wits, South Africa

E-mail address: florian.luca@wits.ac.za

SCUOLA NORMALE SUPERIORE, PIAZZA DEI CAVALIERI 7, 56126 PISA, ITALY *E-mail address*: emanuele.tron@sns.it