Doctoral Dissertation for the

PhD in Pure and Applied Mathematics

Università degli Studi di Torino and Politecnico di Torino

Arithmetic Properties of Linear Recurrences and Other Topics in Number Theory

Author:

Carlo Sanna

Supervisor:

Prof. Danilo Bazzanella

Torino

2018

Abstract

This thesis is a recollection of several contributions to Number Theory. It is divided into two parts, which are essentially independent from each other.

In the first part, we prove a number of results concerning the arithmetic properties of terms of linear recurrences. Given two linear recurrences F and G over a number field $\mathbb{K}$ satisfying some mild hypotheses, we give an upper bound for the counting function of the set $\mathcal{N}$ of positive integers n such that the ratio $F(n)/G(n)$ belongs to a finitely generated subring of $\mathbb{K}$. This makes quantitative a result of Corvaja and Zannier, which states that $\mathcal{N}$ has zero natural density. We investigate also the set of positive integers n which are relatively prime with the nth term of a linear recurrence over the integers, and the set of positive integers n such that the G.C.D. of n and the nth Fibonacci number is equal to a prescribed integer.

The second part of the thesis, on the other hand, consists of two chapters dealing with unrelated topics. In the first chapter, we prove that any sequence $f(n)_{n\geq 0}$, where $f \in \mathbb{Z}[X]$ is a quadratic or cubic polynomial, satisfies a coprimality condition known as *Pillai property*. This extends a result of Evans, who considered the case of f being a linear polynomial, and settles a conjecture of Harrington and Jones. In the second chapter, we study the set of positive integers n which are relatively prime with the nth central binomial coefficient $\binom{2n}{n}$, and we improve a result of Pomerance regarding the upper density of such set.

Acknowledgements

I express my sincere gratitude to my advisor Prof. Danilo Bazzanella, for supporting me during my last two years of PhD studies and related research. I am also grateful to Prof. Umberto Cerruti, who has been my tutor at the fist year of PhD studies, for pointing me out to some really interesting topics of research; and to Prof. Umberto Zannier, for a fruitful conversation on the results of Chapter 1. I thank the referees of my thesis, Prof. Giuseppe Molteni and Prof. Alberto Perelli, for their careful reading and helpful comments. My sincere thanks also goes to my coauthors Paolo Leonetti, Márton Szikszai, and Emanuele Tron, for their precious collaboration; and to the anonymous referees, for their invaluable work.

I thank all my friends, especially those of Collegio Einaudi and of its orbiting satellites of Casa Funk and Pecetto, for having always been a major source of support.

Last but not the least, I thank my family: my parents, my sister, my grandmothers, and my aunts and uncles, for supporting me during the last 8 years of higher education.

Declaration

I hereby declare that, except where indicated by specific reference in the text, the contents and organization of this thesis constitute my own original work. Work done in collaboration with others is indicated as such. The views expressed in the dissertation are those of the author.

Carlo Sanna

Contents

Introduction

The present thesis is a recollection of several contributions to Number Theory.

The thesis is subdivided into two parts, namely Part I and Part II, which are essentially independent from each other and consist, respectively, of four and two chapters.

In the first part, we prove a number of results concerning the arithmetic properties of terms of linear recurrences. Before presenting the content of Part I, we recall some basic facts about linear recurrences, which we give for granted in the sequel.

A *linear recurrence* is a sequence of complex numbers $(u_n)_{n\geq 0}$ such that

$$u_n = a_1 u_{n-1} + a_2 u_{n-2} + \cdots + a_k u_{n-k},$$

for all integers $n \geq k$, where $a_1, \ldots, a_k$ are fixed complex numbers with $a_k \neq 0$. In turn, this is equivalent to a (unique) expression as *generalized power sum*

$$u_n = \sum_{i=1}^{r} f_i(n)\, \alpha_i^n,$$

for all integers $n \geq 0$, where $f_1, \ldots, f_r \in \mathbb{C}[X]$ are polynomials, which are nonzero unless $(u_n)_{n\geq 0}$ is identically zero, and $\alpha_1, \ldots, \alpha_r \in \mathbb{C}^*$ are all the distinct roots of the *characteristic polynomial*

$$X^k - a_1 X^{k-1} - a_2 X^{k-2} - \cdots - a_k.$$

Classically, $\alpha_1, \ldots, \alpha_r$ and k are called the *roots* and the *order* of the linear recurrence, respectively. Moreover, the linear recurrence is said to be *nondegenerate* if none of the ratios α_i/α_j ($i \neq j$) is a root of unity, and it is said to be *simple* if all the polynomials $f_1, \ldots, f_r$ are constant. This latter terminology comes from the fact that $1 + \deg f_i$ is

equal to the multiplicity of α_i as a root of the characteristic polynomial, so that a linear recurrence is simple if and only if all the roots of its characteristic polynomial are simple. When $k = 2$, $u_0 = 0$, $u_1 = 1$, and a_1, a_2 are relatively prime integers, $(u_n)_{n\geq 0}$ is said to be a *Lucas sequence*. The most famous example of Lucas sequence is undoubtedly the sequence of *Fibonacci numbers*, here denoted by $(F_n)_{n\geq 0}$, for which $a_1 = a_2 = 1$. Of course, replacing the field of complex numbers $\mathbb{C}$ by any ring $\mathcal{R}$, the definition of linear recurrence still makes sense (although the theory of generalized power sums may not), and in such a case we speak of *linear recurrence over* $\mathcal{R}$. We refer the reader to [19] for the general theory of linear recurrences that we have just mentioned.

Part I consists of four chapters. In the first chapter, based on a paper by the author [46], we study the set

$$\mathcal{N} := \big\{n \geq 0 : G(n) \neq 0,\ F(n)/G(n) \in \mathfrak{R}\big\},$$

where $F(n)_{n\geq 0}$ and $G(n)_{n\geq 0}$ are two linear recurrences (satisfying some mild hypothesis) over a number field $\mathbb{K}$, and $\mathfrak{R}$ is a finitely generated subring of $\mathbb{K}$. The classical result known as *Hadamard-quotient Theorem*, conjectured by Pisot and proved by van der Poorten, says that if $\mathcal{N}$ contains all the positive integers but finitely many, then $n \mapsto F(n)/G(n)$ is itself a linear recurrence [44, 61]. Corvaja and Zannier [12] extended the Hadamard-quotient Theorem by proving that if $\mathcal{N}$ is infinite then there exists a nonzero polynomial $P \in \mathbb{C}[X]$ such that both $n \mapsto P(n)F(n)/G(n)$ and $n \mapsto G(n)/P(n)$ are linear recurrences. Also, they showed that if F/G is not a linear recurrence, then $\mathcal{N}$ has zero natural density, and they suggested that their proof could be adapted to show that

$$\#\mathcal{N}(x) \ll_{\mathbb{K}} \frac{x}{(\log x)^{\delta}},$$

for any $\delta < 1$ and all sufficiently large $x > 1$. Our contribution is the proof that the following more precise upper bound holds: If F/G is not a linear recurrence, then

$$\#\mathcal{N}(x) \ll_{F,G} x \cdot \left(\frac{\log\log x}{\log x}\right)^{h}, \tag{1}$$

for all $x \geq 3$, where h is a positive integer effectively computable in terms of F and G. Furthermore, assuming the Hardy–Littlewood h-tuple conjecture, we show how it is

possible to construct linear recurrences F and G such that F/G is not a linear recurrence and

$$\#\mathcal{N}(x) \gg_{F,G} \frac{x}{(\log x)^h},$$

for all $x \geq 3$. This seems to suggest that the upper bound (1) should be optimal, except for the term $\log\log x$. The proof of (1) employs a quantitative version of Chebotarev density theorem, the large sieve inequality, and bounds for the number of zeros of sparse polynomials in finite fields.

The special case in which F is a linear recurrence over the integers and $G(n) = n$, has been studied by many authors. For such F, G the computation of the exponent in bound (1) gives $h = 1$. However, better bounds are known. Precisely, define the set

$$\mathcal{D}_u := \{n \in \mathbb{N} : n \mid u_n\},$$

where $(u_n)_{n\geq 1}$ is a linear recurrence over the integers, which to avoid trivialities we assume to be not identically zero. Alba González, Luca, Pomerance, and Shparlinski [1] proved that if $(u_n)_{n\geq 1}$ is nondegenerate, simple, and of order $k \geq 2$, then

$$\#\mathcal{D}_u(x) \ll_k \frac{x}{\log x},$$

for all sufficiently large $x > 1$. Furthermore, André-Jeannin [4] and Somer [56] studied the arithmetic properties of the elements of $\mathcal{D}_u$ when $(u_n)_{n\geq 0}$ is a Lucas sequence. In such a case, generalizing a previous result of Luca and Tron [36] on Fibonacci numbers, we have proved in [47] that

$$\#\mathcal{D}_u(x) \leq x \Big/ \exp\left(\left(\frac{1}{2} + o(1)\right) \frac{\log x \log\log\log x}{\log\log x}\right),$$

as $x \to \infty$, where the $o(1)$ depends on a_1 and a_2.

In the second chapter, which is based on [48], we study the "dual" set

$$\mathcal{C}_u := \{n \in \mathbb{N} : \gcd(n, u_n) = 1\}.$$

Precisely, we prove that for any nondegenerate linear recurrence over the integers $(u_n)_{n\geq 0}$, the set $\mathcal{C}_u$ has a natural density, and that this density is positive if and only if $(u_n/n)_{n\geq 1}$ is not a linear recurrence.

In the third chapter, we focus on the sequence of Fibonacci numbers. For each positive integer k, we define the set

$$\mathcal{A}_k := \{n \in \mathbb{N} : \gcd(n, F_n) = k\}.$$

Our main result is that $\mathcal{A}_k$ has a natural density which is given by

$$\mathbf{d}(\mathcal{A}_k) = \sum_{d=1}^{\infty} \frac{\mu(d)}{\operatorname{lcm}(d, z(d))}, \tag{2}$$

where the series is absolutely convergent, and $z(d)$ denotes the rank of appearance of d, that is, the smallest positive integer n such that $d \mid F_n$. We also give an effective criterion to establish when the natural density of $\mathcal{A}_k$ is zero and we show that this is the case if and only if $\mathcal{A}_k$ is empty. This chapter relies on the paper [51], a joint work with Emanuele Tron. The methods of proof involve a result on the natural density of a set of multiples, and some combinatorial arguments founded on the inclusion-exclusion principle.

Finally, in Chapter 4, we study the set of positive integers k such that $\mathcal{A}_k \neq \varnothing$ or, equivalently, the set of positive integers of the form $\gcd(n, F_n)$, for some positive integer n. Setting

$$\mathcal{F} := \{k \in \mathbb{N} : \mathcal{A}_k \neq \varnothing\} = \{\gcd(n, F_n) : n \in \mathbb{N}\},$$

we prove that, on the one hand,

$$\#\mathcal{F}(x) \gg \frac{x}{\log x}$$

for all $x \geq 2$, while, on the other hand, $\mathcal{F}$ has zero natural density. The chapter is based on the paper [35], a joint work with Paolo Leonetti. The proof makes use of the Brun–Titchmarsh theorem and a recent result of Cubre and Rouse [13] regarding the relative density of the set of prime numbers p such that $z(p)$ is divisible by a fixed positive integer m.

The second part of the thesis, on the other hand, consists of two chapters dealing with unrelated topics.

In Chapter 5, we study a coprimality condition on consecutive values of integer polynomials. A sequence of integers $s(n)_{n\geq 0}$ is said to have the *Pillai property* if there exists an integer $G \geq 2$ such that for all integers $k \geq G$ there exist infinitely many integers $n \geq 0$ such that none of the integers $s(n+1), s(n+2), \ldots, s(n+k)$ is relatively prime with all the others. In such a case, G_s is defined as the smallest possible G. At a first glance, this property may sound strange, but it was considered by Pillai [39] in an attempt to prove that the product of consecutive integers is never a perfect power. Precisely, he conjectured that the sequence of natural numbers has the Pillai property, and he was able to prove that if G_s exists then $G_s \geq 17$. Then, using this latter information in a nontrivial way, he proved that the product of at most 16 consecutive integers cannot be a perfect power. Although later Erdős and Selfridge [16] proved that the product of consecutive integers is never a perfect power by using a different method, the study of G_s in various sequences attracted attention per se. Erdős [14] was the first to prove the existence of G_s when s is the sequence of natural numbers. Actually, he did that before the work of Pillai, while proving a lower bound for prime gaps. Later, the combined efforts of Pillai [39] and Brauer [6] gave the explicit result $G_s = 17$. Evans [17] proved that every arithmetic progression has the Pillai property, while Ohtomo and Tamari [38] showed that $G_s \leq 384$ when s is the sequence of odd integers. Then, Hajdu and Saradha [21] proved an effective upper bound on G_s depending on the difference of the arithmetic progression, and they also gave a heuristic algorithm to find the exact value of G_s.

We prove that for each quadratic or cubic polynomial $f \in \mathbb{Z}[X]$ the sequence $f(n)_{n\geq 0}$ has the Pillai property. This extends Evans' result on arithmetic progressions, which are linear polynomials in $\mathbb{Z}[X]$. Also, it settles a conjecture of Harrington and Jones [25] regarding quadratic polynomials. This chapter is based on the paper [50], a joint work Márton Szikszai. Our proof relies on elementary properties of the roots of f modulo a prime, a quantitative version of Chebotarev density theorem, results on the p-adic valuations of products of consecutive polynomial values, and a special covering argument for residue classes.

In the last chapter, we improve a result of Pomerance regarding the "index divisibility

problem" for the sequence of central binomial coefficients. Given a sequence of integers $(a_n)_{n\geq 1}$ with some combinatorial or number-theoretic meaning, the "index divisibility problem" for $(a_n)_{n\geq 0}$ is the study of the set of positive integers n such that n divides a_n. This topic has interested several authors. As we have already mentioned, Alba González, Luca, Pomerance, and Shparlinski [1] considered the case of $(a_n)_{n\geq 1}$ being a linear recurrence; while André-Jeannin [4], Luca and Tron [36], Sanna [47], and Somer [56] focused on Lucas sequences. Gottschlich [20], Silverman and Stange [54] studied this problem for elliptic divisibility sequences; and Chen, Gassert and Stange [10] consider the case when $a_n = \phi^{(n)}(0)$ is the nth iterate of a polynomial map $\phi \in \mathbb{Z}[X]$.

We study the case in which $a_n = \binom{2n}{n}$ is the nth central binomial coefficient. Let $\mathcal{A}$ be the set of positive integers n such that n divides the central binomial coefficient $\binom{2n}{n}$. Ulas and Schinzel [60, Theorems 3.2 and 3.4] proved that $\mathcal{A}$ and its complement $\mathbb{N} \setminus \mathcal{A}$ are both infinite. Pomerance [41, Theorem 3] studied the upper density of $\mathcal{A}$ and proved that $\overline{\mathbf{d}}(\mathcal{A}) \leq 1 - \log 2 = 0.30685\ldots$ Actually, probably for aesthetic reasons, Pomerance stated his result with $1/3$ instead of $1-\log 2$, but from the proof it is clear that he proved the bound with the latter quantity. Also, Pomerance [41, end of pag. 7] conjectured that $\mathcal{A}$ has a positive lower density, and indeed numerical experiments [55] seem to suggest that the lower density of $\mathcal{A}$ is at least $1/9$.

We improve Pomerance's result by showing that

$$\overline{\mathbf{d}}(\mathcal{A}) \leq 1 - \log 2 - 0.05551 = 0.25134\ldots$$

Then, similarly to the content of Chapter 2, we consider the "dual" set of $\mathcal{A}$, that is, the set $\mathcal{B}$ of all positive integers n such that $\binom{2n}{n}$ and n are relatively prime. It is easy to see that each odd prime number belongs to $\mathcal{B}$. Hence, by the Prime Number Theorem,

$$\#\mathcal{B}(x) \geq (1 + o(1)) \cdot \frac{x}{\log x},$$

as $x \to +\infty$. Our second result is that

$$\#\mathcal{B}(x) \ll \frac{x}{\sqrt{\log x}},$$

for all $x > 1$.

The chapter is based on the paper [49]. The proofs are based on a classical result of Kummer about the divisibility of binomial coefficients by a given prime number, and an estimate for the number of y-rough numbers not exceeding a certain limit.

Notation

We employ the Landau–Bachmann "Big Oh" and "little oh" notations O and o, as well as the associated Vinogradov symbols $\ll$ and $\gg$, with their usual meanings. Moreover, we write $A \asymp B$ to mean that both $A \ll B$ and $A \gg B$ hold. Any dependence of implied constants is explicitly stated or indicated with subscripts.

Given a set of positive integers $\mathcal{S}$, we put $\mathcal{S}(x) := \mathcal{S} \cap [1, x]$ for any $x \geq 1$, and we recall that the *natural density* of $\mathcal{S}$ is defined as

$$\mathbf{d}(\mathcal{S}) := \lim_{x \to +\infty} \frac{\#\mathcal{S}(x)}{x},$$

whenever this limit exists, while the *upper density* and the *lower density* of $\mathcal{S}$ are defined as

$$\overline{\mathbf{d}}(\mathcal{S}) := \limsup_{x \to +\infty} \frac{\#\mathcal{S}(x)}{x} \quad \text{and} \quad \underline{\mathbf{d}}(\mathcal{S}) := \liminf_{x \to +\infty} \frac{\#\mathcal{S}(x)}{x},$$

respectively.

Throughout, we reserve the letters p and q for prime numbers, and we write ν_p for the p-adic valuation.

For integers a and $m > 0$, we use $(a \bmod m)$ to denote the unique nonnegative integer $r < m$ such that m divides $a - r$.

As usual, $\mu(n)$, $\varphi(n)$, and $\tau(n)$, denote the Möbius function, the Euler's totient function, and the number of divisors of a positive integer n, respectively.

We write $\mathcal{O}_{\mathbb{K}}$ for the ring of integers of a number field $\mathbb{K}$, while $N_{\mathbb{K}}(\alpha)$ denotes the norm of $\alpha \in \mathbb{K}$ over $\mathbb{Q}$.

Part I

Arithmetic properties of linear recurrences

Chapter 1

Distribution of integral values for the ratio of two linear recurrences

Abstract. *Let F and G be linear recurrences over a number field $\mathbb{K}$, and let $\mathfrak{R}$ be a finitely generated subring of $\mathbb{K}$. Furthermore, let $\mathcal{N}$ be the set of positive integers n such that $G(n) \neq 0$ and $F(n)/G(n) \in \mathfrak{R}$. Under mild hypothesis, Corvaja and Zannier proved that $\mathcal{N}$ has zero natural density. I prove that*

$$\#\mathcal{N}(x) \ll x \cdot \left(\frac{\log\log x}{\log x}\right)^h$$

for all $x \geq 3$, where h is a positive integer that can be computed in terms of F and G. Assuming the Hardy–Littlewood k-tuple conjecture, this result is optimal except for the term $\log\log x$. This work is appeared in [46].

1.1 Introduction

Let F and G be linear recurrences and let $\mathfrak{R}$ be a finitely generated subring of $\mathbb{C}$. Assume also that the roots of the characteristic polynomials of F and G generate a multiplicative torsion-free group. This "torsion-free" hypothesis is not a loss of generality. Indeed, if the group generated by such roots has torsion order q, then for each $r = 0, 1, \ldots, q-1$ the roots of the characteristic polynomials of the linear recurrences $F_r(n) = F(qn + r)$ and $G_r(n) = G(qn + r)$ generate a torsion-free group. Therefore, all the results in the following can be extended just by partitioning $\mathbb{N}$ into the arithmetic progressions of modulo q and by studying each pair of linear recurrences F_r, G_r separately. Finally, define the following set of natural numbers

$$\mathcal{N} := \{n \in \mathbb{N} : G(n) \neq 0,\ F(n)/G(n) \in \mathfrak{R}\}.$$

Regarding the condition $G(n) \neq 0$, note that, by the "torsion-free" hypothesis, $G(n)$ is nondegenerate and hence the Skolem–Mahler–Lech Theorem [19, Theorem 2.1] implies that $G(n) = 0$ only for finitely many $n \in \mathbb{N}$. In the sequel, we shall tacitly disregard such integers.

Divisibility properties of linear recurrences have been studied by several authors. A classical result, conjectured by Pisot and proved by van der Poorten, is the Hadamard-quotient Theorem, which states that if $\mathcal{N}$ contains all sufficiently large integers, then F/G is itself a linear recurrence [44, 61].

Corvaja and Zannier [12, Theorem 2] gave the following wide extension of the Hadamard-quotient Theorem (see also [11] for a previous weaker result by the same authors).

Theorem 1.1.1. *If $\mathcal{N}$ is infinite, then there exists a nonzero polynomial $P \in \mathbb{C}[X]$ such that both the sequences $n \mapsto P(n)F(n)/G(n)$ and $n \mapsto G(n)/P(n)$ are linear recurrences.*

The proof of Theorem 1.1.1 makes use of the Schmidt's Subspace Theorem. We refer the reader to [5] (see also [63, 64]) for a survey on several applications of the Schmidt's Subspace Theorem in Number Theory.

Let $\mathbb{K}$ be a number field. For the sake of simplicity, from now on we shall assume that $\mathfrak{R} \subseteq \mathbb{K}$ and that F and G have coefficients and values in $\mathbb{K}$.

Corvaja and Zannier [12, Corollary 2] proved the following theorem about $\mathcal{N}$.

Theorem 1.1.2. *If F/G is not a linear recurrence, then $\mathcal{N}$ has zero natural density.*

Corvaja and Zannier also suggested [12, Remark p. 450] that their proof of this result could be adapted to show that if F/G is not a linear recurrence then

$$\#\mathcal{N}(x) \ll_{\delta,\mathbb{K}} \frac{x}{(\log x)^{\delta}}, \tag{1.1}$$

for any $\delta < 1$ and for all sufficiently large $x > 1$.

In this chapter, we prove a more precise upper bound than (1.1), namely:

Theorem 1.1.3. *If F/G is not a linear recurrence, then*

$$\#\mathcal{N}(x) \ll_{F,G} x \cdot \left(\frac{\log\log x}{\log x}\right)^{h},$$

for all $x \geq 3$, where h is a positive integer depending on F and G.

Both the positive integer h and the implied constant in the bound of Theorem 1.1.3 are effectively computable, we give the details in §1.4. In particular, we have the following corollary.

Corollary 1.1.1. *If F/G is not a linear recurrence, $G \in \mathbb{Z}[X]$, and $\gcd(G, f_1, \ldots, f_r) = 1$, where $f_1, \ldots, f_r$ are the polynomials appearing in the generalized power sum expression of F, then h can be taken as the number of irreducible factors of G in $\mathbb{Z}[X]$ (counted without multiplicity).*

Except for the term $\log\log x$, we believe that Corollary 1.1.1 should be optimal. Indeed, pick a positive integer h and an *admissible h-tuple* $\mathbf{h} = \{n_1, \ldots, n_h\}$, that is, $n_1 < \cdots < n_h$ are positive integers such that for each prime number p there exists a residue class modulo p which does not intersect $\mathbf{h}$. Assuming the Hardy–Littlewood h-tuple conjecture [24, p. 61], we have that the number $T_{\mathbf{h}}(x)$ of positive integers $n \leq x$ such that $n + n_1, \ldots, n + n_h$ are all prime numbers satisfies

$$T_{\mathbf{h}}(x) \sim \mathfrak{S}_{\mathbf{h}} \cdot \frac{x}{(\log x)^{h}},$$

as $x \to +\infty$, where

$$\mathfrak{S}_{\mathbf{h}} := \prod_{p} \left(1 - \frac{\upsilon_{\mathbf{h}}(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-h}$$

and $\upsilon_{\mathbf{h}}(p)$ is the number of residue classes modulo p which intersect $\mathbf{h}$. In particular, since $\upsilon_{\mathbf{h}}(p) < h$, it is easy to see that $\mathfrak{S}_{\mathbf{h}} > 0$. On the other hand, by Fermat's little theorem, if we pick the linear recurrences

$$F(n) = (2^{n+n_1} - 2) \cdots (2^{n+n_h} - 2), \quad G(n) = (n + n_1) \cdots (n + n_h),$$

then $F(n)/G(n)$ is an integer for each $n \in T_{\mathbf{h}}(x)$. Therefore, for $\mathfrak{R} = \mathbb{Z}$, we have

$$\#\mathcal{N}(x) \geq T_{\mathbf{h}}(x) \gg \frac{x}{(\log x)^h},$$

for all sufficiently large $x > 1$. Finally, the ratio $R := F/G$ is not a linear recurrence. In fact, assuming that R is a linear recurrence and looking at the generalized power sums representations of both sides of $F = GR$ we get a contradiction: F is simple but GR is not.

1.2 Preliminaries

First, we need to state a quantitative version of the Chebotarev density theorem. Some notation and preliminary facts are necessary. Let $\mathbb{L}$ be a finite Galois extension of $\mathbb{Q}$, let $\mathcal{O}_{\mathbb{L}}$ be its ring of integers, and let $\mathcal{G} := \mathrm{Gal}(\mathbb{L}/\mathbb{Q})$ be its Galois group. If p is a prime number which does not ramify in $\mathbb{L}$, and $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_{\mathbb{L}}$ over p, then the *Frobenius element* of $\mathfrak{p}$ is the unique $\sigma_{\mathfrak{p}} \in \mathcal{G}$ such that $\sigma_{\mathfrak{p}}(a) \equiv a^p \pmod{\mathfrak{p}}$ for all $a \in \mathcal{O}_{\mathbb{L}}$. The set of all Frobenius elements $\sigma_{\mathfrak{p}}$, with $\mathfrak{p}$ prime ideal of $\mathcal{O}_{\mathbb{L}}$ over p, is a conjugacy class of $\mathcal{G}$, denote by σ_p and called the *Frobenius element* of p. We write $\pi_{\mathcal{C}}(x)$ for the number of prime numbers $p \leq x$ not ramifying in $\mathbb{L}$ and such that $\sigma_p = \mathcal{C}$. In particular, when $\mathbb{L}$ is the splitting field over $\mathbb{Q}$ of a nonconstant polynomial $f \in \mathbb{Z}[X]$, if the elements of $\mathcal{C}$ have cycle pattern $d_1, \ldots, d_s$, when regarded as permutations of the roots of f, then $\pi_{\mathcal{C}}(x)$ is the number of primes $p \leq x$ not dividing the discriminant of f and such that the irreducible factors of f modulo p have degrees $d_1, \ldots, d_s$. (For these facts see, e.g., [57].) At this point, Chebotarev density theorem can be stated as follow [53, Theorem 3.4].

Theorem 1.2.1. *Let $\mathbb{L}$ be a finite Galois extension of $\mathbb{Q}$ with Galois group $\mathcal{G}$, and let $\mathcal{C}$ be a coniugacy class of $\mathcal{G}$. We have*

$$\pi_{\mathcal{C}}(x) = \frac{\#\mathcal{C}}{\#\mathcal{G}} \cdot \mathrm{Li}(x) + O_{\mathbb{L}}\left(\frac{x}{\exp\bigl(C\sqrt{\log x}\bigr)}\right)$$

as $x \to +\infty$, where $\mathrm{Li}(x)$ is the logarithmic integral function and $C > 0$ is a constant depending on $\mathbb{L}$.

For each nonconstant polynomial $f \in \mathbb{Z}[X]$ and for each prime number p, let $\eta_f(p)$ be the number of zeros of f modulo p. The next result states that the average of $\eta_f(p)$, as p runs over the prime numbers, is equal to the number of irreducible factors of f in $\mathbb{Z}[X]$, counted without multiplicity.

Lemma 1.2.2. *For any nonconstant polynomial $f \in \mathbb{Z}[X]$, we have*

$$\sum_{p \le x} \eta_f(p) = h\,\mathrm{Li}(x) + O_f\left(\frac{x}{\exp\bigl(C\sqrt{\log x}\bigr)}\right),$$

for all $x \ge 1$, where h is the number of irreducible factors of f in $\mathbb{Z}[X]$, and $C > 0$ is a constant depending only on f.

Proof. It is enough to prove the claim for irreducible f. Let $\mathbb{L}$ be the splitting field of f over $\mathbb{Q}$ and let $\mathcal{G} := \mathrm{Gal}(\mathbb{L}/\mathbb{Q})$. As we have mentioned before, for any coniugacy class $\mathcal{C}$ of $\mathcal{G}$ we have that $\pi_{\mathcal{C}}(x)$ is the number of prime numbers $p \le x$ which do not ramified in $\mathbb{L}$ and such that $\sigma_p = \mathcal{C}$. Also, if $g \in \mathcal{C}$ has cycle pattern $d_1, \dots, d_s$, when regarded as a permutation of the set X of roots of f, then the irreducible factors of f modulo p have degrees $d_1, \dots, d_s$, for each of the former prime numbers. Hence, f has $\#X^g$ zeros modulo p, where X^g denotes the set of roots of f which are fixed by g. Furthermore, since f is irreducible, $\mathcal{G}$ acts transitively on the roots of f, hence

$$\sum_{g \in \mathcal{G}} \#X^g = \#\mathcal{G},$$

by Burnside's lemma. Therefore, by Theorem 1.2.1, we have

$$\sum_{p\le x}\eta_f(p)=\sum_{\mathcal{C}}\sum_{g\in\mathcal{C}}\frac{\#X^g}{\#\mathcal{C}}\pi_{\mathcal{C}}(x)=\sum_{g\in\mathcal{G}}\frac{\#X^g}{\#\mathcal{G}}\operatorname{Li}(x)+O_{\mathbb{L}}\left(\frac{x}{\exp\bigl(C\sqrt{\log x}\bigr)}\right)$$
$$=\operatorname{Li}(x)+O_{\mathbb{L}}\left(\frac{x}{\exp\bigl(C\sqrt{\log x}\bigr)}\right),$$

for some constant $C>0$ depending on f. □

The next lemma follows from the previous by partial summation. We point out that it is due to Kronecker [28] (see also [57, p. 32]), who proved it with more elementary methods, not relying on Chebotarev density theorem.

Lemma 1.2.3. *For any nonconstant polynomial $f\in\mathbb{Z}[X]$, we have*

$$\sum_{p\le x}\eta_f(p)\cdot\frac{\log p}{p}=h\log x+O_f(1),$$

for all $x\ge 1$, where h is the number of irreducible factors of f in $\mathbb{Z}[X]$.

The following lemma regards the minimum of the multiplicative orders of some fixed algebraic numbers modulo a prime ideal.

Lemma 1.2.4. *Let $\beta_1,\dots,\beta_s\in\mathbb{K}$ such that none of them is zero or a root of unity. Then, for all $x\ge 1$, the number of prime numbers $p\le x$ such that some β_i has order less than $p^{1/4}$ modulo some prime ideal of $\mathcal{O}_{\mathbb{K}}$ lying above p is $O(x^{1/2})$, where the implied constant depends only on $\beta_1,\dots,\beta_s$.*

Proof. Let $\mathcal{P}$ be the set of all prime numbers p such that some β_i has order less than $p^{1/4}$ modulo some prime ideal of $\mathcal{O}_{\mathbb{K}}$ lying above p. Write $\beta_i=\alpha_i/m_i$, where $\alpha_i\in\mathcal{O}_{\mathbb{K}}$ and m_i is a nonzero integer. Hence, for each $p\in\mathcal{P}$ there exists a positive integer $t<p^{1/4}$ such that $p\mid N_{\mathbb{K}}(\alpha_i^t-m_i^t)$. Note that

$$|N_{\mathbb{K}}(\alpha_i^t-m_i^t)|=\prod_{j=1}^{n}|\sigma_j(\alpha_i^t-m_i^t)|=\prod_{j=1}^{n}|\sigma_j(\alpha_i)^t-m_i^t|=\exp(O_{\beta_i}(t)),$$

where $n:=[\mathbb{K}:\mathbb{Q}]$ and $\sigma_1,\dots,\sigma_n$ are all the embeddings $\mathbb{K}\to\mathbb{C}$. Therefore,

$$2^{\#\mathcal{P}(x)}\le\prod_{p\in\mathcal{P}(x)}p\;\Big|\;\prod_{i=1}^{s}\prod_{t<x^{1/4}}|N_{\mathbb{K}}(\alpha_i^t-m_i^t)|=\exp(O_{\beta_1,\dots,\beta_s}(x^{1/2})),$$

and the desired claim follows. □

Given a multiplicative function g, let Λ_g be its *associated von Mangoldt function*, that is, the unique arithmetic function satisfying

$$\sum_{d \mid n} g(n/d)\Lambda_g(d) = g(n)\log n,$$

for all positive integers n (see [27, p. 17]). It is easy to prove that Λ_g is supported on prime powers. Precisely, for a fixed prime number p, the quantities $\Lambda_g(p^s)$ $(s = 0, 1, \ldots)$ satisfy the recursive formulas

$$\begin{aligned} &\Lambda_g(1) = 0, \\ &\Lambda_g(p^s) = sg(p^s)\log p - \sum_{j=1}^{s-1} g(p^{s-j})\Lambda_g(p^j), \quad s \geq 1. \end{aligned}$$

We need the next technical lemma.

Theorem 1.2.5. *For each $y > 0$, let g_y be a multiplicative arithmetic function and let $L_y > 0$. Suppose that*

$$\sum_{n \leq x} \Lambda_{g_y}(n) = h\log x + O(L_y) \tag{1.2}$$

and

$$\sum_{n \leq x} |g_y(n)| \ll (\log x)^h, \tag{1.3}$$

for all $x, y \geq 2$, where $h > 0$ is some absolute constant, and the implied constant does not depend on y. Then

$$\sum_{n \leq x} g_y(n) = (\log x)^h \cdot \left(c_{g_y} + O_h\left(\frac{L_y}{\log x}\right)\right),$$

for all $x, y \geq 2$, where

$$c_{g_y} := \frac{1}{\Gamma(h+1)} \prod_p \left(1 + g_y(p) + g_y(p^2) + \cdots\right)\left(1 - \frac{1}{p}\right)^h$$

and Γ is the Euler's Gamma function.

Proof. The proof proceeds exactly as the proof of [27, Theorem 1.1], but using the error term $O(L_y)$ instead of $O(1)$. □

We need the following form of the "large sieve" inequality [27, Theorem 7.14].

Theorem 1.2.6. *Let $\mathcal{P}$ be a finite set of prime numbers, and for each $p \in \mathcal{P}$ let $\Omega_p \subsetneq \{0, 1, \dots, p-1\}$ be a set of residues modulo p. Then we have*

$$\#\{n \leq x : (n \bmod p) \notin \Omega_p,\ \forall p \in \mathcal{P}\} \leq (x + Q^2) \cdot \left(\sum_{m \leq Q} g(m)\right)^{-1}$$

for all $x, Q \geq 1$, where g is the multiplicative arithmetic function supported on squarefree numbers with prime factors in $\mathcal{P}$ and satisfying

$$g(p) = \frac{\#\Omega_p}{p - \#\Omega_p},$$

for all $p \in \mathcal{P}$.

Now we state a technical lemma about the cardinality of a sieved set of integers.

Lemma 1.2.7. *For each prime number p, let $\Omega_p \subsetneq \{0, 1, \dots, p-1\}$ be a set of residues modulo p. Suppose that there exist constants $c, h > 0$ such that $\#\Omega_p \leq c$ for each prime number p and*

$$\sum_{p \leq x} \#\Omega_p \cdot \frac{\log p}{p} = h \log x + O(1), \tag{1.4}$$

for all $x > 1$. Then we have

$$\#\{n \leq x : (n \bmod p) \notin \Omega_p,\ \forall p \in]y, z]\} \ll_{c,h,\delta_1,\delta_2} x \cdot \left(\frac{\log y}{\log x}\right)^h,$$

for all $\delta_1, \delta_2 > 0$, $x > 1$, $2 \leq y \leq (\log x)^{\delta_1}$, and $z \geq x^{\delta_2}$.

Proof. All the constants in this proof, included the implied ones, may depend on c, h, δ_1, δ_2. Clearly, we can assume $\delta_2 \leq 1/2$. By Theorem 1.2.6, we have

$$\#\{n \leq x : (n \bmod p) \notin \Omega_p,\ \forall p \in]y, z]\} \ll x \cdot \left(\sum_{m \leq w} g_y(m)\right)^{-1}, \tag{1.5}$$

where $w := x^{\delta_2}$ and g_y is the multiplicative arithmetic function supported on squarefree numbers with all prime factors $> y$ and such that

$$g_y(p) = \frac{\#\Omega_p}{p - \#\Omega_p},$$

for any prime number $p > y$.

For sufficiently large x, we have $y \leq w$, and it follows from (1.4) and $\#\Omega_p \leq c_1$ that

$$\sum_{p \leq w} g_y(p) \log p = h \log w + O(\log y),$$

which in turn implies that

$$\sum_{n \leq w} \Lambda_{g_y}(n) = h \log w + O(\log y),$$

since Λ_{g_y} is supported on prime powers p^s, with $p > y$, and $\Lambda_{g_y}(p^s) = -(-g_y(p))^s \log p$.

Furthermore, again from (1.4) and $\#\Omega_p \leq c_1$, we have

$$\prod_{p \leq t} \left(1 - \frac{\#\Omega_p}{p}\right)^{-1} \asymp (\log t)^h, \tag{1.6}$$

for all $t \geq 2$, so that

$$\sum_{n \leq w} |g_y(n)| \leq \prod_{p \leq w} (1 + g_y(p)) \leq \prod_{p \leq w} \left(1 - \frac{\#\Omega_p}{p}\right)^{-1} \ll (\log w)^h.$$

At this point, we have proved that (1.2) and (1.3) hold with $L_y = \log y$. Therefore, by Theorem 1.2.5 we have

$$\sum_{n \leq w} g_y(n) = (\log w)^h \cdot \left(c_{g_y} + O\left(\frac{\log y}{\log w}\right)\right), \tag{1.7}$$

where

$$c_{g_y} = \frac{1}{\Gamma(h+1)} \prod_p (1 + g_y(p)) \left(1 - \frac{1}{p}\right)^h.$$

Now using (1.6) we obtain

$$c_{g_y} = \frac{1}{\Gamma(h+1)} \prod_p \left(1 - \frac{\#\Omega_p}{p}\right)^{-1} \left(1 - \frac{1}{p}\right)^h \prod_{p \leq y} \left(1 - \frac{\#\Omega_p}{p}\right) \gg \frac{1}{(\log y)^h}. \tag{1.8}$$

Hence, recalling that $y \leq (\log x)^{\delta_1}$ and $w = x^{\delta_2}$, by (1.7) and (1.8) we find that

$$\sum_{n \leq w} g_y(n) \gg \left(\frac{\log w}{\log x}\right)^h \gg \left(\frac{\log x}{\log y}\right)^h. \tag{1.9}$$

Putting together (1.5) and (1.9), the desired result follows. □

We need a lemma about the number of zeros of a sparse polynomial in a finite field of q elements $\mathbb{F}_q$ [9, Lemma 7].

Lemma 1.2.8. *Let $c_1, \dots, c_r \in \mathbb{F}_q^*$ ($r \geq 2$) and $t_1, \dots, t_r \in \mathbb{Z}$. Then the number T of solutions of the equation*

$$\sum_{i=1}^{r} c_i x^{t_i} = 0, \quad x \in \mathbb{F}_q^* \tag{1.10}$$

satisfies

$$T \leq 2q^{1-1/(r-1)} D^{1/(r-1)} + O(q^{1-2/(r-1)} D^{2/(r-1)}),$$

where

$$D := \min_{1 \leq i \leq r} \max_{j \neq i} \gcd(t_i - t_j, q-1).$$

We will use the following corollary of Lemma 1.2.8, which concerns the number of zeros of a simple linear recurrence over a finite field.

Corollary 1.2.1. *Let $c_1, \dots, c_r, a_1, \dots, a_r \in \mathbb{F}_q^*$ ($r \geq 2$), and let N be the minimum of the orders of the a_i/a_j ($i \neq j$) in $\mathbb{F}_q^*$. Then the number of integers $m \in [0, q-2]$ such that*

$$\sum_{i=1}^{r} c_i a_i^m = 0 \tag{1.11}$$

is $O\big(qN^{-1/(r-1)}\big)$.

Proof. Let g be a generator of the multiplicative group $\mathbb{F}_q^*$, so that for each $i = 1, \dots, r$ we have $a_i = g^{t_i}$ for some integer t_i. Clearly, m is a solution of (1.11) if and only if g^m is a solution of (1.10). Finally, the order of a_i/a_j ($i \neq j$) is given by $(q-1)/\gcd(t_i - t_j, q-1)$, hence $D \leq (q-1)/N$, and the desidered claim follows. □

Given a finite set S of absolute values of $\mathbb{K}$ containing all the archimedean ones, we write $\mathcal{O}_S$ for the ring of S-integers of $\mathbb{K}$, that is, the set of all $\alpha \in \mathbb{K}$ such that $|\alpha|_v \leq 1$ for all $v \notin S$. We state the following easy lemma.

Lemma 1.2.9. *Let S be a finite set of absolute values of $\mathbb{K}$ containing all the archimedean ones, and let $g_1, \ldots, g_t \in \mathbb{K}[X]$ be polynomials such that $(g_1, \ldots, g_t) = 1$. Then there exists a finite set S' of absolute values of $\mathbb{K}$, such that: $S \subseteq S'$, $g_1, \ldots, g_t \in \mathcal{O}_{S'}[X]$, and $(g_1(n), \ldots, g_t(n)) = 1$ for all positive integers n, that is, the ideal of $\mathcal{O}_{S'}$ generated by $g_1(n), \ldots, g_t(n)$ is the whole $\mathcal{O}_{S'}$.*

Proof. Since $(g_1, \ldots, g_t) = 1$, by the Bézout's identity there exist $b_1, \ldots, b_t \in \mathbb{K}[X]$ such that

$$b_1 g_1 + \cdots + b_t g_t = 1.$$

Clearly, we can pick S' so that $S' \supseteq S$ and $b_i, g_i \in \mathcal{O}_{S'}[X]$ for all $i = 1, \ldots, t$. Hence, for each $n \in \mathbb{N}$, we have

$$b_1(n) g_1(n) + \cdots + b_t(n) g_t(n) = 1,$$

which in turn implies that $(g_1(n), \ldots, g_t(n)) = 1$. □

1.3 Proof of Theorem 1.1.3

The first part of the proof proceeds similarly to the proof of Theorem 1.1.2. If $\mathcal{N}$ is finite, then the claim is trivial, hence we suppose that $\mathcal{N}$ is infinite. Then, by Theorem 1.1.1 it follows that $F/G = H/P$, for some linear recurrence H and some polynomial P. As a consequence, without loss of generality, we shall assume that G is a polynomial.

Suppose that the generalized power sum expression of F is

$$F(n) = \sum_{i=1}^{r} f_i(n)\, \alpha_i^n, \quad n \in \mathbb{N},$$

where $f_1, \ldots, f_r \in \mathbb{C}[X]$ are polynomials and $\alpha_1, \ldots, \alpha_r \in \mathbb{C}^*$ are all the distinct roots of the characteristic polynomial of F.

Let S be a finite set of absolute values of $\mathbb{K}$ containing all the archimedean ones. Enlarging $\mathbb{K}$ and S we may assume that $\alpha_1, \dots, \alpha_r$ are S-units, $f_1, \dots, f_r, G \in \mathcal{O}_S[X]$, and $\mathfrak{R} \subseteq \mathcal{O}_S$.

Since F/G is not a linear recurrence, it follows that G does not divide all the $f_1, \dots, f_r$. Moreover, factoring out the greatest common divisor $(G, f_1, \dots, f_r)$ we can even assume that $(G, f_1, \dots, f_r) = 1$ and that G is nonconstant. In particular, by Lemma 1.2.9 we can enlarge S so that $(G(n), f_1(n), \dots, f_r(n)) = 1$ for all $n \in \mathbb{N}$.

It is easy to prove that there exist a positive integer g and a nonconstant polynomial $\widetilde{G} \in \mathbb{Z}[X]$ such that $N_{\mathbb{K}}(G(n)) = \widetilde{G}(n)/g$ for all $n \in \mathbb{N}$. Let h be the number of irreducible factors of $\widetilde{G}$ in $\mathbb{Z}[X]$. Again by enlarging S, we may assume that g is an S-unit.

Let $\mathcal{P}$ be the set of all prime numbers p which do not make $\widetilde{G}$ vanish identically modulo p, such that $p\mathcal{O}_{\mathbb{K}}$ has no prime ideal factor π_v with $v \in S$, and such that the minimum order of the α_i/α_j $(i \neq j)$ modulo any prime ideal above p is at least $p^{1/4}$. Furthermore, let us define

$$\Omega_p := \left\{ \ell \in \{0, \dots, p-1\} : \widetilde{G}(\ell) \equiv 0 \pmod{p} \right\},$$

for any $p \in \mathcal{P}$, and $\Omega_p := \varnothing$ for any prime number $p \notin \mathcal{P}$.

Let $x \geq 3$, $y := (\log x)^{4rh}$, and $z := x^{1/(d+1)}$, where $d := [\mathbb{K} : \mathbb{Q}]$. We split $\mathcal{N}(x)$ into two subsets:

$$\mathcal{N}_1 := \{ n \in \mathcal{N}(x) : (n \bmod p) \notin \Omega_p,\ \forall p \in]y, z] \},$$
$$\mathcal{N}_2 := \mathcal{N} \setminus \mathcal{N}_1.$$

First, we give an upper bound for $\#\mathcal{N}_1$. Hereafter, all the implied constants may depend on F and G. Clearly, $\Omega_p \subsetneq \{0, 1, \dots, p-1\}$ and $\#\Omega_p \leq \deg(\widetilde{G})$ for all prime number p, while from Lemma 1.2.3 and Lemma 1.2.4 it follows that

$$\sum_{p \leq x} \#\Omega_p \cdot \frac{\log p}{p} = h \log x + O(1).$$

Therefore, applying Lemma 1.2.7, we obtain

$$\#\mathcal{N}_1 \ll x \cdot \left(\frac{\log y}{\log x} \right)^h \ll \left(\frac{\log \log x}{\log x} \right)^h.$$

Now we give an upper bound for $\#\mathcal{N}_2$. If $n \in \mathcal{N}_2$ then there exist $p \in \mathcal{P} \cap]y, z]$ and $\ell \in \Omega_p$ such that $n \equiv \ell \pmod{p}$. In particular, p divides $N_{\mathbb{K}}(G(\ell))$ in $\mathcal{O}_S$ and, since $p\mathcal{O}_{\mathbb{K}}$ has no prime ideal factor π_v with $v \in S$, it follows that there exists some prime ideal π of $\mathcal{O}_S$ lying above p and dividing $G(\ell)$. Let $\mathbb{F}_q := \mathcal{O}_S/\pi$, so that q is a power of p. Write $n = \ell + mp$, for some integer $m \geq 0$. Since π divides $G(n)$ and $F(n)/G(n) \in \mathcal{O}_S$, we have that $F(n)$ is divisible by π too. As a consequence, we obtain that

$$\sum_{i=1}^{r} f_i(\ell)\, \alpha_i^{\ell} \left(\alpha_i^{p}\right)^{m} \equiv \sum_{i=1}^{r} f_i(n)\, \alpha_i^{n} \equiv F(n) \equiv 0 \pmod{\pi}. \tag{1.12}$$

Note that $f_1(\ell), \ldots, f_r(\ell)$ cannot be all equal to zero modulo π, since π divides $G(\ell)$ and $(G(\ell), f_1(\ell), \ldots, f_r(\ell)) = 1$. Note also that the minimum order N of the α_i^p/α_j^p $(i \neq j)$ modulo π is equal to the minimum order of the α_i/α_j $(i \neq j)$ modulo π, since $(p, q-1) = 1$. In particular, $N \geq p^{1/4}$, in light of the definition of $\mathcal{P}$.

Therefore, we can apply Corollary 1.2.1 to the congruence (1.12), getting that the number of possible values of m modulo $q-1$ is $O(q/p^{\gamma})$, where $\gamma := 1/(4r)$. Consequently, the number of possible values of $n \leq x$ is

$$\left(\frac{x}{p(q-1)} + 1\right) \cdot O\left(\frac{q}{p^{\gamma}}\right) = O\left(\frac{x}{p^{1+\gamma}}\right),$$

since $p(q-1) < p^{d+1} \leq z^{d+1} \leq x$. Hence, we have

$$\#\mathcal{N}_2 \ll \sum_{p \in \mathcal{P} \cap]y,z]} \frac{x}{p^{1+\gamma}} \ll \int_{y}^{+\infty} \frac{\mathrm{d}t}{t^{1+\gamma}} \ll \frac{x}{y^{\gamma}} = \frac{x}{(\log x)^{h}}.$$

In conclusion,

$$\#\mathcal{N}(x) = \#\mathcal{N}_1 + \#\mathcal{N}_2 \ll x \cdot \left(\frac{\log\log x}{\log x}\right)^{h}$$

as claimed.

1.4 Effectiveness of Theorem 1.1.3

Let us briefly explain the computation of h. First, we have an effective procedure to test if there exists a nonzero polynomial $P \in \mathbb{C}[X]$ such that the sequences $n \mapsto$

$P(n)F(n)/G(n)$ and $n \mapsto G(n)/P(n)$ are linear recurrences, and in such a case P can be determined (see [12, p. 435, Remark 1]).

On the one hand, if P does not exist, then Theorem 1.1.1 implies that $\mathcal{N}$ is finite, hence h can be any positive integer. Moreover, using any effective bound for the number of zeros of a nondegenerate linear recurrence (see, e.g., [3, 52, 62]) at the end of the proof of [12, Proposition 2.1] (precisely, where it is said: "By the Skolem-Mahler-Lech Theorem again, this relation holds identically..."), it is possible to effectively bound $\#\mathcal{N}$. Therefore, if P does not exist then the implied constant in Theorem 1.1.3 is effectively computable.

On the other hand, if P exists, then we can write the linear recurrences $H = PF/G$ as

$$H(n) = \sum_{i=1}^{s} h_i(n)\,\beta_i^n,$$

for some $\beta_1, \ldots, \beta_s \in \mathbb{C}^*$ and $h_1, \ldots, h_s \in \mathbb{C}[X]$. Setting $Q := P/(P, h_1, \ldots, h_s)$, we have that $\widetilde{Q}(n) = N_{\mathbb{K}}(Q(n))$ is a polynomial in $\mathbb{Q}[X]$ and h can be taken as the number of distinct irreducible factors of $\widetilde{Q}$. Furthermore, all the implied constants of the results used in the proof of Theorem 1.1.3 are effectively computable, hence also when P exists the implied constant in Theorem 1.1.3 is effectively computable.

1.5 Proof of Corollary 1.1.1

Let us follow the instruction (and notation) for the computation of h given in §1.4. Clearly, $P = G$ and, consequently, $H = F$, $s = r$, $h_i = f_i$. Furthermore, we have $Q = G$, since $(G, f_1, \ldots, f_r) = 1$. Finally, recalling that $G \in \mathbb{Z}[X]$, we get that $N_{\mathbb{K}}(G(n)) = G(n)^{[\mathbb{K}:\mathbb{Q}]}$ for all positive integers n, hence $\widetilde{Q}(X) = G(X)^{[\mathbb{K}:\mathbb{Q}]}$. At this point, h can be taken as the number of irreducible factors of $\widetilde{Q}$, which is also the number of irreducible factors of G (recalling that we are counting them without multiplicity). The proof is complete.

Chapter 2

On numbers n coprime to the nth term of a linear recurrence

Abstract. *Let $(u_n)_{n\geq 0}$ be a nondegenerate linear recurrence of integers, and let $\mathcal{C}_u$ be the set of positive integers n such that u_n and n are relatively prime. I prove that $\mathcal{C}_u$ has a natural density, and that this density is positive unless $(u_n/n)_{n\geq 1}$ is a linear recurrence. This work is appeared in* [48].

2.1 Introduction

Let $(u_n)_{n\geq 0}$ be a linear recurrence over the integers, so that

$$u_n = a_1 u_{n-1} + a_2 u_{n-2} + \cdots + a_k u_{n-k},$$

for all integers $n \geq k$, where $a_1, \ldots, a_k \in \mathbb{Z}$ and $a_k \neq 0$. To avoid trivialities, we assume that $(u_n)_{n\geq 0}$ is not identically zero. The set

$$\mathcal{D}_u := \{n \in \mathbb{N} : n \mid u_n\}$$

has been studied by several researchers. Alba González, Luca, Pomerance, and Shparlinski [1] proved that if $(u_n)_{n\geq 1}$ is nondegenerate, simple, and of order $k \geq 2$, then

$$\#\mathcal{D}_u(x) \ll_k \frac{x}{\log x},$$

for all sufficiently large $x > 1$. Furthermore, André-Jeannin [4] and Somer [56] studied the arithmetic properties of the elements of $\mathcal{D}_u$ when $(u_n)_{n\geq 0}$ is a Lucas sequence. In such a case, generalizing a previous result of Luca and Tron [36] on Fibonacci numbers, we have proved [47] that

$$\#\mathcal{D}_u(x) \leq x \Big/ \exp\left(\left(\frac{1}{2} + o(1)\right) \frac{\log x \log\log\log x}{\log\log x}\right),$$

as $x \to \infty$, where the $o(1)$ depends on a_1 and a_2.

On the other hand, the "dual" set

$$\mathcal{C}_u = \{n \in \mathbb{N} : \gcd(n, u_n) = 1\}$$

does not seem to have attracted so much attention. We fill this gap by proving the following result:

Theorem 2.1.1. *Let $(u_n)_{n\geq 0}$ be a nondegenerate linear recurrence over the integers. Then, the natural density $\mathbf{d}(\mathcal{C}_u)$ of $\mathcal{C}_u$ exists. Moreover, if $(u_n/n)_{n\geq 1}$ is not a linear recurrence then $\mathbf{d}(\mathcal{C}_u) > 0$. On the other hand, if $(u_n/n)_{n\geq 1}$ is a linear recurrence then $\mathcal{C}_u$ is finite and, a fortiori, $\mathbf{d}(\mathcal{C}_u) = 0$.*

We remark that given the initial conditions and the coefficients of a linear recurrence $(u_n)_{n\geq 0}$, it is easy to test effectively if $(u_n/n)_{n\geq 1}$ is a linear recurrence or not (see Lemma 2.2.1, in §2.2).

2.2 Preliminaries

In this section we give some definitions and collect some preliminary results needed in the later proofs. Let f_u be the characteristic polynomial of $(u_n)_{n\geq 0}$, i.e.,

$$f_u(X) = X^k - a_1X^{k-1} - a_2X^{k-2} - \cdots - a_k.$$

Moreover, let $\mathbb{K}$ be the splitting field of f_u over $\mathbb{Q}$, so that

$$u_n = \sum_{i=1}^{r} g_i(n)\,\alpha_i^n, \tag{2.1}$$

for all integers $n \geq 0$, where $\alpha_1, \ldots, \alpha_r \in \mathcal{O}_{\mathbb{K}}$ are all the distinct roots of f_u, and $g_1, \ldots, g_r \in \mathbb{K}[X]$. We define B_u as the smallest positive integer such that all the coefficients of the polynomials $B_ug_1, \ldots, B_ug_r$ are algebraic integers.

We have the following easy lemma.

Lemma 2.2.1. *$(u_n/n)_{n\geq 1}$ is a linear recurrence if and only if*

$$g_1(0) = \cdots = g_r(0) = 0. \tag{2.2}$$

In such a case, $\mathcal{C}_u$ is finite.

Proof. The first part of the lemma follows immediately from the fact that any linear recurrence can be written as a generalized power sum like (2.1) in a unique way (assuming the roots $\alpha_1, \ldots, \alpha_r$ are distinct, and up to the order of the addends). For the second part, if (2.2) holds then for all positive integers n we have that

$$\frac{B_uu_n}{n} = \sum_{i=1}^{r} \frac{B_ug_i(n)}{n}\,\alpha_i^n$$

is both a rational number and an algebraic integer, hence it is an integer. Therefore, $n \mid B_uu_n$, and so $\gcd(n, u_n) = 1$ only if $n \mid B_u$, which in turn implies that $\mathcal{C}_u$ is finite. □

For the rest of this section, we assume that $(u_n)_{n\geq 0}$ is nondegenerate and that f_u has only simple roots, hence, in particular, $r = k$. We write Δ_u for the discriminant of the

polynomial f_u, and we recall that Δ_u is a nonzero integer. If $k \geq 2$, then for all integers $x_1, \ldots, x_k$ we set

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

and for any prime number p not dividing a_k we define $T_u(p)$ as the greatest integer $T \geq 0$ such that p does not divide

$$\prod_{1 \leq x_2, \ldots, x_k \leq T} \max\{1, |N_{\mathbb{K}}(D_u(0, x_2, \ldots, x_k))|\},$$

where the empty product is equal to 1. It is known that such T exists [19, p. 88]. If $k = 1$, then we set $T_u(p) := +\infty$ for all prime numbers p not dividing a_1. Note that $T_u(p) = 0$ if and only if $k = 2$ and p divides Δ_u.

Finally, for all $\gamma \in]0, 1[$, we define

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

We are ready to state two important lemmas regarding $T_u(p)$ [1, Lemma 2.1, Lemma 2.2].

Lemma 2.2.2. *For all $\gamma \in]0, 1[$ and $x \geq 2^{1/\gamma}$ we have*

$$\#\mathcal{P}_{u,\gamma}(x) \ll_u \frac{x^{k\gamma}}{\gamma \log x}.$$

Lemma 2.2.3. *Assume that p is a prime number not dividing $a_k B_u \Delta_u$ and relatively prime with at least one term of $(u_n)_{n \geq 0}$. Then, for all $x \geq 1$, the number of positive integers $m \leq x$ such that $u_{pm} \equiv 0 \pmod{p}$ is*

$$O_k\left(\frac{x}{T_u(p)} + 1\right).$$

Actually, in [1] both Lemma 2.2.2 and Lemma 2.2.3 were proved only for $k \geq 2$. However, one can easily check that they are true also for $k = 1$.

2.3 Proof of Theorem 2.1.1

For all integers $n \geq 0$, define

$$v_n := B_u \sum_{i=1}^{r} \frac{g_i(n) - g_i(0)}{n} \alpha_i^n \quad \text{and} \quad w_n := B_u \sum_{i=1}^{r} g_i(0)\, \alpha_i^n.$$

Note that both $(v_n)_{n\geq 0}$ and $(w_n)_{n\geq 0}$ are linear recurrences over the algebraic integers, and that the characteristic polynomial of $(w_n)_{n\geq 0}$ has only simple roots.

Let $\mathcal{G}$ be the Galois group of $\mathbb{K}$ over $\mathbb{Q}$. Since u_n is an integer, for any $\sigma \in \mathcal{G}$ we have that

$$nv_n + w_n = B_u u_n = \sigma(B_u u_n) = \sigma(nv_n + w_n) = n\sigma(v_n) + \sigma(w_n), \tag{2.3}$$

for all integers $n \geq 0$. In (2.3) note that both $n\sigma(v_n)$ and $\sigma(w_n)$ are linear recurrences, and the first is a multiple of n, while the characteristic polynomial of the second has only simple roots. Since the expression of a linear recurrence as a generalized power sum is unique, from (2.3) we get that $w_n = \sigma(w_n)$ for any $\sigma \in \mathcal{G}$, hence w_n is an integer.

Thanks to Lemma 2.2.1, we know that $(w_n)_{n\geq 0}$ is identically zero if and only if $(u_n/n)_{n\geq 1}$ is a linear recurrence, and in such a case $\mathcal{C}_u$ is finite, so that the claim of Theorem 2.1.1 is obvious. Hence, we assume that $(w_n)_{n\geq 0}$ is not identically zero.

For the sake of convenience, put $\mathcal{E}_u := \mathbb{N} \setminus \mathcal{C}_u$. Thus we have to prove that the natural density of $\mathcal{E}_u$ exists and is less than 1. For each $y > 0$, we split $\mathcal{E}_u$ into two subsets:

$$\mathcal{E}_{u,y}^- := \{n \in \mathcal{E}_u : p \mid \gcd(n, u_n) \text{ for some } p \leq y\},$$
$$\mathcal{E}_{u,y}^+ := \mathcal{E}_u \setminus \mathcal{E}_{u,y}^-.$$

It is well known that $(u_n)_{n\geq 0}$ is definitively periodic modulo p, for any prime number p. Therefore, it is easy to see that $\mathcal{E}_{u,y}^-$ is an union of finitely many arithmetic progressions and a finite subset of $\mathbb{N}$. In particular, $\mathcal{E}_{u,y}^-$ has a natural density. If we put $\delta_y := \mathbf{d}(\mathcal{E}_{u,y}^-)$, then it is clear that δ_y is a bounded nondecreasing function of y, hence the limit

$$\delta := \lim_{y\to+\infty} \delta_y \tag{2.4}$$

exists finite. We shall prove that $\mathcal{E}_u$ has natural density δ. Hereafter, all the implied constants may depend on $(u_n)_{n\geq 0}$ and k. If $n \in \mathcal{E}_{u,y}^+(x)$ then there exists a prime $p > y$ such that $p \mid n$ and $p \mid u_n$. Furthermore, $B_u u_n = nv_n + w_n$ implies that $p \mid w_n$. Hence, we can write $n = pm$ for some positive integer $m \leq x/p$ such that $w_{pm} \equiv 0 \pmod{p}$. For sufficiently large y, we have that p does not divide $f_w(0)B_w\Delta_w$ (actually, $B_w = 1$) and is coprime with at least one term of $(w_s)_{s\geq 0}$, since $(w_s)_{s\geq 0}$ is not identically zero.

Therefore, by applying Lemma 2.2.3 to $(w_s)_{s\geq 0}$, we get that the number of possible values of m is at most

$$O\bigg(\frac{x}{pT_w(p)}+1\bigg).$$

As a consequence,

$$\#\mathcal{E}_{u,y}^{+}(x) \ll \sum_{y<p\leq x}\left(\frac{x}{pT_w(p)}+1\right) \ll x\cdot\left(\sum_{p>y}\frac{1}{pT_w(p)}+\frac{1}{\log x}\right), \tag{2.5}$$

where we also used the Chebyshev's bound for the number of primes not exceeding x. Setting $\gamma := 1/(k+1)$, by partial summation and Lemma 2.2.2, we have

$$\begin{aligned}\sum_{\substack{p>y\\ p\in\mathcal{P}_{w,\gamma}}}\frac{1}{pT_w(p)} \leq \sum_{\substack{p>y\\ p\in\mathcal{P}_{w,\gamma}}}\frac{1}{p} &= \bigg[\frac{\#\mathcal{P}_{w,\gamma}(t)}{t}\bigg]_{t=y}^{+\infty}+\int_y^{+\infty}\frac{\#\mathcal{P}_{w,\gamma}(t)}{t^2}\mathrm{d}t\\ &\ll \frac{1}{y^{1-k\gamma}}=\frac{1}{y^{\gamma}}.\end{aligned} \tag{2.6}$$

On the other hand,

$$\sum_{\substack{p>y\\ p\notin\mathcal{P}_{w,\gamma}}}\frac{1}{pT_w(p)} \leq \sum_{\substack{p>y\\ p\notin\mathcal{P}_{w,\gamma}}}\frac{1}{p^{1+\gamma}} \ll \int_y^{+\infty}\frac{\mathrm{d}t}{t^{1+\gamma}} \ll \frac{1}{y^{\gamma}} \tag{2.7}$$

Thus, putting together (2.5), (2.6), and (2.7), we obtain

$$\frac{\#\mathcal{E}_{u,y}^{+}(x)}{x} \ll \frac{1}{y^{\gamma}}+\frac{1}{\log x},$$

so that

$$\begin{aligned}\limsup_{x\to+\infty}\left|\frac{\#\mathcal{E}_u(x)}{x}-\delta_y\right| &= \limsup_{x\to+\infty}\left|\frac{\#\mathcal{E}_u(x)}{x}-\frac{\#\mathcal{E}_{u,y}^{-}(x)}{x}\right|\\ &= \limsup_{x\to+\infty}\frac{\#\mathcal{E}_{u,y}^{+}(x)}{x} \ll \frac{1}{y^{\gamma}},\end{aligned} \tag{2.8}$$

hence, by letting $y\to+\infty$ in (2.8) and by using (2.4), we get that $\mathbf{d}(\mathcal{E}_u)=\delta$.

It remains only to prove that $\delta<1$. Clearly,

$$\mathcal{E}_{u,y}^{-}\subseteq\{n\in\mathbb{N} : p\mid n \text{ for some } p\leq y\},$$

so that, by Eratosthenes' sieve and Mertens' third theorem [58, Ch. I.1, Theorem 11], we have

$$\limsup_{x\to+\infty}\frac{\#\mathcal{E}_{u,y}^{-}(x)}{x}\leq 1-\prod_{p\leq y}\left(1-\frac{1}{p}\right)\leq 1-\frac{c_1}{\log y}, \tag{2.9}$$

for all $y \geq 2$, where $c_1 > 0$ is an absolute constant. Furthermore, the last part of (2.8) says that

$$\limsup_{x \to +\infty} \frac{\#\mathcal{E}^+_{u,y}(x)}{x} \leq \frac{c_2}{y^\gamma}, \tag{2.10}$$

for all sufficiently large y, where $c_2 > 0$ is an absolute constant.

Therefore, putting together (2.9) and (2.10), we get

$$\begin{aligned} \delta = \lim_{x \to +\infty} \frac{\#\mathcal{E}_u(x)}{x} &\leq \limsup_{x \to +\infty} \frac{\#\mathcal{E}^-_{u,y}(x)}{x} + \limsup_{x \to +\infty} \frac{\#\mathcal{E}^+_{u,y}(x)}{x} \\ &\leq 1 - \left(\frac{c_1}{\log y} - \frac{c_2}{y^\gamma} \right), \end{aligned} \tag{2.11}$$

for all sufficiently large y.

Finally, picking a sufficiently large y, depending on c_1 and c_2, the bound (2.11) yields $\delta < 1$. The proof of Theorem 2.1.1 is complete.

Chapter 3

On numbers n having a prescribed G.C.D. with the nth Fibonacci number

Abstract. *For each positive integer k, let $\mathcal{A}_k$ be the set of all positive integers n such that $\gcd(n, F_n) = k$, where F_n denotes the nth Fibonacci number. I prove that the natural density of $\mathcal{A}_k$ exists and is equal to*

$$\sum_{d=1}^{\infty} \frac{\mu(d)}{\operatorname{lcm}(dk, z(dk))}$$

where μ is the Möbius function and $z(m)$ denotes the least positive integer n such that m divides F_n. I also give an effective criterion to establish when the natural density of $\mathcal{A}_k$ is zero and I show that this is the case if and only if $\mathcal{A}_k$ is empty. This is a work in collaboration with Emanuele Tron and appeared in [51].

3.1 Introduction

In this chapter, we focus on the linear recurrence of Fibonacci numbers $(F_n)_{n\geq 1}$, defined as usual by $F_1 = F_2 = 1$ and $F_{n+2} = F_{n+1} + F_n$ for all integers $n \geq 1$. For each positive integer k, define the set

$$\mathcal{A}_k := \{n \geq 1 : \gcd(n, F_n) = k\}.$$

Let $z(m)$ be the *rank of appearance*, or *entry point*, of a positive integer m in the sequence of Fibonacci numbers, that is, the smallest positive integer n such that m divides F_n. It is well known that $z(m)$ exists. Furthermore, set $\ell(m) := \operatorname{lcm}(m, z(m))$.

Our first result establishes the existence of the natural density of $\mathcal{A}_k$ and provides an effective criterion to check whether this natural density is positive.

Theorem 3.1.1. *For each positive integer k, the natural density of $\mathcal{A}_k$ exists. Moreover, $\mathbf{d}(\mathcal{A}_k) > 0$ if and only if $\mathcal{A}_k \neq \varnothing$, and this happens if and only if $k = \gcd(\ell(k), F_{\ell(k)})$.*

Our second result is an explicit formula for the natural density of $\mathcal{A}_k$.

Theorem 3.1.2. *For each positive integer k, we have*

$$\mathbf{d}(\mathcal{A}_k) = \sum_{d=1}^{\infty} \frac{\mu(d)}{\ell(dk)}, \tag{3.1}$$

where μ is the Möbius function, and the series (3.1) converges absolutely.

3.2 Preliminaries

The next lemma summarizes some basic properties of the functions ℓ, z and of the Fibonacci numbers.

Lemma 3.2.1. *For all positive integers m, n and all prime numbers p, we have:*

(i) $F_m \mid F_n$ *whenever* $m \mid n$.

(ii) $m \mid F_n$ *if and only if* $z(m) \mid n$.

(iii) $z(\operatorname{lcm}(m,n)) = \operatorname{lcm}(z(m), z(n))$.

(iv) $z(p) \mid p - \left(\frac{p}{5}\right)$, *where* $\left(\frac{p}{5}\right)$ *is a Legendre symbol.*

(v) $\nu_p(F_n) \geq \nu_p(n)$ *whenever* $z(p) \mid n$.

(vi) $m \mid \gcd(n, F_n)$ *if and only if* $\ell(m) \mid n$.

(vii) $\ell(\operatorname{lcm}(m,n)) = \operatorname{lcm}(\ell(m), \ell(n))$.

(viii) $\ell(p) = pz(p)$ *for* $p \neq 5$, *while* $\ell(5) = 5$.

Proof. Facts (i)–(iv) are well-known (see, e.g., [42]). Fact (v) follows quickly from the formulas for $\nu_p(F_n)$ given by Lengyel [34]. Finally, (vi)–(viii) are easy consequences of (i)–(iv) and the definition of ℓ. □

Now we state an easy criterion to establish if $\mathcal{A}_k \neq \varnothing$.

Lemma 3.2.2. $\mathcal{A}_k \neq \varnothing$ *if and only if* $k = \gcd(\ell(k), F_{\ell(k)})$, *for all positive integers* k.

Proof. If $\mathcal{A}_k \neq \varnothing$ then pick $n \in \mathcal{A}_k$, so that $\gcd(n, F_n) = k$. In turn, by Lemma 3.2.1(vi), this implies that $\ell(k) \mid n$, and consequently, again by Lemma 3.2.1(vi),

$$k \mid \gcd(\ell(k), F_{\ell(k)}) \mid \gcd(n, F_n) = k,$$

so that $k = \gcd(\ell(k), F_{\ell(k)})$.

On the other hand, if $k = \gcd(\ell(k), F_{\ell(k)})$ then, obviously, $\ell(k) \in \mathcal{A}_k$, so that $\mathcal{A}_k$ is not empty. □

If $\mathcal{S}$ is a set of positive integers, we define its *set of nonmultiples* as

$$\mathscr{M}(\mathcal{S}) := \{n \geq 1 : s \nmid n \text{ for all } s \in \mathcal{S}\}.$$

Sets of nonmultiples, or more precisely their complement *sets of multiples*

$$\mathcal{M}(\mathcal{S}) := \{n \geq 1 : s \mid n \text{ for some } s \in \mathcal{S}\},$$

have been studied by several authors, we refer the reader to [23] for a systematic treatment of this topic. We shall need only the following result.

Lemma 3.2.3. *If $\mathcal{S}$ is a set of positive integers such that*

$$\sum_{s\in\mathcal{S}} \frac{1}{s} < +\infty,$$

then $\mathscr{M}(\mathcal{S})$ has a natural density. Moreover, if $1 \notin \mathcal{S}$ then $\mathbf{d}(\mathscr{M}(\mathcal{S})) > 0$.

Proof. The part about the existence of $\mathbf{d}(\mathscr{M}(\mathcal{S}))$ is due to Erdős [15], while the second assertion follows easily from the inequality

$$\mathbf{d}(\mathscr{M}(\mathcal{S})) \geq \prod_{s\in\mathcal{S}} \left(1 - \frac{1}{s}\right)$$

proved by Heilbronn [26] and Rohrbach [43]. □

For any $\gamma > 0$, let us define

$$\mathcal{Q}_\gamma := \{p : z(p) \leq p^\gamma\}.$$

The following is a well-known lemma, which belongs to the folklore.

Lemma 3.2.4. *For all $x, \gamma > 0$, we have $\#\mathcal{Q}_\gamma(x) \ll x^{2\gamma}$.*

Proof. It follows from the definition of $\mathcal{Q}_\gamma(x)$ that every $p \in \mathcal{Q}_\gamma(x)$ divides $\prod_{n\leq x^\gamma} F_n$. Therefore,

$$2^{\#\mathcal{Q}_\gamma(x)} \leq \prod_{p\in\mathcal{Q}_\gamma(x)} p \leq \prod_{n\leq x^\gamma} F_n \leq 2^{\sum_{n\leq x^\gamma} n} = 2^{O(x^{2\gamma})},$$

where we employed the inequality $F_n \leq 2^n$, valid for all positive integers n. □

3.3 Proof of Theorem 3.1.1

We begin by showing that $\mathcal{A}_k$ is a scaled set of nonmultiples.

Lemma 3.3.1. *For each positive integer k such that $\mathcal{A}_k \neq \varnothing$, we have*

$$\mathcal{A}_k = \{\ell(k)m : m \in \mathscr{M}(\mathcal{L}_k)\},$$

where

$$\mathcal{L}_k := \{p : p \mid k\} \cup \{\ell(kp)/\ell(k) : p \nmid k\}.$$

Proof. Thanks to Lemma 3.2.1(vi), we know that $n \in \mathcal{A}_k$ implies $\ell(k) \mid n$, hence it is enough to prove that $\ell(k)m \in \mathcal{A}_k$, for some positive integer m, if and only if $m \in \mathscr{M}(\mathcal{L}_k)$.

Clearly, $\ell(k)m \in \mathcal{A}_k$ for some positive integer m, if and only if

$$\nu_p(\gcd(\ell(k)m, F_{\ell(k)m})) = \nu_p(k) \tag{3.2}$$

for all prime numbers p.

Let p be a prime number dividing k. Then, for all positive integer m, by Lemma 3.2.1(iii), we have $z(p) \mid z(k)$ and, in turn, $z(p) \mid \ell(k)m$. Consequently, by Lemma 3.2.1(v), $\nu_p(F_{\ell(k)m}) \geq \nu_p(\ell(k)m)$, so that

$$\nu_p(\gcd(\ell(k)m, F_{\ell(k)m})) = \nu_p(\ell(k)m) = \nu_p(\ell(k)) + \nu_p(m). \tag{3.3}$$

In particular, recalling that $k = \gcd(\ell(k), F_{\ell(k)})$ since $\mathcal{A}_k \neq \varnothing$ and thanks to Lemma 3.2.2, for $m = 1$ we get

$$\nu_p(k) = \nu_p(\gcd(\ell(k), F_{\ell(k)})) = \nu_p(\ell(k)),$$

which together with (3.3) gives

$$\nu_p(\gcd(\ell(k)m, F_{\ell(k)m})) = \nu_p(k) + \nu_p(m). \tag{3.4}$$

Therefore, (3.2) holds if and only if $p \nmid m$.

Now let p be a prime number not dividing k. Then (3.2) holds if and only if

$$p \nmid \gcd(\ell(k)m, F_{\ell(k)m}),$$

that is, by Lemma 3.2.1(vi), $\ell(p) \nmid \ell(k)m$, which in turn is equivalent to

$$\frac{\ell(kp)}{\ell(k)} = \frac{\operatorname{lcm}(\ell(k), \ell(p))}{\ell(k)} \nmid m,$$

since p and k are relatively prime.

Summarizing, we have found that $\ell(k)m \in \mathcal{A}_k$, for some positive integer m, if and only if $p \nmid m$ for all prime numbers p dividing k, and $\ell(kq)/\ell(k) \nmid m$ for all prime numbers q not dividing k, that is, $m \in \mathscr{M}(\mathcal{L}_k)$. □

We recall that a positive integer n is said to be *y-smooth* if all its prime factors are not exceeding y. For $x, y \geq 0$, let as usual $\Psi(x, y)$ be the number of y-smooth numbers not exceeding x. We need the following estimate for $\Psi(x, y)$.

Theorem 3.3.2. *We have $\Psi(x, y) \ll x^{1-1/(2\log y)}$, for $x \geq y \geq 2$.*

Proof. See [58, Ch. III.5, Theorem 1]. □

Now we show that the series of the reciprocals of the $\ell(n)$'s converges. More precisely, we give a bound for the tail of such series. The methods employed are somehow similar to those used to prove the result of [32].

Lemma 3.3.3. *There exists a constant $C > 0$ such that*

$$\sum_{n>x} \frac{1}{\ell(n)} \ll \exp\Big(-C\sqrt{\log x}\Big),$$

for all $x \geq 1$.

Proof. Let $n > 1$ be an integer and let $p := P(n)$ be the greatest prime factor of n. We have that $\operatorname{lcm}(n, z(p))$ is divisible by both p and $z(p)$, thus it is divisible by $\ell(p) = \operatorname{lcm}(p, z(p))$. Hence, we can write $\operatorname{lcm}(n, z(p)) = \ell(p)m$, where m is a positive integer such that $P(m) \leq p + 1$. Also, if p and $\operatorname{lcm}(n, z(p))$ are known then n can be chosen in at most $\tau(z(p))$ ways. Therefore, for all $y \geq 1$, we have

$$\sum_{P(n)>y} \frac{1}{\ell(n)} \leq \sum_{P(n)>y} \frac{1}{\operatorname{lcm}(n, z(P(n)))} \ll \sum_{p>y} \frac{\tau(z(p))}{pz(p)} \sum_{P(m)\leq p+1} \frac{1}{m},$$

where we also used the fact that $\ell(p) \gg pz(p)$ for each prime number p. By Mertens' formula [58, Chapter I.1, Theorem 11], we have

$$\sum_{P(m)\leq p+1} \frac{1}{m} \leq \prod_{q\leq p+1} \left(1 - \frac{1}{q}\right)^{-1} \ll \log p,$$

for all prime numbers p. Put $\beta := 3/4$ and $\gamma := 1/3$. It is well known [58, Chapter I.5, Corollary 1.1] that $\tau(n) \ll_\varepsilon n^\varepsilon$ for any fixed $\varepsilon > 0$. Hence, $\tau(z(p)) \log p \ll p^{1-\beta}$ for all prime numbers p. Thus we have found that

$$\sum_{P(n)>y} \frac{1}{\ell(n)} \ll \sum_{p>y} \frac{\tau(z(p)) \log p}{pz(p)} \ll \sum_{p>y} \frac{1}{p^\beta z(p)}. \tag{3.5}$$

On the one hand, by partial summation and by Lemma 3.2.4, we have

$$\sum_{\substack{p\in\mathcal{Q}_\gamma\\ p>y}}\frac{1}{p^\beta z(p)}\leq\sum_{\substack{p\in\mathcal{Q}_\gamma\\ p>y}}\frac{1}{p^\beta}=\frac{\#\mathcal{Q}_\gamma(t)}{t^\beta}\bigg|_{t=y}^{+\infty}+\beta\int_y^{+\infty}\frac{\#\mathcal{Q}_\gamma(t)}{t^{\beta+1}}\,\mathrm{d}t\ll\frac{1}{y^\beta},\tag{3.6}$$

since $\beta>2\gamma$. On the other hand, by the definition of $\mathcal{Q}_\gamma$, we have

$$\sum_{\substack{p\notin\mathcal{Q}_\gamma\\ p>y}}\frac{1}{p^\beta z(p)}<\sum_{p>y}\frac{1}{p^{\beta+\gamma}}\ll\frac{1}{y^{\beta+\gamma-1}},\tag{3.7}$$

since $\beta+\gamma>1$. Hence, putting together (3.5), (3.6), and (3.7), we get that

$$\sum_{P(n)>y}\frac{1}{\ell(n)}\ll\frac{1}{y^{1/12}}.\tag{3.8}$$

By Theorem 3.3.2 and by partial summation, we have

$$\begin{aligned}\sum_{\substack{P(n)\leq y\\ n>x}}\frac{1}{\ell(n)}\leq\sum_{\substack{P(n)\leq y\\ n>x}}\frac{1}{n}&=\frac{\Psi(t,y)}{t}\bigg|_{t=x}^{+\infty}+\int_x^{+\infty}\frac{\Psi(t,y)}{t^2}\,\mathrm{d}t\\&\ll\int_x^{+\infty}\frac{\mathrm{d}t}{t^{1+1/(2\log y)}}\ll\frac{\log y}{x^{1/(2\log y)}},\end{aligned}\tag{3.9}$$

for all $x,y\geq 2$. At this point, setting $y=\exp\bigl(\sqrt{6\log x}\bigr)$ and putting together (3.8) and (3.9), we get the desired claim. □

Now we are ready for the proof of Theorem 3.1.1. If k is a positive integer such that $\mathcal{A}_k=\varnothing$ then, obviously, the natural density of $\mathcal{A}_k$ exists and is equal to zero. So we can assume $\mathcal{A}_k\neq\varnothing$, which in turn, by Lemma 3.2.2, implies that $k=\gcd(\ell(k),F_{\ell(k)})$.

By Lemma 3.3.3, we have

$$\sum_{n\in\mathcal{L}_k}\frac{1}{n}\ll\sum_p\frac{1}{\ell(kp)}\leq\sum_p\frac{1}{\ell(p)}<+\infty,$$

while clearly $1\notin\mathcal{L}_k$. Hence, Lemma 3.2.3 tell us that $\mathscr{M}(\mathcal{L}_k)$ has a positive natural density. Finally, by Lemma 3.3.1 we conclude that the natural density of $\mathcal{A}_k$ exists and it is positive. The proof of Theorem 3.1.1 is complete.

3.4 Proof of Theorem 3.1.2

We begin by introducing a family of sets. For each positive integer k, let $\mathcal{B}_k$ be the set of positive integers n such that:

(i) $k \mid \gcd(n, F_n)$;

(ii) if $p \mid \gcd(n, F_n)$ for some prime number p, then $p \mid k$.

The essential part of the proof of Theorem 3.1.2 is the following formula for the natural density of $\mathcal{B}_k$.

Lemma 3.4.1. *For all positive integers k, the natural density of $\mathcal{B}_k$ exists and*

$$\mathbf{d}(\mathcal{B}_k) = \sum_{(d,k)=1} \frac{\mu(d)}{\ell(dk)}, \tag{3.10}$$

where the series is absolutely convergent.

Proof. For all positive integers n and d, let us define

$$\varrho(n,d) := \begin{cases} 1 & \text{if } d \mid F_n, \\ 0 & \text{if } d \nmid F_n. \end{cases}$$

Note that ϱ is multiplicative in its second argument, that is,

$$\varrho(n, de) = \varrho(n,d)\varrho(n,e)$$

for all relatively prime positive integers d and e, and all positive integers n.

Using Lemma 3.2.1(vi), it is easy to see that $n \in \mathcal{B}_k$ if and only if $\ell(k) \mid n$ and $\varrho(n,p) = 0$ for all prime numbers p dividing n but not dividing k. Therefore,

$$\begin{aligned} \#\mathcal{B}_k(x) &= \sum_{\substack{n \le x \\ \ell(k) \mid n}} \prod_{\substack{p \mid n \\ p \nmid k}} (1 - \varrho(n,p)) = \sum_{\substack{n \le x \\ \ell(k) \mid n}} \sum_{\substack{d \mid n \\ (d,k)=1}} \mu(d)\varrho(n,d) \\ &= \sum_{\substack{d \le x \\ (d,k)=1}} \mu(d) \sum_{\substack{m \le x/d \\ \ell(k) \mid dm}} \varrho(dm, d), \end{aligned} \tag{3.11}$$

for all $x > 0$. Moreover, given a positive integer d which is relatively prime with k, we have that $\varrho(dm, d) = 1$ and $\ell(k) \mid dm$ if and only if $\operatorname{lcm}(z(d), \ell(k)) \mid dm$, which in turn is equivalent to m being divisible by

$$\frac{\operatorname{lcm}(d, \operatorname{lcm}(z(d), \ell(k)))}{d} = \frac{\operatorname{lcm}(\ell(d), \ell(k))}{d} = \frac{\ell(dk)}{d},$$

since d and k are relatively prime. Hence,

$$\sum_{\substack{m \le x/d \\ \ell(k) \,|\, dm}} \varrho(dm, d) = \sum_{\substack{m \le x/d \\ \ell(dk)/d \,|\, m}} 1 = \left\lfloor \frac{x}{\ell(dk)} \right\rfloor,$$

for all $x > 0$, which together with (3.11) implies that

$$\#\mathcal{B}_k(x) = \sum_{\substack{d \le x \\ (d,k)=1}} \mu(d) \left\lfloor \frac{x}{\ell(dk)} \right\rfloor = x \sum_{\substack{d \le x \\ (d,k)=1}} \frac{\mu(d)}{\ell(dk)} - R(x), \tag{3.12}$$

for all $x > 0$, where

$$R(x) := \sum_{\substack{d \le x \\ (d,k)=1}} \mu(d) \left\{ \frac{x}{\ell(dk)} \right\}.$$

Now, thanks to Lemma 3.3.3, we have

$$\sum_{(d,k)=1} \frac{|\mu(d)|}{\ell(dk)} \le \sum_{d=1}^{\infty} \frac{1}{\ell(d)} < +\infty,$$

hence the series in (3.10) is absolutely convergent.

It remains only to prove that $R(x) = o(x)$ as $x \to +\infty$, and then the desired result follows from (3.12). Actually, we shall prove that $R(x) \ll x \exp\bigl(-D\sqrt{\log x}\bigr)$, where D is a positive constant.

Let $y < x$ be a positive parameter that we will choose later. We bound with 1 the terms appearing in $R(x)$ with $d \le y$, otherwise we estimate them removing the fractional part. In this way we get

$$|R(x)| \le \sum_{d \le x} \left\{ \frac{x}{\ell(dk)} \right\} \le y + \sum_{y < d \le x} \frac{x}{\ell(dk)}.$$

Since $\ell(dk) \ge \ell(d)$, we obtain

$$|R(x)| \le y + x \sum_{d > y} \frac{1}{\ell(d)} \ll y + x \exp\Bigl(-C\sqrt{\log y}\Bigr),$$

by Lemma 3.3.3. Setting $y = x \exp\bigl(-C\sqrt{\log x}\bigr)$ we get the claim. □

At this point, by the definition of $\mathcal{B}_k$ and by the inclusion-exclusion principle, it follows easily that

$$\#\mathcal{A}_k(x) = \sum_{d \mid k} \mu(d)\, \#\mathcal{B}_{dk}(x),$$

for all $x > 0$. Hence, by Lemma 3.4.1, we get

$$\begin{aligned} \mathbf{d}(\mathcal{A}_k) &= \sum_{d \mid k} \mu(d)\, \mathbf{d}(\mathcal{B}_{dk}) = \sum_{d \mid k} \mu(d) \sum_{(e,dk)=1} \frac{\mu(e)}{\ell(dek)} \\ &= \sum_{d \mid k} \sum_{(e,k)=1} \frac{\mu(de)}{\ell(dek)} = \sum_{f=1}^{\infty} \frac{\mu(f)}{\ell(fk)}, \end{aligned} \tag{3.13}$$

since every squarefree positive integer f can be written in a unique way as $f = de$, where d and e are squarefree positive integers such that $d \mid k$ and $\gcd(e, k) = 1$. Also note that the rearrangement of the series in (3.13) is justified by absolute convergence. The proof of Theorem 3.1.2 is complete.

3.5 Final remarks

In order to simplify the exposition, we chose to give the results of this chapter for the sequence of Fibonacci numbers. However, they can be generalized to every nondegenerate Lucas sequence $(u_n)_{n\geq 0}$ satisfying $u_n = a_1 u_{n-1} + a_2 u_{n-2}$, for all integers $n \geq 2$, where a_1 and a_2 are relatively prime integers. There is just a minor complication that must be handled: The rank of appearance $z_u(m)$ of a positive integer m in the Lucas sequence $(u_n)_{n\geq 0}$, that is, the smallest positive integer n such that m divides u_n, exists if and only if m is relatively prime with a_2. Therefore, the arguments involving $z(m)$ must be adapted to $z_u(m)$ considering only the positive integers m which are relatively prime with a_2. Except for that, everything works the same, since $z_u(m)$ and $\ell_u(m) := \operatorname{lcm}(m, z_u(m))$ satisfy the same properties of $z(m)$ and $\ell(m)$. Note only that Lemma 3.2.1(iv) should be replaced by: $z_u(p) \mid p - (-1)^{p-1} \left(\frac{\Delta_u}{p}\right)$ for all prime numbers p not dividing a_2, where $\Delta_u := a_1^2 + 4a_2$ is the discriminant of the characteristic polynomial $X^2 - a_1 X - a_2$. Also, the analog of Lemma 3.2.1(v), that is, $\nu_p(u_n) \geq \nu_p(n)$ whenever $z_u(p) \mid n$, can be proved

by using the formula for the p-adic valuations of terms of Lucas sequence given in [45]. With these changes, the following generalization can be proved.

Theorem 3.5.1. *Let $(u_n)_{n\geq 0}$ be a nondegenerate Lucas sequence satisfying the recurrence $u_n = a_1 u_{n-1} + a_2 u_{n-2}$ for all integers $n \geq 2$, where a_1 and a_2 are relatively prime integers, and define the set*

$$\mathcal{A}_{u,k} := \{n \geq 1 : \gcd(n, u_n) = k\},$$

for each positive integer k. Then $\mathcal{A}_{u,k}$ is not empty if and only if $\gcd(k, a_2) = 1$ and $k = \gcd(\ell_u(k), u_{\ell_u(k)})$. In such a case, $\mathcal{A}_{u,k}$ has a natural density which is given by

$$\mathbf{d}(\mathcal{A}_{u,k}) = \sum_{(d,a_2)=1} \frac{\mu(d)}{\ell_u(dk)},$$

where the series is absolutely convergent.

Finally, we point out that all the results of this chapter (in particular, the estimate of the error term $R(x)$ in the proof of Lemma 3.4.1), as well as their extension to Lucas sequences, are effective. Hence, in principle, all the implied constants could be computed, and an algorithm to compute $\mathbf{d}(\mathcal{A}_k)$ with arbitrary precision could be implemented. However, doing so would be extremely laborious, and probably the convergence would be very slow.

k	$\#\mathcal{A}_k(10^4)/10^4$	$\#\mathcal{A}_k(10^5)/10^5$	$\#\mathcal{A}_k(10^6)/10^6$
1	0.6418	0.64190	0.641878
2	0.0625	0.06248	0.062499
5	0.1231	0.12280	0.122809
7	0.0072	0.00710	0.007081
10	0.0109	0.01077	0.010766
12	0.0217	0.02153	0.021527
13	0.0060	0.00590	0.005911

Table 3.1: The ratios $\#\mathcal{A}_k(x)/x$ for $x = 10^4, 10^5, 10^6$ and for the first values of k such that $\mathcal{A}_k$ is nonempty.

Chapter 4

On the G.C.D. of n and the nth Fibonacci number

Abstract. *Let $\mathcal{F}$ be the set of all integers of the form $\gcd(n, F_n)$, where n is a positive integer and F_n denotes the nth Fibonacci number. I prove that*

$$\#\mathcal{F}(x) \gg \frac{x}{\log x}$$

for all $x \geq 2$, and that $\mathcal{F}$ has zero natural density. The proofs rely on a recent result of Cubre and Rouse which gives, for each positive integer m, an explicit formula for the density of the set of prime numbers p such that m divides the rank of appearance of p, that is, the smallest positive integer k such that p divides F_k. This is a work in collaboration with Paolo Leonetti and appeared in [35].

4.1 Introduction

In the previous chapter, we have seen that for each positive integer k the set

$$\mathcal{A}_k := \{n \in \mathbb{N} : \gcd(n, F_n) = k\}$$

has a positive natural density or it is empty. Now we are interested in understanding "how often" the first case occurs. For, define the set

$$\mathcal{F} := \{k \in \mathbb{N} : \mathcal{A}_k \neq \varnothing\} = \{\gcd(n, F_n) : n \in \mathbb{N}\},$$

The aim of this chapter is to study the structural properties and the distribution of the elements of $\mathcal{F}$. The first result is a lower bound for the counting function of $\mathcal{F}$.

Theorem 4.1.1. *We have*

$$\#\mathcal{F}(x) \gg \frac{x}{\log x},$$

for all $x \geq 2$.

The second result is that $\mathcal{F}$ has zero natural density.

Theorem 4.1.2. *$\mathcal{F}$ has zero natural density.*

4.2 Preliminaries

As in Chapter 3, for each positive integer n, let $z(n)$ be rank of appearance of n, and put $\ell(n) := \operatorname{lcm}(n, z(n))$. Also, we will use the statements of Lemma 3.2.1 implicitly and without further mention.

The next lemma tells us some important information on $\mathcal{F}$.

Lemma 4.2.1. *For all positive integers n and all prime numbers p, we have:*

(i) $n \in \mathcal{F}$ *if and only if* $n = \gcd(\ell(n), F_{\ell(n)})$.

(ii) $p \in \mathcal{F}$ *if* $p \neq 3$ *and* $\ell(q) \nmid z(p)$ *for all prime numbers* q.

(iii) $p \mid n$ *whenever* $\ell(p) \mid \ell(n)$ *and* $n \in \mathcal{F}$.

Proof. Fact (i) is just Lemma 3.2.2. The claim (ii) is easily seen to hold for $p = 2$. Let us suppose that $p > 3$ is a prime number such that $\ell(q) \nmid z(p)$ for all prime numbers q. In particular, $p \neq 5$ since $\ell(5) = z(5) = 5$. Hence, let us suppose hereafter that $p \geq 7$. Since $z(p) \mid p \pm 1$, it easily follows that $p \,||\, \gcd(\ell(p), F_{\ell(p)})$. At this point, if $q \mid \gcd(\ell(p), F_{\ell(p)})$ for some prime $q \neq p$, then $\ell(q) \mid \ell(p) = pz(p)$. But $\ell(q) \nmid z(p)$, hence $p \mid \ell(q) = \operatorname{lcm}(q, z(q))$ so that $p \mid z(q) \leq q + 1$. Similarly, $q \mid \gcd(\ell(p), F_{\ell(p)}) \mid \ell(p)$ implies $q \mid z(p) \leq p + 1$. Hence $|p - q| \leq 1$, which is impossible since $p \geq 7$. Therefore $q \nmid \gcd(\ell(p), F_{\ell(p)})$, with the consequence that $p = \gcd(\ell(p), F_{\ell(p)})$, i.e., $p \in \mathcal{F}$ by (i). This concludes the proof of (ii). Finally, if $\ell(p) \mid \ell(n)$ and $n \in \mathcal{F}$, then

$$p \mid \gcd(\ell(p), F_{\ell(p)}) \mid \gcd(\ell(n), F_{\ell(n)}) = n,$$

as claimed, and also (iii) is proved. □

It follows from a result of Lagarias [30, 31], that the set of prime numbers p such that $z(p)$ is even has a relative density of $2/3$ in the set of all prime numbers. Bruckman and Anderson [8, Conjecture 3.1] conjectured, for each positive integer m, a formula for the limit

$$Z(m) := \lim_{x \to +\infty} \frac{\#\{p \leq x : m \mid z(p)\}}{x/\log x}.$$

Their conjecture was proved by Cubre and Rouse [13, Theorem 2], who obtained the following result.

Theorem 4.2.2. *For any positive integer m, we have*

$$Z(m) = \rho(m) \prod_{q^e || m} \frac{q^{2-e}}{q^2 - 1},$$

where q^e runs over the prime powers in the factorization of m, while

$$\rho(m) := \begin{cases} 1 & \text{if } 10 \nmid m, \\ 5/4 & \text{if } m \equiv 10 \bmod 20, \\ 1/2 & \text{if } 20 \mid m. \end{cases}$$

Note that the arithmetic function Z is not multiplicative since, for example, we have

$$Z(2 \cdot 5) = \frac{5}{4} \cdot \frac{2}{3} \cdot \frac{5}{24} = \frac{25}{144} \neq \frac{5}{36} = \frac{2}{3} \cdot \frac{5}{24} = Z(2)Z(5).$$

However, it is easy to check that the restriction of Z to the odd positive integers is multiplicative. This fact will be useful later.

We also need the following technical lemma.

Lemma 4.2.3. *We have*

$$\sum_{q>y} \frac{1}{\varphi(\ell(q))} \ll \frac{\log\log y}{y^{1/3}},$$

for all $y > 3$.

Proof. For $\gamma > 0$, put $\mathcal{Q}_\gamma := \{p : z(p) < p^\gamma\}$. Thanks to Lemma 3.2.4, we know that $\mathcal{Q}_\gamma(x) \ll x^{2\gamma}$. Set $\gamma = 1/3$. Since

$$\varphi(n) \gg \frac{n}{\log\log n}$$

for all positive integers n [58, Ch. I.5, Theorem 4], while, $\ell(q) \ll q^2$ for all prime numbers q, we have

$$\sum_{q>y} \frac{1}{\varphi(\ell(q))} \ll \sum_{q>y} \frac{\log\log \ell(q)}{\ell(q)} \ll \sum_{q>y} \frac{\log\log q}{\ell(q)}, \tag{4.1}$$

for all $y > 3$.

On the one hand,

$$\sum_{\substack{q>y \\ q\notin\mathcal{Q}_\gamma}} \frac{\log\log q}{\ell(q)} \ll \sum_{\substack{q>y \\ q\notin\mathcal{Q}_\gamma}} \frac{\log\log q}{qz(q)} \leq \sum_{q>y} \frac{\log\log q}{q^{1+\gamma}} \ll \int_y^{+\infty} \frac{\log\log t}{t^{1+\gamma}}\,\mathrm{d}t \ll \frac{\log\log y}{y^\gamma}. \tag{4.2}$$

On the other hand, by partial summation,

$$\begin{aligned}\sum_{\substack{q>y \\ q\in\mathcal{Q}_\gamma}} \frac{\log\log q}{\ell(q)} \leq \sum_{\substack{q>y \\ q\in\mathcal{Q}_\gamma}} \frac{\log\log q}{q} &= \frac{\#\mathcal{Q}_\gamma(t)\log\log t}{t}\bigg|_{t=y}^{+\infty} + \int_y^{+\infty} \frac{\log\log t - \frac{1}{\log t}}{t^2}\#\mathcal{Q}_\gamma(t)\,\mathrm{d}t \\ &\leq \int_y^{+\infty} \frac{\log\log t}{t^{2-2\gamma}}\,\mathrm{d}t \ll \frac{\log\log y}{y^{1-2\gamma}}.\end{aligned} \tag{4.3}$$

The claim follows by putting together (4.1), (4.2), and (4.3), and by recalling that $\gamma = 1/3$. □

Lastly, for all relatively prime integers a and m, define

$$\pi(x, m, a) := \#\{p \leq x : p \equiv a \bmod m\}.$$

We need the following version of the Brun–Titchmarsh theorem [37, Theorem 2].

Theorem 4.2.4. *If a and m are relatively prime integers and $m > 0$, then*

$$\pi(x, m, a) < \frac{2x}{\varphi(m) \log(x/m)},$$

for all $x > m$.

4.3 Proof of Theorem 4.1.1

First, since $1 \in \mathcal{F}$, it is enough to prove the claim only for all sufficiently large x. Let $y > 5$ be a real number to be chosen later. Define the following sets of primes:

$$\begin{aligned}
\mathcal{P}_1 &:= \{p : q \nmid z(p),\ \forall q \in [3, y]\},\\
\mathcal{P}_2 &:= \{p : \exists q > y,\ \ell(q) \mid z(p)\},\\
\mathcal{P} &:= \mathcal{P}_1 \setminus \mathcal{P}_2.
\end{aligned}$$

We have $\mathcal{P} \subseteq \mathcal{F} \cup \{3\}$. Indeed, since $3 \mid \ell(2)$ and $q \mid \ell(q)$ for each prime number q, it follows easily that if $p \in \mathcal{P}$ then $\ell(q) \nmid z(p)$ for all prime numbers q, which, by Lemma 4.2.1(ii), implies that $p \in \mathcal{F}$ or $p = 3$.

Now we give a lower bound for $\#\mathcal{P}_1(x)$. Let P_y be the product of all prime numbers in $[3, y]$. By using the inclusion-exclusion principle and Theorem 4.2.2, we get that

$$\begin{aligned}
\lim_{x \to +\infty} \frac{\#\mathcal{P}_1(x)}{x/\log x} &= \lim_{x \to +\infty} \sum_{m \mid P_y} \mu(m) \cdot \frac{\#\{p \leq x : m \mid z(p)\}}{x/\log x}\\
&= \sum_{m \mid P_y} \mu(m) Z(m) = \prod_{3 \leq q \leq y} (1 - Z(q))\\
&= \prod_{3 \leq q \leq y} \left(1 - \frac{q}{q^2 - 1}\right),
\end{aligned}$$

where we also made use of the fact that the restriction of Z to the odd positive integers is multiplicative.

As a consequence, for all sufficiently large x depending only on y, let say $x \geq x_0(y)$, we have

$$\#\mathcal{P}_1(x) \geq \frac{1}{2} \prod_{3 \leq q \leq y} \left(1 - \frac{q}{q^2-1}\right) \cdot \frac{x}{\log x} \gg \frac{1}{\log y} \cdot \frac{x}{\log x},$$

where the last inequality follows from Mertens' third theorem [58, Ch. I.1, Theorem 11].

We also need an upper bound for $\#\mathcal{P}_2(x)$. Since $z(p) \mid p \pm 1$ for all primes $p > 5$, we have

$$\#\mathcal{P}_2(x) \leq \sum_{q>y} \#\{p \leq x : \ell(q) \mid z(p)\} \leq \sum_{q>y} \pi(x, \ell(q), \pm 1), \tag{4.4}$$

for all $x > 0$, where, for the sake of brevity, we put

$$\pi(x, \ell(q), \pm 1) := \pi(x, \ell(q), -1) + \pi(x, \ell(q), 1).$$

On the one hand, by Theorem 4.2.4 and Lemma 4.2.3, we have

$$\sum_{y<q<x^{1/2}} \pi(x, \ell(q), \pm 1) \ll \sum_{q>y} \frac{1}{\varphi(\ell(q))} \cdot \frac{x}{\log x} \ll \frac{\log\log y}{y^{1/3}} \cdot \frac{x}{\log x}. \tag{4.5}$$

On the other hand, by the trivial estimate for $\pi(x, \ell(q), \pm 1)$ and Lemma 4.2.3, we get

$$\sum_{q>x^{1/2}} \pi(x, \ell(q), \pm 1) \ll \sum_{q>x^{1/2}} \frac{x}{\ell(q)} \leq \sum_{q>x^{1/2}} \frac{x}{\varphi(\ell(q))} \ll x^{2/3} \log\log x. \tag{4.6}$$

Therefore, putting together (4.4), (4.5), and (4.6), we find that

$$\#\mathcal{P}_2(x) \ll \frac{\log\log y}{y^{1/3}} \cdot \frac{x}{\log x} + x^{2/3} \log\log x.$$

In conclusion, there exist two absolute constants $c_1, c_2 > 0$ such that

$$\begin{aligned} \#\mathcal{F}(x) \gg \#\mathcal{P}(x) &\geq \#\mathcal{P}_1(x) - \#\mathcal{P}_2(x) \\ &\geq \left(\frac{c_1}{\log y} - \frac{c_2 \log\log y}{y^{1/3}} - \frac{c_2 \log x \log\log x}{x^{1/3}}\right) \cdot \frac{x}{\log x}, \end{aligned} \tag{4.7}$$

for all $x \geq x_0(y)$.

Finally, we can choose y to be sufficiently large so that

$$\frac{c_1}{\log y} - \frac{c_2 \log\log y}{y^{1/3}} > 0.$$

Hence, from (4.7) it follows that

$$\#\mathcal{F}(x) \gg \frac{x}{\log x},$$

for all sufficiently large x.

4.4 Proof of Theorem 4.1.2

Fix $\varepsilon > 0$ and pick a prime number q such that $1/q < \varepsilon$. Let $\mathcal{P}$ be the set of prime numbers p such that $\ell(q) \mid z(p)$. By Theorem 4.2.2, we know that $\mathcal{P}$ has a positive relative density in the set of primes. As a consequence, we can pick a sufficiently large $y > 0$ so that

$$\prod_{p \in \mathcal{P}(y)} \left(1 - \frac{1}{p}\right) < \varepsilon.$$

Let $\mathcal{B}$ be the set of positive integers without prime factors in $\mathcal{P}(y)$. We split $\mathcal{F}$ into two subsets: $\mathcal{F}_1 := \mathcal{F} \cap \mathcal{B}$ and $\mathcal{F}_2 := \mathcal{F} \setminus \mathcal{F}_1$. If $n \in \mathcal{F}_2$ then n has a prime factor p such that $\ell(q) \mid z(p)$. Hence, $\ell(q) \mid \ell(n)$ and, by Lemma 4.2.1(iii), we get that $q \mid n$, so all the elements of $\mathcal{F}_2$ are multiples of q. In conclusion,

$$\begin{aligned} \limsup_{x \to +\infty} \frac{\#\mathcal{F}(x)}{x} &\leq \limsup_{x \to +\infty} \frac{\#\mathcal{F}_1(x)}{x} + \limsup_{x \to +\infty} \frac{\#\mathcal{F}_2(x)}{x} \\ &\leq \prod_{p \in \mathcal{P}(y)} \left(1 - \frac{1}{p}\right) + \frac{1}{q} < 2\varepsilon, \end{aligned}$$

and, by the arbitrariness of ε, it follows that $\mathcal{F}$ has zero natural density.

Remark 4.4.1. I have not been able to obtain an upper bound for the counting function of $\mathcal{F}$ better than $\#\mathcal{F}(x) = o(x)$. Probably, in order to do so, it is needed an asymptotic formula for $\#\{p \leq x : m \mid z(p)\}$ holding uniformly for a large range of values of x and m, instead of the asymptotic formula of Theorem 4.2.2, which holds for fixed m.

Part II

Other results

Chapter 5

A coprimality condition on consecutive values of polynomials

Abstract. *Let $f \in \mathbb{Z}[X]$ be a quadratic or cubic polynomial. I prove that there exists an integer $G_f \geq 2$ such that for every integer $k \geq G_f$ one can find infinitely many integers $n \geq 0$ with the property that none of*

$$f(n+1), f(n+2), \ldots, f(n+k)$$

is coprime to all the others. This extends a previous result of Evans on linear polynomials and settles a conjecture of Harrington and Jones. This is a work in collaboration with Márton Szikszai and appeared in [50].

5.1 Introduction

A sequence of integers $s = s(n)_{n\geq 0}$ is said to have the *Pillai property* if there exists an integer $G \geq 2$ such that for all integers $k \geq G$ there exist infinitely many integers $n \geq 0$ such that none of the integers $s(n+1), s(n+2), \ldots, s(n+k)$ is relatively prime with all the others. In such a case, G_s is defined as the minimal possible G. Also, $g_s \geq 2$ is defined as the smallest integer, if it exists, such that one can find g_s consecutive terms of s with the property that none of them is coprime to all the others. For instance, the sequence of positive even integers has $g_s = G_s = 2$, while for the sequence of prime numbers neither exists. Note that the existence of G_s implies that of g_s and one has $g_s \leq G_s$. For less trivial examples see the paper of Hajdu and Szikszai [22].

Erdős [14] was the first to prove that the sequence of natural numbers has the Pillai property. Later, the combined efforts of Pillai [39] and Brauer [6] gave a more explicit result, namely that $g_s = G_s = 17$. We note that interest in such a problem is twofold. On one hand, Pillai aimed to solve a classical Diophantine problem by showing that the product of consecutive integers can never be a perfect power. While a complete solution was given by Erdős and Selfridge [16], Pillai [40] himself proved, using his already mentioned result from [39], that it cannot be so if one takes at most 16 consecutive terms. On the other hand, Brauer [6] made connection with his earlier paper [7] on an old problem, studied already by Legendre [33], concerning prime gaps. In fact, Erdős [14] himself also studied prime gaps.

Gradually, the study of g_s and G_s in various sequences, and their importance in analogous problems as the ones mentioned earlier, attracted an increased attention. Evans [17] proved that any arithmetic progression has the Pillai property. Ohtomo and Tamari [38] derived the same, but they also dealt with numerical aspects by showing that $G_s \leq 384$ for the sequence of odd integers. The most recent progress is due to Hajdu and Saradha [21] who gave an effective upper bound on G_s depending only on the difference of the progression together with a heuristic algorithm to find the exact value of it, whenever the number of prime factors of the difference is “small”.

Observe that both the natural numbers and arithmetic progressions can be considered as consecutive values of linear polynomials. Recently, Harrington and Jones [25] studied quadratic sequences, that is, for some quadratic $f \in \mathbb{Z}[X]$ one has $s(n) = f(n)$ for every $n \geq 1$. They computed the exact value of g_s when f is monic or when it belongs to some special family of nonmonic polynomials. Further, they conjectured that g_s exists and that $g_s \leq 35$ for every quadratic polynomial. However, they did not consider G_s to any extent.

In this chapter, we considerably extend the previous results. Before stating our result we note that throughout the chapter we use the notation $g_f = g_s$ and $G_f = G_s$ and write about consecutive values of the polynomial f instead of consecutive terms of the corresponding sequence s. The main theorem is as follows.

Theorem 5.1.1. *If $f \in \mathbb{Z}[X]$ is a quadratic or cubic polynomial, then $f(n)_{n\geq 0}$ has the Pillai property.*

Observe that Theorem 5.1.1 allows us to immediately settle one part of the conjecture made by Harrington and Jones [25] on g_f.

Corollary 5.1.1. *If $f \in \mathbb{Z}[X]$ is quadratic, then g_f exists.*

We do not consider the absolute boundedness of g_f, but we make some remarks on it instead. For every integer $k \geq 2$, there exists a quadratic polynomial $f \in \mathbb{Z}[X]$, reducible in $\mathbb{Z}[X]$, and such that $k \leq g_f \leq G_f$. This follows easily by taking d to be the product of the first k prime numbers and then looking at the polynomial $f(X) = (1 + dX)^2$. On one hand, we have $g_f = g_{1+dX}$ and $G_f = G_{1+dX}$, while on the other we have $k \leq g_{1+dX} \leq G_{1+dX}$. Neverthless, we could not say anything about the irreducible case and we feel that, despite not stating it anywhere and not excluding reducibles before, Harrington and Jones made their conjecture on this more interesting setting.

Let us conclude this section by discussing the main tools employed in the proof of Theorem 5.1.1. The basic idea is to construct for every quadratic or cubic polynomial f an auxiliary polynomial $\tilde{f}$ that, in some sense, controls the existence of "close" solutions to polynomial congruences $f(X) \equiv 0 \pmod{p}$. Then we show that if k is desirably

large, one has enough prime numbers with such close solutions to "cover" some block of k consecutive numbers $f(n+1), f(n+2), \ldots, f(n+k)$. The success of this construction relies on elementary properties of the roots of f modulo a prime number, results on the p-adic valuations of products of consecutive polynomial values, and lower bounds on the number of certain subsets of prime numbers.

Note that our methods can yield, at least in principle, an effective upper bound on G_f. However, the bound would be too large to be useful in practice. Further, we emphasize that Theorem 5.1.1 implies the existence of G_f for every quartic polynomial $f \in \mathbb{Z}[x]$ that is reducible in $\mathbb{Z}[X]$ (we always have a factor of degree at most 3), but our construction already fails to deal with quartic polynomials in general. We point out this more explicitly in the next section.

5.2 Preliminaries

This section is devoted to the auxiliary results we use in the proof of Theorem 5.1.1. Let

$$f(X) = a_k X^k + a_{k-1} X^{k-1} + \cdots + a_0,$$

be a polynomial of degree $k \geq 1$ with integer coefficients $a_0, \ldots, a_k$. We define

$$\widetilde{f}(X) := a_k^{2k-2} \prod_{\substack{1 \leq i,j \leq k \\ i \neq j}} (X - (\alpha_i - \alpha_j)), \tag{5.1}$$

where $\alpha_1, \ldots, \alpha_k$ are all the roots of f in some algebraic closure. Observe that $\widetilde{f}$ can be computed from the relation

$$\mathrm{Res}_X(f(X), f(X+Y)) = a_k^2 Y^k \widetilde{f}(Y),$$

where Res_X is the resultant of polynomials respect to X. In particular, for $k = 2$

$$\widetilde{f}(X) = a_2^2 X^2 - \Delta_f,$$

while for $k = 3$

$$\widetilde{f}(X) = \left(a_3^2 X^2 + 3a_1 a_3 - a_2^2\right)^2 X^2 - \Delta_f,$$

where Δ_f denotes the discriminant of f. We have the following simple, but useful property.

Lemma 5.2.1. *If $f \in \mathbb{Z}[X]$ is a nonconstant polynomial, then f and $\widetilde{f}$ have the same Galois group over $\mathbb{Q}$.*

Proof. The identity

$$\alpha_i = \frac{1}{k}\left(\sum_{j=1}^{k}(\alpha_i - \alpha_j) - \frac{a_{k-1}}{a_k}\right) \quad i = 1, \ldots, k,$$

implies that f and $\widetilde{f}$ have the same splitting field over $\mathbb{Q}$, and hence the same Galois group. □

The next result deals with another interesting connection between f and $\tilde{f}$, namely it relates $\widetilde{f}$ to "close" solutions of the congruence $f(X) \equiv 0 \pmod{p}$.

Lemma 5.2.2. *Let $f \in \mathbb{Z}[X]$ be of degree $k = 2$ or 3 and suppose that $p \mid \widetilde{f}(r)$ for some prime number $p \nmid 6a_k$ and some positive integer r. Then there exists an integer n such that*

$$f(n) \equiv f(n+r) \equiv 0 \pmod{p}.$$

Proof. Let $\alpha_1, \ldots, \alpha_k$ be the roots of f in the algebraic closure of the finite field $\mathbb{F}_p$. Since $p \mid \widetilde{f}(r)$, by (5.1) we can assume that $\alpha_1 - \alpha_2 = r$, where r is considered as an element of $\mathbb{F}_p$. If $k = 2$, then $\alpha_1 + \alpha_2 \in \mathbb{F}_p$ by Viète's formulas and $p \nmid a_k$, so that $\alpha_1, \alpha_2 \in \mathbb{F}_p$, since $p > 2$, and the claim follows. If $k = 3$, we distinguish two cases. If f has a root in $\mathbb{F}_p$, then it is either one of α_1, α_2, so that $\alpha_1, \alpha_2 \in \mathbb{F}_p$, or it is α_3, in which case α_1 and α_2 are the roots of a quadratic polynomial in $\mathbb{F}_p$, and proceeding as in the case $k = 2$ we get again $\alpha_1, \alpha_2 \in \mathbb{F}_q$. If f is irreducible in $\mathbb{F}_p$, then any Galois automorphism of f over $\mathbb{F}_p$ which sends α_1 to α_2 also sends α_2 to α_3. Therefore,

$$r = \alpha_1 - \alpha_2 = \alpha_2 - \alpha_3$$

and

$$3\alpha_2 = (\alpha_1 + \alpha_2 + \alpha_3) - (\alpha_1 - \alpha_2) + (\alpha_2 - \alpha_3),$$

which implies again $\alpha_1, \alpha_2 \in \mathbb{F}_p$, since $p > 3$. □

Remark 5.2.1. Note that the conclusion of Lemma 5.2.2 is no longer true if the hypothesis on the degree is dropped. Take for instance, $f(X) = X^4 + 1$. We have that $7 \mid \widetilde{f}(3)$, but the congruence $f(X) \equiv 0 \bmod 7$ has no solutions at all. We did not manage to find a simple and nice way to construct a family of irreducible polynomials of degree 4 such that Lemma 5.2.2 fails. However, we can give instead a family of reducible quartic polynomials as follows: Given an irreducible quadratic polynomial $g \in \mathbb{Z}[X]$ and a positive integer r, put $f(X) := g(X)g(X + r)$. Then, all prime numbers divide $\widetilde{f}(r) = 0$, but there are infinitely many prime numbers p such that $f(X) \equiv 0 \bmod p$ has no solutions.

Now for any nonconstant polynomial $f \in \mathbb{Z}[X]$ we define

$$\mathcal{P}_f := \{p : p \mid f(n) \text{ for some } n \in \mathbb{N}\}.$$

It is well-known that $\mathcal{P}_f$ has a positive relative density δ_f in the set of prime numbers. More precisely, the Frobenius density theorem says that $\delta_f = \mathrm{Fix}(\mathcal{G})/\#\mathcal{G}$, where $\mathcal{G}$ is the Galois group of f over $\mathbb{Q}$, and $\mathrm{Fix}(\mathcal{G})$ is the number of elements of $\mathcal{G}$ which have at least one fixed point, when regarded as permutations of the roots of f (see, e.g., [57]). If f has a rational root, then it follows easily that $\delta_f = 1$. However, we remark that it can be $\delta_f = 1$ also if f has no rational roots. An example is given by

$$f(X) = (X^2 - 2)(X^2 - 3)(X^2 - 6),$$

because for a prime number p the product of two nonsquares modulo p is a square modulo p. We need the following asymptotic formula for $\#\mathcal{P}_f(x)$.

Theorem 5.2.3. *For any nonconstant polynomial $f \in \mathbb{Z}[X]$, we have*

$$\#\mathcal{P}_f(x) = \delta_f \,\mathrm{Li}(x) + O_f\left(\frac{x}{\exp(C_f\sqrt{\log x})}\right)$$

for all $x \geq 2$, where $C_f > 0$ is a constant depending on f only.

Proof. Let $\mathcal{G}$ be the Galois group of f over $\mathbb{Q}$. Thanks to the considerations about $\pi_{\mathcal{C}}(x)$ given before Theorem 1.2.1, we have

$$\#\mathcal{P}_f(x) = \sum_{\mathcal{C}} \pi_{\mathcal{C}}(x)$$

for all $x > 0$, where the sum is over all conjugacy classes $\mathcal{C}$ of $\mathcal{G}$ whose elements, when regarded as permutations of the roots of f, have a fixed point. Hence, the claim follows from Theorem 1.2.1. □

The next lemma concerns the p-adic valuation of products consisting of consecutive values of a polynomial.

Lemma 5.2.4. *Let $f \in \mathbb{Z}[X]$ be a polynomial without roots in $\mathbb{N}$, and set*

$$Q_N := \prod_{n=1}^{N} f(n), \tag{5.2}$$

for all positive integers N. Then, for any prime number p, we have

$$\nu_p(Q_N) = \frac{t_f N}{p-1} + O_f\left(\frac{\log N}{\log p}\right),$$

for all integers $N \geq 2$, where t_f is the number of roots of f in the p-adic integers.

Proof. This is [2, Theorem 1.2]. Note that in [2] the error term is written as $O(\log N)$, but looking at the proof one can easily check that it is $O_f(\log N/\log p)$. □

Our last auxiliary result establishes a lower bound for the number of "big" prime factors of an irreducible polynomial.

Lemma 5.2.5. *Let $f \in \mathbb{Z}[X]$ be a nonconstant polynomial. For each positive integers N, let $\mathcal{S}_N$ be the set of all prime numbers p such that $p > N$ and $p \mid f(n)$ for some positive integer $n \leq N$. Then, we have*

$$\#\mathcal{S}_N \gg_f (1-\delta_f)N,$$

for all sufficiently large integers N.

Proof. We proceed similarly to the first part of the proof of [18, Theorem 5.1].

Define Q_N as in (5.2). If $\delta_f = 1$, then the claim follows. Hence we can assume that f has no roots in $\mathbb{N}$. In particular, $Q_N \neq 0$ for every integer $N \geq 1$. Clearly, $\mathcal{S}_N = \{p : p \mid Q_N,\ p > N\}$. Put $\mathcal{S}'_N := \{p : p \mid Q_N,\ p \leq N\}$, so that

$$\log|Q_N| = \sum_{p\in\mathcal{S}_N} \nu_p(Q_N)\log p + \sum_{p\in\mathcal{S}'_N} \nu_p(Q_N)\log p, \tag{5.3}$$

for every positive integer N. For the rest of the proof, all the implied constants may depend on f. By Lemma 5.2.4, we have

$$\nu_p(Q_N) = \frac{t_f N}{p-1} + O\left(\frac{\log N}{\log p}\right),$$

for every integer $N \geq 2$, and thus

$$\sum_{p \in \mathcal{S}_N} \nu_p(Q_N) \log p \ll \sum_{p \in \mathcal{S}_N} \log p \ll \sum_{p \in \mathcal{S}_N} \log |f(N)| \ll \#\mathcal{S}_N \log N. \tag{5.4}$$

Since $\mathcal{S}'_N$ is a subset of the set of all prime numbers up to N, by the Prime Number Theorem (or even Chebyshev's estimates), it follows that

$$\#\mathcal{S}'_N \ll \frac{N}{\log N}.$$

Moreover, since $\mathcal{S}'_N \subseteq \mathcal{P}_f(N)$, by Theorem 5.2.3 and by partial summation, we have

$$\sum_{p \in \mathcal{S}'_N} \frac{\log p}{p-1} \leq \sum_{p \in \mathcal{P}_f(N)} \frac{\log p}{p-1} = \delta_f \log N + O(1),$$

for every integer $N \geq 2$. Therefore,

$$\sum_{p \in \mathcal{S}'_N} \nu_p(Q_N) \log p \leq \sum_{p \in \mathcal{S}'_N} \left(\frac{kN \log p}{p-1} + O(\log N) \right) \leq \delta_f kN \log N + O(N). \tag{5.5}$$

for every integer $N \geq 2$. Finally, by Stirling's formula

$$\log |Q_N| = kN \log N + O(N). \tag{5.6}$$

Putting together (5.3), (5.4), (5.5), and (5.6), we get

$$\#\mathcal{S}_N \gg (1 - \delta_f) kN + O\left(\frac{N}{\log N}\right),$$

and the desired result follows. □

5.3 Proof of Theorem 5.1.1

Let $f \in \mathbb{Z}[X]$ be a nonconstant polynomial of degree 2 or 3. If f is reducible in $\mathbb{Z}[X]$, then there exists a linear polynomial $h \in \mathbb{Z}[X]$ such that $h(n) \mid f(n)$ for all integers n; and

the existence of G_f follows immediately from the existence of G_h proved by Evans [17]. Therefore, we can assume that f is irreducible in $\mathbb{Z}[X]$. Hence the Galois group of f over $\mathbb{Q}$ is precisely one of S_2, S_3, or A_3, and by the Frobenius density theorem δ_f is $1/2$, $2/3$, or $1/3$, respectively. Further, by Lemma 5.2.1 we know that f and $\widetilde{f}$ has the same Galois group over $\mathbb{Q}$, and, consequently, by the Frobenius density theorem $\delta_{\widetilde{f}} = \delta_f$.

Let N be a sufficiently large positive integer. Define $\mathcal{S}_N$ as the set of all prime numbers p such that $p > N/2$ and $p \mid \widetilde{f}(r)$ for some positive integer $r \leq N/2$. Thanks to the previous considerations and Lemma 5.2.5, we have that

$$\#\mathcal{S}_N \geq c_1 N, \tag{5.7}$$

for all sufficiently large N, where $c_1 > 0$ is constant depending only on f. Moreover, Lemma 5.2.2 tell us that for each $p \in \mathcal{S}_N$ there exists two integers z_p^- and z_p^+ such that

$$f(z_p^-) \equiv f(z_p^+) \equiv 0 \bmod p,$$

and $0 < z_p^+ - z_p^- \leq N/2 < p$.

Now since

$$\sum_{p \in \mathcal{P}_f} \frac{1}{p} = +\infty,$$

we can fix $s \geq 1$ elements $p_1 < \cdots < p_s$ of $\mathcal{P}_f$ such that

$$\prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right) < \frac{c_1}{3}. \tag{5.8}$$

Moreover, by the definition of $\mathcal{P}_f$, for each $p \in \mathcal{P}_f$ we can pick an integer z_p such that $f(z_p) \equiv 0 \pmod p$.

Let $h_1 < \ldots < h_{N_1}$ be all the elements of $\{1, \ldots, N\}$ which are not divisible by any of the primes $p_1, \ldots, p_s$, and let $k_1 < \cdots < k_{N_2}$ be all the remaining elements, so that $N = N_1 + N_2$. By the Eratosthenes' sieve and (5.8), we have

$$N_1 \leq N \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right) + 2^s < \frac{c_1}{2} N, \tag{5.9}$$

for all sufficiently large N. Let $q_1 < \cdots < q_t$ be all the elements of $\mathcal{S}_N \setminus \{p_1, \ldots, p_s\}$. From (5.7) and (5.9), we get that

$$t \geq c_1 N - s > \frac{c_1}{2} N > N_1,$$

for all sufficiently large N. As a consequence, for any $j = 1, \ldots, N_1$, we can define $r_j = z_{q_j}^-$ if $h_j \leq N/2$, and $r_j = z_{q_j}^+$ if $h_j > N/2$. Finally, we assume N sufficiently large so that $N \geq 2p_s$.

At this point, note that by construction $p_1, \ldots, p_s$ and $q_1, \ldots, q_{N_1}$ are all pairwise distinct. Thus, by the Chinese Remainder Theorem, the system of congruences:

$$\begin{cases} n \equiv z_{p_i} & (\mathrm{mod}\ p_i) \quad i = 1, \ldots, s \\ n \equiv r_j - h_j & (\mathrm{mod}\ q_j) \quad j = 1, \ldots, N_1 \end{cases}$$

has infinitely many positive integer solutions. If n is a solution, then it is easy to see that none of the integers among

$$f(n+1), f(n+2), \ldots, f(n+N)$$

is relatively prime to all the others.

Indeed, take any $h \in \{1, \ldots, N\}$. On one hand, if h is divisible by some p_i, then

$$f(n+h) \equiv f(n+h \pm p_i) \equiv f(z_{p_i}) \equiv 0 \pmod{p_i},$$

so that

$$\gcd(f(n+h), f(n+h \pm p_i)) > 1,$$

while $h \pm p_i \in \{1, \ldots, N\}$ for the right choice of the sign, since $N \geq 2p_s$.

On the other hand, if h is not divisible by any of $p_1, \ldots, p_s$, then $h = h_j$ for some $j \in \{1, \ldots, N_1\}$. If $h_j \leq N/2$, then

$$f(n+h) \equiv f(z_{q_j}^-) \equiv 0 \pmod{q_j},$$

and

$$f(n+h+z_{q_j}^+ - z_{q_j}^-) \equiv f(z_{q_j}^+) \equiv 0 \pmod{q_j},$$

so that

$$\gcd(f(n+h), f(n+h+z_{q_j}^{+} - z_{q_j}^{-})) > 1,$$

while $h + z_{q_j}^{+} - z_{q_j}^{-} \in \{1, \ldots, N\}$. Similarly, if $h_j > N/2$ then

$$\gcd(f(n+h+z_{q_j}^{-} - z_{q_j}^{+}), f(n+h)) > 1,$$

while $h + z_{q_j}^{-} - z_{q_j}^{+} \in \{1, \ldots, N\}$.

Hence, the existence of G_f has been proved.

Chapter 6

Central binomial coefficients divisible by their indices

Abstract. *Let $\mathcal{A}$ be the set of all positive integers n such that n divides the central binomial coefficient $\binom{2n}{n}$. Pomerance proved that the upper density of $\mathcal{A}$ is at most $1 - \log 2 = 0.30685\ldots$ I improve this bound to $1 - \log 2 - 0.05551 = 0.25134\ldots$ Moreover, let $\mathcal{B}$ be the set of all positive integers n such that n and $\binom{2n}{n}$ are relatively prime. I show that $\#(\mathcal{B} \cap [1, x]) \ll x/\sqrt{\log x}$ for all $x > 1$. This work is appeared in* [49].

6.1 Introduction

Given a sequence of integers $(a_n)_{n\geq 1}$ with some combinatorial or number-theoretic meaning, the study of the set of positive integers n such that n divides a_n has interested several reseachers. For instance, as we have already seen in Chapter 2, Alba González, Luca, Pomerance, and Shparlinski [1] considered the case of $(a_n)_{n\geq 1}$ being a linear recurrence; while André-Jeannin [4], Luca and Tron [36], Sanna [47], and Somer [56] focused on Lucas sequences. Furthermore, Gottschlich [20], Silverman and Stange [54] studied this problem for elliptic divisibility sequences; and Chen, Gassert and Stange [10] consider the case when $a_n = \phi^{(n)}(0)$ is the nth iterate of a polynomial map $\phi \in \mathbb{Z}[X]$.

In this chapter, we study the case in which $a_n = \binom{2n}{n}$ is the nth central binomial coefficient. Let $\mathcal{A}$ be the set of positive integers n such that n divides the central binomial coefficient $\binom{2n}{n}$. Ulas and Schinzel [60, Theorems 3.2 and 3.4] proved that $\mathcal{A}$ and its complement $\mathbb{N} \setminus \mathcal{A}$ are both infinite.

Pomerance [41, Theorem 3] studied the upper density of $\mathcal{A}$ and proved the following result.

Theorem 6.1.1. $\overline{\mathbf{d}}(\mathcal{A}) \leq 1 - \log 2 = 0.30685\ldots$

Note that Pomerance, maybe for aesthetic reasons, stated Theorem 6.1.1 with $1/3$ instead of $1 - \log 2$, but he actually proved the latter bound. Also, Pomerance [41, end of pag. 7] conjectured that $\mathcal{A}$ has a positive lower density, and indeed numerical experiments [55] seem to suggest that the lower density of $\mathcal{A}$ is at least $1/9$.

Our first result is the following improvement of Theorem 6.1.1.

Theorem 6.1.2. $\overline{\mathbf{d}}(\mathcal{A}) \leq 1 - \log 2 - 0.05551 = 0.25134\ldots$

At this point, similarly to what we have done in Chapter 2, we consider the "dual" set of $\mathcal{A}$, that is, the set $\mathcal{B}$ of all positive integers n such that $\binom{2n}{n}$ and n are relatively prime. It is easy to see that each odd prime number belongs to $\mathcal{B}$. Hence, by the Prime Number Theorem, we have

$$\#\mathcal{B}(x) \geq (1 + o(1)) \cdot \frac{x}{\log x}, \tag{6.1}$$

as $x \to +\infty$. Our second result is an upper bound for $\#\mathcal{B}(x)$.

Theorem 6.1.3. *We have*

$$\#\mathcal{B}(x) \ll \frac{x}{\sqrt{\log x}}, \tag{6.2}$$

for all $x > 1$.

6.2 Preliminaries

The key tool of the proofs is the following lemma.

Lemma 6.2.1. *For all prime numbers* p *and all positive integers* n*, we have that* p *does not divide* $\binom{2n}{n}$ *if and only if all the digits of* n *written in base* p *are less that* $p/2$*.*

Proof. It is a corollary of the theorem of Kummer [29] which says that, for positive integers m, n and a prime number p, the exponent of p in the prime factorization of $\binom{m+n}{n}$ is equal to the number of carries in the addition $m + n$ when done in base p. □

The next lemma follows easily from Lemma 6.2.1 by a counting argument.

Lemma 6.2.2. *For all prime numbers* p *and all* $x \geq 2$*, the number of positive integers* $n \leq x$ *such that* p *does not divide* $\binom{2n}{n}$ *is at most* px^{θ_p}*, where* $\theta_p := \log(\frac{1}{2}(p+1))/\log p$*.*

Proof. First, the claim is obvious for $p = 2$, simply because $\binom{2n}{n}$ is even for all positive integers n. Hence, suppose that p is odd. By Lemma 6.2.1, a positive integer $n \leq x$ is such that p does not divide $\binom{2n}{n}$ if and only if $n = \sum_{j=0}^{k} d_j p^j$, where $d_0, \ldots, d_k$ are nonnegative integers not exceeding $(p-1)/2$ and $k := \lfloor \log x / \log p \rfloor + 1$. Therefore, the possible choices for n are at most

$$\left(\frac{p+1}{2}\right)^k \leq \left(\frac{p+1}{2}\right)^{\log x/\log p+1} < px^{\theta_p},$$

as claimed. □

We need also the following well-known theorem of Mertens.

Theorem 6.2.3. *We have*

$$\sum_{p \le x} \frac{1}{p} = \log\log x + M + O\left(\frac{1}{\log x}\right),$$

for all $x \ge 2$, where M is the Meissel–Mertens constant.

Proof. See [58, Chapter I.1, Theorem 9]. □

A consequence of Theorem 6.2.3 is the following technical lemma.

Lemma 6.2.4. *We have*

$$\sum_{\substack{\sqrt[3]{2x} < p \le \sqrt{2x} \\ \sqrt{2x} < q \le x/p}} \frac{1}{pq} < 0.06502 + o(1),$$

as $x \to +\infty$.

Proof. Fix a positive integer n and, for the sake of convenience, put $\alpha_k := 1/3 + k/(6n)$ for each nonnegative integer $k \le n$. Thanks to Theorem 6.2.3, we have

$$\begin{aligned}
\sum_{\substack{\sqrt[3]{2x} < p \le \sqrt{2x} \\ \sqrt{2x} < q \le x/p}} \frac{1}{pq} &= \sum_{k=1}^{n} \sum_{(2x)^{\alpha_{k-1}} < p \le (2x)^{\alpha_k}} \frac{1}{p} \sum_{\sqrt{2x} < q \le x/p} \frac{1}{q} \\
&\le \sum_{k=1}^{n} \left(\sum_{(2x)^{\alpha_{k-1}} < p \le (2x)^{\alpha_k}} \frac{1}{p} \right) \left(\sum_{\sqrt{2x} < q \le x/(2x)^{\alpha_{k-1}}} \frac{1}{q} \right) \\
&= \sum_{k=1}^{n} \log\left(\frac{\alpha_k}{\alpha_{k-1}}\right) \log(2 - 2\alpha_{k-1}) + o(1),
\end{aligned}$$

as $x \to +\infty$. The desired bound follows by taking $n = 10^6$. (The author performed the computation using the PARI/GP [59] computer algebra system.) □

An integer $n > 1$ is said to be a *y-rough number* if all its prime factors are greater than y. For all $x, y \ge 0$, let $\Phi(x, y)$ be the number of y-rough numbers not exceeding x. We will make use of the following estimate for $\Phi(x, y)$.

Theorem 6.2.5. *We have*

$$\Phi(x, y) = \frac{e^{-\gamma} x}{\log y} \cdot \left(1 + O\left(\frac{1}{\log y}\right)\right),$$

for all $x \geq 2$ and $2 \leq y \leq \exp(\log x/10 \log\log x)$, where γ is the Euler–Mascheroni constant.

Proof. By a corollary of Brun's combinatorial sieve [58, Chapter I.4, Theorem 2], we have

$$\Phi(x,y) = x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \left(1 + O\left(\frac{1}{(\log y)^2}\right)\right), \tag{6.3}$$

for all $x \geq 2$ and $2 \leq y \leq \exp(\log x/10 \log\log x)$. On the other hand, Mertens formula [58, Chapter I.1, Theorem 11] says that

$$\prod_{p \leq y} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log y} \left(1 + O\left(\frac{1}{\log y}\right)\right), \tag{6.4}$$

for all $y \geq 2$. Hence, putting (6.3) and (6.4) together we get the claim. □

6.3 Proof of Theorem 6.1.2

Let $\mathcal{C} := \mathbb{N} \setminus \mathcal{A}$ be the complement of $\mathcal{A}$. Clearly, it is enough to prove that $\mathcal{C}$ has lower density at least equal to $\log 2 + 0.05551$.

For $x > 0$, let $\mathcal{C}_1$ be the set of positive integers $n \leq x$ having a prime factor greater than $\sqrt{2x}$. We shall prove that $\mathcal{C}_1 \subseteq \mathcal{C}(x)$ and $\#\mathcal{C}_1 \geq (\log 2 + o(1)) \cdot x$, as $x \to +\infty$. To this aim, we proceed exactly as in the proof of Theorem 6.1.1. We include here the reasonings for the sake of completeness, and also to motivate better the rest of the proof. If $n \in \mathcal{C}_1$, then we can write $n = dp$ for a prime number $p > \sqrt{2x}$ and a positive integer d satisfying

$$d \leq \frac{x}{p} < \frac{x}{\sqrt{2x}} = \frac{1}{2}\sqrt{2x} < \frac{p}{2}.$$

Hence, by Lemma 6.2.1, p does not divide $\binom{2n}{n}$, so that $n \in \mathcal{C}$, and thus $\mathcal{C}_1 \subseteq \mathcal{C}(x)$.

Moreover, for each prime number $p \in]\sqrt{2x}, x]$, the number of possible $n = dp \in \mathcal{C}$ is at least $x/p - 1$. Note also that, since $p > \sqrt{2x}$, to each $n \in \mathcal{C}$ it corresponds exactly one p. Therefore, if $\pi(x)$ denotes the number of prime numbers not exceeding x,

$$\#\mathcal{C}_1 \geq \sum_{\sqrt{2x} < p \leq x} \left(\frac{x}{p} - 1\right) \geq x \cdot \sum_{\sqrt{2x} < p \leq x} \frac{1}{p} - \pi(x) = (\log 2 + o(1)) \cdot x, \tag{6.5}$$

as $x \to +\infty$, where we used Theorem 6.2.3 and the well-known fact that $\pi(x) = o(x)$.

Now let $\mathcal{C}_2$ be the set of positive integers $n \le x$ such that $n = mp$, where $p \in]\sqrt[3]{2x}, \sqrt{2x}]$ is a prime number, and m is a positive integer having no prime factor greater than $\sqrt{2x}$ and satisfying $(m \bmod p) < p/2$. If $n \in \mathcal{C}_2$, then writing $n = mp$ as before we have

$$m \le \frac{x}{p} < \frac{x}{\sqrt[3]{2x}} = \frac{1}{2}(2x)^{2/3} < \frac{1}{2}p^2,$$

which, together with $(m \bmod p) < p/2$, implies that $m = d_1 + d_2 p$ for some nonnegative integers $d_1, d_2 < p/2$. Hence, by Lemma 6.2.1, p does not divide $\binom{2n}{n}$, so that $n \in \mathcal{C}$, and thus $\mathcal{C}_2 \subseteq \mathcal{C}(x)$.

At this point, we want to prove a lower bound for $\#\mathcal{C}_2$. For each prime number $p \in]\sqrt[3]{2x}, \sqrt{2x}]$, the number of positive integers $m \le x/p$ such that $(m \bmod p) < p/2$ is at least

$$\left(\frac{x}{p^2} - 1\right) \cdot \frac{p}{2} - 1 \ge \frac{x}{2p} - p.$$

Furthermore, the number of positive integers $m \le x/p$ such that m has a prime factor $q > \sqrt{2x}$ is at most

$$\sum_{\sqrt{2x} < q \le x/p} \frac{x}{pq}.$$

Therefore, the number of positive integers m such that $mp \in \mathcal{C}_2$ is at least

$$\frac{x}{2p} - p - \sum_{\sqrt{2x} < q \le x/p} \frac{x}{pq}. \tag{6.6}$$

Now summing (6.6) over all prime numbers $p \in]\sqrt[3]{2x}, \sqrt{2x}]$ we get

$$\begin{aligned} &\sum_{\sqrt[3]{2x} < p \le \sqrt{2x}} \left(\frac{x}{2p} - p - \sum_{\sqrt{2x} < q \le x/p} \frac{x}{pq}\right) \\ &\qquad \ge \sum_{\sqrt[3]{2x} < p \le \sqrt{2x}} \left(\frac{1}{2p} - \sum_{\sqrt{2x} < q \le x/p} \frac{1}{pq}\right) \cdot x - \pi(\sqrt{2x}) \cdot \sqrt{2x} \\ &\qquad > \left(\tfrac{1}{2}\log\tfrac{3}{2} - 0.06502 + o(1)\right) \cdot x, \end{aligned} \tag{6.7}$$

as $x \to +\infty$, where we applied Theorem 6.2.3 and Lemma 6.2.4. However, note that (6.7) is not a lower bound for $\#\mathcal{C}_2$, since some $n \in \mathcal{C}_2$ could be written as $n = mp$ for

two (and no more than two, since $p > \sqrt[3]{2x}$) different values of p. The number of those double-counted n's is at most

$$\sum_{\sqrt[3]{2x} < p < q \le \sqrt{2x}} \frac{x}{pq} < \frac{x}{2} \cdot \left(\sum_{\sqrt[3]{2x} < p \le \sqrt{2x}} \frac{1}{p} \right)^2 = \left(\tfrac{1}{2} (\log \tfrac{3}{2})^2 + o(1) \right) \cdot x,$$

as $x \to +\infty$, thanks again to Theorem 6.2.3. Hence,

$$\#\mathcal{C}_2 > \left(\tfrac{1}{2} \log \tfrac{3}{2} - 0.06502 - \tfrac{1}{2} (\log \tfrac{3}{2})^2 + o(1) \right) \cdot x > (0.05551 + o(1)) \cdot x, \tag{6.8}$$

as $x \to +\infty$.

At this point, since $\mathcal{C}_1 \cap \mathcal{C}_2 = \varnothing$ and $\mathcal{C}_1 \cup \mathcal{C}_2 \subseteq \mathcal{C}(x)$, by (6.5) and (6.8) we get

$$\#\mathcal{C}(x) \ge \#\mathcal{C}_1 + \#\mathcal{C}_2 > (\log 2 + 0.05551 + o(1)) \cdot x,$$

as $x \to +\infty$, so that

$$\underline{\mathbf{d}}(\mathcal{C}) \ge \log 2 + 0.05551,$$

and the proof is complete.

6.4 Proof of Theorem 6.1.3

Suppose $x \ge 2$ is sufficiently large, and put $y := \exp\bigl(\sqrt{\log x}/4\bigr)$. We split $\mathcal{B}$ into two subsets:

$$\mathcal{B}_1 := \{n \in \mathcal{B} : p \mid n \text{ for some prime number } p \le y\},$$
$$\mathcal{B}_2 := \mathcal{B} \setminus \mathcal{B}_1.$$

If $n \in \mathcal{B}_1(x)$, then $n = mp$ for a prime number $p \le y$ and a positive integer m. Moreover, since n and $\binom{2n}{n}$ are relatively prime, p does not divide $\binom{2n}{n}$, which in turn, by Lemma 6.2.1, implies that p does not divide $\binom{2m}{m}$. Now Lemma 6.2.2 tell us that for any prime number $p \le y$ there are at most $p(x/p)^{\theta_p}$ positive integers $m \le x/p$ such that p does not divide $\binom{2m}{m}$. Therefore,

$$\begin{aligned} \#\mathcal{B}_1(x) &\le \sum_{p \le y} p \left(\frac{x}{p}\right)^{\theta_p} \le \sum_{p \le y} p \left(\frac{x}{p}\right)^{1 - 1/(4 \log p)} \ll x^{1 - 1/(4 \log y)} \cdot y \\ &= x \cdot \exp\left(-\frac{\log x}{4 \log y} + \log y \right) \ll \frac{x}{\sqrt{\log x}}, \end{aligned} \tag{6.9}$$

where we also used the inequality

$$\theta_p = \frac{\log\left(\frac{1}{2}(p+1)\right)}{\log p} = 1 - \frac{\log(2p/(p+1))}{\log p} < 1 - \frac{1}{4\log p}.$$

On the other hand, thanks to Theorem 6.2.5, we have

$$\#\mathcal{B}_2(x) \leq \Phi(x,y) \ll \frac{x}{\log y} \ll \frac{x}{\sqrt{\log x}}. \tag{6.10}$$

Hence, putting together (6.9) and (6.10), we get

$$\#\mathcal{B}(x) = \#\mathcal{B}_1(x) + \#\mathcal{B}_2(x) \ll \frac{x}{\sqrt{\log x}},$$

as desired.

Bibliography

[1] J. J. Alba González, F. Luca, C. Pomerance, and I. E. Shparlinski, *On numbers n dividing the nth term of a linear recurrence*, Proc. Edinb. Math. Soc. (2) **55** (2012), no. 2, 271–289.

[2] T. Amdeberhan, L. A. Medina, and V. H. Moll, *Asymptotic valuations of sequences satisfying first order recurrences*, Proc. Amer. Math. Soc. **137** (2009), no. 3, 885–890.

[3] F. Amoroso and E. Viada, *On the zeros of linear recurrence sequences*, Acta Arith. **147** (2011), no. 4, 387–396.

[4] R. André-Jeannin, *Divisibility of generalized Fibonacci and Lucas numbers by their subscripts*, Fibonacci Quart. **29** (1991), no. 4, 364–366.

[5] Y. F. Bilu, *The many faces of the subspace theorem [after Adamczewski, Bugeaud, Corvaja, Zannier...]*, Astérisque **317** (2008), Exp. No. 967, vii, 1–38. Séminaire Bourbaki. Vol. 2006/2007.

[6] A. Brauer, *On a property of k consecutive integers*, Bull. Amer. Math. Soc. **47** (1941), 328–331.

[7] A. Brauer and M. Zeitz, *Über eine zahlentheoretische behauptung von legendre*, Sitzungsberichte d. Berliner Mathematischen Gesellschaft **29** (1930), 116–125.

[8] P. S. Bruckman and P. G. Anderson, *Conjectures on the Z-densities of the Fibonacci sequence*, Fibonacci Quart. **36** (1998), no. 3, 263–271.

[9] R. Canetti, J. Friedlander, S. Konyagin, M. Larsen, D. Lieman, and I. Shparlinski, *On the statistical properties of Diffie-Hellman distributions*, Israel J. Math. **120** (2000), no. 1, 23–46.

[10] A. S. Chen, T. A. Gassert, and K. E. Stange, *Index divisibility in dynamical sequences and cyclic orbits modulo p*, New York J. Math. **23** (2017), 1045–1063.

[11] P. Corvaja and U. Zannier, *Diophantine equations with power sums and universal Hilbert sets*, Indag. Math. (N.S.) **9** (1998), no. 3, 317–332.

[12] P. Corvaja and U. Zannier, *Finiteness of integral values for the ratio of two linear recurrences*, Invent. Math. **149** (2002), no. 2, 431–451.

[13] P. Cubre and J. Rouse, *Divisibility properties of the Fibonacci entry point*, Proc. Amer. Math. Soc. **142** (2014), no. 11, 3771–3785.

[14] P. Erdős, *On the difference of consecutive primes*, Q. J. Math. **6** (1935), 124–128.

[15] P. Erdős, *On the density of the abundant numbers*, J. London Math. Soc. **9** (1934), no. 4, 278.

[16] P. Erdős and J. L. Selfridge, *The product of consecutive integers is never a power*, Illinois J. Math. **19** (1975), 292–301.

[17] R. Evans, *On N consecutive integers in an arithmetic progression*, Acta Sci. Math. (Szeged) **33** (1972), 295–296.

[18] G. Everest, S. Stevens, D. Tamsett, and T. Ward, *Primes generated by recurrence sequences*, Amer. Math. Monthly **114** (2007), no. 5, 417–431.

[19] G. Everest, A. van der Poorten, I. Shparlinski, and T. Ward, *Recurrence sequences*, Mathematical Surveys and Monographs, vol. 104, American Mathematical Society, Providence, RI, 2003.

[20] A. Gottschlich, *On positive integers n dividing the nth term of an elliptic divisibility sequence*, New York J. Math. **18** (2012), 409–420.

[21] L. Hajdu and N. Saradha, *On a problem of Pillai and its generalizations*, Acta Arith. **144** (2010), no. 4, 323–347.

[22] L. Hajdu and M. Szikszai, *On the GCD-s of k consecutive terms of Lucas sequences*, J. Number Theory **132** (2012), no. 12, 3056–3069.

[23] R. R. Hall, *Sets of multiples*, Cambridge Tracts in Mathematics, vol. 118, Cambridge University Press, Cambridge, 1996.

[24] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), no. 1, 1–70.

[25] J. Harrington and L. Jones, *Extending a theorem of Pillai to quadratic sequences*, Integers **15A** (2015), Paper No. A7, 22.

[26] H. A. Heilbronn, *On an inequality in the elementary theory of numbers*, Proc. Cambridge Philos. Soc. **33** (1937), 207–209.

[27] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.

[28] L. Kronecker, *Über die Irreductibilität von Gleichungen*, Monatsberichte Königl. Preußisch. Akad. Wissenschaft. Berlin (1880), 155–162.

[29] E. E. Kummer, *Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen*, J. Reine Angew. Math. **44** (1852), 93–146.

[30] J. C. Lagarias, *The set of primes dividing the Lucas numbers has density $2/3$*, Pacific J. Math. **118** (1985), no. 2, 449–461.

[31] J. C. Lagarias, *Errata to: "The set of primes dividing the Lucas numbers has density $2/3$" [Pacific J. Math.* **118** *(1985), no. 2, 449–461]*, Pacific J. Math. **162** (1994), no. 2, 393–396.

[32] K. S. E. Lee, *On the sum of a prime and a Fibonacci number*, Int. J. Number Theory **6** (2010), no. 7, 1669–1676.

[33] A.-M. Legendre, *Théorie des nombres*, Paris, 1830. Tome II.

[34] T. Lengyel, *The order of the Fibonacci and Lucas numbers*, Fibonacci Quart. **33** (1995), no. 3, 234–239.

[35] P. Leonetti and C. Sanna, *On the greatest common divisor of n and the nth Fibonacci number*, Rocky Mountain J. Math. (accepted). `https://arxiv.org/abs/1704.00151`.

[36] F. Luca and E. Tron, *The distribution of self-Fibonacci divisors*, Advances in the theory of numbers, 2015, pp. 149–158.

[37] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika **20** (1973), 119–134.

[38] M. Ohtomo and F. Tamari, *On relative prime number in a sequence of positive integers*, J. Statist. Plann. Inference **106** (2002), no. 1-2, 509–515. Experimental design and related combinatorics.

[39] S. S. Pillai, *On m consecutive integers. I*, Proc. Indian Acad. Sci., Sect. A. **11** (1940), 6–12.

[40] S. S. Pillai, *On m consecutive integers. II*, Proc. Indian Acad. Sci., Sect. A. **11** (1940), 73–80.

[41] C. Pomerance, *Divisors of the middle binomial coefficient*, Amer. Math. Monthly **122** (2015), no. 7, 636–644.

[42] M. Renault, *The period, rank, and order of the (a, b)-Fibonacci sequence* $\bmod m$, Math. Mag. **86** (2013), no. 5, 372–380.

[43] H. Rohrbach, *Beweis einer zahlentheoretischen Ungleichung*, J. Reine Angew. Math. **177** (1937), 193–196.

[44] R. Rumely, *Notes on van der Poorten's proof of the Hadamard quotient theorem. I, II*, Séminaire de Théorie des Nombres, Paris 1986–87, 1988, pp. 349–382, 383–409.

[45] C. Sanna, *The p-adic valuation of Lucas sequences*, Fibonacci Quart. **54** (2016), no. 2, 118–124.

[46] C. Sanna, *Distribution of integral values for the ratio of two linear recurrences*, J. Number Theory **180** (2017), 195–207.

[47] C. Sanna, *On numbers n dividing the nth term of a Lucas sequence*, Int. J. Number Theory **13** (2017), no. 3, 725–734.

[48] C. Sanna, *On numbers n relatively prime to the nth term of a linear recurrence*, Bull. Malays. Math. Sci. Soc. (2017). https://doi.org/10.1007/s40840-017-0514-8.

[49] C. Sanna, *Central binomial coefficients divisible by or coprime to their indices*, Int. J. Number Theory **14** (2018), no. 4, 1135–1141.

[50] C. Sanna and M. Szikszai, *A coprimality condition on consecutive values of polynomials*, Bull. Lond. Math. Soc. **49** (2017), no. 5, 908–915.

[51] C. Sanna and E. Tron, *The density of numbers n having a prescribed G.C.D. with the nth Fibonacci number*, Indag. Math. (N.S.) **29** (2018), no. 3, 972–980.

[52] W. M. Schmidt, *Zeros of linear recurrence sequences*, Publ. Math. Debrecen **56** (2000), no. 3-4, 609–630.

[53] J.-P. Serre, *Lectures on $N_X(p)$*, Chapman & Hall/CRC Research Notes in Mathematics, vol. 11, CRC Press, Boca Raton, FL, 2012.

[54] J. H. Silverman and K. E. Stange, *Terms in elliptic divisibility sequences divisible by their indices*, Acta Arith. **146** (2011), no. 4, 355–378.

[55] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*. Sequence A014847.

[56] L. Somer, *Divisibility of terms in Lucas sequences by their subscripts*, Applications of Fibonacci numbers, Vol. 5 (St. Andrews, 1992), 1993, pp. 515–525.

[57] P. Stevenhagen and H. W. Lenstra Jr., *Chebotarëv and his density theorem*, Math. Intelligencer **18** (1996), no. 2, 26–37.

[58] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995.

[59] The PARI Group, Bordeaux, *PARI/GP*, 2013.

[60] M. Ulas and A. Schinzel, *A note on Erdős-Straus and Erdős-Graham divisibility problems (with an appendix by Andrzej Schinzel)*, Int. J. Number Theory **9** (2013), no. 3, 583–599.

[61] A. J. van der Poorten, *Solution de la conjecture de Pisot sur le quotient de Hadamard de deux fractions rationnelles*, C. R. Acad. Sci. Paris Sér. I Math. **306** (1988), no. 3, 97–102.

[62] A. J. van der Poorten and H. P. Schlickewei, *Zeros of recurrence sequences*, Bull. Austral. Math. Soc. **44** (1991), no. 2, 215–223.

[63] U. Zannier, *Some applications of Diophantine approximation to Diophantine equations. With special emphasis on the Schmidt subspace theorem*, Forum, Udine, 2003.

[64] U. Zannier, *Lecture notes on Diophantine analysis*, Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], vol. 8, Edizioni della Normale, Pisa, 2009. With an appendix by Francesco Amoroso.