Last nonzero digits and p-adic valuations of special sequences

Bartosz Sobolewski

PhD thesis

Primary supervisor: dr hab. Maciej Ulas, prof. UJ
Auxiliary supervisor: dr Jakub Byszewski

Faculty of Mathematics and Computer Science
Jagiellonian University in Kraków
Kraków, November 16, 2020

Contents

Introduction

In this thesis we investigate selected properties of base-b expansions of terms of interesting number sequences, including factorials, linear recurrence sequences, polynomials and p-adic analytic functions evaluated at consecutive integers. Problems of this type have been widely studied in the literature, with emphasis put on calculating the p-adic valuation ν_p for a prime p. (In some cases the investigation can also been extended to the "b-adic valuation" ν_b for any base b, where $\nu_b(n) = \sup\{v : b^v \mid n\}$.) Most likely, the earliest result of this type is the famous formula for $\nu_p(n!)$ due to Legendre [47], dating back to the beginning of the XIX century. Since then, numerous other sequences of combinatorial significance, such as binomial coefficients and Stirling numbers of the first and second kind, have been an object of interest in this regard. The study of p-adic valuations of linear recurrence sequences seems to have grown in popularity recently, starting with the result of Lengyel [48], who fully characterized the p-adic valuation of Fibonacci and Lucas numbers. Among other interesting examples, we would like to mention the paper of Bell [8] on p-regularity of the p-adic valuation of a polynomial evaluated at consecutive integers, and a generalization to p-adic analytic functions by Shu and Yao [68].

Another interesting type of properties describing base-b expansions can be collectively called the "last nonzero digits". In particular, the sequence $(\ell_b(n!))_{n\geq 0}$, describing the last nonzero digit in base-b expansion of $n!$, has recently gained the attention of several authors. The current line of research was initiated by the work of Deshouillers and Ruzsa [29] on the especially interesting case $b = 12$. They proved that the sequence $(\ell_{12}(n!))_{n\geq 0}$ coincides with a 3-automatic sequence on a set of asymptotic density 1, which allowed them to compute how often each possible value occurs in the sequence. The question of whether $(\ell_{12}(n!))_{n\geq 0}$ is automatic itself was answered negatively by Deshouillers [27], while a criterion for a general base b was later provided by Lipka [51].

Similar results on last nonzero digits of expressions other than $n!$ are rather scattered throughout the literature. In this thesis we attempt to develop a systematic approach to the topic. We consider the function $\mathcal{L}_b(n)$, which describes the number obtained by deleting all the trailing zeros in the base-b representation of n, as well as $\ell_{b,d}(n)$, which denotes the integer represented by the last block of d digits in $\mathcal{L}_b(n)$. Since the whole base-b expansion can be retrieved from the knowledge of $\mathcal{L}_b$ and ν_b and vice versa, the investigation of last nonzero digits and valuations can be seen as mutually complementary. As it turns out, the functions $\ell_{b,d}$ and $\mathcal{L}_b$ are also, in a sense, compatible with automatic and regular sequences, respectively. In this thesis

we show that for many naturally occurring sequences $(s_n)_{n\geq 0}$, it is possible (but not trivial) to settle whether $(\ell_{b,d}(s_n))_{n\geq 0}$ is automatic and $(\mathcal{L}_b(s_n))_{n\geq 0}$ is regular. The results usually follow from the interplay between the arithmetic behavior of s_n and the properties of automatic and regular sequences.

The thesis consists of four chapters whose contents can be outlined as follows. Chapter 1 provides theoretical background on the main concepts and tools used throughout the entire thesis. The first part of the chapter is focused on automatic and regular sequences and their basic properties. The second part is devoted to p-adic numbers and elementary p-adic analysis.

The remaining chapters contain original research of the author. In Chapter 2 we introduce precise definitions of the functions $\mathcal{L}_b$ and $\ell_{b,d}$ and survey some of their fundamental properties. Based on the author's paper [70], we then classify the bases b such that $(\ell_{b,d}(n!))_{n\geq 0}$ is an automatic sequence (or coincides with an automatic sequence on a set of density 1). This is done by an explicit construction of a uniform morphism generating the sequence. Using this description, we calculate the frequency with which each possible value occurs in the sequence and its subsequences along arithmetic progressions. In this way we extend and generalize many other results by authors such as Kakutani, Dekking, Dresden, Deshouillers, Luca, Ruzsa, Lipka [43, 25, 31, 28, 29, 26, 27, 51].

Chapter 3 is based on yet unpublished work of the author. For a given base b we consider last nonzero digits of tuples $f = (f_1, \ldots, f_s)$ of p_i-adic analytic functions evaluated at consecutive integers, where p_i ranges over the prime factors of b. In particular, this entails polynomials and linear recurrence sequences. The specific problem and the approach used was directly inspired by the already mentioned work of Bell, Shu, and Yao [8, 68], concerning p-regularity of $(\nu_p(f(n)))_{n\geq 0}$ for p prime. We fully characterize automaticity of $(\ell_{b,d}(f(n)))_{n\geq 0}$ and regularity of $(\mathcal{L}_b(f(n)))_{n\geq 0}$ for any b, d and f. The results are illustrated by a number of various examples.

In Chapter 4 we consider generalized Fibonacci sequences $(t_n(k))_{n\geq 0}$, defined by a linear recurrence of order k. The direct motivation behind our work on this topic are the results by Lengyel and Marques [49, 50] on the 2-adic valuation of so-called "Tribonacci", "Tetranacci", and "Pentanacci" sequences ($k = 3, 4, 5$). Following the paper of the author [69], we extend their method in order to compute $\nu_2(t_n(k))$ for any even $k \geq 4$. We also derive a partial formula in the case of $k \geq 5$ odd, thus giving an alternative elementary proof of the result by Young [73]. Subsequently, we apply the obtained formula to completely solve two Diophantine equations containing the terms $t_n(k)$ with fixed even $k \geq 4$. The first equation involves expressing $m!$ as a product of a fixed number of terms $t_n(k)$, while the second one is the problem of representation of $t_n(k)$ by certain ternary quadratic forms. The approach used to solve the latter equation also demonstrates another application of the methods concerning last nonzero digits, developed in the earlier chapters.

Acknowledgements. I would like to express my deepest gratitude to Maciej Ulas and Jakub Byszewski for their guidance, many inspiring discussions and, most of all, endless patience. Without their support completing this dissertation would not have been possible.

List of symbols

Below we give a list of symbols used throughout the thesis which are potentially ambiguous.

$\mathbb{N}$	the set of nonnegative integers $\{0, 1, \ldots\}$
$\mathbb{Z}_p$	the ring of p-adic integers
$\exp_p$	the p-adic exponential
$\log_p$	the p-adic logarithm
$\#A$	the cardinality of the set A
$\gcd(a_1, \ldots, a_n)$	the greatest common divisor of the integers $a_1, \ldots, a_n$
$\operatorname{lcm}(a_1, \ldots, a_n)$	the least common multiple of the integers $a_1, \ldots, a_n$
$a \bmod m$	the integer r lying in the set $\{0, 1, \ldots, m-1\}$ such that $a \equiv r \pmod{m}$
$\operatorname{ord}_m(a)$	the multiplicative order of a modulo m
$\lfloor x \rfloor$	the floor of $x \in \mathbb{R}$, that is, the largest integer n such that $n \leq x$
$R[X]$	the ring of polynomials over a commutative ring R
$O(\cdot)$	for real-valued nonnegative functions f, g we write $f(x) = O(g(x))$ if there exist positive constants x_0, C such that $f(x) \leq Cg(x)$ for all $x \geq x_0$
$\Omega(\cdot)$	for real-valued nonnegative functions f, g we write $f(x) = \Omega(g(x))$ if $g(x) = O(f(x))$

1. Preliminaries

In this chapter we survey some important notions and results, which will be useful throughout the thesis. The topics covered in the following sections include automatic and regular sequences as well as p-adic analysis.

1.1 Automatic and regular sequences

The notion of an *automatic sequence* first appeared under the name "uniform tag sequence" in the influential paper of Cobham [19] on the theory of computation. Since then, a lot of connections have been discovered between automatic sequences and other areas of mathematics, in particular number theory. A prominent example is the theorem of Christol [16], which relates automatic sequences to algebraic power series in positive characteristic. Among many generalizations of automatic sequences, particularly significant are regular sequences, introduced by Allouche and Shallit in [3] (see also [5]). Below we give a basic overview of the theory of automatic and regular sequences, mainly based on the monograph of Allouche and Shallit [4].

1.1.1 Basic definitions

We begin by introducing some fundamental notation and terminology as in [4, Section 1.1]. Let Σ be a nonempty set of *letters*, called an *alphabet*. The set of all finite *words* created from letters in Σ, together with the empty word ϵ, is denoted by Σ^*. This set equipped with the operation of *concatenation* of words is a monoid. The *length* $|w|$ of a word $w \in \Sigma^*$ is the number of letters in w. If $a \in \Sigma$ and $w \in \Sigma^*$, then $|w|_a$ denotes the number of occurences of a in w. We also consider infinite words of the form $\mathbf{a} = a_0 a_1 \cdots$ with $a_0, a_1, \ldots \in \Sigma$, which can be identified with sequences $(a_n)_{n \geq 0}$.

Following [4, Sections 4.1, 4.3], we proceed to give the definition of a simple model of computation, called a *deterministic finite automaton with output* (DFAO). Consider the following:

- a finite *set of states* Q,
- a finite *input alphabet* Σ,
- a *transition function* $\delta \colon Q \times \Sigma \to Q$,
- an *initial state* $q_0 \in Q$,

- an *output alphabet* Δ,
- an *output function* $\tau\colon Q \to \Delta$.

We assume that the sets Q, Σ, Δ are nonempty.

Definition 1.1. A deterministic finite automaton with output is a sextuple

$$\mathcal{A} = (Q, \Sigma, \delta, q_0, \Delta, \tau).$$

A DFAO $\mathcal{A}$ defines a *finite-state function* $f\colon \Sigma^* \to \Delta$ in the following manner: $\mathcal{A}$ takes as input a word $w \in \Sigma^*$ and, starting from the initial state q_0, we move between states by means of the transition function δ according to the consecutive letters of w read from left to right. After the last letter is read, the DFAO returns the output from the final state reached. In order to write this formally, we extend the definition of δ to $Q \times \Sigma^*$ by $\delta(q, \epsilon) = q$ and

$$\delta(q, w_1 \cdots w_l) = \delta(\cdots \delta(\delta(q, w_1), w_2) \cdots, w_l)$$

for all $q \in Q$ and $w_1, \ldots, w_l \in \Sigma$. Then at input w the DFAO returns the output

$$f(w) = \tau(\delta(q_0, w)).$$

As seen in Example 1.1 below, a DFAO can be represented as a labelled multigraph whose vertices correspond to the states $q \in Q$ and are labelled by their names and output, written in the form $q/\tau(q)$. The edges are defined by the transition function δ and labelled by the letters in Σ. More precisely, there is an edge from q to q' with label $b \in \Sigma$ if and only if $\delta(q, b) = q'$. The initial state q_0 is indicated by an additional arrow labelled "start".

Example 1.1. Let $Q = \{q_0, q_1\}, \Sigma = \{0, 1\}, \Delta = \{1, -1\}$, and define the transition and output functions as follows:

$$\delta(q_0, 0) = \delta(q_1, 1) = q_0, \qquad \delta(q_0, 1) = \delta(q_1, 0) = q_1,$$

$$\tau(q_0) = 1, \qquad \tau(q_1) = -1.$$

The corresponding graph is displayed in Figure 1.1.

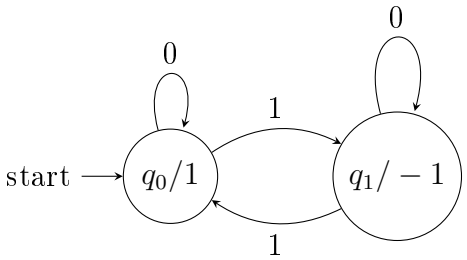

Figure 1.1: An example of a 2-DFAO

For example, we find that at input 011001 the DFAO returns -1. More generally, at input $w \in \Sigma^*$, the DFAO computes the value $(-1)^{|w|_1}$.

In order to define automatic sequences we will now consider a class of finite automata for which the input alphabet Σ is the set of digits in a given base (see [4, Chapter 5]). More precisely, let $k \geq 2$ be an integer and put $\Sigma_k = \{0, 1, \ldots, k-1\}$. We call a DFAO with input alphabet Σ_k a k-DFAO. For $w \in \Sigma_k^*$ let $[w]_k$ denote the integer represented in base k by w, where the leftmost digit is the most significant. More precisely, set $[\epsilon]_k = 0$ and for $w = w_l \cdots w_0$ with $w_0, \ldots, w_l \in \Sigma$ put

$$[w]_k = \sum_{i=0}^{l} w_i k^i.$$

Note that adding any number of leading zeros to w does not affect the value $[w]_k$. Conversely, let $(n)_k$ denote the base-k expansion of $n \geq 0$ without leading zeros. We are now ready to give the main definition of this section.

Definition 1.2. A sequence $(a_n)_{n\geq 0}$ is called *k-automatic* if there exists a k-DFAO $\mathcal{A} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that

$$a_n = \tau(\delta(q_0, w))$$

for all $n \geq 0$ and $w \in \Sigma_k^*$ such that $[w]_k = n$.

We say that a sequence is *automatic* if it is k-automatic for some $k \geq 2$, and otherwise *nonautomatic*. It is clear that an automatic sequence can only take a finite number of values.

Example 1.2. Consider the famous Thue–Morse sequence $(t_n)_{n\geq 0}$, defined by

$$t_n = (-1)^{s_2(n)}, \tag{1.1}$$

where $s_2(n)$ is the sum of binary digits of n (another variant $t_n = s_2(n) \bmod 2$ is often considered). This is a 2-automatic sequence, since it is generated by the 2-DFAO in Figure 1.1 in the following way:

$$t_n = \tau(\delta(q_0, w))$$

for all $n \geq 0$ and $w \in \Sigma_2^*$ such that $[w]_2 = n$.

It is possible to relax the condition in the definition of an automatic sequence and assume that there exists a k-DFAO $\mathcal{A} = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that for all $n \geq 0$ we have

$$a_n = \tau(\delta(q_0, (n)_k)).$$

In fact, this modification yields the same class of sequences, as per [4, Theorem 5.2.1].

1.1.2 Uniform morphisms

Based on [4, Chapter 6], in this subsection we give an equivalent definition of automatic sequences in terms of fixed points of uniform morphisms. Similarly as before, let Σ and Δ be finite alphabets. A map $\varphi : \Sigma^* \to \Delta^*$ is called a *morphism* if

$$\varphi(vw) = \varphi(v)\varphi(w)$$

for all $v, w \in \Sigma^*$. A morphism is uniquely determined by its values $\varphi(a)$ for $a \in \Sigma$ and can be naturally extended to infinite words.

Definition 1.3. Let $k \geq 1$ be an integer. We say that a morphism φ is *k-uniform* if $|\varphi(a)| = k$ for all $a \in \Sigma$. A 1-uniform morphism is called a *coding*.

In this thesis we will only consider uniform morphisms, however the discussion below can also be extended to arbitrary morphisms. Assume that φ is a k-uniform morphism with $k \geq 2$ and that $\Sigma = \Delta$, so that we can iterate φ. Let φ^0 be the identity on Σ^* and define $\varphi^{n+1}(w) = \varphi(\varphi^n(w))$ for all integers $n \geq 0$ and $w \in \Sigma^*$. Then φ is said to be *prolongable* on $a \in \Sigma$ if there exists a word $x \in \Sigma^*$ of length $k-1$ such that $\varphi(a) = ax$. In such a case the sequence of words $a, \varphi(a), \varphi^2(a), \ldots$ converges to the infinite word

$$\varphi^\omega(a) = a\, x\, \varphi(x)\, \varphi^2(x) \cdots$$

in the sense that for each $n \geq 0$ the word $\varphi^n(a)$ is a prefix of $\varphi^\omega(a)$ and the lengths $|\varphi^n(a)|$ tend to infinity. The word $\varphi^\omega(x)$ is the unique fixed point of φ starting with a. The following theorem of Cobham [19] (see also [4, Theorem 6.3.2]) gives a characterization of automatic sequences through uniform morphisms.

Theorem 1.1 (Cobham)**.** *A sequence $\mathbf{a} = (a_n)_{n\geq 0}$ with values in Δ is k-automatic if and only if there exists a k-uniform morphism $\varphi\colon \Sigma^* \to \Sigma^*$ prolongable on some $a \in \Sigma$ and a coding $\tau\colon \Sigma^* \to \Delta^*$ such that*

$$\mathbf{a} = \tau(\varphi^\omega(a)).$$

In short, one can say that $\mathbf{a}$ is the image, under a coding, of a fixed point of a k-uniform morphism. We point out that a sequence $\mathbf{b} = (b_n)_{n\geq 0}$ with values in Σ is a fixed point of a k-uniform morphism φ if and only if for all $n \geq 0$ we have

$$\varphi(b_n) = b_{kn} b_{kn+1} \cdots b_{kn+k-1}, \tag{1.2}$$

as proved in [4, Lemma 6.3.1]. Then by Theorem 1.1 any sequence $\mathbf{a} = (a_n)_{n\geq 0}$ of the form $\mathbf{a} = \tau(\mathbf{b})$, with τ a coding, is a k-automatic sequence.

As a continuation of Examples 1.1 and 1.2 we construct a 2-uniform morphism which yields the Thue–Morse sequence $\mathbf{t} = (t_n)_{n\geq 0}$.

Example 1.3. The formula (1.1) gives $t_0 = 1$ and

$$t_{2n} = t_n, \qquad t_{2n+1} = -t_n \tag{1.3}$$

for all $n \geq 0$. Define the 2-uniform morphism $\varphi\colon \{1,-1\}^* \to \{1,-1\}^*$ by

$$\varphi(1) = 1\ -1, \qquad \varphi(-1) = -1\ 1.$$

Then by (1.2) it is clear that

$$\mathbf{t} = \varphi^\omega(1) = \varphi(\mathbf{t}).$$

Alternatively, the form of φ could be derived from the 2-DFAO in Figure 1.1 by identifying the states q_0, q_1 with their respective output $1, -1$ and putting $\varphi(q) = \delta(q,0)\delta(q,1)$ for $q = q_0, q_1$ (this is in line with the proof of Theorem 1.1).

1.1.3 The k-kernel and regular sequences

In this subsection we give yet another characterization of k-automatic sequences using the notion of a k-kernel, whose natural generalization leads to the definition of regular sequences. Let $\mathbf{a} = (a_n)_{n\geq 0}$ be an infinite sequence and let $k \geq 2$ be an integer.

Definition 1.4. The k-kernel of the sequence $\mathbf{a} = (a_n)_{n\geq 0}$ is the set

$$\mathcal{K}_k(\mathbf{a}) = \{(a_{k^j n+i})_{n\geq 0} : j \geq 0,\ 0 \leq i \leq k^j - 1\}.$$

The following result (see [4, Theorem 6.6.2]) can be used as an equivalent definition of automatic sequences.

Theorem 1.2. *The sequence $\mathbf{a} = (a_n)_{n\geq 0}$ is k-automatic if and only if $\mathcal{K}_k(\mathbf{a})$ is finite.*

Example 1.4. The 2-kernel of the Thue–Morse sequence $\mathbf{t} = (t_n)_{n\geq 0}$ is

$$\mathcal{K}_k(\mathbf{t}) = \{(t_n)_{n\geq 0}, (t_{2n+1})_{n\geq 0}\},$$

which follows straight from the relations (1.3).

Allouche and Shallit [3] generalized this description in order to encompass a number of interesting sequences which take infinitely many distinct values (and thus are not automatic). Let R be a $\mathbb{Z}$-module and let $k \geq 2$ be an integer. We can treat the sequences in $\mathcal{K}_k(\mathbf{a})$ as elements of the $\mathbb{Z}$-module $R^{\mathbb{N}}$.

Definition 1.5. A sequence $(a_n)_{n\geq 0}$ with values in R is *k-regular* if the $\mathbb{Z}$-module generated by its k-kernel is finitely generated.

We say that a sequence is *regular* if it is k-regular for some $k \geq 2$, and otherwise nonregular.

Example 1.5. Let $s_k(n)$ be the sum of base-k digits of an integer $n \geq 0$. Then we have

$$s_k(k^j n + i) = s_k(n) + s_k(i)$$

for all integers $n \geq 0, j \geq 0$, and i such that $0 \leq i \leq k^j - 1$. It follows that the $\mathbb{Z}$-module generated by the k-kernel of $(s_k(n))_{n\geq 0}$ is generated by two elements: $(s_k(n))_{n\geq 0}$ and the constant sequence with all terms equal to 1. Hence, the considered sequence is k-regular.

The next result [3, Theorem 2.3] exhibits a connection between regular and automatic sequences.

Theorem 1.3. *A sequence is k-regular and takes on only finitely many values if and only if it is k-automatic.*

1.1.4 Properties of automatic and regular sequences

We now give some results on automatic and regular sequences, which will be useful in the later chapters. The proofs for automatic sequences can mostly be found in [4, Sections 5.4 and 6.8], and for regular sequences – in [3, Section 2] (if not, we provide a reference). In the results below whenever a sequence is assumed or claimed to be regular, we implicitly assume that the set containing its terms is a ring with the structure of a $\mathbb{Z}$-module.

We start with the trivial case of *eventually periodic* sequences. We call a sequence $(a_n)_{n\geq 0}$ eventually periodic if there exists an integer $T > 0$ such that $a_{n+T} = a_n$ for all n sufficiently large. We refer to any such T as a period of $(a_n)_{n\geq 0}$. Unless specified otherwise, throughout the thesis we do not assume that T is minimal with this property.

Theorem 1.4. *Let $(a_n)_{n\geq 0}$ be an eventually periodic sequence. Then it is k-automatic (and thus k-regular) for all $k \geq 2$.*

In what follows we assume that $k \geq 2$ is a fixed integer.

Theorem 1.5. *Let $t \geq 1$ be an integer. Then the sequence $(a_n)_{n\geq 0}$ is k-automatic (resp. k-regular) if and only if it is k^t-automatic (resp. k^t-regular).*

In the automatic case this is [4, Theorem 6.6.4].

On the other hand, a remarkable result by Cobham [18] shows that when k and l are multiplicatively independent, then only eventually periodic sequences can be simultaneously k- and l-automatic. Recall that two positive integers k, l are said to be *multiplicatively independent* if their only common integer power is 1.

Theorem 1.6 (Cobham)**.** *Let $k, l \geq 2$ be multiplicatively independent integers. If a sequence $(a_n)_{n\geq 0}$ is both k-automatic and l-automatic, then it is eventually periodic.*

This was later generalized by Bell [7] to regular sequences. From his results we can extract the following.

Theorem 1.7. *Let $k, l \geq 2$ be multiplicatively independent integers. If a sequence $(a_n)_{n\geq 0}$ is both k-regular and l-regular, then it is k-regular for all $k \geq 2$ and satisfies a linear recurrence.*

We now focus on simple closure properties of automatic and regular sequences.

Theorem 1.8. *Let $(a_n)_{n\geq 0}$ differ from a k-automatic (resp. k-regular) sequence by only a finite number of terms. Then $(a_n)_{n\geq 0}$ is also k-automatic (resp. k-regular).*

The next result relates k-automaticity (k-regularity) of a sequence and its subsequences along arithmetic progressions.

Theorem 1.9. *If the sequence $(a_n)_{n\geq 0}$ is k-automatic (resp. k-regular), then for all integers $b, c \geq 0$ the sequence $(a_{bn+c})_{n\geq 0}$ is k-automatic (resp. k-regular). Conversely, let $b \geq 0$ be an integer and let $(a_n)_{n\geq 0}$ be a sequence such that $(a_{bn+c})_{n\geq 0}$ is k-automatic (resp. k-regular) for all $c = 0, 1, \ldots, b-1$. Then $(a_n)_{n\geq 0}$ is k-automatic (resp. k-regular).*

Remark 1.10. By [3, Remark on p. 169], if we assign arbitrary values (for example zeros) to the terms a_n with negative indices, then the claim of Theorem 1.9 holds for any $c \in \mathbb{Z}$.

The following result shows that functions of k-automatic sequences are also k-automatic.

Theorem 1.11. *Let $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$ be k-automatic sequences with values in Δ, Δ', respectively, and let Λ be a finite set. Then the following sequences are also k-automatic:*

(i) $(\rho(a_n))_{n\geq 0}$, *where $\rho\colon \Delta^* \to \Lambda^*$ is any coding;*

(ii) $(a_n, b_n)_{n\geq 0}$;

(iii) $(f(a_n, b_n))_{n\geq 0}$, *where $f\colon \Delta \times \Delta' \to \Lambda$ is any function.*

As a corollary, if $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$ take values in the same ring and are k-automatic, then $(a_n + b_n)_{n\geq 0}$ and $(a_n b_n)_{n\geq 0}$ are k-automatic as well. This is also true for k-regular sequences.

Theorem 1.12. *Let $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$ be k-regular sequences taking values in a commutative ring R and let $\lambda \in R$. Then the sequences $(\lambda a_n)_{n\geq 0}$, $(a_n + b_n)_{n\geq 0}$ and $(a_n b_n)_{n\geq 0}$ are also k-regular.*

Treating the terms of the sequence $(n)_{n\geq 0}$ as elements of a commutative ring R, we can easily check that this sequence is k-regular for every $k \geq 2$. Using Theorem 1.12 we can deduce that a similar statement holds for polynomials evaluated at consecutive integers.

Corollary 1.13. *Let R be a commutative ring and $f \in R[X]$. Then the sequence $(f(n))_{n\geq 0}$ is k-regular for every $k \geq 2$.*

We point out that an arbitrary function of two k-regular sequences may not be k-regular (see for example [3, pp. 168–169]).

The following two results are an immediate consequence of the definition of a k-regular sequence, however we include them for lack of a better reference.

Proposition 1.14. *Let $(a_n)_{n\geq 0}$, $(b_n)_{n\geq 0}$ be sequences taking values in $\mathbb{Z}$-modules R, R', respectively. Then the sequence of pairs $(a_n, b_n)_{n\geq 0}$ is k-regular if and only if $(a_n)_{n\geq 0}$, $(b_n)_{n\geq 0}$ are both k-regular.*

Proposition 1.15. *Let R, R' be $\mathbb{Z}$-modules and let $\phi\colon R \to R'$ be a homomorphism. If $(a_n)_{n\geq 0}$ is a k-regular sequence with values in R, then $(\phi(a_n))_{n\geq 0}$ is also a k-regular sequence.*

Proposition 1.15 in conjunction with Theorem 1.3 implies the following corollary.

Corollary 1.16. *Let $(a_n)_{n\geq 0}$ be a k-regular sequence of integers. Then for all integers $m \geq 1$ the sequence $(a_n \bmod m)_{n\geq 0}$ is k-automatic.*

It turns out that regular sequences exhibit at most polynomial growth.

Theorem 1.17. *Let $(a_n)_{n\geq 0}$ be a k-regular sequence with values in $\mathbb{C}$. Then there exists a constant $C > 0$ such that $|a_n| = O(n^C)$.*

Finally, we give an interesting property of automatic sequences taking integer values; this is [4, Corollary 6.9.3].

Theorem 1.18. *Let $(a_n)_{n\geq 0}$ be a k-automatic sequence with values in $\mathbb{Z}$ and let $l \geq 2$ be an integer. Then the following sequences are also k-automatic:*

(i) *the running sum sequence $((\sum_{m=0}^{n} a_m) \bmod l)_{n\geq 0}$;*

(ii) *the running product sequence $((\prod_{m=0}^{n} a_m) \bmod l)_{n\geq 0}$.*

1.1.5 The frequencies of letters

Following the description in [4, Chapter 8], we demonstrate how to compute the frequencies of letters in automatic sequences. Let $\mathbf{a} = (a_n)_{n\geq 0}$ be an infinite word made from letters in Σ and let $c \in \Sigma$ be a fixed letter. We define the *frequency* of c in $\mathbf{a}$ as the limit

$$\mathrm{Freq}_{\mathbf{a}}(c) = \lim_{n\to\infty} \frac{|a_0 a_1 \cdots a_{n-1}|_c}{n} = \lim_{n\to\infty} \frac{\#\{0 \leq m \leq n-1 : a_m = c\}}{n},$$

if it exists. In other words, $\mathrm{Freq}_{\mathbf{a}}(c)$ is the natural density of the set $\{n \geq 0 : a_n = c\}$. The frequency of a letter in an infinite word or even in an automatic sequence may not exist, as shown in [4, Example 8.1.2]. We will however consider a class of automatic sequences generated by primitive morphisms (defined below) for which all the frequencies exist.

Assume that the sequence $(a_n)_{n\geq 0}$ with values in $\Sigma = \{c_1, \ldots, c_d\}$ is a fixed point of a k-uniform morphism $\varphi\colon \Sigma^* \to \Sigma^*$. The morphism φ is called *primitive* if there exists an integer $n \geq 1$ such that for all $c \in \Sigma$ the word $\varphi^n(c)$ contains all letters from Σ. We associate with φ the *incidence matrix* $M(\varphi) = [m_{i,j}]_{1\leq i,j\leq d}$, where $m_{i,j} = |\varphi(c_j)|_{c_i}$. One can verify that $M(\varphi)^n = M(\varphi^n)$, and thus the morphism φ is primitive if and only if $M(\varphi)$ is *primitive*, i.e., there exists an integer $n \geq 1$ such that $M(\varphi)^n$ has all entries positive. Under the assumption of primitivity, the matrix $M(\varphi)$ has a positive eigenvalue r of multiplicity one such that any other complex eigenvalue λ satisfies $|\lambda| \leq r$ and there exists an eigenvector with all entries positive

corresponding to r (see [42, Theorem 8.4.4]). This eigenvalue is called the Perron–Frobenius eigenvalue. In particular, if a primitive matrix is row- or column-stochastic, then by [4, Theorem 8.3.13] its Perron–Frobenius eigenvalue equals 1. The following result ([4, Theorem 8.4.7]) provides a simple way to compute the frequencies of letters using the Perron–Frobenius eigenvalue of $M(\varphi)$.

Theorem 1.19. *Let $\mathbf{a} = (a_n)_{n\geq 0}$ be a fixed point of a primitive uniform morphism φ. Then the frequencies of all letters exist and are nonzero. Furthermore, the vector of frequencies of the letters in $\mathbf{a}$ is the positive normalized eigenvector corresponding to the Perron–Frobenius eigenvalue of the incidence matrix $M(\varphi)$.*

As a consequence, if $M(\varphi)$ is a row-stochastic matrix multiplied by a positive constant, then $\mathrm{Freq}_{\mathbf{a}}(c_i) = 1/d$ for all $i = 1, \ldots, d$.

We point out that if $\mathbf{b} = (b_n)_{n\geq 0}$ is a coding of $\mathbf{a}$, then the computation of the frequencies of letters in $\mathbf{b}$ is almost immediate once the frequencies of letters in $\mathbf{a}$ are known.

1.2 p-adic numbers

p-adic numbers were first introduced by Hensel near the end of the nineteenth century. They have since then evolved into a powerful tool in number theory, allowing the study of congruences and divisibility properties by means of analytic methods. p-adic numbers have also found applications in many other areas of mathematics and even physics. The following presentation is based on the books of Gouvêa [37], Koblitz [45], Robert [64], and the notes of Conrad [22, 23, 21].

1.2.1 The field of p-adic numbers

Let p be a fixed prime number. We start with the basic definitions of p-adic valuation and p-adic norm.

Definition 1.6. The *p-adic valuation* ν_p is defined for $n \in \mathbb{Z}$ by

$$\nu_p(n) = \begin{cases} \max\{v \geq 0 : p^v | n\} & \text{if } n \neq 0, \\ +\infty & \text{if } n = 0. \end{cases}$$

For $n/m \in \mathbb{Q}$ with $n, m \in \mathbb{Z} \setminus \{0\}$, we define

$$\nu_p\left(\frac{n}{m}\right) = \nu_p(n) - \nu_p(m).$$

Definition 1.7. The *p-adic norm* $|\cdot|_p$ is defined for $x \in \mathbb{Q}$ by

$$|x|_p = \begin{cases} p^{-\nu_p(x)} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

It can be checked that $|\cdot|_p$ is indeed a norm. In the following two results we give some elementary properties of the p-adic valuation and p-adic norm.

Theorem 1.20. *For all $x, y \in \mathbb{Q}$ we have the following:*

(i) $\nu_p(xy) = \nu_p(x) + \nu_p(y)$;

(ii) $\nu_p(x+y) \geq \min\{\nu_p(x), \nu_p(y)\}$;

(iii) *if $\nu_p(x) \neq \nu_p(y)$, then $\nu_p(x+y) = \min\{\nu_p(x), \nu_p(y)\}$.*

Theorem 1.21. *For all $x, y \in \mathbb{Q}$ we have the following:*

(i) $|xy|_p = |x|_p |y|_p$;

(ii) $|x+y|_p \leq \max\{|x|_p, |y|_p\}$;

(iii) *if $|x|_p \neq |y|_p$, then $|x+y|_p = \max\{|x|_p, |y|_p\}$.*

Property (ii) of Theorem 1.21 says that $|\cdot|_p$ is a *non-Archimedean* norm on $\mathbb{Q}$.

The field of rational numbers equipped with the metric $d_p(x,y) = |x-y|_p$ is not complete. We can however take the completion of $\mathbb{Q}$ with respect to d_p and (uniquely) extend ν_p and $|\cdot|_p$ to the resulting field while retaining their properties stated above.

Definition 1.8. The completion of $\mathbb{Q}$ with respect to d_p is called the *field of p-adic numbers* $\mathbb{Q}_p$. The set $\mathbb{Z}_p = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}$ is the *ring of p-adic integers*.

The family of clopen balls

$$\overline{B}(a,r) = \{x \in \mathbb{Q}_p : |x-a|_p \leq r\}$$

with $a \in \mathbb{Q}_p$ and $r > 0$ forms a basis of the topology on $\mathbb{Q}_p$. In particular, we have $\overline{B}(0,1) = \mathbb{Z}_p$. The ring of p-adic integers $\mathbb{Z}_p$ is a compact set and can be equivalently defined as the completion of $\mathbb{Z}$ with respect to d_p. The set $\mathbb{N}$ of nonnegative integers is dense in $\mathbb{Z}_p$.

The elements of $\mathbb{Q}_p$ are conveniently described through their *p-adic expansions*. We can write $x \in \mathbb{Q}_p$ in the form

$$x = \cdots + x_2 p^2 + x_1 p + x_0 + x_{-1} p^{-1} + \cdots + x_{-l} p^{-l},$$

where $l \geq 0$ and the coefficients x_i are digits from the set $\{0, 1, \ldots, p-1\}$. The least v such that $x_v \neq 0$ is equal to $\nu_p(x)$. If $x \in \mathbb{Z}_p$, then the p-adic expansion of x only contains terms with nonnegative powers of p. Reduction modulo p^j with $j \geq 0$ an integer is defined for elements of $x = \cdots + x_1 p + x_0 \in \mathbb{Z}_p$ by

$$x \bmod p^j = x_{j-1} p^{j-1} + \cdots + x_1 p + x_0.$$

Arithmetic operations involving p-adic expansions can be performed in a standard way using carries. A number $x \in \mathbb{Q}_p$ is rational if and only if its p-adic expansion is eventually periodic (see [23, Theorem 3.1]).

We will mostly be dealing with the analytic properties of power series defined on $\mathbb{Z}_p$ with coefficients in finite extensions of $\mathbb{Q}_p$, thus it will be convenient to have a complete and algebraically closed field containing $\mathbb{Q}_p$. We can fix an algebraic closure $\mathbb{Q}_p^{\mathrm{alg}}$ of $\mathbb{Q}_p$ and further extend the definition of $|\cdot|_p$ and ν_p to $\mathbb{Q}_p^{\mathrm{alg}}$ (in a unique way). It turns out that $\mathbb{Q}_p^{\mathrm{alg}}$ is, again, not complete. However, by taking the completion of $\mathbb{Q}_p^{\mathrm{alg}}$ with respect to d_p we obtain the field $\mathbb{C}_p$, which is both complete and algebraically closed. The construction of this field is described in [64, Chapter 3] and [45, Chapter 3] (in the latter book this field is denoted by Ω). The topology of $\mathbb{C}_p$ can be described in the same way as before – in terms of a basis of clopen balls.

We now discuss a generalization of the above construction of $\mathbb{Z}_p$ and $\mathbb{Q}_p$ with p replaced by any integer $b \geq 2$. The description is loosely based on the lecture notes by Ershov [32, 33].

Any nonzero rational number x can be written in the form

$$x = b^v \frac{q}{r}, \tag{1.4}$$

where v, q, r are integers such that r is coprime to both q and b, and q is not divisible by b. Moreover, such a representation is unique up to changing the sign of q and r. We define the *b-adic valuation* of x, denoted by $\nu_b(x)$, to be the number v in (1.4) and also put $\nu_b(0) = +\infty$. In particular, for any integer n we have

$$\nu_b(n) = \begin{cases} \max\{v \geq 0 : b^v | n\} & \text{if } n \neq 0, \\ +\infty & \text{if } n = 0. \end{cases} \tag{1.5}$$

We remark that ν_b is not a valuation in the strict sense in the case of b composite, since it satisfies properties (ii) and (iii) of Theorem 1.20 but only a weaker version of property (i), namely $\nu_b(xy) \geq \nu_b(x) + \nu_b(y)$. Nevertheless, ν_b gives rise to the b-adic norm $|\cdot|_b$ defined by $|x|_b = b^{-\nu_b(x)}$. Using this as a starting point we may construct in the same fashion the rings $\mathbb{Z}_b$ of b-adic integers and $\mathbb{Q}_b$ of b-adic numbers as completions of $\mathbb{Z}$ and $\mathbb{Q}$ with respect to $|\cdot|_b$. We point out that unless b is a prime power, $\mathbb{Q}_b$ is not a field (in particular, it contains zero divisors). The elements of $\mathbb{Z}_b$ and $\mathbb{Q}_b$ can be represented in the form of (possibly infinite) b-adic expansions with digits $0, 1, \ldots, b-1$. If n is a nonnegative integer, then its b-adic expansion is finite and coincides with the usual base-b expansion. We may reduce the elements of $\mathbb{Z}_b$ modulo b^j with $j \geq 0$ an integer in the same way as in the p-adic case.

It turns out that b-adic numbers can also be represented by tuples of p-adic numbers with p running over the set of prime divisors of b. Let

$$b = p_1^{l_1} \cdots p_s^{l_s},$$

be the prime factorization of b and write $b_i = p_i^{l_i}$ for $i = 1, \ldots, s$. In [33] a natural isomorphism π_b between $\mathbb{Z}_b$ and the product ring $\mathbb{Z}_{p_1} \times \cdots \times \mathbb{Z}_{p_s}$ is constructed. We can extend it to an isomorphism between $\mathbb{Q}_b$ and $\mathbb{Q}_{p_1} \times \cdots \times \mathbb{Q}_{p_s}$ by setting

$$\pi_b(x) = b^{\nu_b(x)} \pi(b^{-\nu_b(x)} x)$$

for any $x \in \mathbb{Q}_b$, using the fact that $b^{-\nu_b(x)}x \in \mathbb{Z}_b$.

Below we list a few important properties of π_b, which follow directly from the aforementioned construction:

(i) if $x \in \mathbb{Q}$, then $\pi_b(x) = (x, \ldots, x)$;

(ii) if $\pi_b(x) = (y_1, \ldots, y_s)$ for some $x \in \mathbb{Q}_b$, then

$$\nu_b(x) = \min_{1 \le i \le s} \nu_{b_i}(y_i) = \min_{1 \le i \le s} \left\lfloor \frac{\nu_{p_i}(y_i)}{l_i} \right\rfloor; \tag{1.6}$$

(iii) if $\pi_b(x) = (y_1, \ldots, y_s)$ for some $x \in \mathbb{Z}_b$, then $x \bmod b^j$ is the integer $n \in \{0, 1, \ldots, b^j - 1\}$ satisfying

$$n \equiv y_i \pmod{b_i^j}$$

for $i = 1, \ldots, s$.

We point out that the property (iii) can in fact be used to construct π_b.

In the following chapters we will identify $\mathbb{Z}_b$ and $\mathbb{Q}_b$ with $\mathbb{Z}_{p_1} \times \cdots \times \mathbb{Z}_{p_s}$ and $\mathbb{Q}_{p_1} \times \cdots \times \mathbb{Q}_{p_s}$, respectively, through π_b. All the operations will be performed on elements of the form $(y_1, \ldots, y_s)$ with $y_i \in \mathbb{Q}_{p_i}$, while b-adic expansions will come up in relation to the interpretation of certain functions describing digits. In particular, we use (1.6) above as the definition of ν_b on the product ring $\mathbb{Q}_{p_1} \times \cdots \times \mathbb{Q}_{p_s}$ and (iii) as the definition of reduction modulo b^j on $\mathbb{Z}_{p_1} \times \cdots \times \mathbb{Z}_{p_s}$. Moreover, due to property (i) we can treat $\mathbb{Q}$ as the subring of $\mathbb{Q}_{p_1} \times \cdots \times \mathbb{Q}_{p_s}$ via the diagonal embedding, and simply write x instead of $(x, \ldots, x)$ to denote x rational. We do not pursue this topic further and focus on p-adic numbers with p prime.

1.2.2 p-adic analysis

We now proceed to recall some facts from p-adic analysis, which will be used in the later chapters. To begin with, we give a general result concerning the existence of zeros of continuous functions in $\mathbb{Z}_p$, which is a consequence of the Bolzano–Weierstrass Theorem.

Theorem 1.22. *Let $f\colon \mathbb{Z}_p \to \mathbb{C}_p$ be a continuous function. Then f has a zero in $\mathbb{Z}_p$ if and only if f has a zero modulo p^m for all integers $m \ge 0$, i.e., for each $m \ge 0$ there exists $x \in \mathbb{Z}_p$ such that $|f(x)|_p \le p^{-m}$.*

Since $\mathbb{N}$ is dense in $\mathbb{Z}_p$, it follows that a continuous function f has a root in $\mathbb{Z}_p$ if and only if the sequence $(\nu_p(f(n)))_{n \ge 0}$ is not bounded from above. By arguments similar as in classical analysis, p-adically continuous functions include polynomials and power series (within their disc of convergence).

Below we state a version of the famous Hensel's Lemma (adapted from [21, Theorem 4.1]), which gives an effective condtition for the existence of a root of a polynomial with p-adic coefficients. Here for a polynomial $f \in \mathbb{Z}_p[X]$ we let f' denote the derivative of f, given by the usual formula.

Theorem 1.23 (Hensel's Lemma)**.** *Let $f \in \mathbb{Z}_p[X]$ and assume that $x_0 \in \mathbb{Z}_p$ satisfies*

$$|f(x_0)|_p < |f'(x_0)|_p^2.$$

Then there exists a unique root $x \in \mathbb{Z}_p$ of f such that $|x - x_0|_p < |f'(x_0)|_p$.

We now turn our attention to functions given by power series. Let

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n, \tag{1.7}$$

where $a, a_0, a_1, \ldots \in \mathbb{C}_p$. We restrict ourselves to the case where $x \in \mathbb{Z}_p$; a similar discussion can be carried out when the variable x belongs to $\mathbb{C}_p$. We define the *radius of convergence* $R \in [0, +\infty) \cup \{+\infty\}$ by the formula

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|_p^{1/n}}.$$

The series $f(x)$ converges for $x \in \mathbb{Z}_p$ such that $|x-a|_p < R$, and diverges when $|x-a|_p > R$. Unlike in classical analysis, $f(x)$ either converges for all x such that $|x-a|_p = R$, or does not converge for any such x. The *disc of convergence* D of f is either $D = \{x \in \mathbb{Z}_p : |x-a|_p \leq R\}$ or $D = \{x \in \mathbb{Z}_p : |x-a|_p < R\}$, depending on which of the above conditions holds. These considerations lead to the following definition.

Definition 1.9. Let $B \subset \mathbb{Z}_p$ be an open set. We say that f is *strictly analytic* on B if it is given by a power series of the form (1.7) and the disc of convergence of f contains B. We say that f is *locally analytic* on B if for every $b \in B$ there exists a neighborhood B_b of b such that f is strictly analytic on B_b.

The following summary of properties of p-adic power series is based mainly on [22, Section 7]. Assume that $f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$ has a positive radius of convergence R. Then f is uniformly continuous on any ball $\overline{B}(a, r)$ contained within its disc of convergence D. For power series strictly analytic on $\mathbb{Z}_p$ we can reformulate this fact in the following way.

Proposition 1.24. *Let f be strictly analytic on $\mathbb{Z}_p$. Then for every integer $M \geq 0$ there exists an integer $T \geq 0$ such that for all $x, y \in \mathbb{Z}_p$ we have*

$$\nu_p(f(p^T y + x) - f(x)) \geq M.$$

We define the derivatives $f^{(k)}$ of f as in the classical case, namely $f^{(0)} = f$,

$$f'(x) = f^{(1)}(x) = \sum_{n=1}^{\infty} n a_n (x-a)^{n-1},$$

and inductively

$$f^{(k+1)} = (f^{(k)})'$$

for $k \geq 1$. Each derivative has the same disc of convergence as the original series. The coefficients of f can be expressed through the derivatives $f^{(k)}$ in the following way:

$$a_n = \frac{f^{(n)}(a)}{n!}. \tag{1.8}$$

For $b \in D$ we may write f as a power series centered at b:

$$f(x) = \sum_{n=0}^{\infty} b_n (x-b)^n,$$

for some coefficients $b_0, b_1, \ldots \in \mathbb{C}_p$, which can be written explicitly using (1.8) with a replaced by b. This new power series has the same disc of convergence D. As a consequence, the center of a power series which is strictly analytic on B may be chosen to be any point of B. We note that the definition of the derivative does not depend on the choice of the center of the disc of convergence.

We now recall an important result of Strassman concerning the number of zeros of a strictly analytic function on $\mathbb{Z}_p$. a proof for power series centered at 0 with coefficients in $\mathbb{Q}_p$ can be found in [37, Theorem 4.4.6] and is easily extended to power series centered at any $a \in \mathbb{Z}_p$ with coefficients in $\mathbb{C}_p$.

Theorem 1.25 (Strassman)**.** *Let f be a nonzero power series strictly analytic on $\mathbb{Z}_p$. Then f has only finitely many zeros in $\mathbb{Z}_p$.*

By [37, Corollary 4.4.7], the roots $\alpha_1, \ldots, \alpha_m \in \mathbb{Z}_p$ of f can be extracted so that

$$f(x) = \left(\prod_{i=1}^{m} (x - \alpha_i) \right) g(x), \tag{1.9}$$

where g is strictly analytic on $\mathbb{Z}_p$ and has no roots in $\mathbb{Z}_p$.

Following [45, pp. 78–81], we now take a closer look at two special power series: the *p-adic exponential* $\exp_p$ and the *p-adic logarithm* $\log_p$. Define

$$\exp_p(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

This series has the disc of convergence $\{x \in \mathbb{Z}_p : |x|_p < p^{-1/(p-1)}\}$; this is a consequence of the following formula proved by Legendre [47].

Theorem 1.26 (Legendre)**.** *Let p be a prime number. Then for all integers $n \geq 0$ we have*

$$\nu_p(n!) = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor = \frac{n - s_p(n)}{p-1},$$

where $s_p(n)$ is the sum of the base-p digits of n.

Also let

$$\log_p(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(x-1)^n.$$

The disc of convergence of $\log_p$ is $\{x \in \mathbb{Z}_p : |x-1|_p < 1\}$. For any x, y in the discs of convergence of the respective functions we have the usual identities

$$\begin{aligned} \exp_p(x+y) &= \exp_p(x)\exp_p(y), \\ \log_p(xy) &= \log_p(x) + \log_p(y). \end{aligned}$$

Moreover, for $x \in \mathbb{Z}_p$ such that $|x|_p < p^{-1/(p-1)}$, we have

$$\begin{aligned} \exp_p(\log_p(1+x)) &= 1+x, \\ \log_p(\exp_p(x)) &= x. \end{aligned}$$

Writing $D_p = \{x \in \mathbb{Z}_p : |x|_p < p^{-1/(p-1)}\}$, we can sum up these properties as follows: $\exp_p$ and $\log_p$ are mutually inverse isomorphisms of the additive group D_p and the multiplicative group $1 + D_p$.

One may also consider exponentiation in other bases. For a fixed $a \in 1 + D_p$ we can define a^x as the power series in x given by

$$a^x = \exp_p(x \log_p(a)). \tag{1.10}$$

This series converges for all $x \in \mathbb{Z}_p$. When $x = n$ is a nonnegative integer, then the value of (1.10) coincides with raising a to the power n.

1.2.3 Interpolation of linear recurrence sequences

To conclude this chapter we are going to recall how to express the terms of integer linear recurrence sequences by means of p-adic power series, which will be a recurring theme in the thesis. The description below is based on the proofs of the famous Skolem–Mahler–Lech Theorem found in the paper by Hansel [41] and the book of Everest, van der Poorten, Shparlinski, and Ward [34, Section 2.1].

First, we recall some standard terminology, as in [34, Section 1.1]. Consider a sequence $(s_n)_{n\geq 0}$ of integers satisfying

$$s_{n+k} = \sum_{i=0}^{k-1} a_i s_{n+i} \tag{1.11}$$

for all $n \geq 0$, where $k \geq 1$ and $a_0, \ldots, a_{k-1}$ are fixed integers. We say that (1.11) is a *linear recurrence relation* and $(s_n)_{n\geq 0}$ is a *linear recurrence sequence*. The relation (1.11) is said to be of *order* k. In the sequel we only consider relations such that a_0 is nonzero.

The polynomial

$$P(x) = x^k - \sum_{i=0}^{k-1} a_i x^i$$

is called the *characteristic polynomial* of the relation (1.11). If this is the linear recurrence relation of the minimal order satisfied by a sequence $(s_n)_{n\geq 0}$, then we say that P is the *minimal polynomial* of this sequence. The sequence $(s_n)_{n\geq 0}$ is said to be *degenerate* if its minimal polynomial has a pair of distinct roots whose ratio is a root of unity. Otherwise, the sequence is called *nondegenerate*.

We now begin the construction of the interpolating functions. Let $(s_n)_{n\geq 0}$ be a linear recurrence sequence satisfying the relation (1.11) with characteristic polynomial P. Let p be a prime not dividing a_0 and let $\alpha_1, \ldots, \alpha_r$ be the (distinct) roots of P in $\mathbb{C}_p$ with respective multiplicities $l_1, \ldots, l_r$, summing up to k. Then for all integers $n \geq 0$ we have

$$s_n = \sum_{j=1}^{r} \beta_j(n)\alpha_j^n,$$

for some $\beta_1, \ldots \beta_r, \in \mathbb{C}_p[X]$, where $\deg \beta_j = l_j - 1$.

We cannot yet use (1.10) in order to express s_n in terms of a p-adic analytic function, since in general $\alpha_j \notin 1 + D_p$. We will however see that this can be done for subsequences of $(s_n)_{n\geq 0}$ along certain arithmetic progressions.

Consider the *companion matrix* of the relation (1.11), namely

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{k-1} \end{bmatrix}.$$

Clearly, P is the characteristic polynomial of C and for all $n \geq 0$ we have

$$C \begin{bmatrix} s_n \\ s_{n+1} \\ \vdots \\ s_{n+k-1} \end{bmatrix} = \begin{bmatrix} s_{n+1} \\ s_{n+2} \\ \vdots \\ s_{n+k} \end{bmatrix}.$$

The matrix C is invertible modulo p, so there exists an integer $m > 0$ such that $C^m \equiv I_k \pmod{p}$, where I_k is the $k \times k$ identity matrix. This implies that $(s_n)_{n\geq 0}$ is periodic modulo p with period m. The characteristic polynomial of the matrix $C^m - I_k$ is

$$\prod_{j=1}^{r}(x - (\alpha_j^m - 1))^{l_j} = x^k + \sum_{i=0}^{k-1} b_i x^i, \tag{1.12}$$

for certain $b_0, b_1, \ldots, b_{k-1} \in \mathbb{Z}$. Observe that $\nu_p(b_i) \geq k - i$. This implies that the inequalities $\nu_p(\alpha_j^m - 1) \geq 1$ hold for $j = 1, \ldots, r$, since otherwise we would have

$$\nu_p((\alpha_j^m - 1)^k) < \nu_p\left(\sum_{i=0}^{k-1} b_i(\alpha_j^m - 1)^i\right),$$

which contradicts the fact that $\alpha_j^m - 1$ is a root of the polynomial in (1.12). Thus, $\alpha_j^m \in 1 + D_p$ if $p \neq 2$. In the case $p = 2$ it may occur that $\nu_2(\alpha_j^m - 1) = 1$ for some j, and hence $\alpha_j^m \notin 1 + D_2$. Nevertheless, we have $C^{2m} \equiv I_k \pmod 4$, and $\alpha_j^{2m} \in 1 + D_2$ follows in a similar fashion.

In either case, we have established the existence of an integer $\pi > 0$ such that $\alpha_j^\pi \in 1 + D_p$ for $j = 1, \ldots, r$. Define the p-adic analytic functions $f_i \colon \mathbb{Z}_p \to \mathbb{Q}_p$ for $i = 0, 1, \ldots, \pi - 1$ by

$$f_i(x) = \sum_{j=1}^{r} \beta_j(x) \alpha_j^{\pi x + i} = \sum_{j=1}^{k} \beta_j(x) \alpha_j^i \exp_p(x \log_p(\alpha_j^\pi)).$$

We then have

$$s_{\pi n + i} = f_i(n)$$

for all $i = 0, 1, \ldots \pi - 1$ and $n \geq 0$. Therefore, the functions $f_1, \ldots, f_{\pi - 1}$ evaluated at nonnegative integers collectively interpolate the sequence $(s_n)_{n \geq 0}$.

2. Last nonzero digits of factorials

This chapter, mainly based on the author's paper [70], is devoted to the study of the properties of last nonzero digits in base-b expansions of consecutive factorials. We give a necessary and sufficient condition on b for automaticity of the resulting sequence. Moreover, we explicitly compute how often each finite string of d digits not ending with 0 is the final such string in the base b-expansion of $n!$.

2.1 Introduction

Let $b \geq 2$ and $d \geq 1$ be fixed integers. We introduce the function $\ell_{b,d}\colon \mathbb{Z} \to \{0, 1, \ldots, b^d - 1\}$, defined by

$$\ell_{b,d}(m) = \begin{cases} b^{-\nu_b(m)} m \bmod b^d & \text{if } m \neq 0, \\ 0 & \text{if } m = 0, \end{cases}$$

where ν_b denotes the b-adic "valuation" as in (1.5). To simplify the notation, in the case $d = 1$ we write ℓ_b instead of $\ell_{b,1}$. The value of $\ell_{b,d}(m)$ for $m \geq 0$ is precisely the integer whose base-b expansion is given by the last block of d digits not ending with zero in the base-b expansion of m. For example, we have

$$\ell_{6,2}(2400) = \ell_{6,2}([15040]_6) = [04]_6 = 4.$$

In short, we refer to the value $\ell_{b,d}(m)$ as *d last nonzero base-b digits of m* or simply *last nonzero digits of m* if the parameters b, d are not specified.

A number of authors have shown interest in studying last nonzero digits of factorials. Kakutani [43] considered the sequence $(\ell_{10}(n!))_{n\geq 0}$ in the setting of ergodic theory and essentially showed it is 5-automatic. Later, Dekking [25] constructed a 9-uniform morphism such that the sequence $(\ell_3(n!))_{n\geq 0}$ is its fixed point, thus proving it to be 3-automatic. The sequence $(\ell_{10}(n!))_{n\geq 0}$, together with $(\ell_{10}(n^n))_{n\geq 0}$ and $(\ell_{10}(F_n))_{n\geq 0}$, where $(F_n)_{n\geq 0}$ is the Fibonacci sequence, was again studied by Dresden [30, 31], who proved that the real numbers

$$\sum_{n=1}^{\infty} \frac{\ell_{10}(n!)}{10^n}, \qquad \sum_{n=1}^{\infty} \frac{\ell_{10}(n^n)}{10^n}, \qquad \sum_{n=1}^{\infty} \frac{\ell_{10}(F_n)}{10^n}$$

are transcendental. In his proof for factorials, Dresden implicitly used 5-automaticity of the sequence $(\ell_{10}(n!))_{n\geq 0}$. Deshouillers and Luca [28] investigated the sequence

$(\ell_{4,2}(n!))_{n\geq 0}$ (in a disguised form) and proved that a certain coding of this sequence is 2-automatic. Using its description in terms of uniform morphisms and Legendre's three-square theorem [47], they were able to conlude that

$$\#\{n \leq x : n! \text{ is a sum of three squares}\} = \frac{7}{8}x + O(x^{2/3}).$$

Deshouillers and Ruzsa [29] pointed out that the reasoning behind automaticity of $(\ell_{10}(n!))_{n\geq 0}$ can be extended to most bases b. In the smallest case where their observation does not apply, namely $b = 12$, they proved that the sequence $(\ell_{12}(n!))_{n\geq 0}$ coincides with a 3-automatic sequence on a set of n of asymptotic density 1. Moreover, they computed that for $a \in \{1, \ldots, 11\}$ the frequency of a in $(\ell_{12}(n!))_{n\geq 0}$ is $1/2$ if $a = 4, 8$, and otherwise 0. Further work by Deshouillers [26, 27] and Byszewski and Konieczny [13] revealed that the characteristic sequences of the sets $\{n : \ell_{12}(n!) = a\}$ for $a \in \{3, 4, 6, 8, 9\}$, and thus also $(\ell_{12}(n!))_{n\geq 0}$ itself, are not automatic. The question of automaticity of $(\ell_b(n!))_{n\geq 0}$ for general b was finally answered by Lipka [51]. In order to state his result, let

$$b = p_1^{l_1} \cdots p_s^{l_s}$$

be the prime factorization of b, where $p_1, \ldots, p_s$ are distinct primes and $l_1, \ldots, l_s$ are positive integers. If there are at least two prime factors, we additionally reorder the primes so that

$$l_1(p_1 - 1) \geq l_2(p_2 - 1) \geq \cdots \geq l_s(p_s - 1) \tag{2.1}$$

and

$$p_1 = \max\{p_i : l_i(p_i - 1) = l_1(p_1 - 1)\}. \tag{2.2}$$

We define the set

$$\mathcal{B} = \{b \geq 2 : s = 1 \text{ or } l_1(p_1 - 1) > l_2(p_2 - 1)\}.$$

Lipka proved that the condition $b \in \mathcal{B}$, hinted at in [29], is not only sufficient but also necessary for automaticity of $(\ell_b(n!))_{n\geq 0}$. He also provided an explicit construction of a $p_1^{l_1}$-DFAO generating this sequence in the automatic case.

Theorem 2.1 (Lipka)**.** *The sequence $(\ell_b(n!))_{n\geq 0}$ is p_1-automatic if $b \in \mathcal{B}$ and nonautomatic otherwise.*

We stress that being nonautomatic is a property much more difficult to establish than not being p_1-automatic.

In the present chapter we will extend Theorem 2.1 (and many other discussed results) to any number $d \geq 1$ of last nonzero digits of $n!$ and also prove that even if $(\ell_{b,d}(n!))_{n\geq 0}$ is not automatic, it coincides with an automatic sequence on a set of asymptotic density 1.

Theorem 2.2. *The sequence $(\ell_{b,d}(n!))_{n\geq 0}$ is p_1-automatic if $b \in \mathcal{B}$. It is not automatic if $b \notin \mathcal{B}$, however it coincides with a p_1-automatic sequence on a set of asymptotic density* 1.

A characterization of the considered sequence will be given in terms of a p_1-uniform morphism. This in turn allows us to use the method described in Section 1.1.5 to compute the frequencies of the letters $a = 1, \ldots, b^d - 1$ in the sequence $(\ell_{b,d}(n!))_{n\geq 0}$. It turns out that these frequencies are exactly the same in the subsequences along arithmetic progressions.

Theorem 2.3. *Let $a \in \{1, \ldots, b^d - 1\}$ be such that $b \nmid a$ and let $k \geq 1, l \geq 0$ be integers. The frequency of a in the sequence $(\ell_{b,d}((kn + l)!))_{n\geq 0}$ is*

$$\frac{1}{(p_1 - 1)l_1} p_1^{\nu_{p_1}(a) - l_1 d + 1}$$

if $(b/p_1^{l_1})^d \mid a$, and otherwise 0.

The results on frequencies in [28] (except for the error term) and [29] are special cases of Theorem 2.3. The proofs of Theorem 2.2 and Theorem 2.3 are provided in Section 2.3 and Section 2.5, respectively.

Remark 2.4. One may ask why we study last *nonzero* digits rather than, simply, last digits of base-b expansions of integer sequences. The main reason is that many integer sequences of interest are eventually periodic modulo b, and thus their last digits form sequences trivially k-automatic for every $k \geq 2$. This is true for linear recurrence sequences, $(n^n)_{n\geq 0}$ (proven by Hampel [40]), $(n!)_{n\geq 0}$, $(\binom{n}{m})_{n\geq 0}$ with m fixed (proven by Fray [36]), $(f(n))_{n\geq 0}$ for $f \in \mathbb{Z}[X]$, etc.

2.2 Properties of last nonzero digits

Fix the base $b \geq 2$ and the number of digits $d \geq 1$. In this section we describe some fundamental properties of last nonzero digits of b-adic numbers (see the end of Section 1.2.1). Although this degree of generality is only needed in the next chapter, we have decided to discuss the general case here for the sake of a more consistent description.

Let

$$b = p_1^{l_1} \cdots p_s^{l_s}$$

be the prime factorization of b, where $p_1, \ldots, p_s$ are distinct primes and $l_1, \ldots, l_s$ are positive integers. For each $i = 1, \ldots, s$ denote

$$b_i = p_i^{l_i},$$
$$q_i = \frac{b}{b_i},$$

and let r_i be an integer such that

$$q_i r_i \equiv 1 \pmod{b_i^d},$$

i.e., a multiplicative inverse of q_i modulo b_i^d. Then for any $x = (x_1, \ldots, x_s) \in \mathbb{Z}_b$ we can write explicitly

$$x \bmod b^d = \sum_{i=1}^{s} q_i^d r_i^d x_i \bmod b^d. \tag{2.3}$$

We now define in full generality the functions being the main object of interest in this and the following chapter. For any $x \in \mathbb{Q}_b$ we let

$$\mathcal{L}_b(x) = \begin{cases} b^{-\nu_b(x)}x & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$

where ν_b is defined according to the formula (1.6), namely

$$\nu_b(x) = \min_{1 \leq i \leq s} \nu_{b_i}(x_i) = \min_{1 \leq i \leq s} \left\lfloor \frac{\nu_{p_i}(x_i)}{l_i} \right\rfloor.$$

Furthermore, for any integer $d \geq 1$ we define

$$\ell_{b,d}(x) = \mathcal{L}_b(x) \bmod b^d. \tag{2.4}$$

It is immediate from the definition that $\mathcal{L}_b$ takes values in $\mathbb{Z}_b$, while $\ell_{b,d}$ — in the set

$$\{0\} \cup \{1 \leq a \leq b^d - 1 : b \nmid a\}.$$

Moreover $\mathcal{L}_b(x) = 0$ and $\ell_{b,d}(x) = 0$ if and only if $x = 0$. The functions $\mathcal{L}_b(x)$ and $\ell_{b,d}(x)$ describe last nonzero digits of the b-adic expansion of $x \in \mathbb{Q}_b$. More precisely, $\mathcal{L}_b(x)$ is the b-adic integer obtained by deleting all the trailing zeros in the b-adic expansion of x, while $\ell_{b,d}(x)$ is the integer represented by the last block of d digits not ending with 0. This is clear if x is a nonnegative integer and in the general case follows from the discussion in Section 1.2.1.

We now proceed to prove various properties of $\mathcal{L}_b$ and $\ell_{b,d}$. The first lemma helps to compute their values evaluated at products and sums of integers.

Lemma 2.5. *Let $x, y \in \mathbb{Q}_b$. The functions $\mathcal{L}_b$ and $\ell_{b,d}$ have following properties:*

(i) $\mathcal{L}_b(bx) = \mathcal{L}_b(x)$ *and* $\ell_{b,d}(bx) = \ell_{b,d}(x)$;

(ii) *if* $\nu_b(xy) = \nu_b(x) + \nu_b(y)$, *then*

$$\mathcal{L}_b(xy) = \mathcal{L}_b(x)\mathcal{L}_b(y)$$

and

$$\ell_{b,d}(xy) = \ell_{b,d}(x)\ell_{b,d}(y) \bmod b^d;$$

(iii) *if* $\nu_b(x) > \nu_b(y)$, *then* $\mathcal{L}_b(x + y) = b^{-\nu_b(y)}(x + y)$;

(iv) *if* $\nu_b(x) \geq \nu_b(y) + d$, *then* $\ell_{b,d}(x + y) = \ell_{b,d}(y)$.

Proof. All the properties follow straight from the definitions. □

The next lemma shows how $\ell_{b,d}$ can be written in terms of simpler expressions, involving prime factors of b.

Lemma 2.6. *For every nonzero $x = (x_1, \ldots, x_s) \in \mathbb{Q}_b$ we have*

(i) $\ell_{b,d}(x) = \sum_{i=1}^{s} b_i^{\nu_{b_i}(x_i)-\nu_b(x)} q_i^d r_i^{\nu_b(x)+d} \ell_{b_i,d}(x_i) \bmod b^d$;

(ii) $\mathcal{L}_{b_i}(x_i) = \mathcal{L}_{p_i}(x_i) p_i^{\nu_{p_i}(x_i) \bmod l_i}$ *and*
$\ell_{b_i,d}(x_i) = \ell_{p_i,l_i d}(x_i) p_i^{\nu_{p_i}(x_i) \bmod l_i} \bmod b_i^d$.

Proof. For each $i = 1, \ldots, s$ the i-th component of $\mathcal{L}_b(x)$ can be written in the form

$$b_i^{\nu_{b_i}(x_i)-\nu_b(x)} q_i^{-\nu_b(x)} \mathcal{L}_{b_i}(x_i).$$

Therefore, $\ell_{b,d}(x)$ satisfies for each $i = 1, \ldots, s$ the congruence

$$\ell_{b,d}(x) \equiv b_i^{\nu_{b_i}(x)-\nu_b(x)} r_i^{\nu_b(x)} \ell_{b_i,d}(x_i) \pmod{b_i^d}. \tag{2.5}$$

As a consequence of (2.3), we get (i).

Moving on to (ii), write $\nu_{p_i}(x_i) = \nu_{b_i}(x_i) l_i + u_i$, where $0 \le u_i \le l_i - 1$. We have

$$\mathcal{L}_{b_i}(x_i) = \frac{x_i}{p_i^{\nu_{b_i}(x_i) l_i + u_i}} p_i^{u_i} = \mathcal{L}_{p_i}(x_i) p^{u_i}.$$

The result for $\ell_{b_i,d}(x_i)$ follows by reducing both sides modulo b_i^d. □

We remark that the exponent $\nu_b(x_i) + d$ of r_i in (i) contributes only through its residue modulo

$$t_i = \operatorname{ord}_{b_i^d}(q_i).$$

As a consequence of the congruence (2.5), we may recover some information about the valuations $\nu_{b_i}(x_i)$ knowing only last nonzero digits of x and vice versa.

Corollary 2.7. *Let $x = (x_1, \ldots, x_s) \in \mathbb{Q}_b$ be nonzero and fix $i \in \{1, \ldots, s\}$. Then*

$$\nu_{b_i}(\ell_{b,d}(x)) < d \text{ if and only if } \nu_{b_i}(x_i) - \nu_b(x) < d.$$

Moreover, in either case we have

$$\nu_{b_i}(\ell_{b,d}(x)) = \nu_{b_i}(x_i) - \nu_b(x).$$

Proof. The claim follows directly from the congruence (2.5). □

We also exhibit a special case where the summands in Lemma 2.6(i) all vanish except for one. Obviously, this is always true when b is a prime power.

Corollary 2.8. *Assume that b has at least two distinct prime factors and let $x = (x_1, \ldots, x_i) \in \mathbb{Q}_b$ be nonzero. If $j \in \{1, \ldots, s\}$ is such that*

$$\nu_{b_i}(x_i) \ge d + \nu_{b_j}(x_j)$$

for all $i \ne j$, then

$$\ell_{b,d}(x) = \ell_{b_j,d}(x_j) q_j^d r_j^{\nu_{b_j}(x_j)+d} \bmod b^d.$$

Proof. The equality is an immediate consequence of Lemma 2.6. □

To conclude this section, we show that the functions $\mathcal{L}_b$ and $\ell_{b,d}$ have good properties from the point of view of regular and automatic sequences. Throughout this thesis we are going to prove a number of variations of this result.

Proposition 2.9. *The sequence $(\mathcal{L}_b(n))_{n\geq 0}$ is b-regular and $(\ell_{b,d}(n))_{n\geq 0}$ is b-automatic.*

Proof. We have $\mathcal{L}_b(bn) = \mathcal{L}_b(n)$ and $\mathcal{L}_b(bn+a) = bn+a$ for all $n \geq 0$ and $a = 1, \ldots, b-1$. Hence, the b-kernel of the sequence $(\mathcal{L}_b(n))_{n\geq 0}$ can be written as

$$\mathcal{K}_b((\mathcal{L}_b(n))_{n\geq 0}) = \{(\mathcal{L}_b(n))_{n\geq 0}\} \cup \bigcup_{a=1}^{b-1} \mathcal{K}_b((bn+a)_{n\geq 0}).$$

The sequences $(bn+a)_{n\geq 0}$ are b-regular as polynomials evaluated at consecutive integers. For each $a = 1, \ldots, b-1$ we thus have a finite set $S_a \subset \mathbb{Z}^{\mathbb{N}}$ generating the $\mathbb{Z}$-module generated by $\mathcal{K}_b((bn+a)_{n\geq 0})$. Therefore, the $\mathbb{Z}$-module generated by $\mathcal{K}_b((\mathcal{L}_b(n))_{n\geq 0})$ is generated by the finite set

$$\{(\mathcal{L}_b(n))_{n\geq 0}\} \cup \bigcup_{a=1}^{b-1} S_a,$$

which gives b-regularity of $(\mathcal{L}_b(n))_{n\geq 0}$.

Consequently, by Corollary 1.16 the sequence $(\ell_{b,d}(n))_{n\geq 0}$ is b-automatic as the reduction modulo b^d of a b-regular sequence. □

2.3 Automaticity of last nonzero digits of factorials

Our goal in this section is to prove Theorem 2.2. We retain all the notation from the previous sections. Assume from now on that the prime factors of b are ordered according to the conditions (2.1) and (2.2).

We first give a sketch of the reasoning. As a consequence of Lemma 2.10 below, if $b \in \mathcal{B}$, then $p_j = p_1$ is the distinguished prime factor in Corollary 2.8 with $x = n!$ for all n sufficiently large. Therefore, for all n sufficiently large the value $\ell_{b,d}(n!)$ depends only on $\ell_{b_1,d}(n!)$ and $\nu_{b_1}(n!) \bmod t_1$. These values can in turn be expressed by $\ell_{p_1,l_1 d}(n!)$ and $\nu_{p_1}(n!) \bmod l_1 t_1$. If we are able to show that the two latter expressions form p_1-automatic sequences, then p_1-automaticity of $(\ell_{b,d}(n!))_{n\geq 0}$ will follow. In the case $d = 1$ this was already observed (more or less) by Deshouillers and Ruzsa, as mentioned in Section 2.1. Even if $b \notin \mathcal{B}$, the same line of reasoning remains true for a set of n of natural density 1. Nonautomaticity in the case $b \notin \mathcal{B}$ will be a simple consequence of the result of Lipka (Theorem 2.1).

Let $S_{b,d} = \mathbb{N}$ if b is a prime power, and otherwise

$$S_{b,d} = \{n \geq 0 : \nu_{b_i}(n!) \geq d + \nu_{b_1}(n!) \text{ for } i = 2, \ldots, s\}.$$

The following result shows that this set is large.

Lemma 2.10. *The set $S_{b,d}$ has natural density 1. Furthermore, if $b \in \mathcal{B}$, then it contains all but a finite number of nonnegative integers.*

Proof. The result is trivial if b is a prime power. If b has at least two distinct prime factors, we consider two cases.

First, assume that $l_1(p_1-1) > l_2(p_2-1)$. Legendre's formula (Theorem 1.26) implies

$$\lim_{n\to\infty} \frac{\nu_{b_i}(n!)}{\nu_{b_1}(n!)} = \frac{l_1(p_1-1)}{l_i(p_i-1)} > 1,$$

for all $i = 2, \ldots, s$. Hence, all sufficiently large n belong to $S_{b,d}$.

Now let $l_1(p_1-1) = l_i(p_i-1)$ for $i = 2, \ldots, t$, where $2 \le t \le s$. By [29, Lemma 3] for each such i there exists $\delta_i > 0$ such that the set of nonnegative integers n satisfying

$$s_{p_i}(n) \le s_{p_1}(n) - \delta_i \log n$$

has natural density 1. For each n in this set we get

$$\nu_{b_i}(n!) = \left\lfloor \frac{n - s_{p_i}(n)}{l_i(p_i-1)} \right\rfloor \ge \left\lfloor \frac{n - s_{p_1}(n) + \delta_i \log n}{l_1(p_1-1)} \right\rfloor \ge \nu_{b_1}(n!) + \left\lfloor \frac{\delta_i \log n}{l_1(p_1-1)} \right\rfloor.$$

Therefore, for n sufficiently large we again have $n \in S_{b,d}$. Since the intersection of a finite number of sets of natural density 1 still has natural density 1, our claim follows. □

The key role in the proof of Theorem 2.2 and in the later sections will be played by the sequence $(\beta_{b,d}(n))_{n\ge 0}$, defined by

$$\beta_{b,d}(n) = \left(\ell_{p_1,l_1d}(n) p_1^{\nu_{p_1}(n) \bmod l_1} q_1^d r_1^{\lfloor \nu_{p_1}(n)/l_1 \rfloor + d}\right) \bmod b^d. \tag{2.6}$$

Indeed, Corollary 2.8 implies that

$$\ell_{b,d}(n!) = \beta_{b,d}(n!)$$

if and only if $n \in S_{b,d}$. It remains to verify that the sequence $(\beta_{b,d}(n!))_{n\ge 0}$ is p_1-automatic.

Proof of Theorem 2.2. Put $a_0 = 1$ and $a_n = \ell_{p_1,l_1d}(n)$ for $n \ge 1$. Since changing a finite number of terms does not affect automaticity, the sequence $(a_n)_{n\ge 0}$ is p_1-automatic. For all $n \in \mathbb{N}$ we have by Lemma 2.5(ii) the equality

$$\ell_{p_1,l_1d}(n!) = \left(\prod_{m=0}^{n} a_m\right) \bmod p_1^{l_1 d},$$

and thus Theorem 1.18(ii) implies that $(\ell_{p_1,l_1d}(n!))_{n\ge 0}$ is also p_1-automatic.

At the same time, the sequence $(\nu_{p_1}(n!))_{n\ge 0}$ is p_1-regular by Legendre's formula and Example 1.5. Hence, both $(\nu_{p_1}(n!) \bmod l_1)_{n\ge 0}$ and $(\lfloor \nu_{p_1}(n!)/l_1 \rfloor \bmod t_1)_{n\ge 0}$ are p_1-automatic sequences due to Corollary 1.16. The terms

$$\beta_{b,d}(n!) = \ell_{p_1,l_1d}(n!) p_1^{\nu_{p_1}(n!) \bmod l_1} q_1^d r_1^{\lfloor \nu_{p_1}(n!)/l_1 \rfloor + d} \bmod b^d$$

are thus expressed as a function of the terms of p_1-automatic sequences, and therefore $(\beta_{b,d}(n!))_{n\geq 0}$ is also a p_1-automatic sequence (recall that the exponent of r_1 in the formula (2.6) only contributes through its residue modulo t_1).

Now, the sequence $(\ell_{b,d}(n!))_{n\geq 0}$ agrees with $(\beta_{b,d}(n!))_{n\geq 0}$ on the set $S_{b,d}$ of natural density 1. According to Lemma 2.10, if $b \in \mathcal{B}$, then this set contains all but a finite number of positive integers, hence $(\ell_{b,d}(n!))_{n\geq 0}$ is p_1-automatic itself.

On the other hand, if for some $b \notin \mathcal{B}$ the sequence $(\ell_{b,d}(n!))_{n\geq 0}$ were automatic, then so would be its image under the coding $a \mapsto a \bmod b$, which is precisely $(\ell_b(n!))_{n\geq 0}$. However, this is not the case by Theorem 2.1. □

2.4 Generating the sequence

The proof of Theorem 2.2 provides little insight into how a p_1-uniform morphism (or a p_1-DFAO) generating the sequence $(\ell_{b,d}(n!))_{n\geq 0}$ might look like. The form of such a morphism for each pair (b, d) can be recovered by following step-by-step the rather complicated construction behind the proof of Theorem 1.18. This renders it difficult to give a recipe for a suitable morphism, let alone calculate the frequencies of $a = 1, \ldots, b^d - 1$ in $(\ell_{b,d}(n!))_{n\geq 0}$, other than on a case-by-case basis.

In this section we shall derive an explicit form of a p_1-uniform morphism generating $(\beta_{b,d}(n!))_{n\geq 0}$. In the case $b \in \mathcal{B}$ a simple modification yields a morphism generating $(\ell_{b,d}(n!))_{n\geq 0}$ as well, since these two sequences differ only by a finite number of terms.

Examining the formula (2.6) defining $\beta_{b,d}(n)$, it is clear that $(\ell_{p,l_1 d}(n!))_{n\geq 0}$ and $(\nu_{p_1}(n!) \bmod l_1 t_1)_{n\geq 0}$ are two sequences of interest. As we will see shortly, certain recurrence relations involving the terms of these sequences will also depend on the residue of n modulo some integer. This motivates us to consider the family of sequences

$$(\alpha_{p,\delta,u,v}(n))_{n\geq 0} = ((\ell_{p,\delta}(n!), \nu_p(n!) \bmod u, n \bmod v))_{n\geq 0},$$

where p is a prime number, and $\delta, u, v \geq 1$ are integers. For technical reasons we assume that v is divisible by $\operatorname{lcm}(p^{\delta-1}, u, 2)$, unless $p = 2, \delta = 2$, in which case we additionally require that 4 divides v. In order to ease the notation we will consider the set Θ of such quadruples $\theta = (p, \delta, u, v)$ of parameters and integchangeably write α_θ. In particular, for the choice of parameters $\theta = (p_1, l_1 d, l_1 t_1, v)$ with any appropriate v the terms $\beta_{b,d}(n!)$ are the image of $\alpha_\theta(n)$ under a coding. The main benefit of this more general approach is that it enables the computation of the frequencies of letters in subsequences along any arithmetic progression.

For now, our aim is to find for $\theta = (p, \delta, u, v) \in \Theta$ a p-uniform morphism generating $(\alpha_\theta(n))_{n\geq 0}$. We begin by deriving recurrence relations governing the behavior of $\ell_{p,\delta}(n!)$ and $\nu_p(n!)$. The first of these relations involves so-called *Gauss factorial* (see for example [24]), defined for integers $n \geq 0, m \geq 2$ by

$$n_m! = \begin{cases} 1 & \text{if } n = 0, \\ \displaystyle\prod_{\substack{1\leq k\leq n \\ \gcd(k,m)=1}} k & \text{if } n > 0. \end{cases}$$

We have the following result.

Lemma 2.11. *Let p be a prime and $\delta \geq 1$ an integer. Then for all integers $n \geq 0$ and $i = 0, \ldots, p-1$ we have*

(i) $\ell_{p,\delta}\left((pn+i)!\right) = \ell_{p,\delta}(n!)(pn+i)_p! \bmod p^\delta$;

(ii) $\nu_p((pn+i)!) = \nu_p(n!) + n$.

Proof. Since none of the numbers $pn+1, \ldots, pn+p-1$ is divisible by p, it is sufficient to prove both identities for $i = 0$.

The first equality holds trivially for $n = 0$ and $i = 0$. By induction on n we have

$$\begin{aligned}
\ell_{p,\delta}\left((p(n+1))!\right) &\equiv \ell_{p,\delta}\left((pn)!\right) \left(\prod_{j=1}^{p-1} \ell_{p,\delta}(pn+j)\right) \ell_{p,\delta}(pn+p) \\
&\equiv \ell_{p,\delta}(n!)(pn)_p! \left(\prod_{j=1}^{p-1} (pn+j)\right) \ell_{p,\delta}(n+1) \\
&\equiv \ell_{p,\delta}((n+1)!)(p(n+1))_p! \pmod{p^\delta},
\end{aligned}$$

The equality for $\nu_p((pn)!)$ follows immediately from Legendre's formula (in either form). □

In order to handle the expression $(pn+i)_p!$ modulo a power of p we will use the following standard lemma.

Lemma 2.12. *Let p be a prime and $\delta \geq 1$ an integer. We have*

$$(p^\delta)_p! \equiv \begin{cases} 1 \pmod{p^\delta} & \text{if } p = 2 \text{ and } \delta \neq 2, \\ -1 \pmod{p^\delta} & \text{otherwise .} \end{cases}$$

Proof. The claim follows from the fact that the product of all elements in a finite abelian group is equal to the product of the elements of order two. □

We are now ready to construct a p-uniform morphism such that $(\alpha_\theta(n))_{n\geq 0}$ is its fixed point. It acts on the alphabet

$$\Lambda_\theta = (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+ \times \{0, 1, \ldots, v-1\},$$

where $(\mathbb{Z}/p^\delta\mathbb{Z})^\times$ and $(\mathbb{Z}/u\mathbb{Z})^+$ denote the multiplicative group modulo p^δ and the additive group modulo u, respectively. Let $\psi_\theta \colon \Lambda_\theta^* \to \Lambda_\theta^*$ be defined by

$$\psi_\theta(x, y, z) = (x_0, y_0, z_0)(x_1, y_1, z_1) \cdots (x_{p-1}, y_{p-1}, z_{p-1}),$$

where

$$\begin{aligned}
x_i &= x(pz+i)_p! \bmod p^\delta, \\
y_i &= y + z \bmod u, \\
z_i &= pz + i \bmod v
\end{aligned}$$

for $i = 0, 1, \ldots, p-1$.

Proposition 2.13. *Let $\theta = (p, \delta, u, v) \in \Theta$. The sequence $(\alpha_\theta(n))_{n\geq 0}$ is the fixed point of ψ_θ starting with $(1, 0, 0)$.*

Proof. We have $\alpha_\theta(0) = (1, 0, 0)$ for any choice of parameters. We need to show that for any integer $n \geq 0$ a recurrence of the form (1.2) is satisfied, that is

$$\psi_\theta(\alpha_\theta(n)) = \alpha_\theta(pn)\alpha_\theta(pn + 1) \cdots \alpha_\theta(pn + p - 1).$$

Fix $n \geq 0$ and let

$$(x, y, z) = \alpha_\theta(n) = (\ell_{p,\delta}(n!), \nu_p(n!) \bmod u, n \bmod v).$$

Write $n = mv+z$, where $m \geq 0$ is some integer. By Lemma 2.12 and the first equality in Lemma 2.11 we obtain for each $i = 0, 1, \ldots, p - 1$ the congruence

$$\begin{aligned}\ell_{p,\delta}((pn + i)!) &\equiv \ell_{p,\delta}(n!)(pn + i)_p! \equiv x(pmv + pz + i)_p! \\ &\equiv x((p^\delta)_p!)^{mv/p^{\delta-1}}(pz + i)_p! \equiv x_i \pmod{p^\delta}.\end{aligned}$$

Here we have used the assumptions on v guaranteeing that $v/p^{\delta-1}$ is an even integer whenever $(p^\delta)_p! \equiv -1 \pmod{p^\delta}$.

The second equality of Lemma 2.11 immediately gives $\nu_p((pn + i)!) \equiv y_i \pmod u$ for all $i = 0, 1, \ldots, p - 1$, for which it is important that $u \mid v$.

Obviously, $pz + i \equiv z_i \pmod v$, which completes the proof. □

As we have mentioned before, by choosing suitable values of the parameters $\theta = (p, \delta, u, v) \in \Theta$ we can describe the sequence $(\beta_{b,d}(n!))_{n\geq 0}$ as the image of $(\alpha_\theta(n))_{n\geq 0}$ under a coding. Let $\theta = (p_1, l_1 d, l_1 t_1, v)$, where $v = \operatorname{lcm}(p_1^{l_1 d-1}, l_1 t_1, 2)$ if $p_1^{l_1 d} \neq 4$, and otherwise $v = 4$. We define

$$\begin{aligned}\varphi_{b,d} &= \psi_\theta, \\ \Sigma_{b,d} &= \Lambda_\theta.\end{aligned}$$

By (2.6) the values $\beta_{b,d}(m)$ with m nonzero belong to the alphabet

$$\Delta_{b,d} = \{0 \leq a \leq b^d - 1 : b \nmid a \text{ and } q_1^d \mid a\}.$$

We also define the coding $\tau_{b,d} \colon \Sigma_{b,d}^* \to \Delta_{b,d}^*$, derived from (2.6), by

$$\tau_{b,d}(x, y, z) = x p_1^{y \bmod l_1} q_1^d r_1^{\lfloor y/l_1 \rfloor + d} \bmod b^d.$$

Putting together our considerations so far, we have proven the following theorem.

Theorem 2.14. *We have*

$$(\beta_{b,d}(n!))_{n\geq 0} = \tau_{b,d}(\varphi_{b,d}^\omega((1, 0, 0))).$$

This completes the main goal of this section and may serve as an alternative way of proving that the sequence $(\beta_{b,d}(n!))_{n\geq 0}$ (and $(\ell_{b,d}(n!))_{n\geq 0}$ for $b \in \mathcal{B}$) is p_1-automatic.

While the method described above covers all pairs (b, d), its generality comes at the expense of computational effectiveness (measured by the cardinality of $\Sigma_{b,d}$). As we shall see, in the special case $l_1 = 1$ (and maybe other cases) it is possible to significantly reduce the size of the alphabets and simplify the overall construction. As a starting point, we derive a recurrence relation directly involving $\beta_{b,d}(n!)$. This corrects [70, Proposition 16], where it is erroneously claimed that a similar relation holds for any l_1.

Lemma 2.15. *Assume that $l_1 = 1$. Then for all integers $n \geq 0$ and $i = 0, \ldots, p-1$ we have*

$$\beta_{b,d}((p_1 n + i)!) = \beta_{b,d}(n!)(p_1 n + i)_{p_1}! r_1^n \bmod b^d.$$

Proof. By the definition of $\beta_{b,d}$ and the relations of Lemma 2.11 with $p = p_1, \delta = d$, we obtain

$$\begin{aligned}\beta_{b,d}((p_1 n + i)!) &\equiv \ell_{p_1,d}((p_1 n + i)!) q_1^d r_1^{\nu_{p_1}((p_1 n+i)!)+d} \\ &\equiv \ell_{p_1,d}(n!)(p_1 n + i)_{p_1}! q_1^d r_1^{\nu_{p_1}(n!)+n+d} \\ &\equiv \beta_{b,d}(n!)(p_1 n + i)_{p_1}! r_1^n \pmod{b^d}.\end{aligned}$$

□

Based on this relation, it is not hard to see how the construction should progress. Consider the alphabet

$$\widehat{\Sigma}_{b,d} = \Delta_{b,d} \times \{0, \ldots, v-1\},$$

where $v = \operatorname{lcm}(p_1^{d-1}, t_1, 2)$, unless $p_1 = 2, d = 2$, in which case $v = 4$.

There is a bijection between $\Delta_{b,d}$ and $(\mathbb{Z}/p_1^d\mathbb{Z})^\times$ (given by $a \mapsto a/q_1^d \bmod p_1^d$). Hence, $\widehat{\Sigma}_{b,d}$ is essentially the same as $\Sigma_{b,d}$, but with the factor $(\mathbb{Z}/l_1 t_1 \mathbb{Z})^+$ omitted in the Cartesian product defining the alphabet. The assumption $l_1 = 1$ gives $\#\widehat{\Sigma}_{b,d} = \#\Sigma_{b,d}/t_1$, which is a major improvement from the computational point of view.

We define a p_1-uniform morphism $\widehat{\varphi}_{b,d} \colon \widehat{\Sigma}_{b,d}^* \to \widehat{\Sigma}_{b,d}^*$ by

$$\widehat{\varphi}_{b,d}(x, z) = (x_0, z_0)(x_1, z_1) \cdots (x_{p_1-1}, z_{p_1-1}),$$

where for $i = 0, 1, \ldots, p_1 - 1$, we put

$$\begin{aligned} x_i &= x(p_1 z + i)_{p_1}! r_1^z \bmod b^d, \\ z_i &= p_1 z + i \bmod v. \end{aligned}$$

Finally, we define the coding $\widehat{\tau}_{b,d} \colon \widehat{\Sigma}_{b,d}^* \to \Delta_{b,d}^*$ by

$$\widehat{\tau}_{b,d}(x, z) = x.$$

The following proposition relies on an almost identical argument as Proposition 2.13, thus we omit the proof.

Proposition 2.16. *Assume that $l_1 = 1$. Then*

$$(\beta_{b,d}(n!))_{n\geq 0} = \widehat{\tau}_{b,d}(\widehat{\varphi}^{\omega}_{b,d}(q_1^d r_1^d \bmod b^d, 0)).$$

In the case $d \in \{1, 2\}$ the expression $(p_1 z + i)_{p_1}!$ in the definition of $\widehat{\varphi}_{b,d}$ can be simplified. If $d = 1$, then

$$(p_1 z + i)_{p_1}! \equiv (-1)^z i! \pmod{p_1}.$$

If $d = 2$, we can use the following lemma.

Lemma 2.17. *Let $p \geq 3$ be a prime. Then for all integers $m \geq 0$ we have*

$$(p-1)! \equiv \prod_{j=1}^{p-1} (pm + j) \pmod{p^2}.$$

Therefore,

$$(pm)_p! \equiv ((p-1)!)^m \pmod{p^2}.$$

Proof. Consider the polynomial

$$f(z) = \prod_{j=1}^{p-1} (z + j).$$

In particular, $f(0) = (p-1)!$. Since $1, 2, \ldots, p-1$ are the roots of f modulo p, we can write

$$f(z) = z^{p-1} - 1 + pg(z)$$

where g is a polynomial with integer coefficients. Then

$$\prod_{j=1}^{p-1} (pm + j) = f(pm) \equiv -1 + pg(pm) \equiv -1 + pg(0) \pmod{p^2},$$

which does not depend on m. □

This leads to further simplifications if $q_1 = b/p_1$ is congruent to $(p_1 - 1)!$ modulo p_1^d. Below we summarize these cases.

Lemma 2.18. *Assume that $l_1 = 1$. Let $d \in \{1, 2\}, (x, z) \in \widehat{\Sigma}_{b,d}$, and let*

$$x_i = x(p_1 z + i)_{p_1}! r_1^z \bmod b^d,$$

for $i = 0, 1, \ldots, p_1 - 1$, as in the definition of $\widehat{\varphi}_{b,d}$. Then

$$x_i = \begin{cases} x(-r_1)^z i! & \text{if } d = 1, \\ x((p_1-1)! r_1)^z \prod_{j=1}^{i} (p_1 z + j) & \text{if } d = 2 \text{ and } p_1 \geq 3, \\ x(-1)^{\lfloor (2z+i+1)/4 \rfloor} & \text{if } d = 2 \text{ and } p_1 = 2 \end{cases} \bmod b^d.$$

If additionally $q_1 \equiv (p_1 - 1)! \pmod{p_1^d}$, then

$$x_i = \begin{cases} xi! & \text{if } d = 1, \\ x \prod_{j=1}^{i} (p_1 z + j) & \text{if } d = 2 \text{ and } p_1 \geq 3 \end{cases} \bmod b^d.$$

Proof. Since x is always divisible by q_1^d, the case $d = 1$ follows from $(p_1 - 1)! \equiv -1 \pmod{p_1}$. In the case $d = 2$ and $p_1 \geq 3$ the desired formulas follow from Lemma 2.17. If $p_1 = 2$, then by $l_1 = 1$ and the condition (2.1) we must have $b = 2$. One can check by hand that $(m)_2! \equiv (-1)^{\lfloor (m+1)/4 \rfloor} \pmod 4$ for $m = 0, 1, \ldots, 7$. Finally, $q_1 r_1 \equiv 1 \pmod{p_1^d}$ gives the claim in the case $q_1 \equiv (p_1 - 1)! \pmod{p_1^d}$ □

As a consequence, if $l_1 = 1, d = 1$, and $q_1 \equiv (p_1 - 1)! \equiv -1 \pmod{p_1}$, then x_i in the definition of $\widehat{\varphi}_{b,1}$ does not depend on z at all. In this case we may further reduce the alphabet to $\Delta_{b,1}$. It is the best we can achieve, since by the results in the next section $\beta_{b,1}(n!)$ takes on all values in $\Delta_{b,1}$. Define a p_1-uniform morphism $\widetilde{\varphi}_{b,1}$ acting on $\Delta_{b,1}$ by

$$\widetilde{\varphi}_{b,1}(x) = x_0 x_1 \ldots x_{p_1 - 1}$$

for $x \in \Delta_{b,1}$, where

$$x_i = x \cdot i! \bmod b.$$

By the above discussion we immediately obtain the following characterization.

Corollary 2.19. *Assume that $l_1 = 1$ and $q_1 \equiv -1 \pmod{p_1}$. Then*

$$(\beta_{b,1}(n!))_{n \geq 0} = \widetilde{\varphi}_{b,1}^{\omega}(q_1 r_1 \bmod b).$$

Analogously, if $l_1 = 1, d = 2$, and $q_1 \equiv (p_1 - 1)! \pmod{p_1^2}$, then z only contributes to x_i through its residue modulo p_1 rather than modulo $v = \operatorname{lcm}(p_1, t_1, 2)$. Example 2.1 below illustrates this situation and also shows how to construct a morphism generating the original sequence $(\ell_{b,d}(n!))_{n \geq 0}$.

Example 2.1. Consider two last nonzero digits of $n!$ in base $b = 6$. The condition (2.1) implies $p_1 = 3, l_1 = 1$ and $p_2 = 2, l_2 = 1$. We have $6 \in \mathcal{B}$, so the sequence $(\ell_{6,2}(n!))_{n \geq 0}$ is 3-automatic.

We define a 3-uniform morphism $\widetilde{\varphi}_{6,2}$ and coding $\widetilde{\tau}_{6,2}$ which generate the corresponding 3-automatic sequence $(\beta_{6,2}(n!))_{n \geq 0}$. We have $q_1 = b/p_1 = 2$ and put $r_1 = 5$, so that $q_1 r_1 \equiv 1 \pmod{3^2}$. By definition $\beta_{6,2}(1) = q_1^2 r_1^2 \mod 6^2 = 28$. Since $(p_1 - 1)! = 2 = q_1$, by Lemma 2.18 obtain

$$\beta_{6,2}((3n + i)!) = \beta_{6,2}(n!) \prod_{j=1}^{i} (3z + j) \bmod 6^2$$

for all $n \geq 0$ and $i = 0, 1, 2$, where $z = n \bmod 3$.

The 3-uniform morphism $\widetilde{\varphi}_{6,2}$ is defined on the alphabet

$$\widetilde{\Sigma}_{6,2} = \{4, 8, 16, 20, 28, 32\} \times \{0, 1, 2\}$$

by the formula

$$\widetilde{\varphi}_{6,2}(x, z) = (x_0, 0)(x_1, 1)(x_2, 2),$$

where

$$x_0 = x, \qquad x_1 = x(3z + 1) \bmod 6^2, \qquad x_2 = x(3z + 1)(3z + 2) \bmod 6^2.$$

	x					
z	4	8	16	20	28	32
0	$4,4,8$	$8,8,16$	$16,16,32$	$20,20,4$	$28,28,20$	$32,32,28$
1	$4,16,8$	$8,32,16$	$16,28,32$	$20,8,4$	$28,4,20$	$32,20,28$
2	$4,28,8$	$8,20,16$	$16,4,32$	$20,32,4$	$28,16,20$	$32,8,28$

Table 2.1: The values x_1, x_2, x_3 in $\widetilde{\varphi}_{6,2}(x,z)$

In Table 2.1 below we give explicit values x_1, x_2, x_3 for each argument (x,z).

The sequence $(\beta_{6,2}(n!), n \bmod 3)_{n\geq 0}$ is the fixed point of $\widetilde{\varphi}_{6,2}$ starting with $(28,0)$ and $(\beta_{6,2}(n!))_{n\geq 0}$ is its image under the coding $\widetilde{\tau}_{6,2}(x,y) = x$.

In order to generate the sequence $(\ell_{b,d}(n!))_{n\geq 0}$ we need to modify the above construction to handle the terms such that $\ell_{b,d}(n!) \neq \beta_{b,d}(n!)$. It is easily verified that $\nu_2(n!) \geq 2 + \nu_3(n!)$ for all $n \geq 4$, thus these are precisely the four initial terms. We add four new symbols to $\widetilde{\Sigma}_{6,2}$, say A, B, C, D, and define

$$\begin{aligned} \widetilde{\varphi}_{6,2}(A) &= ABC, & \widetilde{\varphi}_{6,2}(B) &= D(4,1)(20,2), \\ \widetilde{\varphi}_{6,2}(C) &= (20,0)(32,1)(4,2), & \widetilde{\varphi}_{6,2}(D) &= (28,0)(28,1)(20,2), \end{aligned}$$

so that the first coordinate of $\widetilde{\varphi}^{\omega}_{6,2}(A)$ forms the sequence

$$A, B, C, D, \beta_{6,2}(4!), \beta_{6,2}(5!), \ldots.$$

Finally, we put

$$\widetilde{\tau}_{6,2}(A) = 1, \qquad \widetilde{\tau}_{6,2}(B) = 1, \qquad \widetilde{\tau}_{6,2}(C) = 2, \qquad \widetilde{\tau}_{6,2}(D) = 1,$$

which yields

$$(\ell_{6,2}(n!))_{n\geq 0} = \widetilde{\tau}_{6,2}(\widetilde{\varphi}^{\omega}_{6,2}(A)),$$

as desired.

2.5 Frequencies of letters

This section is devoted to the computation of the frequencies of letters in the sequence $(\ell_{b,d}(n!))_{n\geq 0}$ and its subsequences along arithmetic progressions by employing the method outlined in Section 1.1.5.

Let us start with an example.

Example 2.2. Consider once again the sequence $(\ell_{6,2}(n!))_{n\geq 0}$. The frequencies of letters in this sequence are the same as in $(\beta_{6,2}(n!))_{n\geq 0}$. Using Table 2.1, we can retrieve the form of the 18×18 incidence matrix $M(\widetilde{\varphi}_{6,2})$. Calculations show that $M(\widetilde{\varphi}_{6,2})^4$ has all entries positive, and thus $M(\widetilde{\varphi}_{6,2})$ is primitive. Moreover, $\frac{1}{3}M(\widetilde{\varphi}_{6,2})$ is a row-stochastic matrix. This could be also deduced without computing the incidence matrix directly, from the observation that for each fixed z the functions of x defining x_0, x_1, x_2 are bijective. As a consequence, each $(x,z) \in \widetilde{\Sigma}_{6,2}$ appears in $\widetilde{\varphi}^{\omega}_{6,2}(28,0)$

with frequency 1/18. Since for all $n \geq 0$ the first coordinate of the nth symbol in this infinite word is $\beta_{6,2}(n!)$, we obtain that each $a \in \{4, 8, 16, 20, 28, 32\}$ appears in $(\ell_{6,2}(n!))_{n\geq 0}$ with frequency 1/6.

In the general case we are going to proceed in the same way. The proof that for any $\theta = (p, \delta, u, v) \in \Theta$ the morphism ψ_θ (and thus also $\varphi_{b,d}$ as a special case) is primitive will however require considerably more effort than in Example 2.2, since a direct computation of the powers of the incidence matrix $M(\psi_\theta)$ depending on θ seems rather futile. Utilizing certain symmetries exhibited by ψ_θ, we will also show that $\frac{1}{p}M(\psi_\theta)$ is a row-stochastic matrix. Then it will follow that $(\alpha_\theta(n))_{n\geq 0}$ takes all values in Λ_θ with frequency $1/\#\Lambda_\theta$. Finding the frequencies of letters in $(\ell_{b,d}(n!))_{n\geq 0}$ will then rely on a simple calculation. First, we give a useful observation concerning ψ_θ.

Recall that we have defined the first and second components of $(x, y, z) \in \Lambda_\theta$ as elements of the multiplicative group $(\mathbb{Z}/p^\delta\mathbb{Z})^\times$ and the additive group $(\mathbb{Z}/u\mathbb{Z})^+$, respectively. When performing operations on these components we will thus omit the moduli.

Lemma 2.20. *Let $\theta = (p, \delta, u, v) \in \Theta$ and let m, i, z be integers such that $m \geq 1$, $0 \leq i \leq p^m - 1$, and $0 \leq z \leq v - 1$. Then there exists $(e, f, g) \in \Lambda_\theta$, depending only on m, i, z, such that for all $(x, y) \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ the ith symbol in $\psi_\theta^m(x, y, z)$ (counting from 0) is equal to $(xe, y + f, g)$.*

Proof. We use induction on m. For $m = 1$ we deduce straight from the definition of ψ_θ that

$$(e, f, g) = ((pz + i)_p! \bmod p^\delta, z \bmod u, (pz + i) \bmod v).$$

Now assume that the result is satisfied up to m. Write

$$i = pi_1 + i_0,$$

where i_1, i_0 are integers such that $0 \leq i_1 \leq p^m - 1$ and $0 \leq i_0 \leq p - 1$. Also let $(e_1, f_1, g_1) \in \Lambda_\theta$ be the triple corresponding to m, i_1, z. Observe that the ith symbol in $\psi_\theta^{m+1}(x, y, z)$ is obtained by taking i_0th symbol in the word

$$\psi_\theta(xe_1, y + f_1, g_1),$$

where the argument of ψ_θ is precisely the i_1th symbol of $\psi_\theta^m(x, y, z)$. Hence, the triple (e, f, g) corresponding to m, i, z is equal to

$$(e, f, g) = (e_1(pg_1 + i_1)_p! \bmod p^\delta, (f_1 + g_1) \bmod u, (pg_1 + i_1) \bmod v),$$

and the result follows. □

Our main tool in studying the properties of ψ_θ is the relation R_θ on Λ_θ such that

$$(x, y, z)R_\theta(x', y', z')$$

if and only if (x', y', z') appears in $\psi_\theta^m(x, y, z)$ (at any position) for some integer $m \geq 1$.

From Lemma 2.20, we immediately obtain a corollary concerning R_θ.

Corollary 2.21. *Let $\theta = (p, \delta, u, v) \in \Theta$. Assume that $(x, y, z)R_\theta(x', y', z')$ for some $(x, y, z), (x', y', z') \in \Lambda_\theta$. Then for any $(X, Y) \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ we have*

$$(xX, y + Y, z)R_\theta(x'X, y' + Y, z').$$

Proof. Let m, i be such that (x', y', z') appears at position i in $\psi_\theta^m(x, y, z)$. Hence, the triple $(e, f, g) \in \Lambda_\theta$ in Lemma 2.20 applied to m, i, z, is given by $(x', y', z') = (ex, y + f, g)$. Replacing (x, y) with $(xX, y + Y)$ in Lemma 2.20 and substituting (x', y', z'), we obtain the assertion. □

We would like to prove that all the elements of Λ_θ are related through R_θ, which is a condition close to primitivity of ψ_θ. This will be done by first verifying that R_θ is an equivalence relation and then showing that there exists only one equivalence class.

Lemma 2.22. *For any $\theta \in \Theta$ the relation R_θ is an equivalence relation on Λ_θ.*

Proof. Transitivity of R_θ follows from the definition.

Moving on to symmetry, assume that $(x, y, z)R_\theta(x', y', z')$. From the formulas defining ψ_θ or Proposition 2.13 we see that for any $m \geq 1$ the projection of $\psi_\theta^m(x, y, z)$ onto the third coordinate is a sequence of p^m consecutive integers modulo v. Hence, there exist $x'' \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times, y'' \in (\mathbb{Z}/u\mathbb{Z})^+$ such that $(x', y', z')R_\theta(xx'', y + y'', z)$, and thus also $(x, y, z)R_\theta(xx'', y + y'', z)$ by transitivity. Corollary 2.21 and transitivity of R_θ imply that for all integers $n \geq 1$ we have $(x, y, z)R_\theta(x(x'')^n, y + ny'', z)$. Again, by transitivity we get $(x', y', z')R_\theta(x(x'')^n, y + ny'', z)$. Choosing n to be the order of (x'', y'') in $(\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$, we obtain $(x', y', z')R_\theta(x, y, z)$.

Obviously, each element $(x, y, z) \in \Lambda_\theta$ is related with at least one element. By transitivity and symmetry we thus obtain $(x, y, z)R_\theta(x, y, z)$. □

To prove that R_θ has only one equivalence class it is now sufficient to check that $(1, 0, 0)R_\theta(x, y, z)$ for every $(x, y, z) \in \Lambda_\theta$, or equivalently, that the terms $\alpha_\theta(n)$ take all values in Λ_θ. This is not obvious and the proof will rely on an identity involving the terms $\ell_{p,\delta}(n!)$.

Lemma 2.23. *Let p be a prime and $\delta \geq 1$ an integer. For any integers $n, m \geq 0$ not divisible by p, and integers $s \geq 0, t \geq \delta + \lfloor \log m / \log p \rfloor$ we have*

$$(-1)^{pm-1}\ell_{p,\delta}((p^{s+t}n)!)\, m \equiv \ell_{p,\delta}((p^{s+t}n - p^s m)!)\, \ell_{p,\delta}((p^s m)!)\, n \pmod{p^\delta}.$$

Proof. Expanding $\ell_{p,\delta}((p^{s+t}n)!)$, we obtain

$$\begin{aligned}\ell_{p,\delta}((p^{s+t}n)!) \equiv &\ell_{p,\delta}((p^{s+t}n - p^s m)!) \\ &\left(\prod_{j=1}^{p^s m - 1} \ell_{p,\delta}(p^{s+t}n - j)\right) n \pmod{p^\delta}. \end{aligned} \tag{2.7}$$

Since $t \geq \delta + \lfloor \log m / \log p \rfloor$, we have $\nu_p(j) \leq s + t - \delta$ for each $j = 1, \ldots, p^s m - 1$. By Lemma 2.5(iii) we obtain

$$\ell_{p,\delta}(np^{s+t} - j) \equiv -\ell_{p,\delta}(j) \pmod{p^\delta}.$$

Our claim follows by multiplying both sides of (2.7) by $\ell_{p,\delta}(p^s m) \equiv m \pmod{p^\delta}$. □

The next lemma is the key part of our reasoning towards proving that ψ_θ is primitive.

Lemma 2.24. *Let $\theta = (p, \delta, u, v) \in \Theta$. Then for every $(x, y) \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ we have*

$$(1, 0, 0)R_\theta(x, y, 0).$$

Proof. We want to show that the set

$$H = \{(x, y) \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+ : (1, 0, 0)R_\theta(x, y, 0)\}$$

is the whole group $(\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$. This will be done by first verifying that H is a subgroup and then exhibiting a set of generators of $(\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ in H.

Let $(x, y), (x', y') \in H$, so that $(1, 0, 0)R_\theta(x, y, 0)$ and $(1, 0, 0)R_\theta(x', y', 0)$. By Corollary 2.21 we obtain $((x')^{-1}, -y', 0)R_\theta(1, 0, 0)$, and thus symmetry of R_θ gives $(1, 0, 0)R_\theta((x')^{-1}, -y', 0)$. Combined with another application of Corollary 2.21, this yields

$$(1, 0, 0)R_\theta(x, y, 0)R_\theta(x(x')^{-1}, y - y', 0),$$

and hence $(xx'^{-1}, y - y') \in H$.

We now move on to find elements in H generating $(\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$. Let H_1 be the projection of H onto the first coordinate, a subgroup of $(\mathbb{Z}/p^\delta\mathbb{Z})^\times$. Lemma 2.23 with $s = \nu_p(v), m = v/p^s, t \geq \delta + \lfloor \log m / \log p \rfloor$, and $n = mx$, where x is any positive integer not divisible by p, gives

$$-\ell_{p,\delta}((p^t xv)!) \equiv \ell_{p,\delta}((p^t xv - v)!)\ell_{p,\delta}(v!)x \pmod{p^\delta}, \tag{2.8}$$

where we have used the fact that pm is even (this is guaranteed by the assumptions on v). We have $\ell_{p,\delta}((p^t xv)!), \ell_{p,\delta}((p^t xv - v)!), \ell_{p,\delta}(v!) \in H_1$. Considering $-x$ as an element of the ring $\mathbb{Z}/p^\delta\mathbb{Z}$, we also get $-x \in H_1$. Since x was arbitrary, we obtain $H_1 = (\mathbb{Z}/p^\delta\mathbb{Z})^\times$.

Now put $x = 1$ in the congruence (2.8), so that

$$-\ell_{p,\delta}((p^t v)!) \equiv \ell_{p,\delta}((p^t v - v)!)\ell_{p,\delta}(v!) \pmod{p^\delta}. \tag{2.9}$$

Applying t times the first relation of Lemma 2.11, we get

$$\ell_{p,\delta}((p^t v)!) \equiv \ell_{p,\delta}(v!) \prod_{j=1}^{t} (p^j v)_p! \equiv \ell_{p,\delta}(v!) \pmod{p^\delta},$$

where the second congruence follows from Lemma 2.12 and the assumptions on v. Hence, (2.9) becomes

$$\ell_{p,\delta}((p^t v - v)!) \equiv -1 \pmod{p^\delta} \tag{2.10}$$

and this holds for any $t \geq \delta + \lfloor \log m / \log p \rfloor$.

We now compute $\nu_p((p^t v - v)!) \bmod u$ by Legendre's formula. Since $p^t > m$, we obtain

$$\begin{aligned} s_p(p^t v - v) &= s_p(p^t m - m) = s_p(p^t(m-1) + p^t - m) \\ &= s_p(p^t(m-1)) + s_p(p^t - m) = s_p(m-1) + s_p(p^t - m). \end{aligned}$$

We can further calculate

$$s_p(p^t - m) = s_p(p^t - 1 - (m-1)) = t(p-1) - s_p(m-1),$$

which gives $s_p(p^t v - v) = t(p-1)$. Hence,

$$\nu_p((p^t v - v)!) = \frac{p^t - 1}{p-1} v - t \equiv -t \pmod{u}.$$

Taking $t \equiv 0 \pmod{u}$ and $t \equiv -1 \pmod{u}$, the above congruence together with (2.10) give $(-1, 0) \in H$ and $(-1, 1) \in H$, respectively, so also $(1, 1) \in H$. Taking into account that for every $x \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times$ there exists $y \in (\mathbb{Z}/u\mathbb{Z})^+$ such that $(x, y) \in H$, as we have proved earlier, we obtain the claim. □

A simple argument now shows that ψ_θ is a primitive morphism.

Proposition 2.25. *For any $\theta \in \Theta$ the morphism ψ_θ is primitive.*

Proof. Let $(x, y, z) \in \Lambda_\theta$. From the definition of ψ_θ we see that there exists a pair $(x', y') \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ such that $(x', y', 0) R_\theta (x, y, z)$. By Lemma 2.24 and transitivity of R_θ we get $(1, 0, 0) R_\theta (x, y, z)$. Therefore, by Lemma 2.22 R_θ has only one equivalence class. In other words, the incidence matrix $M(\psi_\theta)$ is *irreducible*. Recall that a nonnegative square matrix M of dimension k is called irreducible if for any $i, j \in \{1, \ldots, k\}$ there exists a positive integer n such that the entry of M^n at position (i, j) is positive. Because our matrix $M(\psi_\theta)$ has a positive diagonal entry corresponding to $(1, 0, 0)$, it is primitive by [42, Theorem 8.5.9] □

We are finally ready to compute the frequencies of letters in the sequence $(\alpha_\theta(n))_{n \geq 0}$.

Theorem 2.26. *Let $\theta = (p, \delta, u, v) \in \Theta$. Each letter from Λ_θ appears in the sequence $(\alpha_\theta(n))_{n \geq 0}$ with frequency*

$$\frac{1}{p^{\delta-1}(p-1)uv}.$$

Equivalently, for any integers $k \geq 1, l \geq 0$, each letter from $(\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ appears in the sequence $(\ell_{p,\delta}((kn+l)!), \nu_p((kn+l)!) \bmod u)_{n \geq 0}$ with frequency

$$\frac{1}{p^{\delta-1}(p-1)u}.$$

Proof. To begin, we argue that the two statements are equivalent. The first part of our claim holds if and only if for any $z \in \{0, 1, \ldots, v-1\}$ each symbol $(x, y) \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ occurs in the sequence $(\ell_{p,\delta}((vn+z)!), \nu_p((vn+z)!) \bmod u)_{n \geq 0}$ with frequency

$$\frac{1}{p^{\delta-1}(p-1)u}.$$

Hence, the second part applied to $k = v, l = z$ implies the first part. The converse is also true, since we can choose the parameter v to be divisible by k and express

the arithmetic progression $(kn+l)_{n\geq 0}$ as the union (up to finitely many terms) of arithmetic progressions of the form $(vn+z)_{n\geq 0}$ with $z \equiv l \pmod{k}$.

We now proceed to prove the first part of the statement. By Proposition 2.25 it is sufficient to show that the incidence matrix of ψ_θ is a scalar multiple of a row-stochastic matrix. Since ψ_θ is a p-uniform morphism, this is equivalent to saying that there are exactly p occurrences of each triple $(x,y,z) \in \Lambda_\theta$ in the words $\psi_\theta(x',y',z')$ with $(x',y',z') \in \Lambda_\theta$.

Let $N(x,y,z)$ denote the number of such occurrences. By Lemma 2.20 the value of $N(x,y,z)$ only depends on z. For any x',y' fixed as (i,z') varies over $\{0,1,\ldots,p-1\}\times\{0,1,\ldots,v-1\}$ the values $pz'+i \bmod v$, in the third coordinate of $\psi_\theta(x',y',z')$, are equally distributed among residue classes modulo v. Hence, each $z \in \{0,1,\ldots,v-1\}$ occurs in the third coordinate the same amount of times, or in other words, the sum $\sum_{(x,y)} N(x,y,z)$ with (x,y) running over $(\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ has the same value for each z. Since the summands are all equal too, we must have $N(x,y,z) = p$ for all $(x,y,z) \in \Lambda_\theta$, as desired. □

Roughly speaking, Theorem 2.26 asserts that for any $(p,\delta,u,v) \in \Theta$ the components of α_θ behave like independent uniformly distributed random variables. To conclude the present chapter, we apply the result just proved to verify that the frequencies of letters in the sequence $(\ell_{b,d}(n!))_{n\geq 0}$ are as stated in Theorem 2.3.

Proof of Theorem 2.3. Since $(\ell_{b,d}(n!))_{n\geq 0}$ and $(\beta_{b,d}(n!))_{n\geq 0}$ coincide on a set of density 1, it is sufficient to compute the frequencies of letters in the latter sequence. Let $a \in \{1,\ldots,b^d-1\}$ be such that $b \nmid a$. If $q_1^d \nmid a$, or equivalently $a \notin \Delta_{b,d}$, then the frequency of a is zero.

If $q_1^d \nmid a$, we use the description of $(\ell_{b,d}(n!))_{n\geq 0}$ as the image of $(\alpha_\theta(n))_{n\geq 0}$ under the coding $\tau_{b,d}$, where $\theta = (p_1, l_1 d, l_1 t_1, v)$ with an appropriate v. For any $(x,y,z) \in \Lambda_\theta$ we have $\tau_{b,d}(x,y,z) = a$ if and only if the congruences

$$\begin{aligned} x p_1^{\nu_{p_1}(a)} r_1^{\lfloor y/l_1 \rfloor} &\equiv a \pmod{p_1^{l_1 d}}, \\ y &\equiv \nu_{p_1}(a) \pmod{l_1} \end{aligned}$$

are satisfied. Among $p_1^{l_1 d-1}(p_1-1)l_1 t_1$ possible pairs $(x,y) \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times \times (\mathbb{Z}/u\mathbb{Z})^+$ there are exactly $p_1^{\nu_{p_1}(a)} t_1$ solutions to this system of congruences. Thus, by Theorem 2.26 the symbol a appears in $(\beta_{b,d}((kn+l)!))_{n\geq 0}$ with frequency

$$\frac{1}{l_1(p_1-1)} p_1^{\nu_{p_1}(a)-l_1 d+1}.$$

□

As an immediate corollary, in the case $l_1 = 1$ all the frequencies are equal.

Corollary 2.27. *Assume that $l_1 = 1$. Let $a \in \{1,\ldots,b^d-1\}$ be such that $b \nmid a$ and let $k \geq 1, l \geq 0$ be integers. The frequency of a in the sequence $(\ell_{b,d}((kn+l)!))_{n\geq 0}$ is*

$$\frac{1}{p_1^{d-1}(p_1-1)}.$$

3. Last nonzero digits of polynomials and p-adic analytic functions

In this chapter we investigate last nonzero digits of polynomials and p-adic analytic functions evaluated at integers. We shall give necessary and sufficient conditions for these sequences to be automatic or regular. Unsurprisingly, these results also turn out to be closely related to regularity of p-adic valuations of polynomials and p-adic analytic functions.

3.1 Introduction

The properties of last nonzero digits of sequences other than $(n!)_{n\geq 0}$ have also been studied by several authors (though usually not referred to in this way). In Section 2.1 we have already mentioned the results of Dresden [30, 31] on the last nonzero digit of F_n and n^n in base 10. Periodicity of the sequence $(\ell_b(n^n))_{n\geq 0}$ for any base b was studied by Grau and Oller-Marcén [38]. It can be extracted from their reasoning that for b prime this sequence is b-automatic. Robbins [63] and later Latushkin and Ushakov [46] determined precisely which Fibonacci and Lucas numbers can be expressed as a sum of three squares, which is closely connected to the study of the sequences $(\ell_{4,2}(F_n))_{n\geq 0}$ and $(\ell_{4,2}(L_n))_{n\geq 0}$ via Legendre's three-square theorem. In the case of the numbers $H_d(n)$ counting permutations on n symbols being products of disjoint d-cycles, for certain primes p the sequence $(\ell_{p,w}(H_p(n)))_{n\geq 0}$ was proved to be periodic for all w by Miska and Ulas [59, Corollary 4.8]. From the results of Ulas and Żmija [71, Theorems 3.5 and 4.1] one can deduce that for any fixed prime $p \geq 3$ and integer $m \geq 1$ the sequence $(\ell_p(d_m(n)))_{n\geq 0}$ is p-automatic, where $d_m(n)$ counts certain mp-colored p-ary partitions of n.

In this chapter we study regularity and automaticity of last nonzero digits of polynomials and p-adic analytic functions evaluated at integers. The main inspiration behind considering this particular family of functions comes from the following result by Shu and Yao [68] (earlier proved by Bell [8] for polynomials).

Theorem 3.1 (Shu, Yao). *Let $f\colon \mathbb{Z}_p \to \mathbb{C}_p$ be a p-adic locally analytic function on $\mathbb{Z}_p$ which does not have any root in $\mathbb{N}$. Then the sequence $(\nu_p(f(n)))_{n\geq 0}$ is p-regular if and only if all the roots of f in $\mathbb{Z}_p$ are contained in $\mathbb{Q}$.*

Their results rely in particular on examining the p-adic expansions of the roots of the considered functions. We observed that a similar approach also works in the case of last nonzero digits. Our study reveals a direct connection between p-regularity of the sequences of p-adic valuations and last nonzero digits of $f(n)$ in base p. We also extend the scope of our investigation to last nonzero digits in an arbitrary base $b \geq 2$. As we will see in the later sections, the results for prime power bases are quite different from those for bases having two or more distinct prime factors.

We now state the main goals for this chapter. Let $b \geq 2$ be an integer base and $d \geq 1$ the number of considered digits. Let $p_1, \ldots, p_s$ denote the prime factors of b. We study s-tuples $f = (f_1, \ldots, f_s)$, where $f_i \colon \mathbb{Z}_{p_i} \to \mathbb{Q}_{p_i}$ is locally analytic on $\mathbb{Z}_{p_i}$ for $i = 1, \ldots, s$. We consider each such f as a function defined on the set

$$\mathbb{Q} \cap \mathbb{Z}_b = \{x \in \mathbb{Q} : \nu_{p_i}(x) \geq 0 \text{ for } i = 1, \ldots, s\},$$

and write $f(x) = (f_1(x), \ldots, f_s(x))$.

This is a quite general setting, which entails many naturally occuring examples. In particular, if f_i are all equal to the same polynomial with rational coefficients, then we can identify f with this polynomial. Moreover, one can study s-tuples of $(f_1, \ldots, f_s)$ such that all f_i interpolate the same linear recurrence sequence p_i-adically.

We would like to answer the following questions.

Questions:

1. When is $(\ell_{b,d}(f(n)))_{n\geq 0}$ a k-automatic sequence?
2. When is $(\mathcal{L}_b(f(n)))_{n\geq 0}$ a k-regular sequence?

Observe that Proposition 2.9 already answers both of these questions in the case when $k = b$ and $f(x) = x$.

In view of Cobham's Theorem and its generalization by Bell (Theorems 1.6 and 1.7) there are a priori three possible answers to both questions, namely that the considered sequence is:

(a) k-automatic (resp. k-regular) for all $k \geq 2$;

(b) k-automatic (resp. k-regular) for some $k \geq 2$ and not l-automatic (resp. l-regular) for l multiplicatively independent with k;

(c) not automatic (resp. not regular).

We shall see that for both questions every case (a) – (c) occurs for a vast class of functions.

3.2 Some basic reductions

Before attempting to answer Questions 1 and 2, in this section we argue that without loss of generality we can narrow down the class of considered functions. We let

$$b = p_1^{l_1} \cdots p_s^{l_s}$$

be the prime factorization of b, where $p_1, \ldots, p_s$ are distinct primes and $l_1, \ldots, l_s$ are positive integers. For $i = 1, \ldots, s$ put $b_i = p_i^{l_i}$.

First, we explain why it is sufficient to consider $f_1, \ldots, f_s$ strictly analytic rather than locally analytic (following an argument by Shu and Yao [68]). This is convenient, as it allows us to use Strassman's Theorem together with the formula (1.9).

Proposition 3.2. *Let f_i be locally analytic on $\mathbb{Z}_{p_i}$ for $i = 1, \ldots, s$. Then there exists an integer $R \geq 0$ such that for each $i = 1, \ldots, s$ and $a = 0, 1, \ldots, b^R - 1$ the function $f_i(b^R x + a)$ of $x \in \mathbb{Z}_{p_i}$ is strictly analytic on $\mathbb{Z}_{p_i}$.*

Proof. Starting with the case $s = 1$, let f be locally analytic on $\mathbb{Z}_p$. Then for every $x \in \mathbb{Z}_p$ we can find an integer $r(x) \geq 0$ such that f is strictly analytic on the clopen ball $\overline{B}(x, p^{-r(x)})$. By compactness of $\mathbb{Z}_p$ there exist $x_1, \ldots, x_m$ such that the balls $\overline{B}(x_m, p^{-r(x_m)})$ cover $\mathbb{Z}_p$. Let r be an integer such that $r \geq \max_{1 \leq j \leq m} r(x_j)$. Then for any $a \in \mathbb{Z}$ there exists j such that we have $p^r x + a \in \overline{B}(x_j, p^{-r(x_j)})$ for all $x \in \mathbb{Z}_p$. Hence, the functions $f(p^r x + a)$ are all strictly analytic on $\mathbb{Z}_p$ as functions of x.

In the case where b has $s \geq 2$ distinct prime factors $p_1, \ldots, p_s$ we repeat the above reasoning for each p_i, obtaining a corresponding integer r_i. The assertion follows by taking R such that $R \geq r_i / l_i$ for all $i = 1, \ldots, s$. □

Theorem 1.9 says that k-regularity of the sequence $(\mathcal{L}_b(f(n)))_{n \geq 0}$ is equivalent to k-regularity of $(\mathcal{L}_b(f(b^R n + a)))_{n \geq 0}$ for all $a = 0, 1, \ldots, b^R - 1$. The same observation also applies to k-automaticity of $(\ell_{b,d}(f(n)))_{n \geq 0}$. Hence, rather than directly consider s-tuples of locally analytic functions, by virtue of Proposition 3.2 we can answer our questions "locally" for each s-tuple $f(b^R n + a)$ of strictly analytic functions.

Moreover, we can further reduce the problem to the nondegenerate case where none of the functions $f_1, \ldots, f_s$ is identically zero. Clearly, if $f_1 = \cdots = f_s = 0$, then $\mathcal{L}_b(f(n)) = \ell_{b,d}(f(n)) = 0$, so the resulting sequences are constant, and thus k-automatic for every $k \geq 2$. If not all f_i are equal to 0, we may use the following proposition.

Proposition 3.3. *Let $f = (f_1, \ldots, f_s)$, where f_i is a strictly analytic function on $\mathbb{Z}_{p_i}$ for $i = 1, \ldots, s$. Assume that $f_s = 0$. Let $\overline{b} = b/b_s$ and $\overline{f} = (f_1, \ldots, f_{s-1})$. Then for any integer $k \geq 2$ we have the following:*

(i) *the sequence $(\mathcal{L}_b(f(n)))_{n \geq 0}$ is k-regular if and only if $(\mathcal{L}_{\overline{b}}(\overline{f}(n)))_{n \geq 0}$ is k-regular;*

(ii) *the sequence $(\ell_{b,d}(f(n)))_{n \geq 0}$ is k-automatic if and only if $(\ell_{\overline{b},d}(\overline{f}(n)))_{n \geq 0}$ is k-automatic.*

Proof. Because $\nu_{b_s}(f_s(n)) = +\infty$, we have $\nu_b(f(n)) = \nu_{\overline{b}}(\overline{f}(n))$ for all integers $n \geq 0$. Hence, we can write

$$\mathcal{L}_b(f(n)) = \left(\mathcal{L}_{\overline{b}}(\overline{f}(n)), 0\right).$$

The first assertion follows from Proposition 1.14 which says that a sequence of pairs $(u_n, v_n)_{n \geq 0}$ is k-regular if and only if both sequences $(u_n)_{n \geq 0}$ and $(v_n)_{n \geq 0}$ are k-regular.

Part (ii) uses the fact that the isomorphism between the rings $\mathbb{Z}/b^d\mathbb{Z}$ and $\mathbb{Z}/b_1\mathbb{Z} \times \cdots \times \mathbb{Z}/b_s\mathbb{Z}$ maps $\ell_{b,d}(f(n))$ to the s-tuple of values whose ith component is equal to

$$b_i^{\nu_{b_i}(f_i(n))-\nu_b(f(n))} r_i^{\nu_b(f(n))} \ell_{b_i,d}(f_i(n)) \bmod b_i^d \tag{3.1}$$

for $i = 1, \ldots, s$, where r_i denotes a multiplicative inverse of b/b_i modulo b_i^d. This is a direct consequence of the congruence relations (2.5). Theorem 1.11 implies that $(\ell_{b,d}(f(n)))_{n\geq 0}$ is k-automatic if and only if for each $i = 1, \ldots, s$ the sequence consisting of the terms (3.1) is k-automatic. But for $i = s$ this is the zero sequence, which is l-automatic for any $l \geq 2$. The result follows. □

Propositions 3.2 and 3.3 encourage us to consider the family of functions

$$\mathcal{A}_b = \{f = (f_1, \ldots, f_s) : f_i \colon \mathbb{Z}_{p_i} \to \mathbb{Q}_{p_i} \text{ is strictly analytic on } \mathbb{Z}_{p_i}, \\ f_i \neq 0 \text{ for } i = 1, \ldots, s\}.$$

We point out that $\mathcal{A}_b$ only depends on the set of prime factors of b.

The following result shows that a further reduction can be made when studying regularity of $(\mathcal{L}_b(f(n)))_{n\geq 0}$.

Proposition 3.4. *Let $f = (f_1, \ldots, f_s) \in \mathcal{A}_b$ be such that the sequence $(\mathcal{L}_b(f(n)))_{n\geq 0}$ is k-regular for some $k \geq 2$. Then for each $i = 1, \ldots, s$ the function f_i is a polynomial.*

Proof. We first prove that if $g \in \mathcal{A}_p$ for some prime p and $(g(n))_{n\geq 0}$ is k-regular for some $k \geq 2$, then g must be a polynomial. Write

$$g(x) = \sum_{m=0}^{\infty} a_m x^m,$$

where $a_m \in \mathbb{Q}_p$ for $m = 0, 1, \ldots$. Since $(g(n))_{n\geq 0}$ is k-regular, the $\mathbb{Z}$-submodule generated by the subsequences $(g(k^j n))_{n\geq 0}$ with $j \geq 0$ is finitely generated. Choose an integer $J \geq 0$ such that the subsequences $(g(k^j n))_{n\geq 0}$ with $j = 0, 1, \ldots, J$ generate this submodule. Let $\alpha_0, \alpha_1, \ldots, \alpha_J \in \mathbb{Z}$ be such that

$$g(k^{J+1} n) = \sum_{j=0}^{J} \alpha_j g(k^j n)$$

for all $n \geq 0$. Using the fact that $\mathbb{N}$ is dense in $\mathbb{Z}_p$, we can equate the coefficients of the power series on both sides and obtain

$$a_m \left(k^{(J+1)m} - \sum_{j=0}^{J} \alpha_j k^{jm} \right) = 0$$

for all $m \geq 0$. However, the expression in the parentheses tends to infinity as $m \to \infty$, so only finitely many a_m can be nonzero.

Moving on to the general case, by Strassman's Theorem each f_i has only finitely many roots in $\mathbb{Z}_{p_i}$. Therefore, we can find integers $t \geq 0$ and $c \geq 0$ such that for each

$i = 1, \ldots, s$ the function $f_i(b^t x + c)$ of $x \in \mathbb{Z}_{p_i}$ has at most one root in $\mathbb{Z}_{p_1}$. Lemma 3.20 below shows that the sequence $(\nu_b(f(b^t n + c)))_{n \geq 0}$ is periodic with period being a power of b. Hence, there exists an integer $T \geq t$ such that $(\nu_b(f(b^T n + c)))_{n \geq 0}$ is a constant sequence with all terms equal to, say, v. This means that

$$\mathcal{L}_b(f(b^T n + c)) = b^{-v} f(b^T n + c)$$

for all $n \geq 0$. Since we assume that $(\mathcal{L}_b(f(n)))_{n \geq 0}$ is k-regular, according to Proposition 1.14 and Theorem 1.9, the sequences $(b^{-v} f_i(b^T n + c))_{n \geq 0}$ for $i = 1, \ldots, s$ are all k-regular. By our earlier reasoning we deduce that each of the functions $b^{-v} f_i(b^T x + c)$ is a polynomial in x, and thus also f_i is a polynomial. The result follows. □

This result motivates us to distinguish the subset $\mathcal{P}_b \subset \mathcal{A}_b$, defined by

$$\mathcal{P}_b = \{f = (f_1, \ldots, f_s) : f_i \in \mathbb{Q}_{p_i}[X], f_i \neq 0 \text{ for } i = 1, \ldots, s\}.$$

Taking into account the results proved so far, we reformulate Questions 1 and 2 by adding certain assumptions on the considered functions.

Questions:

1′. Let $f \in \mathcal{A}_b$. When is $(\ell_{b,d}(f(n)))_{n \geq 0}$ a k-automatic sequence?

2′. Let $f \in \mathcal{P}_b$. When is $\mathcal{L}_b(f(n))_{n \geq 0}$ a k-regular sequence?

Our main focus from now on is to answer Questions 1′ and 2′.

3.3 Prime power bases

In this section we address Questions 1′ and 2′ in the case where $b = p^l$ is a prime power.

We let $\mathcal{R}_f$ denote the (finite) set of roots of a function $f \in \mathcal{A}_p$. For each root $\theta \in \mathcal{R}_f$ let m_θ be its multiplicity and $g_\theta \in \mathcal{A}_p$ – the function such that

$$f(x) = (x - \theta)^{m_\theta} g_\theta(x)$$

for all $x \in \mathbb{Z}_p$.

We now state the main results of the present section. The first theorem fully answers Question 1′.

Theorem 3.5. *Let $f \in \mathcal{A}_p$, $d \geq 1$ and let $\mathcal{R}'_f$ be the subset of $\mathcal{R}_f$ defined by*

$$\mathcal{R}'_f = \{\theta \in \mathcal{R}_f : l \nmid m_\theta \text{ or } \lambda(p^{ld}) \nmid p^{\nu_p(g_\theta(\theta)) \bmod l} m_\theta\}.$$

Then the sequence $(\ell_{p^l,d}(f(n)))_{n \geq 0}$ is

(a) *periodic, if $\mathcal{R}'_f = \varnothing$;*

(b) *p-automatic and not k-automatic for k multiplicatively independent with p, if* $\varnothing \neq \mathcal{R}'_f \subset \mathbb{Q}$;

(c) *not k-automatic for any* $k \geq 2$ *if* $\mathcal{R}'_f \setminus \mathbb{Q} \neq \varnothing$.

We note that in the case $l = 1$ the form of $\mathcal{R}'_f$ simplifies to

$$\mathcal{R}'_f = \{\theta \in \mathcal{R}_f : \lambda(p^d) \nmid m_\theta\}.$$

As mentioned in Section 3.2, a corresponding result for p-adic locally analytic functions can be derived from Theorem 3.5. The statement becomes more complicated though, as in this case all the conditions need to be considered locally.

The second main theorem provides a complete classification of all cases corresponding to Question 2′. Note that in its statement $\mathcal{R}_f$ plays exactly the same role as $\mathcal{R}'_f$ in Theorem 3.5.

Theorem 3.6. *Let* $f \in \mathcal{P}_p$. *Then the sequence* $(\mathcal{L}_{p^l}(f(n)))_{n\geq 0}$ *is*

(a) *k-regular for every* $k \geq 2$ *if* $\mathcal{R}_f = \varnothing$;

(b) *p-regular and not k-regular for k multiplicatively independent with p, if* $\varnothing \neq \mathcal{R}_f \subset \mathbb{Q}$;

(c) *not k-regular for any* $k \geq 2$ *if* $\mathcal{R}_f \setminus \mathbb{Q} \neq \varnothing$.

In particular, the above conditions are independent of l.

We defer the proofs of both theorems to the end of this section.

Theorems 3.5 and 3.6 combined with the results of Bell, Shu and Yao reveal an interesting connection between p-adic valuations and last nonzero digits of polynomials and p-adic locally analytic functions evaluated at integers.

Corollary 3.7. *Let* $f \in \mathcal{P}_p$ *be such that f has no root in* $\mathbb{N}$. *Then the sequence* $(\nu_p(f(n)))_{n\geq 0}$ *is p-regular if and only if the sequence* $(\mathcal{L}_{p^l}(f(n)))_{n\geq 0}$ *is p-regular.*

Proof. By Theorem 3.1 p-regularity of $(\nu_p(f(n)))_{n\geq 0}$ is equivalent to f having no roots in $\mathbb{Z}_p \setminus \mathbb{Q}$. The result follows from Theorem 3.6. □

Corollary 3.8. *Let f be locally analytic on* $\mathbb{Z}_p$ *with no root in* $\mathbb{N}$. *Then the sequence* $(\nu_p(f(n)))_{n\geq 0}$ *is p-regular if and only if for all* $d \geq 1$ *the sequence* $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ *is p-automatic.*

Proof. By Proposition 3.2, we can find and integer $R \geq 0$ such that for all $a = 0, 1, \ldots, p^R - 1$ the functions $f_a(x) = f(p^R x + a)$ of $x \in \mathbb{Z}_p$ are strictly analytic on $\mathbb{Z}_p$. Moreover, the sequence $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ is p-automatic if and only if the sequence $(\ell_{p^l,d}(f_a(n)))_{n\geq 0}$ is p-automatic for every a.

Now, for fixed a the latter sequence is p-automatic if and only if f_a it satisfies the condition (a) or (b) of Theorem 3.5 for all $d \geq 1$, which is further equivalent to f_a having no roots in $\mathbb{Z}_p \setminus \mathbb{Q}$. This, by Theorem 3.1 is a necessary and sufficient condition for p-regularity of $(\nu_p(f_a(n)))_{n\geq 0}$.

Our claim follows from the fact that $(\nu_p(f(n)))_{n\geq 0}$ is p-regular if and only if all of the above subsequences are p-regular. □

In particular, Corollary 3.8 can be applied to linear recurrence sequences whose p-adic valuation is known to be p-regular (such examples are discussed in more detail in Chapter 4) to deduce that their last nonzero digits in the base p^l form a p-automatic sequence. We note that although the family of strictly analytic functions collectively interpolating a linear recurrence sequence cannot always be "glued" into a locally analytic function, a similar argument involving subsequences along arithmetic progressions also works in this case. Unfortunately, it seems that a reasoning in the opposite direction is hard to perform in practice, as determining last nonzero digits usually requires computing the valuation first.

We now head towards proving Theorems 3.5 and 3.6. To begin, we give an overview of the main tools used in the proofs. First of all, by (1.9), for $f \in \mathcal{A}_p$ we have the factorization

$$f(x) = \left(\prod_{\theta \in \mathcal{R}_f} (x-\theta)^{m_\theta} \right) g(x), \tag{3.2}$$

where $\mathcal{R}_f$ denotes the (finite) set of roots of f in $\mathbb{Z}_p$, the number m_θ is the multiplicity of a root θ, and $g \in \mathcal{A}_p$ has no root in $\mathbb{Z}_p$. Obviously, if f is a polynomial with coefficients in $\mathbb{Q}_p$, then so is g.

Dealing with the case $l \geq 2$, much of the time we will be relying on the factorizations

$$\mathcal{L}_{p^l}(x) = p^{\nu_p(x) \bmod l} \mathcal{L}_p(x), \tag{3.3}$$

$$\ell_{p^l,d}(x) \equiv p^{\nu_p(x) \bmod l} \ell_{p,ld}(x) \pmod{p^{ld}}, \tag{3.4}$$

valid for all nonzero $x \in \mathbb{Q}_p$, as proved in Section 2.2. These two formulas will often be combined with the multiplicative properties of $\mathcal{L}_p$ and $\ell_{p,\delta}$, namely

$$\mathcal{L}_p(xy) = \mathcal{L}_p(x)\mathcal{L}_p(x), \tag{3.5}$$

$$\ell_{p,\delta}(xy) \equiv \ell_{p,\delta}(x)\ell_{p,\delta}(y) \pmod{p^\delta}, \tag{3.6}$$

for all $x, y \in \mathbb{Q}_p$. Here and in the sequel we write δ to denote the number of considered digits in base p, whereas d plays the same role with respect to the base p^l (usually we will take $\delta = ld$).

We are not going to apply the above formulas directly to $f(n)$ factorized according to (3.2). Instead, we will study last nonzero digits of p-adic strictly analytic functions having at most one distinct root in $\mathbb{Z}_p$ (and satisfying certain additional conditions). This is justified by Proposition 3.10 below. First, we give an auxiliary result, which is an extension of a lemma by Shu and Yao [68, Lemma on p. 949].

Proposition 3.9. *Let p be a prime and let $f \in \mathcal{A}_p$ be such that f has no root in $\mathbb{Z}_p$. Then for any integer $\delta \geq 1$ there exists an integer $T \geq 0$ such that we have*

$$\nu_p(f(p^T x + y)) = \nu_p(f(y))$$

and

$$\ell_{p,\delta}(f(p^T x + y)) = \ell_{p,\delta}(f(y))$$

for all $x, y \in \mathbb{Z}_p$. In particular, the sequences $(\nu_p(f(n)))_{n\geq 0}$ and $(\ell_{p,\delta}(f(n))_{n\geq 0}$ are periodic with a period being a power of p.

Proof. Theorem 1.22 implies that there exists an integer $V \geq 0$ such that $\nu_b(f(x)) \leq V$ for all $x \in \mathbb{Z}_p$. Letting T be as in Proposition 1.24 applied to $M = V + \delta$, we immediately get both equalities. Periodicity follows by plugging in $x = n, y = 1$. □

As a sidenote, Medina, Moll and Rowland [57] computed the minimal period of $(\nu_p(f(n)))_{n\geq 0}$ in the special case when f is a polynomial with integer coefficients irreducible over $\mathbb{Z}_p$.

In the proof of Proposition 3.10 (and several further results) for $\theta \in \mathbb{Z}_p$ and $j \geq 0$ integer we will be using the notation $\theta[j] = \theta \bmod p^j$ and $\theta\{j\} = p^{-j}(\theta - \theta[j])$. In other words, $\theta[j]$ is represented by j initial digits in the p-adic expansion of θ, while $\theta\{j\}$ is obtained from θ by deleting these digits.

Proposition 3.10. *Let $f \in \mathcal{A}_p$ and let $\delta \geq 1$ be an integer. Then for each sufficiently large integer $T \geq 0$ and for each $a = 0, 1, \ldots, p^T - 1$ the function $f_a \in \mathcal{A}_p$, defined by*

$$f_a(x) = f(p^T x + a)$$

for $x \in \mathbb{Z}_p$, satisfies one of the following conditions:

(i) *if $a \neq \theta[T]$ for all $\theta \in \mathcal{R}_f$, then f_a has no root in $\mathbb{Z}_p$;*

(ii) *if $a = \theta[T]$ for some $\theta \in \mathcal{R}_f$, then*

$$f_a(x) = p^{Tm_\theta} \left(x - \theta\{T\}\right)^{m_\theta} g_\theta(p^T x + \theta[T]), \tag{3.7}$$

where $g_\theta(p^T x + \theta[T])$ additionally satisfies

$$\nu_p(g_\theta(p^T x + \theta[T])) = \nu_p(g_\theta(\theta)), \tag{3.8}$$

$$\ell_{p,\delta}(g_\theta(p^T x + \theta[T])) = \ell_{p,\delta}(g_\theta(\theta)) \tag{3.9}$$

for all $x \in \mathbb{Z}_p$.

Proof. If f has no roots in $\mathbb{Z}_p$, we can take $T = 0$ so that the condition (i) is satisfied.

If f has a root in $\mathbb{Z}_p$, we first find an integer $t \geq 0$ such that for each $\theta \in \mathbb{R}_f$ the function $g_\theta(p^t x + \theta[t])$ of x has no root in $\mathbb{Z}_p$. In the case when f has precisely one distinct root, we can choose arbitrary $t \geq 0$. Otherwise, let t be any such that

$$t > \max\{\nu_p(\theta - \sigma) : \theta, \sigma \in \mathcal{R}_f, \theta \neq \sigma\}.$$

For each $\theta \in \mathcal{R}_f$ let T_θ be the integer obtained from Proposition 3.9 applied to the function $g_\theta(p^t x + \theta[t])$. Choosing any integer $T \geq t + \max_{\theta\in\mathcal{R}_f} T_\theta$, we see that for all $\theta \in \mathcal{R}_f$ the values $\nu_p(g_\theta(p^T x + \theta[T]))$ and $\ell_{p,\delta}(g_\theta(p^T x + \theta[T]))$ are constant with respect to $x \in \mathbb{Z}_p$. By continuity of g, we get (3.8) and (3.9). As a result, for each $a = 0, 1, \ldots, p^T - 1$ one of the conditions (i),(ii) is satisfied. □

We now discuss how Proposition 3.10 affects our approach. First, let $f \in \mathcal{A}_p$ and consider the sequence $(\ell_{p^l,d}(f(n)))_{n\geq 0}$. It is k-automatic if and only if for each $a = 0, 1, \ldots, p^T - 1$ the sequence $(\ell_{p^l,d}(f(p^T n + a)))_{n\geq 0}$ is k-automatic. Moreover, if $a \neq \theta[T]$ for any $\theta \in \mathbb{R}_f$, then by (3.4) and Proposition 3.9 $(\ell_{p^l,d}(f(p^T n + a)))_{n\geq 0}$ is a

periodic sequence, hence k-automatic for every $k \geq 2$. For the same reason, the factor $g_\theta(p^T x+\theta[T])$ in (3.7) does not affect automaticity of $(\ell_{p^l,d}(f(p^T n+a)))_{n\geq 0}$ in the case when $a \neq \theta[T]$. An analogous observation regarding regularity of $(\mathcal{L}_{p^l,d}(f(n)))_{n\geq 0}$ for $f \in \mathcal{P}_p$ also turns out to be true (this is explained in more detail in the proof of Theorem 3.6).

All things considered, in essence it remains to study last nonzero digits of the expression $p^{Tm_\theta}\,(x-\theta\{T\})^{m_\theta}$ in (3.7), which can be written in a general form $p^v(n-\theta)^m$ with m, v integers such that $m \geq 1, v \geq 0$ and $\theta \in \mathbb{Z}_p$. We investigate its properties in a series of lemmas. In the proofs of Theorems 3.5 and 3.6 we will combine all the intermediate results into a statement concerning a general function $f \in \mathcal{A}_p$.

We start with a useful technical result which more or less says that a p-adic integer θ is uniquely determined by the values $\nu_p(n-\theta) \bmod l$ as well as $\ell_{p,\delta}(n-\theta)$ with $n \in \mathbb{N}$.

Lemma 3.11. *Let $\theta, \sigma \in \mathbb{Z}_p$ be such that $\theta \neq \sigma$ and let $l \geq 2, \delta \geq 1$ be integers. Then*

(i) *there exist infinitely many integers $n \geq 0$ such that $\nu_p(n-\theta) \not\equiv \nu_p(n-\sigma) \pmod{l}$;*

(ii) *for any $x \in (\mathbb{Z}/p^\delta\mathbb{Z})^\times$ such that $x \neq \ell_{p,\delta}(\theta-\sigma)$ there exist infinitely many integers $n \geq 0$ such that $\ell_{p,\delta}(n-\theta) = x$ and $\ell_{p,\delta}(n-\sigma) = \ell_{p,\delta}(\theta-\sigma)$.*

Proof. Let $v = \nu_p(\theta-\sigma)$. In part (i) we put $n = p^{v+1}(1+pt)+\theta[v+2]$, where $t \geq 1$ is an arbitrary integer. This yields $\nu_p(n-\theta) = v+1$ and $\nu_p(n-\sigma) = v$.

Part (ii) is vacuously true if $p = 2$ and $\delta = 1$. If this is not the case, we choose $n = p^{v+\delta}(x+p^\delta t)+\theta[v+2\delta]$, obtaining

$$\ell_{p,\delta}(n-\theta) = \ell_{p,\delta}(x+p^\delta(\theta\{v+2\delta\}+t)) = x$$

and

$$\ell_{p,\delta}(n-\sigma) = \ell_{p,\delta}(p^\delta x+p^{2\delta}(\theta\{v+2\delta\}+t)+p^{-v}(\theta-\sigma)) = \ell_{p,\delta}(\theta-\sigma),$$

as desired. □

We let λ denote the Carmichael function, which assigns to each positive integer n the least positive integer m such that $a^m \equiv 1 \pmod{n}$ for all integers a coprime with n. In particular, when $n = p^l$ is a prime power, we have

$$\lambda(p^l) = \begin{cases} p^{l-1}(p-1) & \text{if } p \neq 2 \text{ or } p = 2, l \leq 2, \\ 2^{l-2} & \text{if } p = 2, l \geq 3. \end{cases}$$

In the following two lemmas we study periodicity of last nonzero digits in base p^l of expressions of the form $p^v(n-\theta)^m$. It turns out that the conditions $l \mid m$ and $\lambda(p^{ld-v}) \mid m$, which almost immediately imply periodicity of $(\ell_{p^l,d}(p^v(n-\theta)^m))_{n\geq 0}$, are also necessary.

Lemma 3.12. *Let $m \geq 1, l \geq 1$ be integers and let $\theta \in \mathbb{Z}_p \setminus \mathbb{N}$. The following conditions are equivalent:*

(i) *the sequence* $(\nu_p((n-\theta)^m) \bmod l)_{n\geq 0}$ *is eventually periodic;*

(ii) $\nu_p((n-\theta)^m) \equiv 0 \pmod{l}$ *for all* $n \geq 0$*;*

(iii) $l \mid m$*.*

Proof. If $l = 1$, then all three conditions are always satisfied, so in what follows we assume that $l \geq 2$.

Obviously, (iii) implies (ii) and (ii) implies (i).

It remains to show the implication from (i) to (iii). For the sake of contradiction, suppose that $(\nu_p((n-\theta)^m) \bmod l)_{n\geq 0}$ is eventually periodic with period $T > 0$, but $l \nmid m$. Recall that the image under a coding of a k-automatic sequence is again k-automatic. Since there is a bijection mapping the values $\nu_p(x^m) = m\nu_p(x)$ modulo l to the values $\nu_p(x)$ modulo $l/\gcd(m,l)$, without loss of generality we may consider $m = 1$.

Lemma 3.11(i) with $\sigma = \theta - T$ shows that the congruence $\nu_p(n+T-\theta) \equiv \nu_p(n-\theta) \pmod{l}$ does not hold for infinitely many n, which gives a contradiction with eventual periodicity of the sequence $(\nu_p((n-\theta)^m) \bmod l)_{n\geq 0}$. □

Lemma 3.13. *Let* d, m, l, v *be integers such that* $d \geq 1$, $m \geq 1$, *and* $0 \leq v < l$*. Let* $\theta \in \mathbb{Z}_p \setminus \mathbb{N}$*. The following conditions are equivalent:*

(i) *the sequence* $(\ell_{p^l,d}(p^v(n-\theta)^m))_{n\geq 0}$ *is eventually periodic;*

(ii) $\ell_{p^l,d}(p^v(n-\theta)^m) = p^v$ *for all* $n \geq 0$*;*

(iii) $l \mid m$ *and* $\lambda(p^{ld-v}) \mid m$*.*

Proof. The implication from (ii) to (i) is trivial.

If the condition (iii) holds, then $\nu_p(p^v(n-\theta)^m) \bmod l = v$ and

$$\ell_{p,ld}((n-\theta)^m) \equiv 1 \pmod{p^{ld-v}}$$

for all $n \geq 0$. Using (3.4) and (3.6), we obtain (ii).

Now assume that the sequence $(\ell_{p^l,d}(p^v(n-\theta)^m))_{n\geq 0}$ is eventually periodic with period T. Then $(\nu_p(p^v(n-\theta)^m) \bmod l)_{n\geq 0}$ is also eventually periodic, so $l \mid m$ by Lemma 3.12.

Suppose that $\lambda(p^{ld-v}) \nmid m$. Lemma 3.11(ii) applied to $x \in (\mathbb{Z}/p^{ld}\mathbb{Z})^\times$ such that $x^m \not\equiv (\ell_{p,ld}(T))^m \pmod{p^{ld-v}}$ and $\sigma = \theta - T$ gives arbitrarily large n satisfying

$$\ell_{p^l,d}(p^v(n+T-\theta)^m) \neq \ell_{p^l,d}(p^v(n-\theta)^m).$$

We get a contradiction, so $\lambda(p^{ld-v}) \mid m$, and thus (i) implies (iii). □

We now move on to the question of automaticity and regularity of last nonzero digits of the expression $p^v(n-\theta)^m$ for consecutive n. First, we show that the resulting sequences are p-regular or p-automatic when θ is rational.

Lemma 3.14. *Let $\theta \in \mathbb{Q} \cap \mathbb{Z}_p$. Let m, l, v be integers such that $m \geq 1$ and $0 \leq v < l$. Then the sequence $(\mathcal{L}_{p^l}(p^v(n-\theta)^m))_{n\geq 0}$ is p-regular and for all $d \geq 1$ the sequence $(\ell_{p^l,d}(p^v(n-\theta)^m))_{n\geq 0}$ is p-automatic.*

Proof. Theorem 1.9 together with Remark 1.10 show that without loss of generality we may consider $\theta \notin \mathbb{N}$ (by replacing n with $n+\theta+1$ if $\theta \in \mathbb{N}$). The equality (3.3) yields

$$\mathcal{L}_{p^l}(p^v(n-\theta)^m) = p^{(v+m\nu_p(n-\theta)) \bmod l}\mathcal{L}_p((n-\theta)^m). \tag{3.10}$$

By Theorem 3.1 the sequence $(\nu_p(n-\theta))_{n\geq 0}$ is p-regular, and thus becomes p-automatic when reduced modulo l. This means that the first factor on the right-hand side of (3.10) is p-automatic, hence p-regular.

We now consider the second factor. Write $\theta = c/a$ in lowest terms with a positive. We have

$$\mathcal{L}_p(n-\theta) = \mathcal{L}_p\left(\frac{an-c}{a}\right) = \mathcal{L}_p(an-c)\mathcal{L}_p\left(\frac{1}{a}\right).$$

By Proposition 2.9 and Theorem 1.9 the sequence $(\mathcal{L}_p(an-c))_{n\geq 0}$ is p-regular. Due to Remark 1.10 this is true even if $c > 0$, in which case $an-c$ can take negative values. Hence, $(\mathcal{L}_p(n-\theta))_{n\geq 0}$ is also p-regular. By the multiplicative property (3.5) of $\mathcal{L}_p$ we obtain $\mathcal{L}_p((n-\theta)^m) = (\mathcal{L}_p(n-\theta))^m$, which means that $(\mathcal{L}_p((n-\theta)^m))_{n\geq 0}$ is p-regular as the termwise product of p-regular sequences. Our claim follows. □

On the other hand, if θ is an irrational p-adic integer, the resulting sequence of last nonzero digits turns out to be nonautomatic (or nonregular), unless the condition in Lemma 3.13(iii) is satisfied. This is exhibited in the next two lemmas.

Lemma 3.15. *Let $\theta \in \mathbb{Z}_p \setminus \mathbb{Q}$ and let $m \geq 1, l \geq 2$ be integers such that $l \nmid m$. Then the sequence $(\nu_p((n-\theta)^m) \bmod l)_{n\geq 0}$ is not automatic.*

Proof. Similarly as in the proof of Lemma 3.12 it is sufficient to prove the assertion in the case $m = 1$ and $l \geq 2$.

Let $k \geq 2$ and write $k = p^e c$, where $e \geq 0$ and $\nu_p(c) = 0$. By Theorem 1.5 a sequence is k-automatic if and only if it is k^t-automatic, so without loss of generality we can assume that $l \mid e$. We will show that for all integers $j \geq 0$ the subsequences $(\nu_p(k^j n + \theta[ej] - \theta) \bmod l)_{n\geq 0}$ from the k-kernel are distinct. Their terms may be simplified to the form

$$\nu_p(p^{ej}c^j n + \theta[ej] - \theta) \equiv ej + \nu_p(n - c^{-j}\theta\{ej\}) \equiv \nu_p(n - c^{-j}\theta\{ej\}) \pmod{l}.$$

Hence, for any fixed $i, j \geq 0$ with $j \neq i$ it is enough to find $n \geq 0$ such that

$$\nu_p(n - c^{-j}\theta\{ej\}) \not\equiv \nu_p(n - c^{-i}\theta\{ei\}) \pmod{l}. \tag{3.11}$$

Because θ is irrational, we have $c^{-j}\theta\{ej\} \neq c^{-i}\theta\{ei\}$, and thus Lemma 3.11(i) gives a value of n for which (3.11) holds. Consequently, $(\nu_p((n-\theta)^m) \bmod l)_{n\geq 0}$ is not a k-automatic sequence. □

Lemma 3.16. *Let d, m, l, v be integers such that $d \geq 1$, $m \geq 1$, and $0 \leq v < l$. Let $\theta \in \mathbb{Z}_p \setminus \mathbb{Q}$. If $l \nmid m$ or $\lambda(p^{ld-v}) \nmid m$, then the sequence $(\ell_{p^l,d}(p^v(n-\theta)^m))_{n\geq 0}$ is not automatic.*

Proof. If $l \nmid m$, then the result follows by Lemma 3.15, as k-automaticity of the sequence $(\ell_{p^l,d}(p^v(n-\theta)^m))_{n\geq 0}$ entails k-automaticity of $(\nu_p((n-\theta)^m))_{n\geq 0}$.

Otherwise, if $l \mid m$, then we must have $\lambda(p^{ld-v}) \nmid m$. Again, take $k \geq 2$ and write $k = p^e c$, where $e \geq 0$ and $\nu_p(c) = 0$. Replacing k by $k^{\lambda(p^{ld})}$ if necessary, we may assume that $c \equiv 1 \pmod{p^{ld}}$. We want to show that for $j \geq 0$ the subsequences $(\ell_{p^l,d}(p^v(k^j n + \theta[ej] - \theta)^m))_{n\geq 0}$ from the k-kernel are all distinct.

By $l \mid m$ we obtain

$$\ell_{p^l,d}(p^v(k^j n + \theta[ej] - \theta)^m) \equiv p^v(\ell_{p,ld}(n - c^{-j}\theta\{ej\}))^m \pmod{p^{ld}}.$$

Hence, it is sufficient to prove that for $j \geq 0$ there are infinitely many distinct sequences of the form $((\ell_{p,ld-v}(n - c^{-j}\theta\{ej\}))^m)_{n\geq 0}$.

Take any $i, j \geq 0$, where $i \neq j$. Let $x \in (\mathbb{Z}/p^{ld-v}\mathbb{Z})^\times$ be such that

$$x^m \not\equiv (\ell_{p,ld-v}(c^{-j}\theta\{ej\} - c^{-i}\theta\{ei\}))^m \pmod{p^{ld-v}}.$$

Then from Lemma 3.11(ii) we get an integer $n \geq 0$ satisfying

$$(\ell_{p,ld-v}(n - c^{-j}\theta\{ej\}))^m \not\equiv (\ell_{p,ld-v}(n - c^{-i}\theta\{ei\}))^m \pmod{p^{ld-v}},$$

and the result follows. □

We are now ready to prove Theorems 3.5 and 3.6.

Proof of Theorem 3.5. The case where f has no roots in $\mathbb{Z}_p$ falls under (a). Then Proposition 3.9 ensures that both factors on the right side of

$$\ell_{p^l,d}(f(n)) \equiv p^{\nu_p(f(n)) \bmod l}\ell_{p,ld}(f(n)) \pmod{p^{ld}}$$

are periodic with respect to n, and the result follows.

Now assume that f has at least one root in $\mathbb{Z}_p$. Similarly as in the proof of Lemma 3.14, without loss of generality we may assume that the roots of f do not lie in $\mathbb{N}$ (by replacing n with $n + M$ for some sufficiently large integer M).

Let T be as in Proposition 3.10 and divisible by l. We are going to examine the subsequences $(\ell_{p^l,d}(f(p^T n + a)))_{n\geq 0}$ with a ranging over $\{0, 1, \ldots, p^T - 1\}$.

If there is no $\theta \in \mathcal{R}_f$ such that $a = \theta[T]$, then the function $f(p^T x + a)$ has no root in $\mathbb{Z}_p$, so the sequence $(\ell_{p^l,d}(f(p^T n + a)))_{n\geq 0}$ is periodic, as has been already shown. Hence, the sequence $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ is eventually periodic (resp. k-automatic) if and only if all the remaining subsequences $(\ell_{p^l,d}(f(p^T n + \theta[T])))_{n\geq 0}$ with $\theta \in \mathcal{R}_f$ are eventually periodic (resp. k-automatic).

For $\theta \in \mathcal{R}_f$ denote $v_\theta = \nu_p(g_\theta(\theta)) \bmod l$ and $\ell_\theta = \ell_{p,ld}(g_\theta(\theta))$. Then for all $n \geq 0$ we have

$$\ell_{p^l,d}(p^{-v_\theta} g(p^T n + \theta[T])) = \ell_\theta,$$

as in the statement of Proposition 3.10.

Therefore, applying the function $\ell_{p^l,d}$ to both sides of the equality (3.7), we obtain

$$\ell_{p^l,d}(f(p^T n + \theta[T])) \equiv \ell_{p^l,d}(p^{v_\theta}(n - \theta\{T\})^{m_\theta})\ell_\theta \pmod{p^{ld}} \tag{3.12}$$

(here we used $l \mid T$ and property (ii) of Lemma 2.5). We point out that $\theta \notin \mathbb{N}$ also implies $\theta\{T\} \notin \mathbb{N}$.

Now we move on to consider the cases (a) – (c) separately. In the case (a) we have $l \mid m_\theta$ and $\lambda(p^{ld}) \mid p^{v_\theta} m_\theta$ for every $\theta \in \mathcal{R}_f$. The congruence (3.12) then becomes

$$\ell_{p^l,d}(f(p^T n + \theta[T])) \equiv p^{v_\theta}\ell_\theta \pmod{p^{ld}},$$

and periodicity of $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ follows.

In the case (b), if $\theta \in \mathcal{R}_f$ is irrational, then again $l \mid m_\theta, \lambda(p^{ld}) \mid p^{v_\theta} m_\theta$ and the reasoning is exactly the same as in the previous case. For each $\theta \in \mathbb{Q} \cap \mathcal{R}_f$ the first factor on the right-hand side of (3.12) is p-automatic by Lemma 3.14. Hence, the sequence $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ is p-automatic.

Now if this sequence were k-automatic for some k multiplicatively independent with p, then by Cobham's Theorem it would be eventually periodic. Choose $\theta \in \mathcal{R}'_f$. If $l \nmid m_\theta$, then eventual periodicity of (3.12) is ruled out by Lemma 3.12. If $l \mid m_\theta$, the we must have $\lambda(p^{ld}) \nmid p^{v_\theta} m_\theta$ and again, (3.12) cannot be eventually periodic due to Lemma 3.13.

In the case (c) we choose $\theta \in \mathcal{R}'_f \setminus \mathbb{Q}$. In the same way as above, Lemmas 3.15 and 3.16 imply that the right-hand side of (3.12) is not automatic. □

Proof of Theorem 3.6. In the case (a) the sequence $(p^{-\nu_{p^l}(f(n))})_{n\geq 0}$ is periodic by Lemma 3.9, and thus k-regular for every k. This is also the case for $(f(n))_{n\geq 0}$ by Corollary 1.13. Therefore $(\mathcal{L}_{p^l}(f(n)))_{n\geq 0}$ is also k-regular as the termwise product of k-regular sequences.

In the case (b), by the same argument as in the proof of Theorem 3.5 we may assume that the roots of f do not lie in $\mathbb{N}$. We have

$$\mathcal{L}_{p^l}(f(n)) = p^{\nu_p(f(n)) \bmod l}\mathcal{L}_p(f(n))$$

for all $n \geq 0$. The first factor on the right hand-side is p-regular because of Theorem 3.1.

In order to prove that the second factor is p-regular as well, write f in the form

$$f(x) = \left(\prod_{\theta\in\mathcal{R}_f} (x-\theta)^{m_\theta}\right) g(x),$$

where g is a polynomial irreducible over $\mathbb{Z}_p$. Multiplicativity of $\mathcal{L}_p$ yields

$$\mathcal{L}_p(f(n)) = \left(\prod_{\theta\in\mathcal{R}_f} \mathcal{L}_p((n-\theta)^{m_\theta})\right) \mathcal{L}_p(g(n)).$$

The sequence $(\mathcal{L}_p(g(n)))_{n\geq 0}$ is p-regular, as in the case (a). Such claim is also true for each of the sequences $(\mathcal{L}_p((n-\theta)^{m_\theta}))_{n\geq 0}$ due to Lemma 3.14. In consequence, $(\mathcal{L}_{p^l}(f(n)))_{n\geq 0}$ is p-regular.

Now let k be multiplicatively independent with p and take d large enough so that case (b) of Theorem 3.5 is satisfied. Then $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ is not k-automatic. Using the fact that a k-regular sequence of p-adic integers reduced modulo a power of p is necessarily k-automatic (this follows from Theorem 1.3 and Proposition 1.15), we obtain the result.

Case (c) follows in the same way from case (c) of Theorem 3.5. □

3.4 Bases with several prime factors

In this section we turn to the case when b is a positive integer base with $s \geq 2$ distinct prime factors. We write the prime factorization of b in the form

$$b = p_1^{l_1} \cdots p_s^{l_s},$$

where $p_1, \ldots, p_s$ are distinct primes and $l_1, \ldots, l_s$ are positive integers. For $i = 1, \ldots, s$ we put $b_i = p_i^{l_i}$ and let r_i be an integer such that

$$r_i \frac{b}{b_i} \equiv 1 \pmod{b_i^d}$$

(this is consistent with the notation in Section 2.2).

To begin, we state the main results of the present section, which provide complete answers to Questions $1'$ and $2'$ for such b. Their proofs are given at the end of the section.

Theorem 3.17. *Let $f = (f_1, \ldots, f_s) \in \mathcal{A}_b$. We have the following.*

(a) *If for some i the function f_i has no root in $\mathbb{Z}_{p_i}$, then for all $d \geq 1$ the sequence $(\ell_{b,d}(f(n)))_{n\geq 0}$ is periodic.*

(b) *Assume that there exists $\theta \in \mathbb{Q} \cap \mathbb{Z}_b$ such that for each $i = 1, \ldots, s$ the number θ is the only root of f_i in $\mathbb{Z}_{p_i}$ and has multiplicity $m_i \geq 1$. Let $w_1, \ldots, w_s$ be positive integers satisfying $m_1 w_1 = \ldots = m_s w_s$ and put $k = b_1^{w_1} \cdots b_s^{w_s}$. Then for all $d \geq 1$ the sequence $(\ell_{b,d}(f(n)))_{n\geq 0}$ is k-automatic and not l-automatic for any $l \geq 2$ multiplicatively independent with k.*

(c) *Otherwise, the sequence $(\ell_{b,d}(f(n)))_{n\geq 0}$ is not k-automatic for any $d \geq 1$ and $k \geq 2$.*

Theorem 3.18. *Let $f = (f_1, \ldots, f_s) \in \mathcal{P}_b$. We have the following.*

(a) *If for some i the polynomial f_i has no root in $\mathbb{Z}_{p_i}$, then the sequence $(\mathcal{L}_b(f(n)))_{n\geq 0}$ is k-regular for every $k \geq 2$.*

(b) *Assume that there exists* $\theta \in \mathbb{Q} \cap \mathbb{Z}_b$ *and an integer* $m \geq 1$ *such that for each* $i = 1, \ldots, s$ *the number* θ *is the only root of* f_i *in* $\mathbb{Z}_{p_i}$ *and has multiplicity* m. *Then the sequence* $(\mathcal{L}_b(f(n)))_{n\geq 0}$ *is* b*-regular and not* l*-regular for any* $l \geq 2$ *multiplicatively independent with* b.

(c) *Otherwise, the sequence* $(\mathcal{L}_b(f(n)))_{n\geq 0}$ *is not* k*-regular for any* $k \geq 2$.

A comparison of the above results with Theorems 3.5 and 3.6, shows that the case when b has several prime factors is not a simple generalization of the case $b = p^l$. For instance, in contrast to the case where b is a prime power, the number of digits d does not affect k-automaticity of the sequence $(\ell_{b,d}(f(n)))_{n\geq 0}$. Moreover, for f fixed, changing the exponents l_i in the prime factorization of b has no effect on which of the cases (a)–(c) occurs in either of Theorems 3.17 and 3.18. A more detailed discussion on this topic is carried out in Section 3.5.

The overall structure of our reasoning is largely the same as in the previous section, however the details are rather different, as expected. For convenience we split the argument into shorter, more manageable parts.

Regularity of $(\mathcal{L}_b(f(n)))_{n\geq 0}$ for $f \in \mathcal{P}_b$, as well as automaticity (or periodicity) of $(\ell_{b,d}(f(n)))_{n\geq 0}$ for $f \in \mathcal{A}_b$ satisfying appropriate conditions will follow from a direct calculation.

On the other hand, proving nonautomaticity and nonregularity (or nonperiodicity) requires more effort. Rather than directly studying last nonzero digits of $f(n)$, we will use their characterization given in Corollary 2.7. For $d = 1$ this result says that $\nu_{b_i}(\ell_b(f(n))) = 0$ if and only if $\nu_{b_i}(f_i(n)) = \nu_b(f(n))$. In other words, the characteristic sequence of the set

$$\{n \geq 0 : \nu_{b_i}(f_i(n)) = \nu_b(f(n))\}$$

(or its complement in $\mathbb{N}$) is for each $i = 1, \ldots, s$ the image under a coding of $(\ell_b(f(n)))_{n\geq 0}$. Hence, if this characteristic sequence is not k-automatic (or not eventually periodic), then neither is $(\ell_b(f(n)))_{n\geq 0}$. Since $\nu_b(x) = \min_{1\leq i\leq s} \nu_{b_i}(x_i)$ for $x = (x_1, \ldots, x_i) \in \mathbb{Q}_b$, in order to make use of the above observation we need to investigate inequalities between the values $\nu_{b_i}(f_i(n))$. In this regard, the following technical lemma can be viewed as a counterpart of Lemma 3.11 from the previous section.

Lemma 3.19. *Let* $\rho_i, \sigma_i \in \mathbb{Z}_{p_i}$ *for* $i = 1, \ldots, s$. *Let* $m_1, \ldots, m_s$ *be nonnegative integers and* A, B *arbitrary integers. If* $\rho_j \neq \sigma_j$ *for some* $j \in \{1, \ldots, s\}$, *then there exist infinitely many nonnegative integers* n *such that*

$$\nu_{b_j}((n - \rho_j)^{m_j}) + A \leq \min_{i\neq j} \nu_{b_i}((n - \rho_i)^{m_i})$$

and

$$\nu_{b_j}((n - \sigma_j)^{m_j}) + B > \min_{i\neq j} \nu_{b_i}((n - \sigma_i)^{m_i}).$$

Proof. Without loss of generality let $j = 1$. It is sufficient to find infinitely many n satistying the following system of inequalities:

$$\begin{cases} \nu_{b_2}((n-\sigma_2)^{m_2}) < \nu_{b_1}((n-\sigma_1)^{m_1}) + B, \\ \nu_{b_1}((n-\rho_1)^{m_1}) + A \le \nu_{b_2}((n-\rho_2)^{m_2}), \\ \nu_{b_1}((n-\rho_1)^{m_1}) + A \le \nu_{b_3}((n-\rho_3)^{m_3}), \\ \vdots \\ \nu_{b_1}((n-\rho_1)^{m_1}) + A \le \nu_{b_s}((n-\rho_s)^{m_s}). \end{cases} \tag{3.13}$$

For each $i = 1, \dots, s$ the set of possible values of $\nu_{b_i}(x^{m_i})$ with $x \in \mathbb{Z}_{p_i}, x \neq 0$ is $V_i = \{\lfloor m_i k / l_i \rfloor : k \in \mathbb{N}\}$. An application of the Chinese Remainder Theorem shows that for any choice $v_1 \in V_1, \dots, v_s \in V_s$ there exist infinitely many integers $n \ge 0$ such that

$$\begin{cases} \nu_{b_1}((n-\sigma_1)^{m_1}) = v_1, \\ \nu_{b_2}((n-\rho_2)^{m_2}) = v_2, \\ \vdots \\ \nu_{b_s}((n-\rho_s)^{m_s}) = v_s. \end{cases}$$

If $\rho_2 \neq \sigma_2$, then the left side of each of the inequalities (3.13) remains bounded for sufficiently large $v_1, \dots, v_s$, while the right side grows to infinity as $v_1, \dots, v_s$ increase. Hence, for $v_1, \dots, v_s$ large enough we obtain a suitable value of n.

Otherwise, if $\rho_2 = \sigma_2$, then the first two inequalities in (3.13) can be written as

$$\nu_{b_1}((n-\rho_1)^{m_1}) + A \le \nu_{b_2}((n-\rho_2)^{m_2}) < \nu_{b_1}((n-\sigma_1)^{m_1}) + B.$$

As before, we take large $v_1, \dots, v_s$, where additionally $v_2 < v_1 + B$. □

In the following lemma we study periodicity of the sequence $(\ell_{b,d}(f(n)))_{n\ge 0}$ for $f = (f_1, \dots, f_s) \in \mathcal{A}_b$. Here and in a few later results we assume that each function f_i has at most one distinct root in $\mathbb{Z}_{p_i}$ (and possibly some additional properties). Due to an analogue of Proposition 3.10, we will eventually be able to combine these special cases into to a result for general f in the proofs of Theorems 3.17 and 3.18.

Lemma 3.20. *Let $f \in \mathcal{A}_b$ be such that each of the functions f_i has at most one root in $\mathbb{Z}_{p_i}$ (possibly with multiplicity). The following conditions are equivalent:*

(i) *for all $d \ge 1$ the sequence $(\ell_{b,d}(f(n)))_{n\ge 0}$ is periodic;*

(ii) *for all $d \ge 1$ the sequence $(\ell_{b,d}(f(n)))_{n\ge 0}$ is eventually periodic;*

(iii) *for some $i \in \{1, \dots, s\}$ the function f_i has no root in $\mathbb{Z}_{p_i}$;*

(iv) *the sequence $(\nu_b(f(n)))_{n\ge 0}$ is periodic.*

Furthermore, a power of b can be chosen as a period in (i), (ii), *and* (iv).

Proof. Obviously, (i) implies (ii).

Next, we show by contraposition that (ii) implies (iii). By the discussion before Lemma 3.19 it is enough to prove that the characteristic sequence of the set $\{n \geq 0 : \nu_{b_1}(f_1(n)) = \nu_b(f(n))\}$ is not eventually periodic.

Assume that for each $i = 1, \ldots, s$ the function f_i has precisely one root $\theta_i \in \mathbb{Z}_{p_i}$ and let $m_i \geq 1$ denote its multiplicity. Write

$$f_i(x) = (x - \theta_i)^{m_i} g_i(x),$$

where $x \in \mathbb{Z}_{p_i}$, the function g_i is strictly analytic on $\mathbb{Z}_{p_i}$ and $g_i(\theta_i) \neq 0$. Let $G \geq 1$ be an integer such that $|\nu_{b_i}(g_i(x))| < G$ for all $i = 1, \ldots, s$ and $x \in \mathbb{Z}_{p_i}$. Then for each $i = 1, \ldots, s$ and $n \geq 0$ (with the exception of $n = \theta_i$ in the case $\theta_i \in \mathbb{N}$) we obtain

$$|\nu_{b_i}(f_i(n)) - \nu_{b_i}((n - \theta_i)^{m_i})| \leq G.$$

Here we used the fact that $\nu_{b_i}(x) + \nu_{b_i}(y) \leq \nu_{b_i}(xy) \leq \nu_{b_i}(x) + \nu_{b_i}(y) + 1$ for any $x, y \in \mathbb{Q}_{p_i}$.

Therefore, the inequality

$$\nu_{b_1}((n - \theta_1)^{m_1}) + 2G \leq \min_{2 \leq i \leq s} \nu_{b_i}((n - \theta_i)^{m_i})$$

implies $\nu_{b_1}(f_1(n)) = \nu_b(f(n))$, while

$$\nu_{b_1}((n - \theta_1)^{m_1}) - 2G > \min_{2 \leq i \leq s} \nu_{b_i}((n - \theta_i)^{m_i})$$

implies $\nu_{b_1}(f_1(n)) > \nu_b(f(n))$.

Suppose that $(\ell_{b,d}(f(n)))_{n \geq 0}$ is eventually periodic with period $T > 0$. Lemma 3.19 applied to $j = 1, \rho_i = \theta_i, \sigma_i = \theta_i - T, A = 2G, B = -2G$ shows that there exists arbitrarily large n such that $\nu_{b_1}(f_1(n)) = \nu_b(f(n))$ but $\nu_{b_1}(f_1(n+T)) > \nu_b(f(n+T))$. Thus, we obtain a contradiction and (iii) follows.

We now show that (iii) implies (i). Take any integer $d \geq 1$. By (2.5) for every $i = 1, \ldots, s$ we have

$$\ell_{b,d}(f(n)) \equiv b_i^{\nu_{b_i}(f_i(n)) - \nu_b(f(n))} r_i^{\nu_b(f(n))} \ell_{b_i,d}(f_i(n)) \pmod{b_i^d}. \tag{3.14}$$

It is enough to prove that for each $i = 1, \ldots, s$ the expression on the right in (3.14) reduced modulo b_i^d is periodic with respect to n.

By the assumption there exists an integer V such that $\nu_b(f(n)) \leq V$ for all $n \geq 0$. Choose $i \in \{1, \ldots, s\}$. Proposition 1.24 applied to $M = l_i(V + d)$ provides an integer $T_i \geq 0$ such that for all $n \geq 0$ we have

$$\nu_{b_i}(f_i(n + p_i^{T_i}) - f_i(n)) \geq V + d.$$

Letting $\pi = p_1^{T_1} \cdots p_s^{T_s}$, we can write

$$f_i(n + \pi) = f_i(n) + b_i^{V+d} h_i(n),$$

where $h_i(n) \in \mathbb{Z}_{p_i}$. From this we see that the expression $\min\{\nu_{b_i}(f_i(n)), V+d\}$ is periodic with period π. Thus, so is $\nu_b(f(n)) = \min_{1\le i\le s} \nu_{b_i}(f_i(n))$ and also the factor $r_i^{\nu_b(f(n))}$ reduced modulo b_i^d.

Now let $\gamma_i(n)$ be the product of the remaining factors in (3.14):

$$\gamma_i(n) = b_i^{\nu_{b_i}(f_i(n))-\nu_b(f(n))}\ell_{b_i,d}(f_i(n)).$$

By the earlier considerations, the expression $\min\{\nu_{b_i}(f_i(n)) - \nu_b(f(n)), d\}$ is periodic with period π.

Hence, if $\nu_{b_i}(f_i(n)) - \nu_b(f(n)) \ge d$, then

$$\gamma_i(n) \equiv 0 \equiv \gamma_i(n+\pi) \bmod b_i^d.$$

Otherwise, if $\nu_{b_i}(f_i(n))-\nu_b(f(n)) < d$, then we have $\nu_{b_i}(f_i(n)) < V+d$, so $\nu_{b_i}(f_i(n)) = \nu_{b_i}(f_i(n+\pi))$. It follows that

$$\ell_{b_i,d}(f_i(n+\pi)) \equiv \ell_{b_i,d}(f_i(n)) + b_i^{V+d-\nu_{b_i}(f_i(n))}h_i(n) \pmod{b_i^d}.$$

Multiplying both sides of this congruence by

$$b_i^{\nu_{b_i}(f_i(n+\pi))-\nu_b(f(n+\pi))} = b_i^{\nu_{b_i}(f_i(n))-\nu_b(f(n))},$$

we again obtain

$$\gamma_i(n) \equiv \gamma_i(n+\pi) \bmod b_i^d.$$

Therefore, $(\gamma_i(n))_{n\ge 0}$ is periodic and our claim follows.

Finally, we prove the equivalence of (iii) and (iv). As we have already seen in the proof of the previous implication, (iv) follows from (iii).

Now, for the sake of contradiction assume that $(\nu_b(f(n)))_{n\ge 0}$ is periodic but each of the functions f_i has a root in $\theta_i \in \mathbb{Z}_{p_i}$. Let $(n_l)_{n\ge 0}$ be a sequence of nonnegative integers such that $n_l \equiv \theta_i \pmod{b_i^l}$ for all $i = 1, \dots, s$. Then the values $\nu_b(f(n_l))$ are unbounded, thus we obtain a contradiction.

In the above considerations, the number π is a common period in (i), (ii), and (iv). By taking its suitable multiple, we see that a power of b is also a period. □

The following result is a partial analogue of Lemma 3.20 for $f \in \mathcal{P}_b$.

Lemma 3.21. *Let $f = (f_1, \dots, f_s) \in \mathcal{P}_b$. If f_i has no root in $\mathbb{Z}_{p_i}$ for some $i \in \{1, \dots, s\}$, then the sequence $(\mathcal{L}_b(f(n)))_{n\ge 0}$ is k-regular for every $k \ge 2$.*

Proof. By Lemma 3.20, the sequence $(\nu_b(f(n)))_{n\ge 0}$ is periodic. Let $\pi > 0$ be a period of this sequence. Then for each $a = 0, 1, \dots, \pi - 1$ the function

$$\mathcal{L}_b(f(\pi x + a)) = b^{-\nu_b(f(a))}f(\pi x + a)$$

of $x \in \mathbb{Q} \cap \mathbb{Z}_b$ is an element of $\mathcal{P}_b$. Therefore, each component of $(\mathcal{L}_b(f(\pi n + a)))_{n\ge 0}$ is a k-regular sequence for every k as a sequence of values of a polynomial evaluated at consecutive integers. This also implies k-regularity of $(\mathcal{L}_b(f(\pi n + a)))_{n\ge 0}$ and the result follows. □

We now proceed to prove k-automaticity of $(\ell_{b,d}(f(n)))_{n\geq 0}$ with a special value of k for certain $f \in \mathcal{A}_b$.

Lemma 3.22. *Let $\theta \in \mathbb{Q} \cap \mathbb{Z}_b$ and $d \geq 1$. Let $f = (f_1, \ldots, f_s) \in \mathcal{A}_b$ and assume that for each $i = 1, \ldots, s$ the function f_i is of the form*

$$f_i(x) = (x - \theta)^{m_i} g_i(x),$$

where $m_i \geq 1$ is an integer and g_i is strictly analytic on $\mathbb{Z}_{p_i}$ and such that $\nu_{p_i}(g_i(x))$, $\ell_{b_i,d}(g_i(x))$ are constant with respect to $x \in \mathbb{Z}_{p_i}$. Let $k = b_1^{w_1} \cdots b_s^{w_s}$, where $w_1, \ldots, w_s$ are positive integers satisfying

$$m_1 w_1 = \ldots = m_s w_s.$$

Then the sequence $(\ell_{b,d}(f(n)))_{n\geq 0}$ is k-automatic.

Proof. Writing $\theta = t/u$ in lowest terms with $u > 0$, we obtain

$$f_i(n) = \frac{1}{u^{m_i}}(un - t)^{m_i} g_i(n) = h_i(un - t),$$

where

$$h_i(x) = \frac{1}{u^{m_i}} y^{m_i} g_i\left(\frac{x}{u} + \theta\right)$$

for $x \in \mathbb{Z}_{p_i}$. By Theorem 1.9 (and Remark 1.10) the problem can be reduced to studying k-automaticity of the sequence $(\ell_{b,d}(h(n)))_{n\geq 0}$, where $h = (h_1, \ldots, h_s) \in \mathcal{A}_b$. Replacing h by f, from now on we assume that

$$f_i(x) = x^{m_i} g_i(x).$$

Put

$$c_i = p_i^{\nu_{p_i}(g_i(x))} \ell_{b_i,d}(g_i(x)),$$

which does not depend on y for $i = 1, \ldots, s$, and let D be the common value of $m_i w_i$. Recall that a sequence is k-automatic if and only if it is k^t-automatic. Raising k to a suitable power, we may thus assume that $\ell_{b_i,d}(k) = 1$ and D is such that $r_i^D \equiv 1 \pmod{b_i^d}$ for all $i = 1, \ldots, s$.

We consider the subsequences $(\ell_{b,d}(f_i(kn + a)))_{n\geq 0}$ with $a = 0, 1, \ldots, k - 1$. For $a = 0$ we have

$$\nu_{b_i}(f_i(kn)) = D + \nu_{b_i}(f_i(n)),$$

and thus also

$$\nu_b(f(kn)) = D + \nu_b(f(n)).$$

As a result, we obtain

$$r_i^{\nu_{b_i}(f_i(kn))} \equiv r_i^{\nu_{b_i}(f_i(n))} \pmod{b_i^d}$$

and

$$\ell_{b_i,d}(f_i(kn)) \equiv \ell_{b_i,d}(k^{m_i}) \ell_{b_i,d}\left(c_i n^{m_i}\right) \equiv \ell_{b_i,d}(f_i(n)) \pmod{b_i^d}.$$

By the above identities combined with Lemma 2.6 applied to $f(n)$ and $f(kn)$, we get

$$\ell_{b,d}(f(kn)) = \ell_{b,d}(f(n)).$$

Now choose $a \in \{1, \ldots, k-1\}$. Then for some i we have $\nu_{p_i}(a) < \nu_{p_i}(k)$, so the function $f_i(ky+a)$ of y has no root in $\mathbb{Z}_{p_i}$. Lemma 3.20 implies that $(\ell_{b,d}(f(kn+a)))_{n\geq 0}$ is periodic. Consequently, $(\ell_{b,d}(f(n)))_{n\geq 0}$ is k-automatic by an argument similar as in Proposition 2.9. □

A similar assertion can also be made about $(\mathcal{L}_b(f(n)))_{n\geq 0}$ when the components of $f \in \mathcal{P}_b$ admit a simple form. Note that unlike in Lemma 3.22 we only consider b-regularity. We will see later (in Lemma 3.26 below) why this is the case.

Lemma 3.23. *Let $\theta \in \mathbb{Q} \cap \mathbb{Z}_b$. Let $f = (f_1, \ldots, f_s) \in \mathcal{P}_b$ and assume that for each $i = 1, \ldots, s$ the function f_i is of the form*

$$f_i(x) = c_i(x-\theta)^m, \tag{3.15}$$

where $m \geq 1$ is an integer and $c_i \in \mathbb{Q}_{p_i}$. Then the sequence $(\mathcal{L}_b(f(n)))_{n\geq 0}$ is b-regular.

Proof. For the same reason as in the proof Lemma 3.22, without loss of generality we can set $\theta = 0$.

We examine the subsequences $(\mathcal{L}_b(f(bn+a)))_{n\geq 0}$ with $a = 0, 1, \ldots, b-1$. In the case $a = 0$ we obtain

$$\mathcal{L}_b(f(bn)) = \mathcal{L}_b(f(n)).$$

On the other hand, if $a \neq 0$, then for some i the polynomial $f_i(bx+a)$ has no root in $\mathbb{Z}_{p_i}$, so b-regularity of the sequence $(\mathcal{L}_b(f(bn+a)))_{n\geq 0}$ follows by Lemma 3.21.

We deduce that $(\mathcal{L}_b(f(n)))_{n\geq 0}$ is b-regular precisely in the same way as in Proposition 2.9. □

Lemma 3.24. *Let $f = (f_1, \ldots, f_s) \in \mathcal{A}_b$ be such that that for each $i = 1, \ldots, s$ the function f_i is of the form*

$$f_i(x) = (x-\theta_i)^{m_i} g_i(x),$$

where $\theta_i \in \mathbb{Z}_{p_i}$, $m_i \geq 1$ is an integer, and g_i is strictly analytic with no root in $\mathbb{Z}_{p_i}$. If the roots $\theta_1, \ldots, \theta_s$ are not all equal, then the sequence $(\ell_b(f(n)))_{n\geq 0}$ is not k-automatic for any $k \geq 2$.

Proof. Without loss of generality we can assume that none of the roots θ_i lies in $\mathbb{N}$ by replacing x with $x+M$ if necessary, where M is a sufficiently large integer.

Let $k \geq 2$ be an integer. By raising k to a suitable power, we may assume that it is of the form

$$k = b_1^{w_1} \cdots b_s^{w_s} c,$$

where $w_1, \ldots, w_s, c$ are nonnegative integers and $\gcd(b, c) = 1$.

We simultaneously approximate the p_i-adic integers θ_i by the sequence of integers $(n_l)_{l\geq 0}$ given by

$$\begin{cases} n_l \equiv \theta_1 \pmod{b_1^{w_1 l}}, \\ \quad \vdots \\ n_l \equiv \theta_s \pmod{b_s^{w_s l}}, \end{cases}$$

where $0 \leq n_l < (b_1^{w_1} \cdots b_s^{w_s})^l \leq k^l$. We will show that there exist infinitely many distinct subsequences of the form $(\ell_b(f(k^l n + n_l)))_{n\geq 0}$ with $l \geq 0$. In order to do this, we study inequalities between the valuations $\nu_{b_i}(f_i(k^l n + n_l))$.

Let $G \geq 1$ be an integer such that $|\nu_{b_i}(g_i(x))| < G$ for all $i = 1, \ldots, s$ and $x \in \mathbb{Z}_{p_i}$. Also write $\theta_i\{l\} = (\theta_i - n_l)/k^l$, analogously to the notation used in the previous section. Then for every $n \geq 0$ we have

$$\left|\nu_{b_i}(f_i(k^l n + n_l)) - \nu_{b_i}\left((n - \theta_i\{l\})^{m_i}\right) - m_i w_i l\right| \leq G. \tag{3.16}$$

Choose any pair of integers l, m such that $m > l \geq 0$. We have $\theta_i\{l\} = \theta_i\{m\}$ if and only if

$$\theta_i = \frac{k^l n_m - k^m n_l}{k^l - k^m}$$

and this expression does not depend on i. By the assumption that $\theta_1, \ldots, \theta_s$ are not all equal, for some index $j \in \{1, \ldots, s\}$ (depending on l, m) we have $\theta_j\{l\} \neq \theta_j\{m\}$.

Just like in the proof of Lemma 3.20, we use the inequality (3.16) in conjunction with Lemma 3.19 (with $\rho_i = \theta_i\{l\}, \sigma_i = \theta_i\{m\}$ and appropriate A, B) to conclude that there exists an integer $n \geq 0$ such that

$$\nu_{b_j}(f_j(k^l n + n_l)) = \nu_b(f(k^l n + n_l))$$

and

$$\nu_{b_j}(f_j(k^m n + n_m)) > \nu_b(f(k^m n + n_m)).$$

This implies that

$$\ell_b(f(k^l n + n_l)) \neq \ell_b(f(k^m n + n_m))$$

and our claim follows. □

We now state a technical result, which essentially says that for any base k the k-adic expansion of a rational number is eventually periodic. A proof for a prime base is given in the notes of Conrad [23] and can be adapted to the general case. Nevertheless, for completeness we provide an alternative (shorter) proof. First, we set some notation, also used in a few subsequent results. Let $\theta \in \mathbb{Q} \cap \mathbb{Z}_k$. For each integer $l \geq 0$ we write $\theta[l, k] = \theta \bmod k^l$ and $\theta\{l, k\} = (\theta - \theta[l, k])/k^l$. For simplicity, k will be suppressed from the notation whenever it does not cause ambiguity.

Lemma 3.25. *Let $k \geq 2$ be an integer and $\theta \in \mathbb{Q} \cap \mathbb{Z}_k$. Then*

(i) *for all integers $l \geq 0$, $m \geq 0$ we have $(\theta\{l + m, k\}) = (\theta\{l, k\})\{m, k\}$;*

(ii) *the sequence $(\theta\{l, k\})_{l\geq 0}$ is eventually periodic.*

Proof. Considering k fixed, we omit it from the notation. Part (i) is shown by induction on m with $l \geq 0$ fixed. If $m = 0$ then the assertion holds. We also provide a proof for $m = 1$, as it will be needed either way. Write $\theta[l+1] = k^l a + \theta[l]$ for some $a \in \{0, 1, , \ldots, k-1\}$. We have

$$(\theta\{l\})[1] = \left(\frac{\theta - \theta[l+1] + k^l a}{k^l}\right)[1] = a = \frac{\theta[l+1] - \theta[l]}{k^l}.$$

Therefore,

$$(\theta\{l\})\{1\} = \frac{\theta\{l\} - (\theta\{l\})[1]}{k} = \frac{\frac{\theta - \theta[l]}{k^l} - \frac{\theta[l+1] - \theta[l]}{k^l}}{k} = \theta\{l+1\}.$$

Now if (i) holds for some $m \geq 0$, then

$$(\theta\{l\})\{m+1\} = ((\theta\{l\})\{m\})\{1\} = (\theta\{l+m\})\{1\} = \theta\{l+m+1\},$$

as desired.

In order to prove (ii) we write $\theta = t/u$ in lowest terms. Since $u\theta[l] \equiv t \pmod{k^l}$, the number $u\theta\{l\} = (t - u\theta[l])/k^l$ is an integer for each l. Moreover, we have the bound

$$\left|\frac{t - u\theta[l]}{k^l}\right| \leq |t| + |u|\frac{\theta[l]}{k^l} < |t| + |u|.$$

Therefore, there exist indices $m > l$ such that

$$\frac{\theta - \theta[m]}{k^m} = \frac{\theta - \theta[l]}{k^l},$$

which yields $\theta\{m\} = \theta\{l\}$. The result follows from (i). □

We now proceed to give the final auxiliary lemma, which complements Lemma 3.23. More precisely, it shows that if we allow the exponents in (3.15) to vary, then the resulting sequence of last nonzero digits is not regular.

Lemma 3.26. *Let $f = (f_1, \ldots, f_s) \in \mathcal{P}_b$ and assume that for each $i = 1, \ldots, s$ the function f_i is of the form*

$$f_i(x) = (x - \theta)^{m_i} g_i(x),$$

where $m_i \geq 1$ is an integer, $\theta \in \mathbb{Q} \cap \mathbb{Z}_b$, and $g_i \in \mathcal{P}_{p_i}$ is such that the values $\nu_{p_i}(g_i(x))$, $\ell_{b_i}(g_i(x))$ are constant with respect to $x \in \mathbb{Z}_{p_i}$. If not all m_i are equal, then the sequence $(\mathcal{L}_b(f(n)))_{n \geq 0}$ is not regular.

Proof. We first reduce the problem to showing that if the sequence $(\mathcal{L}_b(f(n)))_{n \geq 0}$ is k-regular, then k must be of the form

$$k = b_1^{w_1} \cdots b_s^{w_s},$$

where $w_1, \ldots, w_s$ are positive integers satisfying

$$m_1 w_1 = \ldots = m_s w_s.$$

Indeed, by Lemma 3.22 the sequence $(\ell_b(f(n)))_{n\geq 0}$ is k-automatic. At the same time, if the sequence $(\mathcal{L}_b(f(n)))_{n\geq 0}$ were l-regular for some l multiplicatively independent with k, then $(\ell_b(f(n)))_{n\geq 0}$ would simultaneously be l-automatic. But then Cobham's Theorem would imply that this sequence is eventually periodic, which is ruled out by Lemma 3.20.

Therefore, it remains to prove that $(\mathcal{L}_b(f(n)))_{n\geq 0}$ is not k-regular. We begin with some preparatory steps. By Lemma 3.25 the sequence $(\theta\{l,k\})_{l\geq 0}$ is eventually periodic with period T. Replacing k with k^T, we can assume that $T=1$. Again, we suppress k in the notation and from now on write $\theta[l]=\theta[l,k]$ and $\theta\{l\}=\theta\{l,k\}$ for all $l\geq 0$. Let $L\geq 0$ be such that $\theta\{l\}=\theta\{L\}$ for all $l\geq L$. It is sufficient to prove that the subsequence $(\mathcal{L}_b(f(k^Ln+\theta[L])))_{n\geq 0}$ is not k-regular. Since for $i=1,\ldots,s$ we have

$$f_i(k^Ln+\theta[L])=(n-\theta\{L\})^{m_i}k^{Lm_i}g_i(k^Ln+\theta[L])),$$

without loss of generality we can assume that $L=0$, which means that $\theta\{l\}=\theta$ for all $l\geq 0$.

Using this fact, we obtain

$$f_i(k^ln+\theta[l])=k^{lm_i}(n-\theta)^{m_i}g_i(k^ln+\theta[l]))$$

for all $i=1,\ldots,s$, integers $l\geq 0$ and $n\geq 0$. Let D denote the common value of m_iw_i. By the assumption that the p_i-adic valuation of $g(y)$ is constant, we get

$$\nu_{b_i}(f_i(k^ln+\theta[l]))=lD+\nu_{b_i}(f_i(n)),$$

and thus also

$$\nu_b(f(k^ln+\theta[l]))=lD+\nu_b(f_i(n)). \tag{3.17}$$

Renumber the primes so that $m_1<m_2$. By Proposition 1.14, it is enough to show that the sequence of values lying in the first coordinate of $\mathcal{L}_b(f(n))$, that is $(b^{-\nu_b(f(n))}f_1(n))_{n\geq 0}$, is not k-regular. For all $l\geq 0$ and $n\geq 0$ define

$$\beta_l(n)=b^{-\nu_b(f(k^ln+\theta[l]))}f_1(k^ln+\theta[l]).$$

We will prove that the $\mathbb{Z}$-submodule generated by the family $\{(\beta_l(n))_{n\geq 0}:l\geq 0\}$ is not finitely generated.

Suppose that this is not the case and for some $t\geq 0$ the sequences $(\beta_l(n))_{n\geq 0}$ with $l=0,1,\ldots,t$ generate said $\mathbb{Z}$-module. In particular there exist integers $\alpha_0,\alpha_1,\ldots,\alpha_t$ such that for all $n\geq 0$ we have

$$\sum_{l=0}^{t}\alpha_l\beta_l(n)=\beta_{t+1}(n).$$

By the definition of $\beta_l(n)$ and (3.17) this equality can be written (after some simplification) as

$$\sum_{l=0}^{t}\alpha_lC^lg_1(k^ln+\theta[l])=C^{t+1}g_1(k^{t+1}n+\theta[t+1]), \tag{3.18}$$

where $C = k^{m_1}/b^D$. In fact, since $\mathbb{N}$ is dense in $\mathbb{Z}_{p_1}$, the equality (3.18) also holds with n replaced by $x \in \mathbb{Z}_{p_1}$. Plugging in $x = \theta$ and using the fact that $\theta = \theta\{l\}$ for all $l \geq 0$, we obtain

$$\sum_{l=0}^{t} \alpha_l C^l g_1(\theta) = C^{t+1} g_1(\theta).$$

But $g_1(\theta) \neq 0$, so we get

$$\sum_{l=0}^{t} \alpha_l C^l = C^{t+1}. \tag{3.19}$$

However, by the choice of m_1 we have $\nu_{b_2}(C) = m_1 w_2 - D < m_2 w_2 - D = 0$. This means that (3.19) cannot hold and we arrive at a contradiction. □

We now have all the necessary tools to prove Theorems 3.17 and 3.18.

Proof of Theorem 3.17. As in the proof of Theorem 3.5, we focus our attention on subsequences along arithmetic progressions. For each $i = 1, \ldots, s$ let $T_i \geq 0$ be the integer obtained from Proposition 3.10 applied to f_i. Take an integer T such that $T \geq T_i/l_i$ for all $i = 1, \ldots, s$. We will consider the subsequences $(\ell_{b,d}(f(b^T n + a)))_{n \geq 0}$ with $a = 0, 1, \ldots, b^T - 1$.

For each $i = 1, \ldots, s$ and $a = 0, 1, \ldots, b^T - 1$ let f_{ia} denote the function defined by

$$f_{ia}(x) = f_i(b^T x + a),$$

where $x \in \mathbb{Z}_{p_i}$. Then, analogously as in Proposition 3.10, each function f_{ia} satisfies one of the following conditions:

(i) if there is no root $\theta_i \in \mathbb{Z}_{p_i}$ of f_i such that $a \equiv \theta_i \pmod{b_i^T}$, then f_{ia} has no root $x \in \mathbb{Z}_{p_i}$;

(ii) if $a \equiv \theta_i \pmod{b_i^{T_i}}$ for some root $\theta_i \in \mathbb{Z}_{p_i}$ of f_i of multiplicity m_i, then

$$f_{ia}(x) = b^{Tm_i} (x - \theta_{ia})^{m_i} g_{ia}(x), \tag{3.20}$$

where $\theta_{ia} = (\theta_i - a)/b^T$ and g_{ia} is a function strictly analytic on $\mathbb{Z}_{p_i}$ such that $\nu_{p_i}(g_{ia}(x))$, $\ell_{b_i,d}(g_{ia}(x))$ are constant with respect to $x \in \mathbb{Z}_{p_i}$.

We now proceed to consider each case separately. Case (a) follows by the implication from (iii) to (i) in Lemma 3.20, applied to each s-tuple $(f_{1a}, \ldots, f_{sa})$.

In the case (b), since b is fixed, we will use the simplified notation $\theta[T]$, $\theta\{T\}$ instead of $\theta[T, b]$, $\theta\{T, b\}$. If $a \neq \theta[T]$, then $a \not\equiv \theta \pmod{b_i^T}$ for some i, so the sequence $(\ell_{b,d}(f(b^T n + a)))_{n \geq 0}$ is periodic (thus also k-automatic).

Otherwise, if $a = \theta[T]$, then for each $i = 1, \ldots, s$ of the function f_{ia} admits the form (3.20), where $\theta_{ia} = \theta\{T\}$. In this case $(\ell_{b,d}(f(b^T n + a)))_{n \geq 0}$ is k-automatic due to Lemma 3.22. It follows that $(\ell_{b,d}(f(n)))_{n \geq 0}$ is k-automatic as well.

At the same time, Lemma 3.20 applied to the function $f(b^T x + \theta[T])$ shows that the sequence $(\ell_{b,d}(f(b^T n + \theta[T])))_{n \geq 0}$ is not eventually periodic. Hence, by Cobham's

Theorem $(\ell_{b,d}(f(n)))_{n\geq 0}$ cannot be l-automatic for any l multiplicatively independent with k.

Finally, in part (c) we can pick an s-tuple of roots $\theta_1, \ldots, \theta_s$ of $f_1, \ldots, f_s$, respectively, such that not all θ_i are equal. Let $a \in \{0, 1, \ldots, b^T - 1\}$ be such that $a \equiv \theta_i \pmod{b_i^T}$ for all $i = 1, \ldots, s$. Since $\theta_{1a}, \ldots, \theta_{sa}$ are not all equal either, Lemma 3.24 applied to the s-tuple $(f_{1a}, \ldots, f_{sa})$ implies that $(\ell_b(f(b^T n+a))_{n\geq 0}$ is not k-automatic for any $k \geq 2$. Consequently, $(\ell_{b,d}(f(n)))_{n\geq 0}$ is not automatic. □

Proof of Theorem 3.18. Part (a) is Lemma 3.21.

In the remaining two cases let $T \geq 0$ be the integer obtained precisely as in the proof of Theorem 3.17. In part (b) for $i = 1, \ldots, s$ we write

$$f_i(x) = (x - \theta)^m g_i(x),$$

where $g_i \in \mathbb{Q}_{p_i}[X]$ has no root in $\mathbb{Z}_{p_i}$. Then by the choice of T we have, in particular,

$$\nu_{p_i}(g_i(x + b^T y)) = \nu_{p_i}(g_i(x))$$

for all $i = 1, \ldots, s$ and $x, y \in \mathbb{Z}_{p_i}$. We study the subsequences $(\mathcal{L}_b(f(b^T n + a)))_{n\geq 0}$ for $a = 0, 1, \ldots, b^T - 1$. We will again write $\theta[T] = \theta[T, b]$ and $\theta\{T\} = \theta\{T, b\}$.

If $a \neq \theta[T]$, then for some i the polynomial $f_i(b^T x + a)$ has no root in $\mathbb{Z}_{p_i}$, so $(\mathcal{L}_b(f(b^T n + a))_{n\geq 0}$ is b-regular as in part (a).

For $a = \theta[T]$ we have

$$f_i(b^T n + \theta[T]) = b^{Tm}(n - \theta\{T\})^m g_i(b^T n + \theta[T])$$

for each $i = 1, \ldots, s$. Letting $g = (g_1, \ldots, g_s)$, we obtain

$$\begin{aligned}\mathcal{L}_b(f(b^T n + \theta[T])) &= \mathcal{L}_b\left(b^{Tm} g(\theta[T])(n - \theta\{T\})^m \frac{g(b^T n + \theta[T])}{g(\theta[T])}\right) \\ &= \mathcal{L}_b(b^{Tm} g(\theta[T])(n - \theta\{T\})^m)\, \frac{g(b^T n + \theta[T])}{g(\theta[T])}, \qquad (3.21)\end{aligned}$$

where we used the property (ii) of Lemma 2.5.

The first factor in (3.21) is b-regular due to Lemma 3.23. At the same time, each component of the second factor is a polynomial in n, so $(g(b^T n + \theta[T])/g(\theta[T]))_{n\geq 0}$ is a b-regular sequence by Corollary 1.13 and Proposition 1.14. It follows that $(\mathcal{L}_b(f(b^T n + \theta[T]))_{n\geq 0}$, and therefore also $(\mathcal{L}_b(f(n)))_{n\geq 0}$ is b-regular.

If this sequence were l-regular for some l multiplicatively independent with b, then $(\ell_b(f(n)))_{n\geq 0}$ would be l-automatic, but this is impossible because of Theorem 3.17(b).

Part (c) can be split into two subcases. First, we have the case where we can pick an s-tuple of roots $\theta_1, \ldots, \theta_s$ of $f_1, \ldots, f_s$, respectively, such that not all θ_i are equal. Then our claim follows in a similar fashion as before from part (c) of Theorem 3.17.

The second case occurs when all f_i have a common root $\theta \in \mathbb{Q} \cap \mathbb{Z}_b$ and no other p_i-adic integer roots but the multiplicities of θ vary with i. For $i = 1, \ldots, s$ write

$$f_i(x) = (x - \theta)^{m_i} h_i(x),$$

where the polynomials $h_i \in \mathbb{Q}_{p_i}[X]$ are irreducible over $\mathbb{Z}_{p_i}$ and $m_i \geq 1$ are integers, not all equal.

It is sufficient to prove that the subsequence $(\mathcal{L}_b(f(b^T n+\theta[T])))_{n\geq 0}$ is not regular. We have

$$f_i(b^T x + \theta[T]) = b^{Tm_i}(x - \theta\{T\})^{m_i} h_i(b^T x + \theta[T]).$$

The choice of T ensures that the values $\nu_{p_i}(h_i(b^T x + \theta[T]))$, $\ell_{b_i}(h_i(b^T x + \theta[T]))$ are constant with respect to $x \in \mathbb{Z}_{p_i}$. Hence, the function $f(b^T x + \theta[T])$ satisfies the assumptions of Lemma 3.26, and the result follows. □

3.5 Further discussion and examples

In this section we discuss some implications of the results proved in this chapter and a number of related examples. It will be helpful to recall how squares in $\mathbb{Q}_p \setminus \{0\}$ look (see [66, pp. 17–18]). Write $\theta \in \mathbb{Q}_p \setminus \{0\}$ in the form $\theta = p^v\sigma$, where $\nu_p(\sigma) = 0$. In the case $p \neq 2$ we have that θ is a square if and only if v is even and $\sigma \bmod p$ is a square in the finite field $\mathbb{F}_p$. If $p = 2$, the latter condition is replaced by $\sigma \equiv 1 \pmod 8$. This fact can also be derived directly from Hensel's Lemma.

To begin, we investigate whether replacing b with its factors may affect regularity of $(\mathcal{L}_b(f(n)))_{n\geq 0}$ or automaticity of $(\ell_{b,d}(f(n)))_{n\geq 0}$. First, let p be a prime. By Theorem 3.6 k-regularity of $(\mathcal{L}_{p^l}(f(n)))_{n\geq 0}$ for $f \in \mathbb{Q}_p[X]$ is equivalent to k-regularity of $(\mathcal{L}_p(f(n)))_{n\geq 0}$. However, in the following two examples we demonstrate that automaticity of $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ neither implies, nor is implied by automaticity of $(\ell_{p,ld}(f(n)))_{n\geq 0}$.

Example 3.1. Let $p = 5, l = 2, d = 1$, and $f(x) = 5(x^2+1)^4$. By the above discussion -1 is a square in $\mathbb{Z}_5$, so there exists a root $\theta \in \mathbb{Z}_5$ of f. By Theorem 3.5 the sequence $(\ell_{5,2}(f(n)))_{n\geq 0}$ is not automatic. However, the same result shows that $(\ell_{5^2}(f(n)))_{n\geq 0}$ is periodic, and thus k-automatic for every $k \geq 2$. In fact, it is easy to check that $\ell_{5^2}(f(n)) = 5$ for all $n \geq 0$.

Example 3.2. Let $p = 2, l = 3, d = 1$, and $g(x) = (x^2 + 15)^2$. Again, g has a root in $\mathbb{Z}_2$. The sequence $(\ell_{2,3}(g(n)))_{n\geq 0}$ is constant with all terms equal to 1. At the same time, according to Theorem 3.5, $(\ell_{2^3}(g(n)))_{n\geq 0}$ is not automatic, as the exponent $m = 2$ does not divide $l = 3$.

Examples of such type can also be produced for other primes p and in the case where f is a power series with infinitely many nonzero coefficients.

We now consider a similar problem for a base $b \geq 2$ having $s \geq 2$ distinct prime factors. As before, write

$$b = p_1^{l_1} \cdots p_s^{l_s}$$

and let $b_i = p_i^{l_i}$ for $i = 1, \ldots, s$. Theorems 3.6 and 3.18 together show that for $f = (f_1, \ldots, f_s) \in \mathcal{P}_b$ regularity of $(\mathcal{L}_b(f(n)))_{n\geq 0}$ implies regularity of $(\mathcal{L}_{b_i}(f_i(n)))_{n\geq 0}$ for at least one $i \in \{1, \ldots, s\}$. The following example demonstrates that this assertion cannot be strengthened.

Example 3.3. Let $b = 10$ and consider the polynomial $f(x) = x^2 + 1$. Then f has a root in $\mathbb{Z}_5$, but no root in $\mathbb{Z}_2$. Hence, $(\mathcal{L}_{10}(f(n)))_{n\geq 0}$ is k-regular for every $k \geq 2$. However, out of the two sequences $(\mathcal{L}_2(f(n)))_{n\geq 0}$ and $(\mathcal{L}_5(f(n)))_{n\geq 0}$, only the former is regular.

It turns out that the converse also fails, even under the assumption that the sequences $(\mathcal{L}_{b_i}(f_i(n)))_{n\geq 0}$ are regular for all $i = 1, \ldots, s$.

Example 3.4. Let $g(x) = x(x+1)$. Then $(\mathcal{L}_{p^l}(g(n)))_{n\geq 0}$ is p-regular for every prime p and exponent $l \geq 1$, but by Theorem 3.18 $(\mathcal{L}_b(g(n)))_{n\geq 0}$ is not regular whenever $b \geq 2$ has two or more prime factors.

A similar argument also works for the function $\ell_{b,d}$. In particular, the reasoning in Examples 3.3 and 3.4 remains valid after replacing each $\mathcal{L}_b$ with ℓ_b for all the considered bases.

It is also interesting to consider the relation between k-regularity of $(\mathcal{L}_b(f(n)))_{n\geq 0}$ and k-automaticity of $(\ell_{b,d}(f(n)))_{n\geq 0}$. Since a k-regular sequence of b-adic integers becomes k-automatic when reduced modulo b^d (by Proposition 1.15), the former property implies the latter. The converse implication is not true in general, as demonstrated by the following examples.

Example 3.5. Let $p = 5$ and $f(x) = (x^2+1)^4$. Then f has an irrational root $\theta \in \mathbb{Z}_5$, but $\ell_5(f(n)) = 1$ for all $n \geq 0$ due to the multiplicity, However, Theorem 3.6 shows that $(\mathcal{L}_5(f(n)))_{n\geq 0}$ is not regular.

When the base is a prime power, in order to ensure that $(\mathcal{L}_{p^l}(f(n)))_{n\geq 0}$ is a regular sequence, we thus need to assume that the sequence $(\ell_{p^l,d}(f(n)))_{n\geq 0}$ is automatic for all $d \geq 1$. By Theorem 3.5 this rules out the possibility that f has an irrational root, which means that case (a) or (b) of Theorem 3.6 occurs. The next example shows that in the case when b has several prime factors, such a strong assumption does not guarantee regularity of $(\mathcal{L}_b(f(n)))_{n\geq 0}$.

Example 3.6. Let $b = 6$ and let $f(x) = (x, x^2)$ for $x \in \mathbb{Q} \cap \mathbb{Z}_6$, where the first component is considered 2-adically and the second one – 3-adically. Theorem 3.17 asserts that the sequence $(\ell_{6,d}(f(n)))_{n\geq 0}$ is 12-automatic regardless of d. On the other hand, by Theorem 3.18 the sequence $(\mathcal{L}_6(f(n)))_{n\geq 0}$ is not regular due to different exponents in the components of $f(x)$.

In the following two examples we showcase the computation of the last nonzero digits of linear recurrence sequences of integers.

Example 3.7. Consider the Fibonacci sequence $(F_n)_{n\geq 0}$, given by $F_0 = 0, F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$ for $n = 0, 1, \ldots$. We are going to derive an explicit formula for the last decimal digit of F_n in an elementary way.

It is convenient to compute ℓ_{10} evaluated at the subsequences $(F_{30n+i})_{n\geq 0}$ for $i = 0, 1, \ldots, 29$. For any $i \neq 0$ it can be easily checked, by examining the Fibonacci sequence modulo 100, that $(\ell_{10}(F_{30n+i}))_{n\geq 0}$ is periodic and 10 is a common period of all such subsequences.

In the case $i = 0$, we will use equality (i) of Lemma 2.6, which for $b = 10$ and x rational takes the form

$$\ell_{10}(x) \equiv 2^{\nu_2(x)-\nu_{10}(x)} \cdot 5 + 5^{\nu_5(x)-\nu_{10}(x)} \cdot 6 \cdot 3^{\nu_{10}(x)} \ell_5(x) \pmod{10}. \tag{3.22}$$

In order to apply this formula, we need a characterization of $\nu_p(F_n)$ and $\ell_p(F_n)$ for $p = 2, 5$.

The p-adic valuations are given by $\nu_5(F_n) = \nu_5(n)$ and

$$\nu_2(F_n) = \begin{cases} 0 & \text{if } n \equiv 1, 2 \pmod 3, \\ 1 & \text{if } n \equiv 3 \pmod 6, \\ \nu_2(n) + 2 & \text{if } n \equiv 0 \pmod 6, \end{cases}$$

as proved by Lengyel [48].

Now, $\ell_2(F_0) = 0$ and $\ell_2(F_n) = 1$ for all $n \geq 1$. In the case $p = 5$ we use the Binet formula

$$F_n = \frac{\alpha^n - \beta^n}{\alpha - \beta},$$

where $\alpha = (1 + \sqrt{5})/2, \beta = (1 - \sqrt{5})/2$ are the roots of the characteristic polynomial of $x^2 - x - 1$. It follows that

$$2^{n-1} F_n = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1} 5^k,$$

for all $n \geq 1$. As in [48, Lemma 1], for $k \neq 0$ we have

$$\nu_5 \left(\binom{n}{2k+1} 5^k \right) > \nu_5(n),$$

which in turn implies

$$\ell_5(F_n) \equiv 3^{n-1} \ell_5(n) \pmod 5. \tag{3.23}$$

We apply the formula (3.22) to $x = F_{30n}/40n$ rather than $x = F_{30n}$. This is because $\nu_2(F_{30n}/40n) = \nu_5(F_{30n}/40n) = 0$ so all the exponents in (3.22) vanish. Using (3.23), we obtain

$$\begin{aligned} \ell_{10}\left(\frac{F_{30n}}{40n}\right) &\equiv 5 + 6\ell_5\left(\frac{30n \cdot 3^{30n-1}}{40n}\right) \\ &\equiv \begin{cases} -1 & \text{if } n \equiv 0 \pmod 2, \\ 1 & \text{if } n \equiv 1 \pmod 2 \end{cases} \pmod{10}. \end{aligned}$$

We can now recover the value of $\ell_{10}(F_{30n})$, writing

$$\ell_{10}(F_{30n}) = \ell_{10}\left(\frac{F_{30n}}{40n} 40n\right) \equiv \ell_{10}\left(\frac{F_{30n}}{40n}\right) \ell_{10}(4n) \pmod{10}.$$

Taking all into account, we have

$$\ell_{10}(F_n) = \begin{cases} \ell_{10}(F_i) & \text{if } n \equiv i \pmod{300} \\ & \text{for some } i \not\equiv 0 \pmod{30}, \\ 10 - \ell_{10}(4n) & \text{if } n \equiv 0 \pmod{60}, \\ \ell_{10}(4n) & \text{if } n \equiv 30 \pmod{60}. \end{cases}$$

Proposition 2.9 implies that $(\ell_{10}(F_n))_{n\geq 0}$ is 10-automatic.

We remark that the same conclusion can also be established analytically. Following the method described in Section 1.2.3, one can verify that each the subsequences $(F_{60n+m})_{n\geq 0}$ for $m = 0, 1, \ldots, 59$ can be interpolated by a 2-adic and a 5-adic strictly analytic function. Let $(f_{2,m}, f_{5,m}) \in \mathcal{A}_{10}$ denote such a pair of functions. The expressions for $\nu_2(F_n)$ and $\nu_5(F_n)$ show that that unless $m = 0$, at least one of $f_{2,m}, f_{5,m}$ has no root in $\mathbb{Z}_2, \mathbb{Z}_5$, respectively. If $m = 0$, then for $p = 2, 5$ the function $f_{p,0}$ has exactly one root in $\mathbb{Z}_p$, which is equal to 0 and has multiplicity 1. Hence, 10-automaticity of $(\ell_{10}(F_n))_{n\geq 0}$ follows by Theorem 3.17.

In the following example we construct a sequence of integers $(u_n)_{n\geq 0}$ such that $(\ell_b(u_n))_{n\geq 0}$ is k-automatic but b, k are multiplicatively independent. We also give a formula for the terms of this sequence by employing a more analytical approach than in the previous example.

Example 3.8. Let $b = 15$ and define $u_n = (2n-1)(16^n - 4)$ for $n = 0, 1, \ldots$. The sequence $(u_n)_{n\geq 0}$ can be intepolated by 3- and 5-adic analytic functions f_3 and f_5. More precisely, for $p = 3, 5$ let

$$g_p(x) = 16^x - 4 = \exp_p(x \log_p 16) - 4$$

and

$$f_p(x) = (2x - 1)g_p(x),$$

where $x \in \mathbb{Z}_p$. Then for all $n \geq 0$ we have

$$u_n = f_3(n) = f_5(n).$$

We now investigate the roots of f_3 and f_5 in $\mathbb{Z}_3$ and $\mathbb{Z}_5$, respectively. By the well-known Lifting the Exponent Lemma (usually attributed to Lucas [55] and Carmichael [14]), we obtain

$$\nu_3(g_3(n)) = \nu_3(4^{2n-1} - 1) = \nu_3(2n - 1) + 1.$$

At the same time $\nu_5(g_5(n)) = 0$ for all $n \geq 0$. Therefore, $x = 1/2$ is a double root of f_3 and a single root of f_5, and these functions have no other p-adic integer roots. Theorem 3.17 implies that the sequence $(\ell_{15}(u_n))_{n\geq 0}$ is 75-automatic.

In order to compute its terms, we examine the behavior of $\ell_p(f_p(n))$ for $p = 3, 5$. For $p = 5$ it is immediate that

$$\ell_5(f_5(n)) \equiv 2\ell_5(2n - 1) \pmod{5}.$$

In the case $p = 3$, letting $g_3(x) = (2x-1)h_3(x)$, we obtain

$$\ell_3(f_3(n)) \equiv (\ell_3(2n-1))^2 \ell_3(h_3(n)) \pmod 3.$$

By Lemma 3.9 we have that $(\ell_3(h_3(n)))_{n\geq 0}$ is periodic. In order to find a period and, consequently, calculate the terms $\ell_3(h_3(n))$, we need to take a closer look at the coefficients of the power series h_3.

Write

$$h_3(x) = \sum_{i=0}^{\infty} c_i x^i.$$

Comparing the coefficients of $g_3(x) = (2x-1)h_3(x)$, we obtain $c_0 = 3$ and

$$c_i = 2c_{i-1} - \frac{\log_3^i 16}{i!}$$

for $i \geq 1$. Expanding the logarithm, we see that $\nu_3(\log_3 16) = 1$ and $\nu_3(c_1) \geq 2$. It also follows that for $i \geq 2$ we have $\nu_3(\log_3^i 16) = i \geq \nu_3(i!) + 2$. Indeed, this is true for $i = 2$, while for $i \geq 3$ Legendre's formula yields

$$\nu_3(i!) + 2 = \frac{i - s_3(i)}{2} + 2 \leq \frac{i+3}{2} \leq i,$$

where s_3 denotes the ternary sum-of-digits function. Thus, $\nu_3(c_i) \geq 2$ for all $i \geq 1$. As a result, we can write

$$h_3(x) = 3 + 9xt(x),$$

where $\nu_3(t(x)) \geq 0$ for all $x \in \mathbb{Z}_3$. This means that $\ell_3(h_3(x))$ is in fact constant and equal to 1. Therefore,

$$\ell_3(f_3(n)) = \ell_3((2n-1)^2).$$

Finally, we obtain

$$\ell_{15}(u_n) = \ell_{15}(3(2n-1)^2, 2(2n-1)), \tag{3.24}$$

where the first component is the 3-adic part and the second – the 5-adic part. We remark that replacing the value $h_3(n)$ by 3 is valid, as $\nu_3(h_3(n)) = \nu_3(3)$ and $\ell_3(h_3(n)) = \ell_3(3)$.

In the following example, which is a continuation of Example 3.8, we derive a recursive formula which allows us to compute the terms (3.24).

Example 3.9. Let

$$\alpha_n = (3n^2, 2n) \in \mathbb{Z}_3 \times \mathbb{Z}_5,$$

so that $\ell_{15}(u_n) = \ell_{15}(\alpha_{2n-1})$ for all $n \geq 1$ (and $\ell_{15}(u_0) = 3$). We are going to determine the values $\ell_{15}(\alpha_{75n+a})$ for $a = 0, 1, \ldots, 74$, depending on $\ell_{15}(\alpha_n)$ but not n. This is essentially the same as constructing a 75-uniform morphism such that the sequence $(\ell_{15}(\alpha_n))_{n\geq 0}$ is its fixed point. We distinguish a few cases, depending on the value of $\nu_5(a)$ and $\nu_3(a)$.

If $\nu_5(a) = 0$, then $\nu_{15}(\alpha_{75n+a}) = 0$, so

$$\ell_{15}(\alpha_{75n+a}) \equiv \alpha_{75n+a} \equiv (0, 2a) \pmod{15}.$$

Recall that $(x_3, x_5) \bmod 15$ is the integer m from the set $\{0, 1, \ldots, 14\}$ such that $m \equiv x_p \pmod{p}$ for $p = 3, 5$.

If $\nu_5(a) = 1$ or $\nu_5(a) = 2, \nu_3(a) = 0$, then $\nu_{15}(\alpha_{75n+a}) = 1$ and

$$\begin{aligned}\ell_{15}(\alpha_{75n+a}) &\equiv \frac{\alpha_{75n+a}}{15} \equiv \left(\frac{1}{5}(75n+a)^2, \frac{2}{3}\left(15n + \frac{a}{5}\right)\right) \\ &\equiv \left(\frac{1}{5}, \frac{2}{3}\right)\left(a^2, \frac{a}{5}\right) \equiv \left(-a^2, -\frac{a}{5}\right) \pmod{15}.\end{aligned}$$

The only remaining case is $a = 0$. We have

$$\ell_{15}(\alpha_{75n}) \equiv \ell_{15}\left(15^2\left(75n^2, \frac{2}{3}n\right)\right) \equiv \left(25, \frac{1}{3}\right)\ell_{15}(3n^2, 2n) \equiv 7\ell_{15}(\alpha_n) \pmod{15}.$$

To sum up, for all $n \geq 0$ and $a = 0, 1, \ldots, 74$ we have

$$\ell_{15}(\alpha_{75n+a}) \equiv \begin{cases} (0, 2a) & \text{if } \nu_5(a) = 0, \\ \left(-a^2, -\frac{a}{5}\right) & \text{if } \nu_5(a) = 1 \\ & \text{or } \nu_5(a) = 2, \nu_3(a) = 0, \\ 7\ell_{15}(\alpha_n) & \text{if } a = 0 \end{cases} \pmod{15}.$$

4. The 2-adic valuation of generalized Fibonacci sequences

This chapter is mainly based on the paper [69] by the author of the thesis. We derive a formula for the 2-adic valuation for a family of linear recurrence sequences $(t_n(k))_{n\geq 0}$ of order k. As an application, we effectively solve a class of Diophantine equations involving factorials and the terms of these sequences. Moreover, we use the main result to study last nonzero digits of $t_n(k)$ and determine exactly which terms are represented by certain ternary quadratic forms.

4.1 Introduction

The p-adic valuation of linear recurrence sequences has been studied by several authors. Lengyel [48] was the first to completely characterize the p-adic valuation of the Fibonacci sequence $(F_n)_{n\geq 0}$ for all primes p. Medina and Rowland [58] noted that the sequence $(\nu_p(F_{n+1}))_{n\geq 0}$ is p-regular and computed its rank, i.e., the rank of the $\mathbb{Z}$-module generated by the p-kernel of this sequence. Sanna [65] generalized the result of Lengyel to nondegenerate Lucas sequences of the first kind $(u_n)_{n\geq 0}$. We recall that Lucas sequences of the first kind are defined by $u_0 = 0, u_1 = 1$ and for $n \geq 0$ by the recurrence relation $u_{n+2} = au_{n+1} + bu_n$, where a, b are given integers. More precisely, Sanna proved the following:

Theorem 4.1 (Sanna). *If p is a prime number such that $p \nmid b$, then*

$$\nu_p(u_n) = \begin{cases} \nu_p(n) + \nu_p(u_p) - 1 & \text{if } p \mid \Delta, p \mid n, \\ 0 & \text{if } p \mid \Delta, p \nmid n, \\ \nu_p(n) + \nu_p(u_{p\tau(p)}) - 1 & \text{if } p \nmid \Delta, \tau(p) \mid n, p \mid n, \\ \nu_p(u_{p\tau(p)}) & \text{if } p \nmid \Delta, \tau(p) \mid n, p \nmid n, \\ 0 & \text{if } p \nmid \Delta, \tau(p) \nmid n, \end{cases}$$

for each positive integer n, where $\Delta = a^2 + 4b$ and $\tau(p) = \min\{n > 0 : p \mid u_n\}$.

Murru and Sanna [60] further generalized the above results and determined for each integer $k \geq 2$ relatively prime to b the k-adic valuation $\nu_k(u_n)$. In the same paper they observed that the sequence $(\nu_k(u_{n+1}))_{n\geq 0}$ is k-regular and computed its rank in the case when k is a prime not dividing b.

The theorem by Shu and Yao (Theorem 3.1) can be viewed as a quite general result on regularity of p-adic valuations of linear recurrence sequences. To see why this result is relevant to our discussion, consider a linear recurrence sequence $(s_n)_{n\geq 0}$ of nonzero integers. By the construction in Section 1.2.3, for almost all primes p we can find an integer $\pi(p) > 0$ and functions $f_0, f_1 \ldots, f_{\pi(p)-1}$ strictly analytic on $\mathbb{Z}_p$ such that the subsequences $(s_{\pi(p)n+i})_{n\geq 0}$ coincide with $(f_i(n))_{n\geq 0}$ for $i = 0, 1, \ldots, \pi(p) - 1$. The result of Shu and Yao together with Theorem 1.9 imply that $(\nu_p(s_n))_{n\geq 0}$ is p-regular if and only if none of f_i have a root in $\mathbb{Z}_p \setminus \mathbb{Q}$. In the same paper, Shu and Yao applied their result to prove p-regularity of p-adic valuations of a broad class of linear recurrence sequences of order two (without giving explicit formulas).

Much less is known in the case of linear recurrence sequences of higher order. Lengyel and Marques [49] studied the "Tribonacci" numbers $(t_n(3))_{n\geq 0}$, defined by $t_0(3) = 0, t_1(3) = t_2(3) = 1$, and $t_{n+3}(3) = t_{n+2}(3) + t_{n+1}(3) + t_n(3)$ for all $n \geq 0$. They determined the 2-adic valuation of $t_n(3)$ and used this result to find all the solutions of the equation $t_n(3) = m!$ in nonnegative integers n, m. In a later paper [50] these authors considered generalized Fibonacci sequences $(t_n(k))_{n\geq 0}$ of any order $k \geq 2$, defined by $t_0(k) = 0, t_1(k) = \cdots = t_{k-1}(k) = 1$, and for all $n \geq 0$ by the recurrence

$$t_{n+k}(k) = \sum_{i=0}^{k-1} t_{n+i}(k). \tag{4.1}$$

They completely characterized the 2-adic valuation of the "Tetranacci" sequence $(t_n(4))_{n\geq 0}$ and also gave an incomplete formula for the 2-adic valuation of the "Pentanacci" sequence $(t_n(5))_{n\geq 0}$ (these are special cases of Theorems 4.2 and 4.4 below).

The main aim of this chapter is to derive by an elementary method the following formula from [69] (with slightly changed notation).

Theorem 4.2. *If $k \geq 4$ is even, then for all $n \geq 0$ we have*

$$\nu_2(t_n(k)) = \begin{cases} 0 & \text{if } n \equiv 1, 2, \ldots, k \pmod{k+1}, \\ 1 & \text{if } n \equiv k+1 \pmod{2(k+1)}, \\ \nu_2(n) + \nu_2(k-2) + 1 & \text{if } n \equiv 0 \pmod{2(k+1)}. \end{cases}$$

It follows that for even $k \geq 4$ the sequence $(\nu_2(t_{n+1}(k)))_{n\geq 0}$ is 2-regular.

Remark 4.3. In a recent paper, Young [74] gave a short proof of Theorem 4.2 using 2-adic analytic functions.

We shall extend our argument to provide an alternative proof of a result by Young [73], which covers the case of odd $k \geq 5$ (obtained by 2-adic analytic approach). Its statement is also given for the terms $t_n(k)$ with n negative, defined by reversing the relation (4.1).

Theorem 4.4 (Young). *If $k \geq 5$ is odd, then for all integers n we have*

$$\nu_2(t_n(k)) = \begin{cases} 0 & \text{if } n \not\equiv 0, k \pmod{k+1}, \\ \nu_2(k-1) & \text{if } n \equiv k \pmod{2(k+1)}, \\ \nu_2(k-3) & \text{if } n \equiv 2k+1 \pmod{2(k+1)}, \\ \nu_2(n-k-1) & \text{if } n \equiv k+1 \pmod{2(k+1)} \text{ and} \\ & \quad \nu_2(n-k-1) < \nu_2(k^2-1), \\ \nu_2(n-2)+1 & \text{if } n \equiv k+1 \pmod{2(k+1)} \text{ and} \\ & \quad \nu_2(n-k-1) > \nu_2(k^2-1), \\ \nu_2(n)-\nu_2(k+1)+1 & \text{if } n \equiv 0 \pmod{2(k+1)}. \end{cases}$$

Remark 4.5. We note that in the case $n \equiv k+1 \pmod{2(k+1)}$ and $\nu_2(n-k-1) > \nu_2(k^2-1)$ of Theorem 4.4 it is straightforward to check that $\nu_2(n-2) = \nu_2(k-1)$, and therefore one may write the result in a simpler form $\nu_2(t_n(k)) = \nu_2(k-1)+1$ for such n.

These two theorems confirm for all $k \geq 4$ a conjecture of Lengyel and Marques [50, Conjecture 2], which can be equivalently stated as follows:

Conjecture 4.6. *If $k \geq 2$, then for all $n \equiv 0 \pmod{2(k+1)}$ we have*

$$\nu_2\left(t_n(k)\right) = \nu_2(n) + \begin{cases} 2 & \text{if } k = 2, \\ \nu_2(k-2) - \nu_2(k+1) + 1 & \text{if } k \geq 3. \end{cases}$$

In the cases $k = 2, 3$ this equality was already known to be true as a consequence of the mentioned results on Fibonacci and Tribonacci sequences.

The formula in Theorem 4.4 is incomplete since it does not cover the case $n \equiv k+1 \pmod{2(k+1)}$ and $\nu_2(n-k-1) = \nu_2(k^2-1)$. Young gave a somewhat explicit characterization in this case by showing that for each odd $k \geq 5$ there exists $z(k) \in \mathbb{Z}_2$ such that $\nu(t_n(k)) = \nu_2(m - z(k)) + 2$ for all n of the form $n = 2(k+1)m + k + 1$. In particular, this means that the 2-adic analytic function interpolating the subsequence $(t_{(2k+2)m+k+1}(k))_{m \geq 0}$ has $z(k)$ as the only root in $\mathbb{Z}_2$ and this root is of multiplicity one. However, Young's results do not settle whether or not this root is rational, and therefore whether or not $(\nu_2(t_n(k)))_{n \geq 0}$ is 2-regular. While we cannot provide a convincing argument for either possiblity, it seems likely that this sequence is not regular at all.

4.2 The 2-adic valuation of $t_n(k)$

In this section we assume that $k \geq 2$ is fixed. To simplify the notation we will thus write t_n (not to be confused with the Thue–Morse sequence considered in Chapter 1) instead of $t_n(k)$. We start our investigation of $\nu_2(t_n)$ by summarizing some basic properties of the sequence $(t_n)_{n \geq 0}$.

It is periodic modulo 2 with period $k+1$, which follows from the relation

$$t_{n+k+1} = 2t_{n+k} - t_n.$$

It is easy to check that if k is even, then $\nu_2(t_n) > 0$ precisely when n is of the form $n = (k+1)m$ with $m \geq 0$ an integer. Similarly, if k is odd, then $\nu_2(t_n) > 0$ if and only if $n = (k+1)m$ or $n = (k+1)m + k$. The companion matrix of the recurrence (4.1) defining $(t_n)_{n\geq 0}$ has the form

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}.$$

We also introduce for $n \geq 0$ the notation:

$$T_n = \begin{bmatrix} t_n \\ t_{n+1} \\ \vdots \\ t_{n+k-1} \end{bmatrix}, \qquad U_n = \begin{bmatrix} t_n & t_{n+1} & \cdots & t_{n+k-1} \\ t_{n+1} & t_{n+2} & \cdots & t_{n+k} \\ \vdots & \vdots & & \vdots \\ t_{n+k-1} & t_{n+k} & \cdots & t_{n+2k-2} \end{bmatrix}.$$

Clearly, $CT_n = T_{n+1}$ and $CU_n = U_{n+1}$, thus for any nonnegative integers m, n there holds

$$C^m T_n = T_{m+n}, \tag{4.2}$$

$$C^m U_n = U_{m+n}. \tag{4.3}$$

The derivation of the formulas for $\nu_2(t_n)$ given in Theorem 4.2 and Theorem 4.4 is essentially split into two main parts. The first part culminates in Proposition 4.11 below, which provides an explicit expression for t_n modulo 2^{k+1} or 2^k for indices n such that potentially $\nu_2(t_n) > 0$ (it depends on the parity of k whether this inequality holds). This in turn is enough to determine $\nu_2(t_n)$ when this valuation is at most k or $k-1$. The second part is focused on calculating greater valuations. To this end, in Proposition 4.14 below we establish a result which resembles the Lifting the Exponent Lemma with a geometric sequence replaced by $(t_n)_{n\geq 0}$.

To begin, we provide the form of the matrix C^{k+1}, which allows to conveniently compute the vector T_{n+k+1}, given T_n.

Lemma 4.7. *We have*

$$C^{k+1} = 2\begin{bmatrix} 1 & 1 & \ldots & 1 \\ 2 & 2 & \ldots & 2 \\ \vdots & \vdots & & \vdots \\ 2^{k-2} & 2^{k-2} & \ldots & 2^{k-2} \\ 2^{k-1} & 2^{k-1} & \ldots & 2^{k-1} \end{bmatrix} - \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 2 & 1 & \ddots & & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 2^{k-2} & 2^{k-3} & \ldots & 1 & 0 \\ 2^{k-1} & 2^{k-2} & \ldots & 2 & 1 \end{bmatrix}. \tag{4.4}$$

Proof. Induction on $i \geq 0$ together with the identity $t_{n+k+1} = 2t_{n+k} - t_n$ give

$$t_{k+1+i} = 2^{i+1}t_k - \sum_{j=0}^{i} 2^{i-j}t_j = 2^{i+1}\sum_{j=0}^{k-1} t_j - \sum_{j=0}^{i} 2^{i-j}t_j.$$

Let D denote the matrix on the right-hand side of (4.4). Putting $i = 0, 1, \ldots, k-1$ in the above identity and using (4.2) we obtain $T_{k+1} = DT_0 = C^{k+1}T_0$, so $(C^{k+1} - D)T_0 = 0$. Observe that the same holds for every sequence satisfying the recurrence relation (4.1), regardless of the initial terms. Hence, $C^{k+1} = D$, as claimed. □

As an important corollary, we obtain information about periods modulo powers of 2 of $(t_n)_{n\geq 0}$ and other sequences defined by the same recurrence relation.

Corollary 4.8. *Let $(s_n)_{n\geq 0}$ be a sequence of integers satisfying for all $n \geq 0$ the recurrence relation*

$$s_{n+k} = \sum_{i=0}^{k-1} s_{n+i}. \tag{4.5}$$

Then for all integers $l \geq 0$ the sequence $(s_n)_{n\geq 0}$ is periodic modulo 2^{l+1} with (not necessarily minimal) period $2^l(k+1)$.

Proof. Let $D = C^{k+1}$ and let I denote the $k \times k$ identity matrix. By Lemma 4.7 all the entries of $D - I$ are even and an easy induction on $l \geq 0$ shows that all the entries of $D^{2^l} - I$ are divisible by 2^{l+1}.

Let $S_n = [s_n, \ldots, s_{n+k-1}]^T$ for all $n \geq 0$. Since C is the companion matrix of the relation (4.5), we obtain

$$S_{2^l(k+1)} - S_0 = (D^{2^l} - I)S_0,$$

which proves our claim. □

As we will see, calculating $T_{m(k+1)}$ through repeated multplication by C^{k+1} yields sums involving binomial coefficients and powers of 2. We recall two standard identities useful in dealing with these expressions.

Lemma 4.9. *For all integers $l, r \geq 0$ we have*

(i) $\sum_{i=0}^{r} \binom{l+i}{l} = \binom{l+r+1}{l+1}$;

(ii) $\sum_{i=0}^{r} \binom{l+i}{l} 2^i = (-1)^{l+1} + 2^{r+1}\sum_{j=0}^{l} \binom{l+r+1}{l-j}(-2)^j$.

Proof. For any fixed $l \geq 0$ the identity (i) follows from induction on r.

In order to prove (ii) we evaluate in two ways the lth derivative of the function

$$f(x) = \frac{1}{l!}\sum_{i=0}^{r} x^{l+i} = \frac{1}{l!}\,\frac{x^{l+r+1} - x^l}{x-1},$$

at $x = 2$. The left-hand side of (ii) is equal to $f^{(l)}(2)$ calculated for f written in the sum form. The other expression for f gives

$$\begin{aligned} f^{(l)}(x) &= \frac{1}{l!}\sum_{j=0}^{l}\binom{l}{j}\left(\frac{1}{x-1}\right)^{(j)}\left(x^{l+r+1}-x^l\right)^{(l-j)} \\ &= \sum_{j=0}^{l}\frac{(-x)^j}{(x-1)^{j+1}}\left[\binom{l+r+1}{l-j}x^{r+1}-\binom{l}{j}\right] \\ &= x^{r+1}\sum_{j=0}^{l}\binom{l+r+1}{l-j}\frac{(-x)^j}{(x-1)^{j+1}}+\left(\frac{-1}{x-1}\right)^{l+1}. \end{aligned}$$

Substituting $x = 2$, we get the desired result. □

In the following lemma we directly calculate the values of $T_{m(k+1)}$ modulo 2^{k+1}, as well as $t_{m(k+1)+k}$ modulo 2^k.

Lemma 4.10. *For $l \geq 0$ define the following column vectors in $\mathbb{N}^k$:*

$$\begin{aligned} w &= [1,\ldots,1]^T, \\ v_l &= \left[\binom{l}{l}, 2\binom{l+1}{l}, \ldots, 2^{k-1}\binom{l+k-1}{l}\right]^T. \end{aligned}$$

Then for any integer $m \geq 1$ we have

$$T_{m(k+1)} \equiv w + (-1)^{m-1}\left(2(k-2)\sum_{l=0}^{m-1}v_l + v_{m-1}\right) \pmod{2^{k+1}}, \tag{4.6}$$

$$t_{m(k+1)+k} \equiv k - 1 + (-1)^m \cdot 2(k-2)\delta_m \pmod{2^k}, \tag{4.7}$$

where $\delta_m = m \bmod 2$.

Proof. First, we use Lemma 4.7 to calculate how multiplication by C^{k+1} acts on the vectors w and v_l. It is readily checked that

$$C^{k+1}w = 2kv_0 - (2v_0 - w) = w + 2(k-1)v_0, \tag{4.8}$$

while Lemma 4.9 applied to each coordinate yields

$$C^{k+1}v_l \equiv 2\cdot(-1)^{l+1}v_0 - v_{l+1} \pmod{2^{k+1}}. \tag{4.9}$$

In order to prove (4.6) we will now use induction on m. Writing $T_0 = w - [1,0,\cdots,0]^T$, we obtain

$$T_{k+1} = C^{k+1}T_0 = w + 2(k-1)v_0 - v_0 = w + 2(k-2)v_0 + v_0,$$

so the base case $m = 1$ holds.

Now let $m \geq 2$ and assume (4.6) is true for $m-1$, which implies

$$T_{m(k+1)} \equiv C^{k+1}\left[w + (-1)^{m-2}\left(2(k-2)\sum_{l=0}^{m-2} v_l + v_{m-2}\right)\right].$$

If m is even, then by (4.8) and (4.9) we get

$$\begin{aligned} T_{m(k+1)} &\equiv w + 2(k-1)v_0 + 2(k-2)\sum_{l=0}^{m-2}\left(2\cdot(-1)^{l+1}v_0 - v_{l+1}\right) - 2v_0 - v_{m-1} \\ &\equiv w + (2(k-1) - 4(k-2) - 2)v_0 - 2(k-2)\sum_{l=1}^{m-1} v_l - v_{m-1} \\ &\equiv w + (-1)^{m-1}\left(2(k-2)\sum_{l=0}^{m-1} v_l + v_{m-1}\right) \pmod{2^{k+1}}. \end{aligned}$$

A similar computation gives the desired form of $T_{m(k+1)}$ modulo 2^{k+1} for m odd.

The congruence (4.7) follows from the fact that $t_{m(k+1)+k}$ is the sum of the entries of $T_{m(k+1)}$. Indeed, by Lemma 4.9(ii), the sum of the entries of each v_l is congruent to $(-1)^{l+1}$ modulo 2^k. Therefore,

$$\begin{aligned} t_{m(k+1)+k} &\equiv k - 1 + (-1)^{m-1}\left(\cdot 2(k-2)\sum_{l=0}^{m-1}(-1)^{l+1} + (-1)^m\right) \\ &\equiv k - 1 + (-1)^m \cdot 2(k-2)\delta_m \pmod{2^k}. \end{aligned}$$

□

Keeping in mind that $t_{m(k+1)}$ is the first entry of $T_{m(k+1)}$, we can easily deduce the following proposition.

Proposition 4.11. *For all integers $m \geq 0$ we have the following congruence relations:*

$$t_{m(k+1)} \equiv \begin{cases} -2(k-2)m & \text{if } m \text{ is even,} \\ 2((k-2)m+1) & \text{if } m \text{ is odd,} \end{cases} \pmod{2^{k+1}}$$

and

$$t_{m(k+1)+k} \equiv \begin{cases} k-1 & \text{if } m \text{ is even,} \\ -k+3 & \text{if } m \text{ is odd.} \end{cases} \pmod{2^k}.$$

We note that the same result could be obtained direcly by induction, as in the paper of Young [73, Proposition 1]. However, our approach gives congruences for the other entries of $T_{m(k+1)}$ modulo a higher power of 2. As already mentioned, Proposition 4.11 allows to easily compute the values $\nu_2(t_{m(k+1)}) \leq k$ and also $\nu_2(t_{m(k+1)+k})$ in the case $k \neq 3$.

We now proceed to the second part of the reasoning, which deals with the case where $\nu_2(t_{m(k+1)})$ is greater than k. First, we give a technical lemma concerning the matrix U_0.

Lemma 4.12. *The matrix U_0 is invertible and*

$$\nu_2(\det U_0) = \begin{cases} 0 & \text{if } k \text{ is even,} \\ \nu_2(k-1)-1 & \text{if } k \text{ is odd.} \end{cases}$$

Proof. If k is even, then t_n is even if and only if $k+1 \mid n$. Consequently,

$$U_0 \equiv \begin{bmatrix} 0 & 1 & 1 & \dots & & 1 \\ 1 & 1 & \dots & & & 1 \\ 1 & \vdots & & & \iddots & 0 \\ \vdots & & & \iddots & \iddots & 1 \\ & & \iddots & \iddots & \iddots & \vdots \\ 1 & 1 & 0 & 1 & \cdots & 1 \end{bmatrix} \pmod 2,$$

where the latter matrix contains zeros exactly at positions $(1,1)$ and $(i,j) \in \{1,\dots,k\}^2$ such that $i+j = k+3$. Subtracting the second column from all the others allows us to quickly compute that $\det U_0 \equiv 1 \pmod 2$.

The case of k odd is slightly harder. By considering the $2k-1$ initial terms t_n modulo $k-1$, we obtain

$$U_0 \equiv \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & & \iddots & 0 \\ 1 & 1 & 1 & & & \iddots & 0 \\ 1 & 1 & & & & \iddots & -1 \\ 1 & & & & & \iddots & -3 \\ \vdots & \iddots & \iddots & \iddots & \iddots & \iddots & \vdots \\ 1 & 0 & 0 & -1 & -3 & \cdots & 1-2^{k-3} \end{bmatrix} \pmod{(k-1)}, \tag{4.10}$$

where the entry at position (i,j) of the matrix on the right is equal to $1-2^{i+j-k-3}$ when $i+j \geq k+3$. Let U denote this matrix. In order to compute the determinant of U we again perform elementary operations on the columns of U. First, we subtract the first column from all the others. Then, for each $j = k, k-1, ..., 3$ (in this order), we subtract the sum of the second to $(j-1)$-st columns from the j-th column. After these operations we get the matrix

$$\begin{bmatrix} 0 & 1 & 0 & -1 & \cdots & -k+4 & -k+3 \\ 1 & & & & & & -1 \\ 1 & & & & & -1 & \\ \vdots & & & & \iddots & & \\ \vdots & & & \iddots & & & \\ 1 & & -1 & & & & \\ 1 & -1 & & & & & \end{bmatrix}, \tag{4.11}$$

where the blank entries are all zeros. By adding all the other columns in (4.11) to the first one, we compute that

$$\det(U) = \pm\left(-1 + \sum_{j=1}^{k-3} j\right) = \pm\frac{(k-1)(k-4)}{2}.$$

From this and (4.10) we conclude that $\nu_2(\det U_0) = \nu_2(k-1) - 1$. □

The following lemma shows how the term t_{n+m} can be expressed using terms with indices close to n and m.

Lemma 4.13. *For all integers $m, n \geq 0$ we have*

$$t_{m+n} = T_m^T U_0^{-1} T_n.$$

Proof. Identities (4.2) and (4.3) yield

$$T_{m+n} = C^m T_n = U_m U_0^{-1} T_n,$$

and the first entry gives the desired formula. □

As the final step before proving Theorem 4.2 and Theorem 4.4, we apply the last two lemmas to show that certain congruences modulo a fixed power of 2 involving the vectors T_n can be "lifted" to congruences modulo arbitrarily large powers of 2. In effect, this will enable us to determine large values of $\nu_2(t_n)$.

Proposition 4.14. *Let l_0, l_1 be fixed integers such that $l_1 \geq 0$ and $l_0 \geq l_1 + \nu_2(k-1) + 1$. If a column vector $A_0 \in \mathbb{Z}^k$ satisfies the congruence*

$$T_{2^{l_0}(k+1)} \equiv 2^{l_0+1} A_0 + T_0 \pmod{2^{l_0+l_1+2}}, \tag{4.12}$$

then for any $l \geq l_0$ and $s \geq 1$ we also have

$$T_{s2^l(k+1)} \equiv s2^{l+1} A_0 + T_0 \pmod{2^{l+l_1+2}}. \tag{4.13}$$

Proof. We use induction on s and l. First, let $s = 1$ and assume that the congruence (4.13) holds for some $l \geq l_0$ with $A_0 = [a_0, a_1, \ldots, a_{k-1}]^T \in \mathbb{Z}^k$. We can write

$$T_{2^l(k+1)} = 2^{l+l_1+2} B_0 + 2^{l+1} A_0 + T_0,$$

where $B_0 = [b_0, b_1, \ldots, b_{k-1}]^T \in \mathbb{Z}^k$. Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be defined by the same recurrence relation as $(t_n)_{n\geq 0}$, namely

$$a_{n+k} = \sum_{i=0}^{k-1} a_{n+i}, \qquad b_{n+k} = \sum_{i=0}^{k-1} b_{n+i},$$

for all $n \geq 0$. Define $A_n = [a_n, \ldots, a_{n+k-1}]^T$ and $B_n = [b_n, \ldots, b_{n+k-1}]^T$. Then for all $n \geq 0$ we have

$$T_{2^l(k+1)+n} = 2^{l+l_1+2} B_n + 2^{l+1} A_n + T_n. \tag{4.14}$$

To simplify the notation, for $i = 0, 1, ..., k-1$ let $E_i \in \mathbb{Z}^k$ denote the vector with 1 on the i-th position (counting from 0) and 0 on the others. Clearly, $T_i = U_0 E_i$ for $i = 0, 1, ..., k-1$.

We will now consider the elements of the vector $T_{2^{l+1}(k+1)}$. Fix some i such that $0 \leq i \leq k-1$. Using Lemma 4.13 and the equality (4.14), we obtain

$$\begin{aligned} t_{2^{l+1}(k+1)+i} &= T_{2^l(k+1)}^T U_0^{-1} T_{2^l(k+1)+i} \\ &= (2^{l+l_1+2}B_0 + 2^{l+1}A_0 + T_0)^T U_0^{-1}(2^{l+l_1+2}B_i + 2^{l+1}A_i + T_i) \\ &= 2^{2l+2}c + 2^{l+1}\left((2^{l_1+1}B_0 + A_0)^T U_0^{-1} T_i + T_0^T U_0^{-1}(2^{l_1+1}B_i + A_i)\right) \\ &\quad + T_0^T U_0^{-1} T_i \\ &= 2^{2l+2}c + 2^{l+1}\left((2^{l_1+1}B_0 + A_0)^T E_i + E_0^T(2^{l_1+1}B_i + A_i)\right) + T_0^T E_i \\ &= 2^{2l+2}c + 2^{l+l_1+3}b_i + 2^{l+2}a_i + t_i, \end{aligned}$$

where $c = (2^{l_1+1}B_0 + A_0)^T U_0^{-1}(2^{l_1+1}B_i + A_i)$ is a rational number. By Lemma 4.12, regardless of parity of k, we have $\nu_2(\det(U_0)) \leq \nu_2(k-1)$. This implies that

$$\nu_2(2^{2l+2}c) \geq 2l + 2 - \nu_2(\det(U_0)) \geq l + l_0 + 2 - \nu_2(k-1) \geq l + l_1 + 3,$$

and hence (4.13) holds for $l+1$.

Now assume that (4.13) is satisfied for some $s \geq 1$ and all $l \geq l_0$. For any $l \geq l_0$ we obtain

$$\begin{aligned} T_{(s+1)2^l(k+1)} &\equiv 2^{l+1}A_{s2^l(k+1)} + T_{s2^l(k+1)} \\ &\equiv 2^{l+1}A_{s2^l(k+1)} + s2^{l+1}A_0 + T_0 \pmod{2^{l+l_1+2}}. \end{aligned}$$

But Corollary 4.8 shows that $2^l(k+1)$ is a period of the sequence $(a_n)_{n\geq 0}$ modulo 2^{l+1} (and also modulo 2^{l_1+1}, as $l \geq l_0 > l_1$), so

$$2^{l+1}A_{s2^l(k+1)} \equiv 2^{l+1}A_0 \pmod{2^{l+l_1+2}}$$

and the result follows. □

We are now ready to prove the formulas for $\nu_2(t_n)$.

Proof of Theorem 4.2. Let $k \geq 4$ be even. If $n \equiv 1, \ldots, k \pmod{k+1}$ then obviously $\nu_2(t_n) = 0$.

If $n \equiv k+1 \pmod{2(k+1)}$, then Proposition 4.11 immediately implies that $\nu_2(t_n) = 1$.

In the remaining case $n \equiv 0 \pmod{2(k+1)}$, we can write $n = s2^l(k+1)$ with $s \geq 1$ odd and $l \geq 1$. Proposition 4.11 gives

$$t_n \equiv -s2^{l+1}(k-2) \pmod{2^{k+1}}. \tag{4.15}$$

When $l \leq \nu_2(k-2) + 1$ this allows us to conclude that

$$\nu_2(t_n) = l + 1 + \nu_2(k-2) = \nu_2(n) + \nu_2(k-2) + 1.$$

To see this, we need to prove the inequality $2\nu_2(k-2)+2 \le k$ for all even $k \ge 4$. Direct calculation shows that this is an equality for $k = 4, 6$. For $k \ge 8$ it follows from the fact that $\nu_2(k-2) \le \log(k-2)/\log 2$ and the function $f(x) = x - 2\log(x-2)/\log 2$ is increasing if $x \ge 5$.

When $l > \nu_2(k-2)+1$, the congruence (4.15) may not be sufficient to compute $\nu_2(t_n)$. Instead, we apply Proposition 4.14 with $l_0 = \nu_2(k-2)+1$ and $l_1 = \nu_2(k-2)$. Observe that by Corollary 4.8 there exists some $A_0 \in \mathbb{Z}^k$ such that the assumption (4.12) of Proposition 4.14 holds. As we have already seen, $l_0 + l_1 + 2 \le k+1$, so (4.15) with $s = 1, l = l_0$ implies

$$t_{2^{l_0}(k+1)} \equiv -2^{l_0+1}(k-2) \pmod{2^{l_0+l_1+2}}.$$

Hence, we can set $a_0 = -(k-2)$ as the first entry of the vector A_0. By Proposition 4.14 for $l \ge l_0$ we get

$$t_{s2^l(k+1)} \equiv -s2^{l+1}(k-2) \pmod{2^{l+l_1+2}}.$$

This again leads to the equality $\nu_2(t_n) = \nu_2(n) + \nu_2(k-2) + 1$. □

The proof of Theorem 4.4 is very similar, so we omit some of the details.

Proof of Theorem 4.4. Let $k \ge 5$ be odd. We have $C^{2(k+1)} \equiv I \pmod 4$, where I is the $k \times k$ identity matrix. Hence, each of the sequences $(t_{2(k+1)m+i})_{m\ge 0}$ for $i = 0, 1, \ldots, 2k+1$ can be interpolated by a 2-adic analytic function. Since $\mathbb{N}$ is dense in $\mathbb{Z}_p$ and $|\cdot|_p$ is a p-adically continuous function, it is sufficient to prove the result for $n \ge 0$.

If $n \equiv 1, \ldots, k-1 \pmod{k+1}$, then $\nu_2(t_n) = 0$. Proposition 4.11 gives the desired formula for $\nu_2(t_n)$ when $n \equiv k \pmod{(k+1)}$.

The case $n \equiv 0 \pmod{2(k+1)}$ is handled similarly as in the previous proof. In order to apply Proposition 4.14, it suffices to put $l_0 = \nu_2(k-1)+1$, $l_1 = 0$, and $a_0 = -(k-2)$. As a result, for $n = s2^l(k+1)$ with s odd and $l \ge 1$ we obtain

$$\nu_2(t_n) = l + 1 = \nu_2(n) - \nu_2(k+1) + 1.$$

In the case $n \equiv k+1 \pmod{2(k+1)}$ write $n = (s2^l+1)(k+1)$ with $s \ge 1$ odd and $l \ge 1$. It follows from Proposition 4.11 that

$$t_n \equiv 2((k-2)(s2^l+1)+1) \equiv 2((k-1)(s2^l+1) - s2^l) \pmod{2^{k+1}}.$$

Assume that $\nu_2(n-k-1) < \nu_2(k^2-1)$, which is equivalent to $l < \nu_2(k-1)$ and implies $\nu_2(k+1) = 1$. We have

$$\nu_2(t_n) = l + 1 = \nu_2(n-k-1) - \nu_2(k+1) + 1 = \nu_2(n-k-1).$$

On the other hand, if $\nu_2(n-k-1) > \nu_2(k^2-1)$, then $l > \nu_2(k-1)$, and hence

$$\nu_2(t_n) = \nu_2(k-1) + 1,$$

which ends the proof by Remark 4.5. □

Remark 4.15. It is interesting to see why the treatment of the case $n \equiv k+1 \pmod{2(k+1)}$ using Proposition 4.14 does not yield satisfactory results for $k \geq 5$ odd. By applying the recurrence relation (4.1) to (4.12), and (4.13), we find that Proposition 4.14 holds when the indices of all vectors are shifted by $k+1$. This observation and Proposition 4.11 lead to the congruence

$$t_{(2^l s+1)(k+1)} \equiv s2^{l+1}(k-2) + 2(k-1) \pmod{2^{l_0+2}}, \tag{4.16}$$

valid for all integers $l \geq l_0 = \nu_2(k-1)+1$ and $s \geq 1$. However, in the "critical" case $l = \nu_2(k-1)$ the expression $s2^{l+1}(k-2)+2(k-1)$ has unbounded 2-adic valuation as s ranges over odd positive integers, and thus it is not possible to determine $\nu_2(t_{(2^l s+1)(k+1)})$ in one go. This problem did not occur for indices n satisfying $n \equiv 0 \pmod{2(k+1)}$, as the congruence relation of the form (4.16) involved the term $t_0 = 0$ instead of $t_{k+1} = 2(k-1)$.

4.3 Applications

In this section we will show how the formula for $\nu_2(t_n)$ can be applied to effectively solve two families of Diophantine equations.

The first type of equations we are going to consider involves the terms t_n and factorials. So far, this type of problems has been mainly studied for binary recurrence sequences. Luca [52] proved that every nondegenerate binary recurrence sequence contains only finitely many terms expressible as a product of factorials and that the solutions can be effectively computed. He also determined that the only Fibonacci numbers being products of factorials are $F_1, F_2, F_3, F_6, F_{12}$. Luca and Stănică [54] generalized the latter result and showed that $F_1F_2F_3F_4F_5F_6F_8F_{10}F_{12} = 11!$ is the largest product of Fibonacci numbers with distinct indices which is also a product of factorials. The largest Fibonacci number at most one away from a product of factorials was shown to be $F_7 = 2!3! + 1$ by Marques [56].

An additive analogue of these equations was investigated by Grossman and Luca [39], who proved that in a nondegenerate binary recurrence sequence $(s_n)_{n\geq 0}$ there exist only finitely many terms expressible as a $\mathbb{Z}$-linear combination of a given number of factorials with coefficients bounded by a fixed constant. Again, all the solutions can be effectively computed. In the same paper the authors also found that the largest Fibonacci and Lucas numbers being a sum or difference of two factorials are $F_{12} = 5!+4!$ and $L_6 = 4!-3!$. The question of expressing Fibonacci numbers as a sum of three factorials was solved Bollman, Hernández and Luca [12], who found that $F_7 = 1! + 3! + 3!$ is the largest such number. Luca and Siksek [53] studied a related problem of finding all factorials expressible as a sum of at most three Fibonacci numbers and determined the largest solution to be $6! = F_{15}+F_{10}+F_{10} = F_{15}+F_{11}+F_8$.

Many of these works use the famous Primitive Divisor Theorem due to Carmichael [15] or its more general form due to Bilu, Hanrot, and Voutier [10]. Unfortunately, such a result is not known for linear recurrence sequences of order 3 or higher, and thus other methods need to be employed in order to solve Diophantine equations involving the terms of these sequences and factorials. It turns out that computing p-adic

valuation of such a sequence is sometimes sufficient, as shown by Lengyel and Marques [49]. They used the formula for the 2-adic valuation of $t_n(3)$ to find all the factorials among the Tribonacci numbers, the largest being $t_7(3) = 4!$. A characterization of $\nu_2(t_n(3)+1)$ and $\nu_2(t_n(3)-1)$ was also used by Facó and Marques [35] as a means to prove that the Brocard–Ramanujan equation $n!+1 = m^2$ has no solutions with m a Tribonacci number.

We shall use the approach demonstrated in [49, 35] in order to completely solve the equation

$$m! = \prod_{i=1}^{d} t_{n_i} \tag{4.17}$$

in positive integers $m, n_1, \ldots, n_d$, where $k \geq 4$ even and $d \geq 1$ are fixed. In essence, the method used relies on comparing both sides of (4.17) in terms of magnitude and 2-adic valuation. Legendre's formula (Theorem 1.26) and Theorem 4.2 give $\nu_2(m!)$ and $\nu_2(t_n)$, respectively. Factorials may be estimated by an inequality form of Stirling's formula

$$\sqrt{2\pi} m^{m+\frac{1}{2}} e^{\frac{1}{12m+1}-m} < m! < \sqrt{2\pi} m^{m+\frac{1}{2}} e^{\frac{1}{12m}-m},$$

valid for all $m \geq 0$, which was proved by Robbins [62]. Following Lengyel and Marques [49], in order to simplify the calculations we use a less precise upper bound.

Lemma 4.16. *For all $m \geq 6$ we have*

$$m! < \left(\frac{m}{2}\right)^m.$$

Proof. The inequality is easily verified for $m = 6$ and for $m \geq 7$ it follows by induction. □

By a lemma of Wolfram [72, Lemma 3.6] the characteristic polynomial $x^k - x^{k-1} - \cdots - x - 1$ has exactly one real root ϕ lying in the interval $(1, 2)$ (more precisely, $\phi \in (2(1-2^{-k}), 2)$). This allows us to bound t_n from below, as in the following lemma.

Lemma 4.17. *For all $n \geq 1$ we have*

$$t_n \geq \phi^{n-k-1}.$$

Proof. For $n = 1, ..., k-1$ we have $t_n = 1 \geq \phi^{n-k-1}$ and also $t_k = k-1 \geq 2 > \phi$. Then we proceed by induction on k. □

Having all the necessary estimates, we are ready to prove the following result.

Theorem 4.18. *Fix $d \geq 1$ and $k \geq 4$ even. Equation* (4.17) *has only finitely many solutions in positive integers $m, n_1, \ldots, n_d$. Moreover, the solutions can be bounded by an effectively computable constant depending only on d and k.*

Proof. Assume that equation (4.17) is satisfied. The sum $s_2(m)$ of binary digits of m does not exceed $\log m/\log 2+1$, thus by Legendre's formula and Theorem 4.2, we obtain

$$\begin{aligned} m-\frac{\log m}{\log 2}-1 \le \nu_2(m!) &= \sum_{i=1}^{d}\nu_2(t_{n_i}) \le \sum_{i=1}^{d}(\nu_2(n_i)+\nu_2(k-2)+1) \\ &= d(\nu_2(k-2)+1)+\nu_2\left(\prod_{i=1}^{d}n_i\right) \\ &\le d(\nu_2(k-2)+1)+\frac{\log\left(\prod_{i=1}^{d}n_i\right)}{\log 2}. \end{aligned} \tag{4.18}$$

At the same time, Lemmas 4.17 and 4.16 imply that for $m \ge 6$ we have

$$\left(\sum_{i=1}^{d}n_i-d(k-1)\right)\log\phi \le \sum_{i=1}^{d}\log(t_{n_i}) = \log(m!) < m\log\frac{m}{2}, \tag{4.19}$$

where $\phi \in (1,2)$ is a root of $x^k-x^{k-1}-\cdots-x-1$. Combining (4.18) and (4.19) via the inequality between the arithmetic and geometric mean of $n_1,\ldots,n_d$, yields

$$m-\frac{\log m}{\log 2}-1-d(\nu_2(k-2)+1)-\frac{1}{\log 2}\log\left(\frac{m}{\log\phi}\log\frac{m}{2}+d(k-1)\right) < 0. \tag{4.20}$$

The left side of (4.20) tends to infinity as $m\to\infty$, and thus there are only finitely many m satisfying this inequality. Furthermore, the bound on m can be effectively computed and depends only on d and k. □

Numerical calculations based on inequality (4.20) show that the only nontrivial solution of equation (4.17) with k even such that $4 \le k \le 10$ and $1 \le d \le 10$ is $t_5(4) = 3!$.

Remark 4.19. It seems plausible that $t_5(4) = 3!$ is in fact the only nontrivial solution even if we let either d or k (or both) be unbounded. However, in this case the method used in Theorem 4.18 seems insufficient. Indeed, let $m(d,k)$ be the largest positive integer such that the inequality (4.20) holds for all $m \le m(d,k)$. It is straightforward to check that $m(d,k)$ is arbitrarily large if d or k is arbitrarily large.

Theorem 4.18 can be generalized to a broader class of integer sequences.

Theorem 4.20. *Let $(s_n)_{n\ge 0}$ be a sequence of positive integers such that*

$$\log s_n = \Omega(n). \tag{4.21}$$

Assume that for a prime p and some constant $C<1$ we have

$$\nu_p(s_n) = O\left(n^C\right). \tag{4.22}$$

Then for each fixed integer $d \geq 1$ the equation

$$m! = \prod_{i=1}^{d} s_{n_i} \tag{4.23}$$

has only a finite number of solutions in nonnegative integers $m, n_1, n_2, \ldots, n_d$. Moreover, the solutions can be bounded by an effectively computable constant depending only on d, p, and the implied constants in (4.21) *and* (4.22).

Proof. By the assumptions, there exist constants $K_1, K_2 > 0$ and an integer $n_0 \geq 0$ such that $\nu_2(s_n) \leq K_1 n^C$ and $\log s_n \geq K_2 n$ for $n \geq n_0$.

Assume that the equation (4.23) is satisfied for some $m, n_1, n_2, \ldots, n_d$. Without loss of generality we may assume that $n_i \geq n_0$ for $i = 1, 2, \ldots, d$. Indeed, suppose that $n_i < n_0$ for $i = j+1, j+2, \ldots, d$ for some $j < d$ (up to renumbering the n_i). Then $m, n_1, \ldots, n_j$ is a solution to the equation $m! = S \prod_{i=1}^{j} s_{n_i}$ with $S = \prod_{i=j+1}^{d} s_{n_i}$ a constant and the argument provided below can be easily extended to this modified equation.

The reasoning is similar as in the proof of Theorem 4.18. Legendre's formula implies

$$\frac{m - \frac{\log m}{\log p} - 1}{p-1} \leq \nu_p(m!) = \sum_{i=1}^{d} \nu_2(s_{n_i}) \leq dK_1 \left(\max_{1 \leq i \leq d} n_i \right)^C. \tag{4.24}$$

On the other hand, for $m \geq 6$ we have by Lemma 4.16

$$K_2 \max_{1 \leq i \leq d} n_i \leq \sum_{i=1}^{d} \log(s_{n_i}) = \log(m!) < m \log \frac{m}{2}. \tag{4.25}$$

Inequalities (4.24) and (4.25) together give

$$m - \frac{\log m}{\log p} - 1 - dK_1(p-1) \left(\frac{m}{K_2} \log \frac{m}{2} \right)^C < 0.$$

Since $C < 1$, the left side of the above inequality treated as a function of m tends to infinity as $m \to \infty$. Moreover, we can effectively compute a bound on m, which depends only on d, p, K_1, K_2. □

The sequences $(t_n(k))_{n \geq 0}$ with $k \geq 4$ even satisfy the conditions (4.21) (by Lemma 4.17) and (4.22) with $p = 2$ and any $C \in (0, 1)$ (by Theorem 4.2). Hence, Theorem 4.20 implies Theorem 4.18, though the latter usually gives a better bound on the solutions. More generally, we have the following consequence of Theorem 4.20.

Corollary 4.21. *Let $(s_n)_{n \geq 0}$ be a nondegenerate linear recurrence sequence of integers satisfying*

$$s_{n+k} = \sum_{i=0}^{k-1} a_i s_{n+i},$$

where $a_0 \neq 0$, and let p be a prime not dividing a_0. Assume that the minimal polynomial of $(s_n)_{n\geq 0}$ has a root of norm greater than 1. If $(\nu_p(s_n))_{n\geq 0}$ is p-regular, then the equation (4.23) has only finitely many solutions in nonnegative integers $m, n_1, n_2, \ldots, n_d$.

Proof. By the result of Shu and Yao (Theorem 3.1) $(\nu_p(s_n))_{n\geq 0}$ is p-regular if and only if the p-adic strictly analytic functions interpolating the subsequences of $(s_n)_{n\geq 0}$ have no roots in $\mathbb{Z}_p \setminus \mathbb{Q}$. For any such function f, we have $\nu_p(f(n)) = O(\log n)$ as $n \to \infty$, because $\nu_p(n-\theta) = O(\log n)$ when $\theta \in \mathbb{Z}_p \cap \mathbb{Q}$. Hence, $\nu_p(s_n) = O(\log n) = O(n^C)$ for any $C \in (0,1)$. Since $(s_n)_{n\geq 0}$ is nondegenerate, [34, Theorem 2.3] implies $\log |s_n| = \Omega(n)$. Therefore, $(s_n)_{n\geq 0}$ satisfies the assumptions of Theorem 4.20 and the result follows. □

Remark 4.22. In general, the assumption (4.22) of Theorem 4.20 does not hold when $(\nu_p(s_n))_{n\geq 0}$ is regular ora $(s_n)_{n\geq 0}$ is a linear recurrence sequence. Indeed, for $(s_n)_{n\geq 0}$ regular Theorem 1.17 shows that $s_n = O(n^C)$, however usually we cannot choose $C < 1$. Similarly, if $(s_n)_{n\geq 0}$ is a linear recurrence sequence such that one of the associated p-adic analytic functions has a root in $\mathbb{Z}_p \setminus \mathbb{Q}$, then the growth of $\nu_p(s_n)$ is hard to control. More precisely, for any function $a\colon \mathbb{N} \to \mathbb{N}$ it is possible to find $\theta \in \mathbb{Z}_p \setminus \mathbb{Q}$ such that $\nu_p(n-\theta) > a(n)$ infinitely often. For example, let $b_0 = 0$ and for $n \geq 0$ choose b_{n+1} arbitrary such that $b_{n+1} > a(p^{b_0} + \cdots + p^{b_n})$. Put $\theta = \sum_{n=0}^{\infty} p^{b_n}$. Then for all $n \geq 0$ we have $\nu_p(p^{b_0} + \cdots + p^{b_n} - \theta) = b_{n+1} > a(p^{b_0} + \cdots + p^{b_n})$, as desired.

We now move on to the second family of equations, namely the representation of the terms of recurrence sequences by quadratic forms. We are mainly interested in ternary quadratic forms, however for completeness we survey some results for unary and binary quadratic forms. We note that the methods used in each case are totally different.

Cohn [20] found that $F_0 = 0, F_1 = F_2 = 1, F_{12} = 144$ and $L_1 = 1, L_3 = 4$ are the only perfect squares among the Fibonacci and Lucas numbers, respectively. He also determined all Fibonacci and Lucas numbers represented by the form $2x^2$. Since then, many other authors studied the equation $u_n = dx^2$, where $(u_n)_{n\geq 0}$ is a binary recurrence sequence and d a nonzero integer. In particular, Shorey and Stewart [67] proved that for $(u_n)_{n\geq 0}$ nondegenerate there exist only finitely many solutions, which can be bounded by an effectively computable constant. For a survey of related results see [44, 61].

Representation of recurrence sequences by binary quadratic forms has also been studied by several authors. The results have a different flavor than in the case of unary forms. Ballot and Luca [6] showed that for infinitely many $d \in \mathbb{Z}$ the set $\{n > 0 : F_n = |x^2 + dy^2| \text{ for some } x, y \in \mathbb{Z}\}$ has positive lower asymptotic density. Moreover, they provided an upper bound (depending on t) for the number of such d with $|d| \leq t$. Alba González and Luca [2] gave bounds for the number $\#\{0 \leq n \leq t : F_n = x^2 + ny^2 \text{ for some } x, y \in \mathbb{Z}\}$. For each $p \equiv 1 \pmod 4$ prime the equation $F_p = x^2 + py^2$ was shown to have a solution $x, y \in \mathbb{Z}$ by Alba González, Berrizbeitia and Luca [1]. Berrizbeitia, Chapman, Luca, and Mendoza [9] generalized the investigation

to Lucas sequences of the first kind $(u_n)_{n\geq 0}$ and exhibited binary quadratic forms representing u_p or $4u_p$, depending on the residue modulo 4 of a prime p. Ciolan, Luca and Moree [17] showed that if a third-order linear recurrence sequence $(u_n)_{n\geq 0}$ satisfies some technical conditions, the number $\#\{0 \leq n \leq t : u_n = x^2 + ny^2 \text{ for some } x, y \in \mathbb{Z}\}$ can be bounded from above by a certain function of t.

To the best of our knowledge, the only equations of this type involving ternary quadratic forms that have been studied so far are $F_n = x^2 + y^2 + z^2$ and $L_n = x^2 + y^2 + z^2$, as already mentioned in the previous chapter (see [63, 46]). In this case, it turns out that the set of n such that a representation exists can be given explicitly. More precisely, F_n is a sum of three squares of integers if and only if

$$n \notin \{12l + 10 : l \in \mathbb{N}\} \cup \{4^{j+1}(24l + 21) : j, l \in \mathbb{N}\}.$$

Moreover, L_n is a sum of three squares of integers if and only if

$$n \notin \{24l + i : l \in \mathbb{Z}, i \in \{4, 8, 11, 16, 20, 21, 23\}\}.$$

Our goal now is to extend the study to the sequences $(t_n(k))_{n\geq 0}$ with $k \geq 4$ even and other special ternary quadratic forms. We will use the following characterization, which can be extracted from the paper of Blackwell, Durham, Thompson, and Treece [11].

Theorem 4.23 (Blackwell, Durham, Thompson, Treece)**.** *A nonnegative integer m is represented by*

(a) $x^2 + y^2 + 2z^2$ *if and only if* $m \neq 4^j(16l + 14)$*;*

(b) $x^2 + 2y^2 + 2z^2$ *if and only if* $m \neq 4^j(8l + 7)$*;*

(c) $x^2 + 2y^2 + 3z^2$ *if and only if* $m \neq 4^j(16l + 10)$*;*

(d) $x^2 + 2y^2 + 4z^2$ *if and only if* $m \neq 4^j(16l + 14)$*;*

(e) $x^2 + y^2 + 5z^2$ *if and only if* $m \neq 4^j(8l + 3)$.

Observe that in each case the condition on m can be equivalently stated in terms of the last nonzero digits $\ell_{4,2}(m)$. As we will see, a characterization of $\ell_{4,2}(t_n(k))$ with $k \geq 4$ even can be derived using the formula for $\nu_2(t_n(k))$ given in Theorem 4.2. Consequently, for each of the quadratic forms listed in Theorem 4.23 it is possible to determine precisely for which integers $n \geq 0$ the term $t_n(k)$ is represented by the given form. We state this as a theorem.

Theorem 4.24. *Let $k \geq 4$ be an even integer and let q be one of the forms in Theorem 4.23. Then the set*

$$\mathbb{N} \setminus \{n \geq 0 : t_n(k) = q(x, y, z) \text{ for some } x, y, z \in \mathbb{Z}\}$$

is a union of sets of the form

$$\{2^t(k + 1)l + a : l \in \mathbb{N}\}, \tag{4.26}$$

and

$$\{2^{t+2j}(k+1)(8l+b) : j, l \in \mathbb{N}\}, \tag{4.27}$$

where t, a, b *are certain integers such that* $1 \leq t \leq \nu_2(k-2)+7, 0 \leq a \leq 2^t(k+1)-1$, *and* $0 \leq b \leq 7$.

Proof. Considering k fixed, we write t_n instead of $t_n(k)$. We will carry out the proof in the case $q(x,y,z) = x^2+y^2+2z^2$ (only minor details need to be changed for the other quadratic forms). By Theorem 4.23 we see that an integer s is not represented by q if and only if $\ell_{4,2}(s) = 14$, or equivalently, $\nu_2(s)$ is odd and $\ell_{2,3}(s) = 7$. We are going to study which terms of each of the subsequences $(t_{2(k+1)m+i})_{m\geq 0}$ with $i = 0, 1, \ldots, 2k+1$ satisfy these conditions.

For $i \neq 0$ the result is almost immediate. Theorem 4.2 says that the valuation $\nu_2(t_{2(k+1)m+i})$ is odd (and constantly equal to 1) only in the case $i = k+1$. By Corollary 4.8 we see that $(t_{2(k+1)m+k+1})_{m\geq 0}$ is periodic modulo 2^4 with a period of length 4. Therefore, the sequence $(\ell_{2,3}(t_{2(k+1)m+k+1}))_{n\geq 0}$ is also periodic with period 4. This is enough to prove that the set of $n = 2(k+1)m + k + 1$ such that t_n is not represented by q, is a union of arithmetic progressions of the form (4.26) (with $t \leq 3$).

The case $i = 0$ is a bit more delicate. We are going to employ the approach using 2-adic analytic functions, similar to the one in Example 3.8. Let $P(x) = x^k - \sum_{j=0}^{k-1} x^j$ be the characteristic polynomial of the recurrence relation defining $(t_n)_{n\geq 0}$. Its roots $\alpha_1, \ldots, \alpha_k \in \mathbb{C}_2$ all have multiplicity one. To see this we define the polynomial $Q(x) = (x-1)P(x) = x^{k+1} - 2x^k + 1$, which has no common roots with its derivative $Q'(x) = (k+1)x^k - 2kx^{k-1}$. Therefore, P has only simple roots and we can write

$$t_n = \sum_{j=1}^{k} \beta_j \alpha_j^n,$$

where $\beta_1, ..., \beta_k \in \mathbb{C}_2$ satisfy

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_k \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_k^{k-1} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{4.28}$$

The equality $Q(\alpha_j) = 0$ gives $\nu_2(\alpha_j) = 0$, which further implies $\nu_2(\alpha_j^{k+1}+1) = 1$ and $\nu_2(\alpha_j^{k+1}-1) \geq 1$. It follows that

$$\nu_2(\alpha_j^{2(k+1)} - 1) \geq 2,$$

so the function

$$f(x) = \sum_{j=1}^{k} \beta_j \alpha_j^{2(k+1)x} = \sum_{j=1}^{k} \beta_j \exp_2(x \log_2(\alpha_j^{2(k+1)}))$$

is strictly analytic on $\mathbb{Z}_2$. We also have

$$t_{2(k+1)m} = f(m)$$

for all $m \geq 0$.

Expanding the function $\exp_2$ into a power series, write

$$f(x) = \sum_{i=0}^{\infty} c_i x^i,$$

where

$$c_i = \sum_{j=1}^{k} \beta_j \frac{\log_2^i(\alpha_j^{2(k+1)})}{i!}.$$

The function f obviously satisfies $f(0) = 0$ and Theorem 4.2 shows that this root is simple and there are no other roots of f in $\mathbb{Z}_2$. Moreover, writing $f(x) = xg(x)$, we obtain $\nu_2(g(x)) = \nu_2(k-2) + 1$ for all $x \in \mathbb{Z}_2$. The conditions equivalent to $t_{2(k+1)m}$ not being represented by q can thus be restated as

$$\nu_2(m) \equiv \nu_2(k-2) \pmod 2 \tag{4.29}$$

and

$$\ell_{2,3}(m)\ell_{2,3}(g(m)) \equiv 7 \pmod{2^3}. \tag{4.30}$$

We now examine the set of nonnegative integers m such that these congruences are satisfied. From Proposition 3.9 we know that $(\ell_{2,3}(g(m)))_{m\geq 0}$ is a periodic sequence and has a period equal to a power of 2. Suppose for now that 2^T is a period of that sequence, where $T = \nu_2(k-2) + 4$. We consider two cases.

If $m \equiv 0 \pmod{2^T}$, then write $m = 2^T s$ for some integer $s \geq 0$. We have $\nu_2(m) = T + \nu_2(s)$ and $\ell_{2,3}(m) = \ell_{2,3}(s)$. The congruences (4.29) and (4.30) become $\nu_2(s) \equiv 0 \pmod 2$ and $\ell_{2,3}(s) \equiv 7\ell_{2,3}^{-1}(g(2^T)) \pmod 8$. The set of $n = 2(k+1)m = 2^{T+1}(k+1)s$ with s satisfying these conditions is of the form (4.27) (with $t \leq \nu_2(k-2)+5$).

If $m \not\equiv 0 \pmod{2^T}$, write $m = 2^{T+2}l + 2^T u + v$, where l, u, v are integers such that $l \geq 0, 0 \leq u \leq 3$, and $0 < v \leq 2^T - 1$. In this case we have $\nu_2(m) = \nu_2(v)$ and $\ell_{2,3}(m) = \ell_{2,3}(2^T u + v)$. Hence, for each choice of u, v the set of $n = 2(k+1)m$ with m satisfying (4.29) and (4.30) is either empty or an arithmetic progression of the form (4.26) (with $t \leq \nu_2(k-2)+7$).

It remains to prove that $2^{\nu_2(k-2)+4}$ is indeed a period of $(\ell_{2,3}(g(m)))_{m\geq 0}$. To this end, we first show that the coefficients c_i of the function f all have nonnegative 2-adic valuation. We begin with the terms containing $\alpha_j^{2(k+1)x}$. Observe that $\nu_2(\log_2(\alpha_j^{2(k+1)})) = 2$ for $j = 1, \ldots, k$. By Legendre's formula we can then deduce that $\nu_2(\log_2^i(\alpha_j^{2(k+1)})) = 2i \geq \nu_2(i!)$ for all $i = 0, 1, \ldots$.

We move on to the coefficients β_i. The polynomial Q clearly has only simple roots in the algebraic closure of the two-element field $\mathbb{F}_2$, hence so does P. This means that the discriminant D of P is odd. At the same time, D is the square of the determinant of the Vandermonde matrix appearing in the equation (4.28). It follows

that this determinant also has 2-adic valuation equal to zero, and hence $\nu_2(\beta_j) \geq 0$ for $j = 1, \ldots, k$. Consequently, $\nu_2(c_i) \geq 0$ for all $i = 0, 1, \ldots$.

Now, using (1.8), for any $x \in \mathbb{Z}_2$ we obtain

$$g(x + 2^{\nu_2(k-2)+4}) = g(x) + \sum_{i=1}^{\infty} \frac{g^{(i)}(x)}{i!} 2^{i(\nu_2(k-2)+4)}.$$

Since all the coefficients of g have nonnegative 2-adic valuation, this is also true for the functions $g^{(i)}(x)/i!$. As a result, we get

$$\frac{g(x + 2^{\nu_2(k-2)+4})}{2^{\nu_2(k-2)+1}} \equiv \frac{g(x)}{2^{\nu_2(k-2)+1}} \pmod{2^3}.$$

Using $\nu_2(g(x)) = \nu_2(k-2) + 1$ and substituting $x = n$ proves our claim. □

We believe that this theorem could be proved without resorting to p-adic analysis, for example by using the second part of Proposition 4.14. However, the presented approach is directly applicable to other linear recurrence sequences $(s_n)_{n\geq 0}$ whose p-adic valuation is known for some prime p and is p-regular. To conclude this section, we show an example in which we can make Theorem 4.24 more precise.

Example 4.1. We will determine which terms of the Tetranacci sequence $(t_n)_{n\geq 0} = (t_n(4))_{n\geq 0}$ are not represented by the quadratic form $q(x, y, z) = x^2 + 2y^2 + 2z^2$. By Theorem 4.23 these are precisely the terms satisfying $\nu_2(t_n) \equiv 0 \pmod 2$ and $\ell_{2,3}(t_n) = 7$.

Inspecting for $i = 1, \ldots, 9$ the subsequences $(t_{10m+i})_{m\geq 0}$ modulo 8 lets us deduce that t_n is not represented by q if n is of the form $10l + 9$ or $20l + 16$ with $l \geq 0$ an integer.

We now study the remaining subsequence $(t_{10m})_{m\geq 0}$. Following the reasoning in the proof of Theorem 4.24 we can write

$$t_{10m} = mg(m),$$

where g is strictly analytic on $\mathbb{Z}_2$ and $\nu_2(g(x)) = 2$ for all $x \in \mathbb{Z}_2$. Moreover, the argument shows that the sequence $(\ell_{2,3}(g(m)))_{m\geq 0}$ has period 2^5. This is not the minimal period, since a direct computation using the values m, t_{10m} for $m = 1, \ldots, 2^5$ yields

$$\ell_{2,3}(g(m)) = \begin{cases} 5 & \text{if } m \equiv 0 \pmod 4, \\ 3 & \text{if } m \equiv 1 \pmod 4, \\ 1 & \text{if } m \equiv 2 \pmod 4, \\ 7 & \text{if } m \equiv 3 \pmod 4. \end{cases}$$

In each case we are looking for m such that

$$\nu_2(m) \equiv 0 \pmod 2, \tag{4.31}$$

$$\ell_{2,3}(m)\ell_{2,3}(g(m)) \equiv 7 \pmod 8. \tag{4.32}$$

If $m \equiv 0 \pmod 4$, then to satisfy the above congruences it must be of the form $m = 4^{j+1}(8l+3)$ with $j \geq 0, l \geq 0$ integers. In the case $m \equiv 1 \pmod 4$ the condition (4.31) always holds, however $m = 8l+5$ satisfies (4.32), while $m = 8l+1$ does not. The case $m \equiv 2 \pmod 4$ contradicts (4.31). Finally, $m \equiv 3 \pmod 4$ contradicts (4.32).

To sum up, the term $t_n(4)$ is not represented by the form $x^2 + 2y^2 + 2z^2$ if and only if n belongs to the set

$$\{20l + a : l \in \mathbb{N}, a \in \{9, 16, 19\}\} \cup \{80l + 50 : l \in \mathbb{N}\} \cup \{4^{j+1}(80l + 30) : j, l \in \mathbb{N}\}.$$

Bibliography

[1] Juan José Alba González, Pedro Berrizbeitia, and Florian Luca, *On the formula* $F_p = u^2 + pv^2$, Int. J. Number Theory **11** (2015), no. 1, 185–191.

[2] Juan José Alba González and Florian Luca, *On the positive integers n satisfying the equation* $F_n = x^2 + ny^2$, Diophantine methods, lattices, and arithmetic theory of quadratic forms, Contemp. Math., vol. 587, Amer. Math. Soc., Providence, RI, 2013, pp. 95–109.

[3] Jean-Paul Allouche and Jeffrey Shallit, *The ring of k-regular sequences*, Theoret. Comput. Sci. **98** (1992), no. 2, 163–197.

[4] ———, *Automatic sequences: Theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.

[5] ———, *The ring of k-regular sequences. II*, Theoret. Comput. Sci. **307** (2003), no. 1, 3–29.

[6] Christian Ballot and Florian Luca, *On the equation* $x^2 + dy^2 = F_n$, Acta Arith. **127** (2007), no. 2, 145–155.

[7] Jason P. Bell, *A generalization of Cobham's theorem for regular sequences*, Sém. Lothar. Combin. **54A** (2005/07), Art. B54Ap, 15 pp.

[8] ———, *p-adic valuations and k-regular sequences*, Discrete Math. **307** (2007), no. 23, 3070–3075.

[9] Pedro Berrizbeitia, Robin Chapman, Florian Luca, and Alberto Mendoza, *Quadratic forms representing pth terms of Lucas sequences*, J. Number Theory **175** (2017), 134–139.

[10] Yuri Bilu, Guillaume Hanrot, and Paul M. Voutier, *Existence of primitive divisors of Lucas and Lehmer numbers*, J. Reine Angew. Math. **539** (2001), 75–122.

[11] Sarah Blackwell, Gabriel Durham, Katherine Thompson, and Tiffany Treece, *A generalization of a method of Mordell to ternary quadratic forms*, Int. J. Number Theory **12** (2016), no. 8, 2081–2105.

[12] Mark Bollman, Santos Hernández Hernández, and Florian Luca, *Fibonacci numbers which are sums of three factorials*, Publ. Math. Debrecen **77** (2010), no. 1-2, 211–224.

[13] Jakub Byszewski and Jakub Konieczny, *A density version of Cobham's theorem*, Acta Arith. **192** (2020), no. 3, 235–247.

[14] Robert D. Carmichael, *On the Numerical Factors of Certain Arithmetic Forms*, Amer. Math. Monthly **16** (1909), no. 10, 153–159.

[15] ———, *On the numerical factors of the arithmetic forms* $\alpha^n \pm \beta^n$, Ann. of Math. (2) **15** (1913/14), no. 1-4, 30–48, 49–70.

[16] Gilles Christol, *Ensembles presque periodiques k-reconnaissables*, Theoret. Comput. Sci. **9** (1979), no. 1, 141–145.

[17] Alexandru Ciolan, Florian Luca, and Pieter Moree, *Counting terms* U_n *of third order linear recurrences with* $U_n = u^2 + nv^2$, J. Ramanujan Math. Soc. **32** (2017), no. 2, 165–183.

[18] Alan Cobham, *On the base-dependence of sets of numbers recognizable by finite automata*, Math. Systems Theory **3** (1969), 186–192.

[19] ———, *Uniform tag sequences*, Math. Systems Theory **6** (1972), 164–192.

[20] John H. E. Cohn, *Square Fibonacci numbers, etc*, Fibonacci Quart. **2** (1964), 109–113.

[21] Keith Conrad, *Hensel's lemma*, Retrieved November 16, 2020, from `https://kconrad.math.uconn.edu/blurbs/gradnumthy/hensel.pdf`.

[22] ———, *Infinite series in p-adic fields*, Retrieved November 16, 2020, from `https://kconrad.math.uconn.edu/blurbs/gradnumthy/infseriespadic.pdf`.

[23] ———, *The p-adic expansion of rational numbers*, Retrieved November 16, 2020, from `https://kconrad.math.uconn.edu/blurbs/gradnumthy/rationalsinQp.pdf`.

[24] John B. Cosgrave and Karl Dilcher, *An introduction to Gauss factorials*, Amer. Math. Monthly **118** (2011), no. 9, 812–829.

[25] F. Michel Dekking, *Regularity and irregularity of sequences generated by automata*, Seminar on Number Theory, 1979–1980 (French), Univ. Bordeaux I, Talence, 1980, pp. Exp. No. 9, 10.

[26] Jean-Marc Deshouillers, *A footnote to the least non zero digit of* $n!$ *in base* 12, Unif. Distrib. Theory **7** (2012), no. 1, 71–73.

[27] ———, *Yet another footnote to the least non zero digit of* $n!$ *in base* 12, Unif. Distrib. Theory **11** (2016), no. 2, 163–167.

[28] Jean-Marc Deshouillers and Florian Luca, *How often is $n!$ a sum of three squares?*, The legacy of Alladi Ramakrishnan in the mathematical sciences, Springer, New York, 2010, pp. 243–251.

[29] Jean-Marc Deshouillers and Imre Z. Ruzsa, *The least nonzero digit of $n!$ in base 12*, Publ. Math. Debrecen **79** (2011), no. 3-4, 395–400.

[30] Gregory P. Dresden, *Two irrational numbers from the last nonzero digits of $n!$ and n^n*, Math. Mag. **74** (2001), no. 4, 316–320.

[31] ———, *Three transcendental numbers from the last non-zero digits of n^n, F_n, and $n!$*, Math. Mag. **81** (2008), no. 2, 96–105.

[32] Mikhail Ershov, *Completions of rings*, Retrieved November 16, 2020, from `http://people.virginia.edu/~mve2x/7751_Fall2009/lecture26.pdf`.

[33] ———, *l-adic integers (continued)*, Retrieved November 16, 2020, from `http://people.virginia.edu/~mve2x/7751_Fall2009/lecture27.pdf`.

[34] Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward, *Recurrence sequences*, American Mathematical Society, Providence, RI, 2003.

[35] Vinícius Facó and Diego Marques, *Tribonacci numbers and the Brocard-Ramanujan equation*, J. Integer Seq. **19** (2016), no. 4, 16.4.4, 7 pp.

[36] Robert D. Fray, *Congruence properties of ordinary and q-binomial coefficients*, Duke Math. J. **34** (1967), 467–480.

[37] Fernando Q. Gouvêa, *p-adic numbers: An introduction*, second ed., Springer-Verlag, Berlin, 1997.

[38] José María Grau and Antonio M. Oller-Marcén, *On the last digit and the last non-zero digit of n^n in base b*, Bull. Korean Math. Soc. **51** (2014), no. 5, 1325–1337.

[39] George Grossman and Florian Luca, *Sums of factorials in binary recurrence sequences*, J. Number Theory **93** (2002), no. 2, 87–107.

[40] Roman Hampel, *The length of the shortest period of rests of numbers n^n*, Ann. Polon. Math. **1** (1955), 360–366.

[41] Georges Hansel, *A simple proof of the Skolem-Mahler-Lech theorem*, International Colloquium on Automata, Languages, and Programming, Springer, Berlin, 1985, pp. 244–249.

[42] Roger A. Horn and Charles R. Johnson, *Matrix analysis*, second ed., Cambridge University Press, Cambridge, 2013.

[43] Shizuo Kakutani, *Ergodic theory of shift transformations*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 2, Univ. California Press, Berkeley, Calif., 1967, pp. 405–414.

[44] Refik Keskin and Olcay Karaatli, *Generalized Fibonacci and Lucas numbers of the form* $5x^2$, Int. J. Number Theory **11** (2015), no. 3, 931–944.

[45] Neal Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, second ed., Graduate Texts in Mathematics, vol. 58, Springer-Verlag, New York, 1984.

[46] Yaroslav A. Latushkin and Vladimir N. Ushakov, *On the representation of Fibonacci and Lucas numbers as the sum of three squares*, Math. Notes **91** (2012), no. 5-6, 663–670.

[47] Adrien-Marie Legendre, *Théorie des nombres*, Firmin Didot fréres, Paris, 1830.

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

[49] Tamás Lengyel and Diego Marques, *The 2-adic order of the Tribonacci numbers and the equation* $T_n = m!$, J. Integer Seq. **17** (2014), no. 10, Article 14.10.1, 8.

[50] ———, *The 2-adic order of some generalized Fibonacci numbers*, Integers **17** (2017), Paper No. A5, 10 pp.

[51] Eryk Lipka, *Automaticity of the sequence of the last nonzero digits of* $n!$ *in a fixed base*, J. Théor. Nombres Bordeaux **31** (2019), no. 1, 283–291.

[52] Florian Luca, *Products of factorials in binary recurrence sequences*, Rocky Mountain J. Math. **29** (1999), no. 4, 1387–1411.

[53] Florian Luca and Samir Siksek, *On factorials expressible as sums of at most three Fibonacci numbers*, Proc. Edinb. Math. Soc. (2) **53** (2010), no. 3, 747–763.

[54] Florian Luca and Pantelimon Stănică, $F_1F_2F_3F_4F_5F_6F_8F_{10}F_{12} = 11!$, Port. Math. (N.S.) **63** (2006), no. 3, 251–260.

[55] Edouard Lucas, *Theorie des Fonctions Numeriques Simplement Periodiques*, Amer. J. Math. **1** (1878), no. 2, 184–196, 197–240, 289–321.

[56] Diego Marques, *Fibonacci numbers at most one away from a product of factorials*, Notes on Number Theory and Discrete Mathematics **18** (2012), no. 3, 13–19.

[57] Luis A. Medina, Victor H. Moll, and Eric Rowland, *Periodicity in the p-adic valuation of a polynomial*, J. Number Theory **180** (2017), 139–153.

[58] Luis A. Medina and Eric Rowland, *p-regularity of the p-adic valuation of the Fibonacci sequence*, Fibonacci Quart. **53** (2015), no. 3, 265–271.

[59] Piotr Miska and Maciej Ulas, *On some properties of the number of permutations being products of pairwise disjoint d-cycles*, Monatsh. Math. **192** (2020), no. 1, 125–183.

[60] Nadir Murru and Carlo Sanna, *On the k-regularity of the k-adic valuation of Lucas sequences*, J. Théor. Nombres Bordeaux **30** (2018), no. 1, 227–237.

[61] Attila Pethö, *Diophantine properties of linear recursive sequences. I*, Applications of Fibonacci numbers, Vol. 7 (Graz, 1996), Kluwer Acad. Publ., Dordrecht, 1998, pp. 295–309.

[62] Herbert Robbins, *A remark on Stirling's formula*, Amer. Math. Monthly **62** (1955), 26–29.

[63] Neville Robbins, *On Fibonacci and Lucas numbers which are sums of precisely four squares*, Fibonacci Quart. **21** (1983), no. 1, 3–5.

[64] Alain M. Robert, *A course in p-adic analysis*, Graduate Texts in Mathematics, vol. 198, Springer-Verlag, New York, 2000.

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

[66] Jean-Pierre Serre, *A course in arithmetic*, Springer-Verlag, New York-Heidelberg, 1973.

[67] Tarlok N. Shorey and Cameron L. Stewart, *On the Diophantine equation* $ax^{2t} + bx^t y + cy^2 = d$ *and pure powers in recurrence sequences*, Math. Scand. **52** (1983), no. 1, 24–36.

[68] Zhang Shu and Jia-Yan Yao, *Analytic functions over* $\mathbb{Z}_p$ *and p-regular sequences*, C. R. Math. Acad. Sci. Paris **349** (2011), no. 17-18, 947–952.

[69] Bartosz Sobolewski, *The 2-adic valuation of generalized Fibonacci sequences with an application to certain Diophantine equations*, J. Number Theory **180** (2017), 730–742.

[70] ———, *On the last nonzero digits of* $n!$ *in a given base*, Acta Arith. **191** (2019), no. 2, 173–189.

[71] Maciej Ulas and Błażej Żmija, *On p-adic valuations of certain m colored p-ary partition functions*, Ramanujan J. (2020), 34 pp.

[72] David A. Wolfram, *Solving generalized Fibonacci recurrences*, Fibonacci Quart. **36** (1998), no. 2, 129–145.

[73] Paul Thomas Young, *2-adic valuations of generalized Fibonacci numbers of odd order*, Integers **18** (2018), Paper No. A1, 13 pp.

[74] ———, *2-adic properties of generalized fibonacci numbers*, Integers **20** (2020), Paper No. A71, 10 pp.