Magic Numbers in Periodic Sequences[*]

Savinien Kreczman[1][0000−0002−1928−0028],
Luca Prigioniero[2][0000−0001−7163−4965],
Eric Rowland[3][0000−0002−0359−8381], and
Manon Stipulanti[1][0000−0002−2805−2465]

[1] Department of Mathematics, University of Liège, Liège, Belgium
{savinien.kreczman,m.stipulanti}@uliege.be
[2] Dipartimento di Informatica, Università degli Studi di Milano, Milan, Italy
prigioniero@di.unimi.it
[3] Department of Mathematics, Hofstra University, Hempstead, New York, USA
eric.rowland@hofstra.edu

Abstract. In formal languages and automata theory, the magic number problem can be formulated as follows: for a given integer n, is it possible to find a number d in the range $[n, 2^n]$ such that there is no minimal deterministic finite automaton with d states that can be simulated by an optimal nondeterministic finite automaton with exactly n states? If such a number d exists, it is called magic. In this paper, we consider the magic number problem in the framework of deterministic automata with output, which are known to characterize automatic sequences. More precisely, we characterize magic numbers for periodic sequences viewed as either automatic, regular, or constant-recursive.

Keywords: Magic numbers · Periodic sequences · Automatic sequences · Regular sequences · Constant-recursive sequences

1   Introduction

The magic number problem has received wide attention in formal languages and automata theory since its recent formulation [16, 17]. In this setting, it is well known that, given a nondeterministic finite automaton with n states, it is always possible to obtain an equivalent deterministic finite automaton, for example by applying the *powerset construction*. The number of states of the resulting machine may vary between n and 2^n, and the upper bound cannot be reduced in the worst case [23]. Therefore, it is natural to ask whether, given a number n, it is possible to find a number d satisfying $n \leq d \leq 2^n$, such that there is no minimal deterministic finite automaton with exactly d states that can be simulated by an optimal nondeterministic finite automaton with exactly n states. If such a number d exists, it is called *magic* for n. This problem has been studied for finite automata with one- and two-letter alphabets (for which
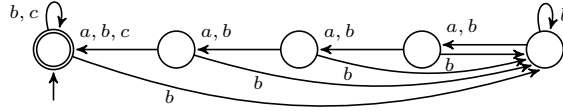
Fig. 1: A nondeterministic finite automaton with 5 states having an equivalent minimal deterministic finite automaton with 2^5 states.

magic numbers have been found [12, 16, 17]), for larger alphabets (for which it has been proved that no magic number exists for each $n \geq 0$ [11, 18, 20]), and for some special cases [14]. The formulation of the magic number problem has been adapted to the cost of operations on regular languages [19]. In that case, the goal is to determine if the whole range of sizes between the lower and the upper bounds can be obtained when applying a given language operation.

Example 1. The nondeterministic finite automaton $\mathcal{A}$ in Figure 1 has 5 states and its corresponding minimal deterministic finite automaton, which can be obtained by minimizing the automaton produced when applying the powerset construction to $\mathcal{A}$, has $2^5 = 32$ states. For each $d \in \{5, \ldots, 32\}$, it is possible to modify $\mathcal{A}$, by adding suitable transitions on the symbol c, to obtain a corresponding minimal deterministic finite automaton with d states. Moreover, this example can be generalized to nondeterministic finite automata with n states, for every $n \geq 0$, to prove that no magic numbers exist in the case of three-letter alphabets [20].

At first, our intention was to investigate the magic number problem in the case of *deterministic finite automata with output*. It is known that these machines, which are simple extensions of deterministic finite automata in which every state is associated with an output, generate the class of *automatic sequences*. More precisely, a sequence $s = s(n)_{n \geq 0}$ is automatic if there is a deterministic finite automaton with output that, for each integer n, outputs $s(n)$ when the digits of n are given as input [3, Chapter 5]. First systematically studied by Cobham [8], automatic sequences have become a central topic in combinatorics on words, with applications in numerous fields, especially number theory [3]. There is a well-established correspondence between a specific minimal deterministic finite automaton with output generating the sequence s and the so-called *kernel* of s, which is a set of particular subsequences extracted from s. Thus, since we are interested in the number of states, it is equivalent to study the kernel sizes of automatic sequences, which is called the *rank*. Specifically, given a period length ℓ, we want to find lower and upper bounds on the possible values of the rank for all periodic sequences with period length ℓ. Moreover, we examine the magic number problem in this context: given an integer $\ell \geq 1$, is it possible to find numbers d within the previous bounds such that there is a periodic sequence with period length ℓ generated by a deterministic finite automaton with output with exactly d states or, equivalently, having rank exactly d? In some specific case, we obtain a full characterization of magic numbers for periodic sequences.

We then extended this study to two other families of sequences, namely *constant-recursive* and *regular*. A sequence is constant-recursive if each of its elements can be obtained from a linear combination of the previous ones. The rank of such a sequence is the smallest degree of all linear recurrences producing it. Motivation to study constant-recursive sequences is that they appear in many areas of computer science and mathematics, such as theoretical biology, software verification, probabilistic model checking, quantum computing, and crystallography; for further details, see [24] or the fine survey [9], and references therein. There are still challenging open problems involving these sequences. For instance, in general, it is not known if it is decidable whether or not a constant-recursive sequence takes the value 0. This question is referred to as the *Skolem problem*. For this class of sequences, we characterize the corresponding magic numbers as sums of the Euler totient function evaluated at divisors of the period length of the sequence.

On the other hand, a sequence s is regular if the vector space generated by its kernel is finitely generated [2]. In this class, the rank of a sequence is the dimension of the vector space. Every automatic sequence, necessarily on a finite alphabet, is regular. In that sense, the notion of regular sequences generalizes that of automatic sequences to infinite alphabets and thus provides a wider framework for the study of sequences. Regular sequences naturally arise in many fields, such as analysis of algorithms, combinatorics, number theory, numerical analysis, and the theory of fractals [2]. In some specific case, we show that the regular case actually boils down to the constant-recursive one.

The paper is organized as follows. In Section 2, we introduce the necessary preliminaries. A specific section is then dedicated to each class of sequences: namely Section 3 to automatic sequences, Section 4 to constant-recursive sequences, and Section 5 to regular sequences. We chose to provide the necessary background for each class of sequences in the corresponding section. As a final comment, proofs omitted due to space constraints may be found in Section 7.

2   Preliminaries

Within a family $\mathcal{F}$ of sequences, we will consider a property $P_{\mathcal{F}}$ on numbers. We say that a number is *magic* if it does not satisfy $P_{\mathcal{F}}$; otherwise, adopting the terminology from [12], it is called *muggle*. The *range for* $P_{\mathcal{F}}$ is the smallest interval containing all muggle numbers with respect to $P_{\mathcal{F}}$.

The *period length* of a periodic sequence s is the smallest integer $\ell \geq 1$ satisfying $s(n + \ell) = s(n)$ for all $n \geq 0$. Let $\mathsf{Per}(\ell)$ denote the set of such periodic sequences with period length ℓ. The *period* of a sequence $s \in \mathsf{Per}(\ell)$ is the tuple $(s(0), s(1), \ldots, s(\ell - 1))$. Note that, in the following, we discard the case $\ell = 1$, which corresponds to constant sequences and for which the three families give a corresponding rank equal to 1.

For a periodic sequence s, subsequences of the form $s(cn + r)_{n \geq 0}$ are also periodic, and their period lengths satisfy the following.

Proposition 1. *Let $\ell \geq 2$, $c, r \geq 0$ be integers. If s is a periodic sequence whose period length divides ℓ, then $s(cn + r)_{n \geq 0}$ is periodic with period length dividing $\frac{\ell}{\gcd(c,\ell)}$.*

3 Automatic Sequences

In this section, we are interested in representations of periodic sequences as automatic sequences. We briefly mention earlier works concerning *periodicity* and *automaticity. Eventual periodicity* and *automatic sets* have been considered in [5, 15, 22]. For $\ell \geq 1$, Alexeev [1] and Sutner [27] studied the automata generating the sequence with period $(1, 0, 0, \ldots, 0)$, with $\ell - 1$ trailing zeroes. Sutner and Tetruashvili determined the rank of the sequence with period $(0, 1, \ldots, \ell-1)$ for specific values of $\ell \geq 1$ [28, Claim 2.1]. Bosma studied some particular cases of several families of periodic sequences on a binary alphabet and considered upper and lower bounds on the rank [6], while Zantema gave an upper bound on the rank of all periodic sequences [29, Theorem 10].

We now introduce the definitions for this section. Let $k, \ell \geq 2$ be integers. Define $\mathrm{ord}_\ell(k)$ to be the smallest integer $m \geq 1$ such that $k^{e+m} \equiv k^e \mod \ell$ for all sufficiently large e. In the special case when k and ℓ are coprime, $\mathrm{ord}_\ell(k)$ is the usual multiplicative order of k modulo ℓ. Let $\mathrm{pre}_\ell(k)$ denote the smallest integer $e \geq 0$ such that $k^{e+\mathrm{ord}_\ell(k)} \equiv k^e \mod \ell$. In other words, $\mathrm{pre}_\ell(k)$ is the length of the preperiod of the sequence $(k^e \mod \ell)_{e \geq 0}$, and $\mathrm{ord}_\ell(k)$ is the (eventual) period length. For every integer n, we let $\mathsf{Div}(n)$ be the set of divisors of n.

Given an integer k, a sequence s is *k-automatic* if there is a deterministic finite automaton with output (DFAO) that, for each integer n, outputs $s(n)$ when fed with the base-k digits of n. To tie the automaton to the base, we call it a *k-automaton*. Here, automata read base-k representations beginning with the least significant digit (note that the other reading direction produces the same set of sequences). There is also a characterization of automatic sequences in terms of the so-called *kernel*. In particular, the *k-kernel* of a sequence s is the set of subsequences $\ker_k(s) = \{s(k^e n + j)_{n \geq 0} \colon e \geq 0 \text{ and } 0 \leq j \leq k^e - 1\}$. A sequence is k-automatic if and only if its k-kernel is finite. We define the *k-rank* (or simply the *rank* when the context is clear) of a k-automatic sequence s to be the quantity $\mathrm{rank}_k(s) = |\ker_k(s)|$. In fact, $\ker_k(s)$ is in bijection with the set of states in the minimal k-automaton that generates s with the property that leading 0's do not affect the output. Every periodic sequence is k-automatic for every $k \geq 2$ [7, p. 88]. Note that, in this case, the alphabet does not matter, so, without loss of generality, the sequences in this section will have nonnegative integer values. For further details on automatic sequences see, e.g., [3].

Example 2. For the sequence $s \in \mathsf{Per}(6)$ with period $(0, 1, 2, 3, 4, 5)$, we illustrate the relationship between the minimal 2-automaton for s and its 2-kernel in Figure 2 and Table 1.

For integers $k, \ell \geq 2$ and the family of periodic sequences $\mathsf{Per}(\ell)$, we consider the following property: an integer satisfies $\mathrm{P_a}(k, \ell)$ if and only if it is equal to the

(e, j)	$s((2^e \bmod 6)n + (j \bmod 6))_{n \geq 0}$	state
$(0, 0)$	$s(n)_{n \geq 0} = 0, 1, 2, 3, 4, 5, 0, 1, 2, 3, 4, 5, \ldots$	00
$(1, 0), (3, 0), (3, 6)$	$s(2n)_{n \geq 0} = 0, 2, 4, 0, 2, 4, \ldots$	10
$(1, 1), (3, 1), (3, 7)$	$s(2n + 1)_{n \geq 0} = 1, 3, 5, 1, 3, 5, \ldots$	11
$(2, 0), (4, 0), (4, 6), (4, 12)$	$s(4n)_{n \geq 0} = 0, 4, 2, 0, 4, 2, \ldots$	20
$(2, 1), (4, 1), (4, 7), (4, 13)$	$s(4n + 1)_{n \geq 0} = 1, 5, 3, 1, 5, 3, \ldots$	21
$(2, 2), (4, 2), (4, 8), (4, 14)$	$s(4n + 2)_{n \geq 0} = 2, 0, 4, 2, 0, 4, \ldots$	22
$(2, 3), (4, 3), (4, 9), (4, 15)$	$s(4n + 3)_{n \geq 0} = 3, 1, 5, 3, 1, 5, \ldots$	23
$(3, 2)$	$s(2n + 2)_{n \geq 0} = 2, 4, 0, 2, 4, 0, \ldots$	32
$(3, 3)$	$s(2n + 3)_{n \geq 0} = 3, 5, 1, 3, 5, 1, \ldots$	33
$(3, 4)$	$s(2n + 4)_{n \geq 0} = 4, 0, 2, 4, 0, 2, \ldots$	34
$(3, 5)$	$s(2n + 5)_{n \geq 0} = 5, 1, 3, 5, 1, 3, \ldots$	35
$(4, 4), (4, 10)$	$s(4n + 4)_{n \geq 0} = 4, 2, 0, 4, 2, 0, \ldots$	44
$(4, 5), (4, 11)$	$s(4n + 5)_{n \geq 0} = 5, 3, 1, 5, 3, 1, \ldots$	45

Table 1: For the sequence $s \in \mathsf{Per}(6)$ with period $(0, 1, 2, 3, 4, 5)$, its 2-kernel contains 13 distinct sequences shown in the second column. The third column contains, for each sequence of the kernel, the corresponding state in the 2-automaton for s depicted in Figure 2.
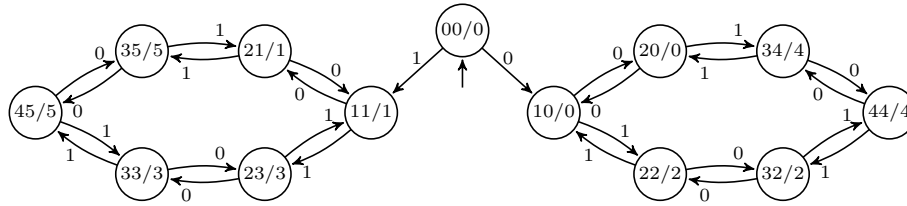

Fig. 2: The minimal 2-automaton outputting the sequence with period $(0, 1, 2, 3, 4, 5)$. Each state has a label of the form α/ρ where α corresponds to a state name in the third column of Table 1 and ρ is the output.

k-rank of some sequence in $\mathsf{Per}(\ell)$. By merging the next two results, we obtain the range for $\mathrm{P_a}(k, \ell)$ as stated in Corollary 1.

Proposition 2. *Let $k, \ell \geq 2$ be two integers. The k-rank of every sequence in $\mathsf{Per}(\ell)$ is at most $B_\ell(k) := \left(\sum_{e=0}^{\mathrm{pre}_\ell(k)-1} \min(k^e, \ell) \right) + \ell \cdot \mathrm{ord}_\ell(k)$. Moreover, the k-rank of the periodic sequence with period $(0, 1, \ldots, \ell - 1)$ is $B_\ell(k)$.*

Proposition 3. *Let $k, \ell \geq 2$ be coprime integers. The k-rank of every sequence in $\mathsf{Per}(\ell)$ is at least ℓ. Moreover, the k-rank of the periodic sequence with period $(0, 0, \ldots, 0, 1)$ is ℓ.*

Corollary 1. *Let $k, \ell \geq 2$ be coprime integers. The interval $[\ell, B_\ell(k)]$ is the range for $\mathrm{P}_a(k, \ell)$.*

The magic and muggle numbers with respect to $P_a(k, \ell)$ when k and ℓ are coprime are characterized in the following theorem.

Theorem 1. *Let $k, \ell \geq 2$ be coprime integers. Let $A = \{d\ell : d \in \mathsf{Div}(\mathrm{ord}_\ell(k))\}$. Then A is the set of muggle numbers in the range for $P_a(k, \ell)$. In particular, the set of magic numbers with respect to $P_a(k, \ell)$ is $\mathbb{N} \setminus A$.*

The following lemmas are useful for the proof of Theorem 1.

Lemma 1. *Let $k, \ell \geq 2$ be coprime integers and let $s \in \mathsf{Per}(\ell)$. Then $\ker_k(s) = \{s(k^e n + j)_{n \geq 0} : e \geq 0 \text{ and } 0 \leq j \leq \ell - 1\}$.*

Proof on page 17

Lemma 2. *Let $k, \ell \geq 2$ be coprime integers. If s belongs to $\mathsf{Per}(\ell)$, then so does $s(kn)_{n \geq 0}$.*

Proof on page 18

We are now ready to give a proof of Theorem 1.

Proof (of Theorem 1). For the sake of conciseness set $T = \{\mathrm{rank}_k(s) : s \in \mathsf{Per}(\ell)\}$. We prove $T = A$ by showing two inclusions.

First inclusion. By Proposition 3, $\mathrm{rank}_k(s) \geq \ell$ for all $s \in \mathsf{Per}(\ell)$. Using Lemma 1, for each $e \geq 0$, let $B_{k,\ell,e}(s) = \{s(k^e n + j)_{n \geq 0} : 0 \leq j \leq \ell - 1\}$ be the set of kernel sequences arising from exponent e. Since k and ℓ are coprime, the sequences in $B_{k,\ell,e}(s)$ are precisely the ℓ shifts $s(k^e(n + i))_{n \geq 0}$ $(0 \leq i \leq \ell - 1)$ of $s(k^e n)_{n \geq 0}$, and they are distinct by Lemma 2, so $|B_{k,\ell,e}(s)| = \ell$. Therefore, if $B_{k,\ell,e}(s) \cap B_{k,\ell,e'}(s) \neq \emptyset$ for some $e, e' \geq 0$, then $B_{k,\ell,e}(s) = B_{k,\ell,e'}(s)$. The sequence $(B_{k,\ell,e}(s))_{e \geq 0}$ of sets is periodic with period length dividing $\mathrm{ord}_\ell(k)$. Therefore $T \subseteq A$.

Second inclusion. Let $d \in \mathsf{Div}(\mathrm{ord}_\ell(k))$. We consider the map $n \mapsto k^d n$ on $\mathbb{Z}/(\ell\mathbb{Z})$, which is a permutation because k^d is invertible modulo ℓ. This permutation has finite order, and the orbit of n under this map is the set $\{n, k^d n, k^{2d} n, \ldots\}$. In particular, the orbit of 0 is $\{0\}$. Let $X_0 = \{0\}, X_1, \ldots, X_m$ be the distinct orbits. These sets partition $\mathbb{Z}/(\ell\mathbb{Z})$ (for an illustration, see Example 3). Define the sequence s on the alphabet $\{0, 1, \ldots, m\}$ by

$$s(n) = i \quad \text{where } (n \bmod \ell) \in X_i. \tag{1}$$

Then s is periodic since $s(n) = s(n + \ell)$. Moreover, s is in $\mathsf{Per}(\ell)$ since 0 only occurs once in the period, namely at positions congruent to 0 modulo ℓ. To prove that $A \subseteq T$, it remains to show that $\mathrm{rank}_k(s) = d\ell$.

First, the sequence $(B_{k,\ell,e}(s))_{e \geq 0}$ defined previously is periodic with period length at most d since $B_{k,\ell,e}(s) = B_{k,\ell,e+d}(s)$ by definition of the orbits. We prove that the period length is exactly d. Let $e, e' \in \{0, 1, \ldots, d - 1\}$ such that $B_{k,\ell,e}(s) = B_{k,\ell,e'}(s)$; we show $e = e'$. Since 0 only occurs once in the period of s, there is precisely one sequence in $B_{k,\ell,e}(s)$ whose initial term is 0, namely $s(k^e n)_{n \geq 0}$. Therefore $s(k^e n)_{n \geq 0} = s(k^{e'} n)_{n \geq 0}$. In particular, $s(k^e) = s(k^{e'})$, so $k^e, k^{e'} \in X_i$ for some $i \neq 0$. By definition of X_i, $k^{e+jd} \equiv k^{e'} \bmod \ell$ for some $j \geq 0$. So $e + jd \equiv e' \bmod \mathrm{ord}_\ell(k)$, which in turn implies $e \equiv e' \bmod d$, yielding in turn $e = e'$. Therefore $\mathrm{rank}_k(s) = d\ell$, as desired.

The second part of the statement follows by Corollary 1. □

d	rank	$3^d \bmod 7$	partition of $\mathbb{Z}/(7\mathbb{Z})$	period
6	42	1	$\{0\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$	$(0, 1, 2, 3, 4, 5, 6)$
3	21	6	$\{0\}, \{1, 6\}, \{2, 5\}, \{3, 4\}$	$(0, 1, 2, 3, 3, 2, 1)$
2	14	2	$\{0\}, \{1, 2, 4\}, \{3, 5, 6\}$	$(0, 1, 1, 2, 1, 2, 2)$
1	7	3	$\{0\}, \{1, 2, 3, 4, 5, 6\}$	$(0, 1, 1, 1, 1, 1, 1)$

Table 2: For each possible value of the rank of periodic sequences in $\mathsf{Per}(7)$, we provide a sequence with such rank.

Notice that the previous proof is constructive, that is, for each divisor d of $\mathrm{ord}_\ell(k)$, Equation (1) defines a sequence whose rank is $d\ell$.

Example 3. We have $\mathrm{ord}_7(3) = 6$, so the muggle numbers are in the set $\{7d : d \in \mathsf{Div}(\mathrm{ord}_7(3))\} = \{7, 14, 21, 42\}$. We provide the period of a periodic sequence in $\mathsf{Per}(7)$ with each rank as shown in Table 2.

Remark 1. Neder [26, A217519] conjectured that if s_ℓ is the sequence with period $(0, 1, \ldots, \ell - 1)$, then "$\mathrm{rank}_2(s_\ell) \leq \ell(\ell - 1)$, with equality if and only if ℓ is a prime with primitive root 2". Recall that a *primitive root* of a prime ℓ is an integer k such that $\mathrm{ord}_\ell(k) = \ell - 1$. Neder's conjecture follows from Theorem 1 and holds even when $k \neq 2$.

4   Constant-Recursive Sequences

An integer sequence s is *constant-recursive* if there exist $c_0, c_1, \ldots, c_{d-1} \in \mathbb{Z}$, with $d \in \mathbb{N}$ and $c_0 \neq 0$ such that

$$s(n + d) = c_{d-1}s(n + d - 1) + \cdots + c_1 s(n + 1) + c_0 s(n) \qquad (2)$$

for all $n \geq 0$. For a constant-recursive sequence s, we let $\mathrm{rank}(s)$ denote its *rank*, which is the smallest integer d such that s satisfies a recurrence relation of the form of Equation (2). In other words, if the set $\{s(n + j)_{n \geq 0} : j \geq 0\}$ containing all shifts of s is called the 1-*kernel* of s, then the rank of s is the dimension of the $\mathbb{Q}$-vector space generated by its 1-kernel.

For an integer $\ell \geq 2$ and the family of periodic sequences $\mathsf{Per}(\ell)$, we consider the following property: an integer satisfies $\mathrm{P}_{\mathrm{cr}}(\ell)$ if and only if it is equal to the rank of some sequence in $\mathsf{Per}(\ell)$. The first few possible values of $\mathrm{rank}(s)$ for $s \in \mathsf{Per}(\ell)$, with $\ell \in \{2, \ldots, 15\}$, are given in Table 3.

We first determine the range for $\mathrm{P}_{\mathrm{cr}}(\ell)$. To do so, we need some definitions and results on *cyclotomic polynomials*. For further details on this topic we address the reader to the nice reference [10].

Definition 1. *Let $m \geq 1$. The mth cyclotomic polynomial is the monic polynomial Φ_m whose roots in $\mathbb{C}$ are the primitive mth roots of unity.*

ℓ	rank$(s), s \in \mathsf{Per}(\ell)$	ℓ	rank$(s), s \in \mathsf{Per}(\ell)$
2	$1, 2$	9	$6, 7, 8, 9$
3	$2, 3$	10	$4, 5, 6, 8, 9, 10$
4	$2, 3, 4$	11	$10, 11$
5	$4, 5$	12	$4, 5, 6, 7, 8, 9, 10, 11, 12$
6	$2, 3, 4, 5, 6$	13	$12, 13$
7	$6, 7$	14	$6, 7, 8, 12, 13, 14$
8	$4, 5, 6, 7, 8$	15	$6, 7, 8, 9, 10, 11, 12, 13, 14, 15$

Table 3: First few values of rank(s) when $s \in \mathsf{Per}(\ell)$ for $\ell \in \{2, \ldots, 15\}$.

The first few cyclotomic polynomials are $\Phi_1(x) = x - 1$, $\Phi_2(x) = x + 1$, $\Phi_3(x) = x^2 + x + 1$, $\Phi_4(x) = x^2 + 1$, $\Phi_5(x) = x^4 + x^3 + x^2 + x + 1$, and $\Phi_6(x) = x^2 - x + 1$. For all $m \geq 2$, $\Phi_m(x)$ is palindromic. We also have the following identity: $x^n - 1 = \prod_{m \in \mathsf{Div}(n)} \Phi_m(x)$ for all $n \geq 1$. The *Euler totient function* ϕ maps every positive integer n to the number of positive integers less than n and relatively prime to n.

Lemma 3 (Garrett [10]). *For all $m \geq 1$, $\Phi_m(x)$ is monic of degree $\phi(m)$.*

As an intermediate step, we characterize the magic and muggle numbers with respect to $\mathrm{P}_{\mathrm{cr}}(\ell)$ in terms of sizes of simple matrices with order ℓ in Theorem 2 (whose proof is postponed). As usual, we let $\mathrm{GL}_d(\mathbb{Z})$ denote the *general linear group of degree d over* $\mathbb{Z}$, made of invertible matrices of size d. Note that for every matrix in $\mathrm{GL}_d(\mathbb{Z})$, its determinant is ± 1. Given a matrix $M \in \mathrm{GL}_d(\mathbb{Z})$, its *order* $\mathrm{ord}\, M$ is the smallest positive integer m such that $M^m = I$, where I is the identity matrix of size d, or $+\infty$ if such an integer does not exist. We say that a matrix M is *simple* if all its eigenvalues (in $\mathbb{C}$) have algebraic multiplicity 1.

Theorem 2. *Let $\ell \geq 2$. The set*

$$R = \{d \in \mathbb{N} : \text{there exists a simple matrix } M \in \mathrm{GL}_d(\mathbb{Z}) \text{ such that } \mathrm{ord}\, M = \ell\}$$

is the set of muggle numbers in the range for $\mathrm{P}_{cr}(\ell)$.

With Theorem 3, Hiller showed that the minimal size of matrices of a given order corresponds to a specific value of the *additive version* ψ of the Euler totient function ϕ, which is defined as follows:
- for every prime p and every $k \geq 1$, $\psi(p^k) = \phi(p^k)$;
- if n is odd, then $\psi(2n) = \phi(n)$; if m and n are relatively prime and not equal to 2, then $\psi(mn) = \phi(m) + \phi(n)$.

Theorem 3 (Hiller [13]). *The smallest integer $d \geq 1$ such that $\mathrm{GL}_d(\mathbb{Z})$ contains a matrix of order ℓ is $\psi(\ell)$.*

Therefore Theorem 3 gives a lower bound for the rank of sequences in $\mathsf{Per}(\ell)$ (see [26, A080737, A152455]).

Proposition 4. *Let $\ell \geq 2$. The interval $[\psi(\ell), \ell]$ is the range for $P_{cr}(\ell)$. Moreover, the sequence $s \in \mathsf{Per}(\ell)$ with period $(0, 0, \ldots, 0, 1)$ satisfies $\mathrm{rank}(s) = \ell$.*

We now give another characterization of the magic and muggle numbers with respect to $P_{cr}(\ell)$ in Theorem 4, whose proof follows from Theorem 2.

Theorem 4. *Let $\ell \geq 2$. Let S be the set of non-empty sets $\{d_1, d_2, \ldots, d_j\}$ of non-negative pairwise distinct integers such that $\mathrm{lcm}(d_1, d_2, \ldots, d_j) = \ell$. The set $R' = \{\sum_{i=1}^{j} \phi(d_i) : \{d_1, d_2, \ldots, d_j\} \in S\}$ is the set of muggle numbers in the range for $P_{cr}(\ell)$. In particular, the set of magic numbers for $P_{cr}(\ell)$ is $\mathbb{N} \setminus R'$.*

From Theorem 4, we directly obtain the following, explaining why only two values are present on prime rows in Table 3.

Corollary 2. *For every prime ℓ, the only muggle numbers with respect to $P_{cr}(\ell)$ are $\ell - 1$ and ℓ.*

We now develop the necessary tools to prove both Theorems 2 and 4. Let s be a constant-recursive sequence satisfying the recurrence relation in Equation (2). The *characteristic polynomial* of the underlying recurrence is $\chi(x) = x^d - c_{d-1}x^{d-1} - \cdots - c_1 x - c_0$. For a monic polynomial $p(x) = x^\ell + a_{\ell-1}x^{\ell-1} + \cdots + a_1 x + a_0$, we let $C(p)$ denote the *companion matrix* of $p(x)$ given by the matrix of size ℓ

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{\ell-1} \end{bmatrix}.$$

The following result links the period length of a periodic sequence satisfying Equation (2) and the order of the companion matrix of the latter recurrence.

Proposition 5. *Let s be a constant-recursive sequence satisfying Equation (2), let $\chi(x)$ be the characteristic polynomial of the underlying recurrence, and let $M \in \mathrm{GL}_d(\mathbb{Z})$ be the companion matrix of $\chi(x)$. Assume that $\mathrm{ord}\, M$ is finite.*
1. *The period length of s is at most $\mathrm{ord}\, M$.*
2. *Moreover, if the initial conditions are $s(0) = s(1) = \cdots = s(d-2) = 0$ and $s(d-1) = 1$, then the period length of s is exactly $\mathrm{ord}\, M$.*

The following result characterizes matrices of finite order.

Theorem 5 (Koo [21, page 147]). *A matrix $M \in \mathrm{GL}_d(\mathbb{Q})$ has finite order if and only if there exist positive integers $r, m_1, \ldots, m_r, n_1, \ldots, n_r$ (where the m_i's are pairwise distinct), an invertible matrix $P \in \mathrm{GL}_d(\mathbb{Q})$, and a block-diagonal matrix*

$$A = \mathrm{diag}(\underbrace{C(\Phi_{m_1}), \ldots, C(\Phi_{m_1})}_{n_1 \ times}, \underbrace{C(\Phi_{m_2}), \ldots, C(\Phi_{m_2})}_{n_2 \ times}, \ldots, \underbrace{C(\Phi_{m_r}), \ldots, C(\Phi_{m_r})}_{n_r \ times}),$$

with $\sum_{i=1}^{r} n_i \deg(\Phi_{m_i}) = d$ such that $M = P^{-1}AP$. Moreover, the order of M is $\mathrm{lcm}(m_1, m_2, \ldots, m_r)$.

Observe that $r \leq \sum_{i=1}^{r} n_i \deg(\Phi_{m_i}) = d$ since $n_i, \deg(\Phi_{m_i}) \geq 1$ for all i.

We state two lemmas useful to prove Theorem 5, describing the characteristic polynomial of the minimal recurrence satisfied by a constant-recursive and periodic sequence.

Lemma 4. *Let $\ell \geq 2$. For every constant-recursive and periodic sequence s whose period length divides ℓ, the characteristic polynomial of the minimal recurrence satisfied by s divides $x^\ell - 1$.*

Lemma 5. *Let $\ell \geq 2$ and let $f(x)$ be a factor of $x^\ell - 1$. The companion matrix M of $f(x)$ has finite order. Moreover, $f(x) = \prod_{i=1}^{r} \Phi_{m_i}(x)$ is a product of distinct cyclotomic polynomials where each index m_i is a divisor of ℓ and $\operatorname{ord} M = \operatorname{lcm}(m_1, m_2, \ldots, m_r)$.*

We are now ready to prove Theorem 2.

Proof (of Theorem 2). For the sake of conciseness, set $T = \{\operatorname{rank}(s) : s \in \mathsf{Per}(\ell)\}$. We prove that $T = R$ by showing two inclusions.

First inclusion. We show $T \subseteq R$. So let $s \in \mathsf{Per}(\ell)$ and let $d = \operatorname{rank}(s)$. We construct a simple matrix $M \in \operatorname{GL}_d(\mathbb{Z})$ such that $\operatorname{ord} M = \ell$. Since $\operatorname{rank}(s) = d$, the sequence s satisfies a recurrence of the form of Equation (2) by definition. Let M be the companion matrix of the characteristic polynomial $\chi(x)$ of this recurrence. By Lemma 4, $\chi(x)$ is a factor of $x^\ell - 1$, so the matrix M is simple. By Lemma 5, M has finite order and $\operatorname{ord} M \leq \ell$. The other inequality, $\ell \leq \operatorname{ord} M$, follows from Proposition 5. Therefore $T \subseteq R$.

Second inclusion. We show $R \subseteq T$. Let $d \geq 1$ and let $M \in \operatorname{GL}_d(\mathbb{Z})$ be a simple matrix such that $\operatorname{ord} M = \ell$. Note that, by minimality, for all $m \in \mathsf{Div}(\ell)$ with $m < \ell$, we have $M^m \neq I$. By Theorem 5, $\chi_M(x) = \prod_{i=1}^{r} \Phi_{m_i}(x)^{n_i}$ for some integers r, m_i, n_i, where each m_i is a divisor of ℓ, the m_i's are distinct, and $\operatorname{lcm}(m_1, m_2, \ldots, m_r) = \ell$. Since M is simple, $n_i = 1$ for all i. Write $\chi_M(x) = x^d - c_{d-1}x^{d-1} - \cdots - c_1 x - c_0$ where $c_0, c_1, \ldots, c_{d-1} \in \mathbb{Z}$. Let s be the sequence satisfying the recurrence $s(n+d) = c_{d-1}s(n+d-1) + \cdots + c_1 s(n+1) + c_0 s(n)$ for all $n \geq 0$ with initial conditions $s(0) = s(1) = \cdots = s(d-2) = 0$ and $s(d-1) = 1$. By Proposition 5, the period length of s is $\operatorname{ord}(M) = \ell$, so $s \in \mathsf{Per}(\ell)$. We claim that $\operatorname{rank}(s) = d$. By definition, s has rank at most d. For all $d' < d$, the only sequence beginning with d' initial zeros and satisfying a recurrence relation of order d' is the constant-zero sequence. Hence $\operatorname{rank}(s) = d$. Consequently, $R \subseteq T$. $\qquad\square$

We illustrate each direction of the proof with an example.

Example 4. Let $s \in \mathsf{Per}(3)$ be the sequence with period $(-1, 0, 1)$. It satisfies $s(n+2) = -s(n+1) - s(n)$ for all $n \geq 0$. The characteristic polynomial of this recurrence is $x^2 + x + 1$, whose companion matrix is

$$M = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

and which is a factor of $x^3 - 1$. It can be checked that $\operatorname{ord} M$ indeed equals 3.

Example 5. Let let $M \in \mathrm{GL}_6(\mathbb{Z})$ be the matrix whose rows are given by the six vectors

$$
\begin{bmatrix} 1 & -1 & 1 & 0 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 1 & 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & -1 & 1 & 1 & 0 & 0 \end{bmatrix},
$$
$$
\begin{bmatrix} -2 & 1 & -2 & -1 & -1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 & -1 & 0 & 0 & -1 \end{bmatrix}.
$$

One can check that M has order 15 and its characteristic polynomial is equal to $\chi_M(x) = x^6 + 2x^5 + 3x^4 + 3x^3 + 3x^2 + 2x + 1 = \Phi_3(x)\Phi_5(x)$. It is not difficult to find an invertible matrix $P \in \mathrm{GL}_6(\mathbb{Q})$ such that $PMP^{-1} = \mathrm{diag}(C(\Phi_3), C(\Phi_5))$. We define s to be the sequence satisfying $s(0) = s(1) = s(2) = s(3) = s(4) = 0$, $s(5) = 1$, and, for all $n \geq 0$,

$$
s(n+6) = -2s(n+5) - 3s(n+4) - 3s(n+3) - 3s(n+2) - 2s(n+1) - s(n).
$$

It can be verified that s has period $(0, 0, 0, 0, 0, 1, -2, 1, 1, -2, 2, -1, -1, 2, -1)$, so $s \in \mathsf{Per}(15)$ and $\mathrm{rank}(s) = 6$.

We now turn to the proof of Theorem 4. Note that the lcm condition in the statement implies that each d_i is a divisor of ℓ.

Proof (of Theorem 4). Let $\ell \geq 2$. By Theorem 2, the sets $\{\mathrm{rank}(s) : s \in \mathsf{Per}(\ell)\}$ and R are equal. By Theorem 5 and Lemma 3, there exists a simple matrix M of size d with order ℓ if and only if $d = \sum_{i=1}^{j} \phi(d_i)$ for some set of integers $\{d_1, d_2, \ldots, d_j\} \in S$. The conclusion of the first part follows and the second one follows by Proposition 4. □

5   Regular Sequences

In this section we look at k-regular sequences [2], which are base-k analogues of constant-recursive sequences. An integer sequence s is *k-regular* if the $\mathbb{Q}$-vector space $V_k(s)$ generated by its k-kernel (defined in Section 3) is finitely generated. The *k-rank* (or simply the *rank* when the context is clear) of s is the dimension of $V_k(s)$. We use, again, the notation $\mathrm{rank}_k(s)$.

Every periodic sequence is k-automatic for every k and therefore also k-regular. However, the corresponding k-ranks of are not necessarily equal.

Example 6. Consider the sequence $s \in \mathsf{Per}(4)$ with period $(0, 1, 1, 1)$. The 2-kernel of s consists of the sequences with periods $(0, 1, 1, 1)$, $(0, 1)$, (1), and (0). Viewed as a 2-automatic sequence, s has rank 4, while, when seen as a 2-regular sequence, it has rank 3 (the sequences with periods $(0, 1, 1, 1)$, $(0, 1)$, and (1) are linearly independent).

For integers $k, \ell \geq 2$ and the family of periodic sequences $\mathsf{Per}(\ell)$, we consider the following property: an integer satisfies $\mathrm{P_r}(k, \ell)$ if and only if it is equal to the k-rank of some sequence in $\mathsf{Per}(\ell)$. To look for magic and muggle numbers with respect to $\mathrm{P_r}(k, \ell)$, we will use the following notion. Let $n \geq 1$ be an integer and let $(a_0, a_1, \ldots, a_{n-1})$ be a sequence of integers. The *circulant matrix*

$C(a_0, a_1, \ldots, a_{n-1})$ is the matrix of size n for which the ith row, $i \in \{1, \ldots, n\}$, is the $(i-1)$st shift of the first row. The polynomial $f(x) = \sum_{i=0}^{n-1} a_i x^i$ is the *associated polynomial* of $C(a_0, a_1, \ldots, a_{n-1})$.

We link the rank of the circulant matrix associated with a given periodic sequence to its constant-recursive and k-regular ranks. The proof of the first result follows from [4, Theorem 2.1.6] and the discussion in [4, Chapter 6, Section 1].

Proposition 6 (Berstel & Reutenauer [4]). *Let $\ell \geq 2$. Let $s \in \mathsf{Per}(\ell)$ and let $M = C(s(0), s(1), \ldots, s(\ell-1))$ be the circulant matrix of size ℓ built on the period of s. The rank of s as a constant-recursive sequence is equal to* $\operatorname{rank} M$.

Proposition 7. *Let $k, \ell \geq 2$ be coprime integers. Let $s \in \mathsf{Per}(\ell)$ and let $M = C(s(0), s(1), \ldots, s(\ell-1))$ be the circulant matrix of size ℓ built on the period of s. Then* $\operatorname{rank}_k(s) = \operatorname{rank} M$.

Theorem 6. *Let $k, \ell \geq 2$ be coprime integers. The range for $\mathrm{P}_r(k, \ell)$ is $[\psi(\ell), \ell]$. Now let S be the set of non-empty sets $\{d_1, d_2, \ldots, d_j\}$ of non-negative pairwise distinct integers with $\operatorname{lcm}(d_1, d_2, \ldots, d_j) = \ell$. The set $R' = \{\sum_{i=1}^{j} \phi(d_i) : \{d_1, d_2, \ldots, d_j\} \in S\}$ is the set of muggle numbers in the range for $\mathrm{P}_r(k, \ell)$. In particular, the set of magic numbers with respect to $\mathrm{P}_r(k, \ell)$ is $\mathbb{N} \setminus R'$.*

Proof. In this proof, for every sequence $s \in \mathsf{Per}(\ell)$, we let M_s denote the circulant matrix $C(s(0), s(1), \ldots, s(\ell-1))$ of size ℓ built on the period of s. As in Section 4, we also let $\operatorname{rank}(s)$ denote the rank of s viewed as constant-recursive sequence. By Propositions 6 and 7, we have that

$$\{\operatorname{rank}_k(s) : s \in \mathsf{Per}(\ell)\} = \{\operatorname{rank} M_s : s \in \mathsf{Per}(\ell)\} = \{\operatorname{rank}(s) : s \in \mathsf{Per}(\ell)\}.$$

In particular, the sets of muggle numbers for $\mathrm{P}_r(k, \ell)$ and $\mathrm{P}_{cr}(\ell)$ are equal. Proposition 4 and Theorem 4 then allow to conclude the proof.   □

6   Conclusion and Future Work

In this paper we studied the magic number problem for three definitions of the rank of periodic sequences. In particular, in the case of constant-recursive sequences, we obtain a full characterization of magic numbers for every period length, while, for k-automatic and k-regular sequences with period length ℓ, we give a characterization provided that k and ℓ are coprime. For such sequences, the general case looks more intricate. Indeed, as Example 6 already shows, the period lengths of sequences in the k-kernel are not always equal to ℓ, but rather divide ℓ. Therefore, a deep understanding of the relationship between the sets $\mathsf{Per}(d)$, where d divides ℓ, is required to characterize the possible ranks, and so in turn the magic and muggle numbers within these families of sequences. Hence, as a natural pursuance of this study, we are currently working on the generalization of the results for automatic and regular sequences to the case where no condition on k and ℓ is specified.

References

1. Alexeev, B.: Minimal DFA for testing divisibility. J. Comput. Syst. Sci. **69**(2), 235–243 (2004). https://doi.org/10.1016/j.jcss.2004.02.001
2. Allouche, J.P., Shallit, J.O.: The ring of k-regular sequences. Theor. Comput. Sci. **98**(2), 163–197 (1992). https://doi.org/10.1016/0304-3975(92)90001-V
3. Allouche, J.P., Shallit, J.O.: Automatic Sequences: Theory, Applications, Generalizations. Cambridge University Press (2003)
4. Berstel, J., Reutenauer, C.: Noncommutative rational series with applications, Encyclopedia of Mathematics and its Applications, vol. 137. Cambridge University Press, Cambridge (2011)
5. Boigelot, B., Mainz, I., Marsault, V., Rigo, M.: An efficient algorithm to decide periodicity of b-recognisable sets using MSDF convention. In: ICALP 2017. LIPIcs, vol. 80, pp. 118:1–118:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2017). https://doi.org/10.4230/LIPIcs.ICALP.2017.118
6. Bosma, W.: Complexity of periodic sequences. Preprint available at https://www.math.ru.nl/ bosma/pubs/periodic.pdf (2019)
7. Büchi, R.J.: Weak second-order arithmetic and finite automata. Mathematical Logic Quarterly **6**(1-6), 66–92 (1960). https://doi.org/https://doi.org/10.1002/malq.19600060105
8. Cobham, A.: Uniform tag sequences. Math. Systems Theory **6**, 164–192 (1972). https://doi.org/10.1007/BF01706087
9. Everest, G., van der Poorten, A.J., Shparlinski, I.E., Ward, T.: Recurrence sequences, Mathematical surveys and monographs, vol. 104. American Mathematical Society (2003)
10. Garrett, P.B.: Abstract algebra. Chapman & Hall/CRC, Boca Raton, FL (2008)
11. Geffert, V.: (Non)determinism and the size of one-way finite automata. In: 7th International Workshop on Descriptional Complexity of Formal Systems - DCFS 2005, Como, Italy, June 30 - July 2, 2005. Proceedings. pp. 23–37. Università degli Studi di Milano, Milan, Italy (2005)
12. Geffert, V.: Magic numbers in the state hierarchy of finite automata. Inform. and Comput. **205**(11), 1652–1670 (2007). https://doi.org/10.1016/j.ic.2007.07.001
13. Hiller, H.: The crystallographic restriction in higher dimensions. Acta Crystallographica Section A **41**(6), 541–544 (1985). https://doi.org/https://doi.org/10.1107/S0108767385001180
14. Holzer, M., Jakobi, S., Kutrib, M.: The magic number problem for subregular language families. Int. J. Found. Comput. Sci. **23**(1), 115–131 (2012). https://doi.org/10.1142/S0129054112400084
15. Honkala, J.: A decision method for the recognizability of sets defined by number systems. RAIRO Theor. Informatics Appl. **20**(4), 395–403 (1986). https://doi.org/10.1051/ita/1986200403951
16. Iwama, K., Kambayashi, Y., Takaki, K.: Tight bounds on the number of states of DFAs that are equivalent to n-state NFAs. Theoret. Comput. Sci. **237**(1-2), 485–494 (2000). https://doi.org/10.1016/S0304-3975(00)00029-3
17. Iwama, K., Matsuura, A., Paterson, M.: A family of NFAs which need $2^n - \alpha$ deterministic states. Theoret. Comput. Sci. **301**(1-3), 451–462 (2003). https://doi.org/10.1016/S0304-3975(02)00891-5
18. Jirásková, G.: Deterministic blow-ups of minimal NFA's. RAIRO Theor. Informatics Appl. **40**(3), 485–499 (2006). https://doi.org/10.1051/ita:2006032

19. Jirásková, G.: On the state complexity of complements, stars, and reversals of regular languages. In: Developments in Language Theory, 12th International Conference, DLT 2008, Kyoto, Japan, September 16-19, 2008. Proceedings. Lecture Notes in Computer Science, vol. 5257, pp. 431–442. Springer (2008). https://doi.org/10.1007/978-3-540-85780-8_34

20. Jirásková, G.: Magic numbers and ternary alphabet. Int. J. Found. Comput. Sci. **22**(2), 331–344 (2011). https://doi.org/10.1142/S0129054111008076

21. Koo, R.: A classification of matrices of finite order over $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Q}$. Mathematics Magazine **76**(2), 143–148 (2003)

22. Marsault, V.: An efficient algorithm to decide periodicity of b-recognisable sets using LSDF convention. Log. Methods Comput. Sci. **15**(3) (2019). https://doi.org/10.23638/LMCS-15(3:8)2019

23. Meyer, A.R., Fischer, M.J.: Economy of description by automata, grammars, and formal systems. In: 12th Annual Symposium on Switching and Automata Theory, East Lansing, Michigan, USA, October 13-15, 1971. pp. 188–191. IEEE Computer Society (1971). https://doi.org/10.1109/SWAT.1971.11

24. Ouaknine, J., Worrell, J.: Decision problems for linear recurrence sequences. In: RP 2012. Lecture Notes in Computer Science, vol. 7550, pp. 21–28. Springer (2012). https://doi.org/10.1007/978-3-642-33512-9_3

25. Rigo, M.: Mathématiques discrètes, notes de cours 2009–2010. Available at http://www.discmath.ulg.ac.be/cours/main_sd.pdf

26. Sloane, N.J.A.: The On-Line Encyclopedia of Integer Sequences. http://oeis.org

27. Sutner, K.: Divisibility and state complexity. The Mathematica Journal **11**(3), 430–445 (2009). https://doi.org/10.3888/tmj.11.3-8

28. Sutner, K., Tetruashvili, S.: Inferring automatic sequences. https://www.cs.cmu.edu/ sutner/papers/auto-seq.pdf (2012)

29. Zantema, H., Bosma, W.: Complexity of automatic sequences. Information and Computation **288**, 104710 (2022). https://doi.org/https://doi.org/10.1016/j.ic.2021.104710, special Issue: Selected Papers of the 14th International Conference on Language and Automata Theory and Applications, LATA 2020

7   Appendix

In this section, we present the proofs that we skipped due to space constraints. We write the statements again for the sake of readability.

7.1   Proof of Proposition 1

Proposition 1. *Let $\ell \geq 2$, $c, r \geq 0$ be integers. If s is a periodic sequence whose period length divides ℓ, then $s(cn + r)_{n \geq 0}$ is periodic with period length dividing $\frac{\ell}{\gcd(c, \ell)}$.*

Proof. For all $n \geq 0$, we have

$$s\left(c\left(n + \frac{\ell}{\gcd(c, \ell)}\right) + r\right) = s\left(cn + \frac{c}{\gcd(c, \ell)}\ell + r\right) = s(cn + r)$$

since $\frac{c}{\gcd(c, \ell)}$ is an integer and the period length of s divides ℓ, which suffices. $\square$

7.2   Proof of Proposition 2

To prove Proposition 2, we need a new notion and two lemmas. We start off with a first one.

Lemma 6. *Let $k, \ell \geq 2$ be integers and let $s \in \mathsf{Per}(\ell)$. Every sequence in $\ker_k(s)$ can be written as $s(cn + r)_{n \geq 0}$ with $c, r \in \{0, 1, \ldots, \ell - 1\}$. In particular, if k and ℓ are coprime, then c is invertible in $\mathbb{Z}_\ell$.*

Proof. Let $e \geq 0$ and $0 \leq j \leq k^e - 1$. Since $s \in \mathsf{Per}(\ell)$, we have $s(k^e n + j)_{n \geq 0} = s((k^e \bmod \ell)n + (j \bmod \ell))_{n \geq 0}$. Now suppose that k and ℓ are coprime. We show that $\gcd(c, \ell) = 1$. To the contrary, suppose that ℓ and c have a common prime factor p, and write $\ell = m_1 p$ and $c = m_2 p$ for some integers $m_1, m_2 \geq 1$. By definition, let $e, m_3 \geq 0$ be such that $k^e = m_3\ell + c$. Then we have $k^e = (m_1 m_3 + m_2)p$, and thus p is a common divisor of k and ℓ, which is impossible. $\square$

For a sequence $s \in \mathsf{Per}(\ell)$, we define two sets of sequences: the (k, ℓ)-*funnel* and the (k, ℓ)-*basin*:

$$\mathrm{fun}_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : 0 \leq e \leq \mathrm{pre}_\ell(k) - 1 \text{ and } 0 \leq j \leq \min(k^e, \ell) - 1\},$$

$$\mathrm{basin}_{k,\ell}(s) = \{s(k^{e+\mathrm{pre}_\ell(k)}n + j)_{n \geq 0} : 0 \leq e \leq \mathrm{ord}_\ell(k) - 1 \text{ and } 0 \leq j \leq \ell - 1\}.$$

It is clear from the definition that $\mathrm{fun}_{k,\ell}(s) \subseteq \ker_k(s)$. We will show that $\mathrm{basin}_{k,\ell}(s) \subseteq \ker_k(s)$ as well. For each kernel sequence $s(k^{e+\mathrm{pre}_\ell(k)}n + j)_{n \geq 0}$ in $\mathrm{basin}_{k,\ell}(s)$, the kernel sequence $s(k^{e+\mathrm{pre}_\ell(k)}(kn+r)+j)_{n \geq 0}$ is also in $\mathrm{basin}_{k,\ell}(s)$, so the basin is closed under this operation. The funnel consists of kernel sequences $s(k^e n + j)_{n \geq 0}$ for which e is too small for k^e to belong to the eventual period of powers of k modulo ℓ. It follows from the definition of the basin that if $s \in \mathsf{Per}(\ell)$ and $k \equiv k' \bmod \ell$, then $\mathrm{basin}_{k,\ell}(s) = \mathrm{basin}_{k',\ell}(s)$.

We have the following description of the kernel.

Lemma 7. *Let $k, \ell \geq 2$ be integers. If $s \in \mathsf{Per}(\ell)$, then $\ker_k(s) = \mathrm{fun}_{k,\ell}(s) \cup \mathrm{basin}_{k,\ell}(s)$.*

Proof. We write $\ker_k(s)$ as the union of the two sets

$$F_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : 0 \leq e \leq \mathrm{pre}_\ell(k) - 1 \text{ and } 0 \leq j \leq k^e - 1\},$$
$$B_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : e \geq \mathrm{pre}_\ell(k) \text{ and } 0 \leq j \leq k^e - 1\}.$$

Since $s \in \mathsf{Per}(\ell)$ (or Lemma 6), we have

$$F_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : 0 \leq e \leq \mathrm{pre}_\ell(k) - 1 \text{ and } 0 \leq j \leq \min(k^e, \ell) - 1\},$$
$$B_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : e \geq \mathrm{pre}_\ell(k) \text{ and } 0 \leq j \leq \min(k^e, \ell) - 1\}.$$

Therefore $F_{k,\ell}(s) = \mathrm{fun}_{k,\ell}(s)$. We show $B_{k,\ell}(s) = \mathrm{basin}_{k,\ell}(s)$. There are two cases. If $k^{\mathrm{pre}_\ell(k)} \geq \ell$, then $\min(k^e, \ell) = \ell$ for all $e \geq \mathrm{pre}_\ell(k)$, so

$$B_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : e \geq \mathrm{pre}_\ell(k) \text{ and } 0 \leq j \leq \ell - 1\}$$
$$= \{s(k^{e+\mathrm{pre}_\ell(k)} n + j)_{n \geq 0} : e \geq 0 \text{ and } 0 \leq j \leq \ell - 1\}.$$

Since $k^{e+\mathrm{ord}_\ell(k)} \equiv k^e \mod \ell$ for all $e \geq \mathrm{pre}_\ell(k)$, we have $B_{k,\ell}(s) = \mathrm{basin}_{k,\ell}(s)$. Now if $k^{\mathrm{pre}_\ell(k)} < \ell$, let $E = \lceil \log_k \ell \rceil$, so that $k^e \geq \ell$ for all $e \geq E$. Since $\mathrm{pre}_\ell(k) < E$ (by definition of E),

$$B_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : \mathrm{pre}_\ell(k) \leq e \leq E - 1 \text{ and } 0 \leq j \leq k^e - 1\}$$
$$\cup \{s(k^e n + j)_{n \geq 0} : e \geq E \text{ and } 0 \leq j \leq \ell - 1\}.$$

We show that the first set in the right-hand side of the previous equality is a subset of the second. Let $\mathrm{pre}_\ell(k) \leq e \leq E - 1$ and $0 \leq j \leq k^e - 1$, so that $s(k^e n + j)_{n \geq 0}$ belongs to the first set. Let $e' \geq E$ such that $k^{e'} \equiv k^e \mod \ell$. Then $s(k^e n + j)_{n \geq 0} = s(k^{e'} n + j)_{n \geq 0}$. Therefore

$$B_{k,\ell}(s) = \{s(k^e n + j)_{n \geq 0} : e \geq E \text{ and } 0 \leq j \leq \ell - 1\}$$
$$= \{s(k^e n + j)_{n \geq 0} : \mathrm{pre}_\ell(k) \leq e \leq \mathrm{pre}_\ell(k) + \mathrm{ord}_\ell(k) - 1 \text{ and } 0 \leq j \leq \ell - 1\}$$
$$= \mathrm{basin}_{k,\ell}(s).$$

$\square$

Proposition 2. *Let $k, \ell \geq 2$ be two integers. The k-rank of every sequence in $\mathsf{Per}(\ell)$ is at most $B_\ell(k) := \left(\sum_{e=0}^{\mathrm{pre}_\ell(k)-1} \min(k^e, \ell) \right) + \ell \cdot \mathrm{ord}_\ell(k)$. Moreover, the k-rank of the periodic sequence with period $(0, 1, \ldots, \ell - 1)$ is $B_\ell(k)$.*

Proof. By Lemma 7,

$$\mathrm{rank}_k(s) \leq |\mathrm{fun}_{k,\ell}(s)| + |\mathrm{basin}_{k,\ell}(s)| \leq \left(\sum_{e=0}^{\mathrm{pre}_\ell(k)-1} \min(k^e, \ell) \right) + \ell \cdot \mathrm{ord}_\ell(k),$$

as desired.

As for the second part of the statement, let g_ℓ be the sequence with period $(0, 1, \ldots, \ell - 1)$. The first part of the statement implies that $\mathrm{rank}_k(g_\ell) \leq B_\ell(k)$. We show that equality holds. Let $s_1, s_2 \in \mathrm{ker}_k(g_\ell)$, and write

$$s_1(n)_{n \geq 0} = g_\ell(k^{e_1} n + j_1)_{n \geq 0} = g_\ell(c_1 n + r_1)_{n \geq 0}$$

and

$$s_2(n)_{n \geq 0} = g_\ell(k^{e_2} n + j_2)_{n \geq 0} = g_\ell(c_2 n + r_2)_{n \geq 0}$$

with $0 \leq e_1, e_2 \leq \mathrm{pre}_\ell(k) + \mathrm{ord}_\ell(k) - 1$, $0 \leq j_1, j_2 \leq \ell - 1$, and $c_1, c_2, r_1, r_2 \in \{0, 1, \ldots, \ell - 1\}$ by Lemma 6. By definition of g_ℓ, we have $g_\ell(n) = g_\ell(m)$ if and only if $n \equiv m \mod \ell$. Therefore $s_1 = s_2$ implies in particular that $r_2 \equiv r_1 \mod \ell$ ($n = 0$) and $c_2 + r_2 \equiv c_1 + r_1 \mod \ell$ ($n = 1$). Since $c_1, c_2, r_1, r_2 \in \{0, 1, \ldots, \ell-1\}$, we obtain $r_1 = r_2$ and $c_1 = c_2$. This in turn implies that $e_1 = e_2$ and $j_1 = j_2$, which concludes the proof.   □

Example 7. Let $k = 2$ and $\ell = 6$. The sequence $(k^e \mod \ell)_{e \geq 0}$ is $1, 2, 4, 2, 4, 2, 4, \ldots$, so $\mathrm{pre}_\ell(k) = 1$ and $\mathrm{ord}_\ell(k) = 2$. For the sequence in Example 2, we obtain $B_\ell(k) = 2^0 + 6 \cdot 2 = 13$. The minimal DFAO generating $g_\ell(n)_{n \geq 0}$ is shown in Figure 2 and has 13 states.

7.3   Proof of Proposition 3

Proposition 3. *Let $k, \ell \geq 2$ be coprime integers. The k-rank of every sequence in $\mathsf{Per}(\ell)$ is at least ℓ. Moreover, the k-rank of the periodic sequence with period $(0, 0, \ldots, 0, 1)$ is ℓ.*

Proof. By Lemma 7, $\mathrm{ker}_k(s) = \mathrm{basin}_{k,\ell}(s)$. Since $\mathrm{pre}_\ell(k) = 0$, then the sequences $s(n+j)_{n \geq 0}$ are elements of $\mathrm{ker}_k(s)$ for $j \geq 0$. There are ℓ distinct such sequences, so $\mathrm{rank}_k(s) \geq \ell$.

For the second part of the statement, let $s \in \mathsf{Per}(\ell)$ be the sequence with period $(0, 0, \ldots, 0, 1)$. By Lemmas 6 and 7, each sequence in $\mathrm{ker}_k(s) = \mathrm{basin}_{k,\ell}(s)$ can be written $s(cn + r)_{n \geq 0}$ where $c, r \in \{0, 1, \ldots, \ell - 1\}$ and $c \neq 0$ is invertible in $\mathbb{Z}_\ell$. The sequence $s(cn+r)_{n \geq 0}$ is periodic with period length at most ℓ. For all $n \geq 0$, $s(cn + r) = 1$ if and only if $cn + r \equiv \ell - 1 \mod \ell$, which is equivalent to $n \equiv c^{-1}(\ell - 1 - r) \mod \ell$. Therefore $s(cn+r)_{n \geq 0}$ is a shift of s, and $\mathrm{rank}_k(s) \leq \ell$. This inequality together with the first part of the statement gives $\mathrm{rank}_k(s) = \ell$.   □

7.4   Proof of Lemma 1

Lemma 1. *Let $k, \ell \geq 2$ be coprime integers and let $s \in \mathsf{Per}(\ell)$. Then $\mathrm{ker}_k(s) = \{s(k^e n + j)_{n \geq 0} : e \geq 0 \text{ and } 0 \leq j \leq \ell - 1\}$.*

Proof. It follows from Lemma 7 as, when k and ℓ are coprime, the set $\mathrm{fun}_{k,\ell}(s)$ is empty.   □

7.5   Proof of Lemma 2

Lemma 2. *Let $k, \ell \geq 2$ be coprime integers. If s belongs to $\mathsf{Per}(\ell)$, then so does $s(kn)_{n \geq 0}$.*

Proof. We proceed by contradiction, and we assume that $s(kn)_{n \geq 0} \in \mathsf{Per}(d)$ with $d \mid \ell$ and $d \neq \ell$. Let j be the inverse of k modulo ℓ which exists because $\gcd(k, \ell) = 1$. Then $s(jkn)_{n \geq 0}$ has period length which divides d. This is impossible since the previous sequence is in fact s, which has period $\ell > d$.   □

7.6   Proof of Proposition 4

Proposition 4. *Let $\ell \geq 2$. The interval $[\psi(\ell), \ell]$ is the range for $\mathrm{P}_{cr}(\ell)$. Moreover, the sequence $s \in \mathsf{Per}(\ell)$ with period $(0, 0, \ldots, 0, 1)$ satisfies $\mathrm{rank}(s) = \ell$.*

Proof. Every sequence s in $\mathsf{Per}(\ell)$ is constant-recursive since $s(n + \ell) = s(n)$ for all $n \geq 0$. Hence its rank is at most ℓ. The lower bound on the rank follows directly from Theorem 3. Rank ℓ is achieved by the sequence with period $(0, 0, \ldots, 0, 1)$.   □

7.7   Proof of Proposition 5

The proof of Proposition 5 was inspired by a similar result for finite fields [25, Proposition IV.5.1]. We start with a lemma stating that powers of the companion matrix give access to further terms of the sequence.

Lemma 8. *Let s be a constant-recursive sequence satisfying Equation (2) and let M be the companion matrix of the corresponding characteristic polynomial. For all $n \geq 0$, we define the vector $\boldsymbol{s}_n = \begin{bmatrix} s(n) \, s(n+1) \cdots s(n+d-1) \end{bmatrix}$. Then for all $m, n \geq 0$, we have $\boldsymbol{s}_n M^m = \boldsymbol{s}_{n+m}$.*

Proof. We proceed by induction on $m \geq 0$. If $m = 0$, the result is clear. By the induction hypothesis, we have $\boldsymbol{s}_n M^{m+1} = \boldsymbol{s}_{n+m} M$, which is equal to

$$\begin{bmatrix} s(n+m+1) \, s(n+m+2) \cdots s(n+m+d-1) \, \sum_{i=0}^{d-1} c_i s(n+m+i) \end{bmatrix}.$$

By assumption, $\sum_{i=0}^{d-1} c_i s(n + m + i) = s(n + m + d)$, which implies that $\boldsymbol{s}_n M^{m+1} = \boldsymbol{s}_{n+m+1}$, as expected.   □

Proposition 5. *Let s be a constant-recursive sequence satisfying Equation (2), let $\chi(x)$ be the characteristic polynomial of the underlying recurrence, and let $M \in \mathrm{GL}_d(\mathbb{Z})$ be the companion matrix of $\chi(x)$. Assume that $\mathrm{ord}\, M$ is finite.*
  1. The period length of s is at most $\mathrm{ord}\, M$.
  2. Moreover, if the initial conditions are $s(0) = s(1) = \cdots = s(d - 2) = 0$ and $s(d - 1) = 1$, then the period length of s is exactly $\mathrm{ord}\, M$.

Proof. Let $\text{ord}\, M = \ell$. For all $n \geq 0$, Lemma 8 gives $\boldsymbol{s}_{n+\ell} = \boldsymbol{s}_n M^\ell = \boldsymbol{s}_n$, so $s(n+\ell) = s(n)$ for all $n \geq 0$. This in turn implies that the period length of s is at most ℓ for all initial conditions. This proves Item (1).

Let us prove Item (2). We show that every period length of the sequence s with initial conditions $s(0) = s(1) = \cdots = s(d-2) = 0$ and $s(d-1) = 1$ is at least ℓ. Let $\{\boldsymbol{e}_i : 1 \leq i \leq d\}$ denote the standard basis of $\mathbb{R}^d$, that is, for all $i \in \{1, 2, \ldots, d\}$, $\boldsymbol{e}_i$ has a 1 in the ith coordinate and 0's elsewhere. By assumption, we have $\boldsymbol{s}_0 = \boldsymbol{e}_d$. We can show by induction that, for all $j \in \{0, 1, \ldots, d-1\}$,

$$\boldsymbol{s}_j = \boldsymbol{s}_0 M^j = \boldsymbol{e}_{d-j} + \sum_{i=d-j+1}^{d} \alpha_{j,i} \boldsymbol{e}_i$$

for some coefficients $\alpha_{i,j} \in \mathbb{Q}$. Let

$$A = \begin{bmatrix} \boldsymbol{s}_{d-1} \\ \boldsymbol{s}_{d-2} \\ \vdots \\ \boldsymbol{s}_0 \end{bmatrix} = \begin{bmatrix} 1 & \alpha_{d-1,2} & \alpha_{d-1,3} & \cdots & \alpha_{d-1,d} \\ 0 & 1 & \alpha_{d-2,2} & \cdots & \alpha_{d-2,d} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \alpha_{1,d} \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \in \text{GL}_d(\mathbb{Q}).$$

Note that $\det(A) = 1$. By Lemma 8 again, we obtain

$$AM^m = \begin{bmatrix} \boldsymbol{s}_{d-1} \\ \boldsymbol{s}_{d-2} \\ \vdots \\ \boldsymbol{s}_0 \end{bmatrix} M^m = \begin{bmatrix} \boldsymbol{s}_{m+d-1} \\ \boldsymbol{s}_{m+d-2} \\ \vdots \\ \boldsymbol{s}_m \end{bmatrix} \tag{3}$$

for all $m \geq 0$. Now suppose that s belongs to $\text{Per}(m)$. We want to show that $m \geq \ell$. By definition of m being a period length, Equation (3) becomes

$$AM^m = \begin{bmatrix} \boldsymbol{s}_{d-1} \\ \boldsymbol{s}_{d-2} \\ \vdots \\ \boldsymbol{s}_0 \end{bmatrix} = A.$$

Since A is invertible, we conclude that $M^m = I$, so m is a multiple of $\text{ord}(M) = \ell$. In particular, $m \geq \ell$, as desired. $\square$

7.8   Proof of Lemma 4

Lemma 4. *Let $\ell \geq 2$. For every constant-recursive and periodic sequence s whose period length divides ℓ, the characteristic polynomial of the minimal recurrence satisfied by s divides $x^\ell - 1$.*

Proof. By assumption, we have $s(n+\ell) = s(n)$ for all $n \geq 0$. The characteristic polynomial of the latter recurrence is $x^\ell - 1$. The conclusion follows by minimality. $\square$

7.9   Proof of Lemma 5

Lemma 5. *Let $\ell \geq 2$ and let $f(x)$ be a factor of $x^\ell - 1$. The companion matrix M of $f(x)$ has finite order. Moreover, $f(x) = \prod_{i=1}^{r} \Phi_{m_i}(x)$ is a product of distinct cyclotomic polynomials where each index m_i is a divisor of ℓ and $\operatorname{ord} M = \operatorname{lcm}(m_1, m_2, \ldots, m_r)$.*

Proof. Since the roots of $f(x)$ are roots of unity, the eigenvalues of M are roots of unity with order dividing ℓ. Therefore the matrix M has finite order by Theorem 5, and its order is $\operatorname{lcm}(m_1, m_2, \ldots, m_r)$.   □

7.10   Proof of Proposition 7

Proposition 7. *Let $k, \ell \geq 2$ be coprime integers. Let $s \in \mathsf{Per}(\ell)$ and let $M = C(s(0), s(1), \ldots, s(\ell - 1))$ be the circulant matrix of size ℓ built on the period of s. Then $\operatorname{rank}_k(s) = \operatorname{rank} M$.*

Proof. By iteratively applying Lemma 2, every sequence $s(k^e n + j)$ with $e \geq 0$ and $0 \leq j \leq k^e - 1$ has period length ℓ, so we have $\operatorname{rank}_k(s), \operatorname{rank} M \leq \ell$. As a first easy case, if the sequences $s(n)_{n \geq 0}, s(n + 1)_{n \geq 0}, \ldots, s(n + \ell - 1)_{n \geq 0}$ are linearly independent, then $\operatorname{rank} M = \ell = \operatorname{rank}_k(s)$ by definition, which suffices. As a second case, assume that there is a smallest set of linearly independent sequences among them, say of size $d < \ell$. By definition, we have $\operatorname{rank} M = d$. Using Proposition 6, every sequence in

$$V = \{s' \in \mathsf{Per}(\ell) : s' \text{ satisfies the same linear recurrence as } s\}$$

can be written as a linear combination of the d linearly independent sequences. We now prove that every sequence in $\ker_k(s)$ belongs to V. By Lemma 6, such a sequence can be written as $s(cn + r)_{n \geq 0}$ with $0 \leq c, r \leq \ell - 1$ (and c is invertible in $\mathbb{Z}_\ell$). Since the sequences s and $s(cn + r)_{n \geq 0}$ both belong to $\mathsf{Per}(\ell)$, let f and g be the respective characteristic polynomials of their minimal recurrences. By Lemma 4 and the factorization of $x^\ell - 1$ into cyclotomic polynomials, there exists a set $S \subseteq \mathsf{Div}(\ell)$ such that $f(x) = \prod_{m \in S} \Phi_m(x)$. The roots of $g(x)$ are the cth powers of those of $f(x)$. Since k and ℓ are coprime, so are c and m, thus the map $x \mapsto x^c$ is a permutation of the roots of $\Phi_m(x)$. Therefore the roots of g are equal to those of f. So $s(cn + r)_{n \geq 0}$ belongs to V. Therefore $\operatorname{rank}_k(s) = d$, as desired.   □

Note that the previous argument does not apply to the case where k and ℓ are not coprime.

Example 8. In Example 6, the kernel sequence s with period $(1, 0)$ is not a linear combination of the other kernel sequences having respective periods $(0, 1, 1, 1)$, (1), and (0). However, s is a linear combination of the shifts of sequence s_0 with period $(0, 1, 1, 1)$ since, if we let s_1, s_2, and s_3 respectively be the first, second, and third shifts of s_0, we have $s = \frac{2}{3} s_0 - \frac{1}{3} s_1 + \frac{2}{3} s_2 - \frac{1}{3} s_3$.