Contents lists available at [ScienceDirect](https://www.elsevier.com/locate/dam)

# Discrete Applied Mathematics

journal homepage: [www.elsevier.com/locate/dam](http://www.elsevier.com/locate/dam)

# On extremal factors of de Bruijn-like graphs

Nicolás Álv[a](#page-0-0)rezª, Verónica Becher ʰ.\*, Martín Mere[b](#page-0-1) <sup>[c](#page-0-3)</sup>, Ivo Pajor <sup>[d](#page-0-4)</sup>, Carlos Miguel Soto<sup>[d](#page-0-4)</sup>

<span id="page-0-0"></span>a *ICC CONICET, Argentina*

<span id="page-0-1"></span><sup>b</sup> *Departamento de Computación, Facultad de Ciencias Exactas y Naturales & ICC, Universidad de Buenos Aires & CONICET, Argentina* <sup>c</sup> *Departamento de Matemática, Facultad de Ciencias Exactas y Naturales & IMAS, Universidad de Buenos Aires &*

<span id="page-0-3"></span>*CONICET, Argentina*

<span id="page-0-4"></span><sup>d</sup> *Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires & Argentina, Argentina*

#### a r t i c l e i n f o

*Article history:* Received 18 October 2023 Received in revised form 27 March 2024 Accepted 9 June 2024 Available online xxxx

*MSC:* 68R01 68R05 05C35

*Keywords:* de Bruijn graph Pure cycling register de Bruijn sequence Perfect necklaces

# A B S T R A C T

In 1972, Mykkeltveit confirmed Golomb's conjecture on the de Bruijn graphs: he proved that the pure cycling register rule yields the maximum number of vertex-disjoint cycles in de Bruijn graphs. We show that this result encompasses the tensor product of the de Bruijn graph for strings of length *n* with a simple cycle of size *k*, when *n* divides *k* or vice versa. Furthermore, we give counting formulas for an array of cycling register rules, which includes Golomb's well-studied linear register rules.

© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

# **1. Introduction and statement of results**

In [[7,](#page-11-0) Chapter VII Conjecture A] Golomb asked what is the maximal number of vertex-disjoint cycles in the de Bruijn graph of strings of length *n*. He conjectured that this is exactly the number of cycles attained using the pure cycling register rule to partition the de Bruijn graph. This conjecture was proved by Mykkeltveit [[15\]](#page-12-0), using Lempel's [\[14\]](#page-12-1) reformulation of the problem, which amounts to determining the minimum number of vertices which, if removed from the graph, will leave it with no cycles.

In this note we consider Golomb's conjecture for a variant of the de Bruijn graph known as the *astute graph*. These graphs are the tensor product of the de Bruijn graph for strings of length *n* and a simple cycle. The Hamiltonian cycles in the (*n*, *k*)-astute graph correspond to the so-called (*n*, *k*)-perfect necklaces introduced in [[1\]](#page-11-1). Recently, perfect necklaces received attention in the theory of uniform distribution modulo 1 [\[2](#page-11-2),[8,](#page-11-3)[9\]](#page-12-2). Nested perfect necklaces were used to construct infinite sequences of symbols in a given alphabet (the expansion of a real number in a given integer base) with the smallest known discrete discrepancy. The discrete discrepancy measures how far the sequence is from having the symbols and the

<span id="page-0-2"></span>Corresponding author.

*E-mail addresses:* [nico.alvarez@gmail.com](mailto:nico.alvarez@gmail.com) (N. Álvarez), [vbecher@dc.uba.ar](mailto:vbecher@dc.uba.ar) (V. Becher), [mmereb@gmail.com](mailto:mmereb@gmail.com) (M. Mereb), [pajorivo@gmail.com](mailto:pajorivo@gmail.com) (I. Pajor), [miguelsotocarlos@gmail.com](mailto:miguelsotocarlos@gmail.com) (C.M. Soto).

<https://doi.org/10.1016/j.dam.2024.06.010>

0166-218X/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.





blocks of symbols equidistributed. It is an open question, that dates back to Korobov in 1955, if there are sequences with a smaller discrete discrepancy.

Along the sequel, we assume a finite alphabet Γ . The elements of Γ *n* are referred to as strings of length *n*. The symbols in a string of length *n* are numbered from 0 to  $n-1$ . The notation *s*[*i*..*j*) for a string  $s = a_0a_1 \ldots a_{n-1}$  denotes the substring  $a_i a_{i+1} \ldots a_{i-1}$ .

**Definition 1** (*Astute graph*). Given *n* and *k* positive integers, the astute graph is defined by  $G_{n,k} = (V_{n,k}, E_{n,k})$ , where  $V_{n,k} = \Gamma^n \times \{0, \ldots, k-1\}$  and  $E_{n,k}$  is the set of all pairs

$$
((s, i), (t, j))
$$

such that  $s[1..n] = t[0..n - 1)$  and  $j = i + 1$  (mod *k*).

**Remark 1.** Notice that *Gn*,<sup>1</sup> is the *n*th de Bruijn graph. In this case, we identify the vertices *Vn*,<sup>1</sup> with Γ *n* . For a presentation of the de Bruijn sequences and the de Bruijn graph, including a historical account, see [\[4](#page-11-4)].

**Definition 2** (2-Factor). A 2-factor of  $G_{n,k}$  is a set of vertex-disjoint cycles (directed circuits) which, together, include all the vertices of  $G_{n,k}$ . The size of a 2-factor is defined as the number of cycles it has. A 2-factor which contains the maximum possible number of cycles is referred to as an *extremal factor*.

To construct 2-factors of de Bruijn and astute graphs, we consider *succession rules*. These are what Golomb in [[7](#page-11-0)] calls Shift Registers.

**Definition 3** (Succession rule). A succession rule is a bijection  $\sigma$  :  $\Gamma^n \to \Gamma^n$  such that for each string  $s = a_0a_1\cdots a_{n-1}$ ,  $\sigma(s) = a_1 a_2 \cdots a_{n-1} a_n$  for some  $a_n \in \Gamma$ .

**Remark 2.** The definition of succession rule implies that in the de Bruijn graph there exists an arc from *s* to σ(*s*), and in the astute graph there exists an arc from vertex (*s*, *i*) to ( $\sigma$ (*s*), *i*+1). This means that the succession rule  $\sigma$  can be thought to act on the vertices of the de Bruijn and astute graphs.

**Definition 4** (*Action of a succession rule on astute graphs*)**.** Given a succession rule σ and a positive integer *k*, we define an action  $A_k(\sigma) : V_{n,k} \to V_{n,k}$  such that  $A_k(\sigma)(s, i) = (\sigma(s), i + 1)$ .

Given a succession rule  $\sigma$  and a positive integer *k*, the subgroup of permutations  $\langle A_k(\sigma) \rangle$  acts on  $V_{n,k}$ . For any vertex  $v \in V_{n,k}$ , the arc  $(v, A_k(\sigma)(v))$  is in the graph  $G_{n,k}$ . This implies that the orbits of this action are simple cycles on the astute graph.

**Definition 5** (2-Factor generated by succession rule). We denote  $F_k(\sigma)$  as the 2-factor composed of all orbits produced by  $A_k(\sigma)$ .

We interpret the alphabet  $\Gamma$  as the quotient ring of integers modulo *b*, where  $b = |\Gamma|$ , therefore we can do linear arithmetic on its symbols.

**Definition 6** (Affine relation). A relation  $R \subseteq \Gamma^{n+1}$  is said to be affine if there exist  $c \in \Gamma$  and coefficients  $\lambda_i, 0 \leq i \leq n$ ,  $\lambda_i \in \Gamma$ , such that

$$
a_0a_1\cdots a_n\in R \iff c=\sum_{0\leq i\leq n}\lambda_i a_i.
$$

**Definition 7** (Affine succession rule). An affine succession rule is a succession rule  $\sigma:\Gamma^n\to\Gamma^n$  constructed from an affine relation R as follows. For each string  $a_0a_1...a_{n-1}$ ,  $\sigma(a_0a_1...a_{n-1})$  is the unique string  $a_1...a_n$  such that  $a_0a_1...a_n$  is in *R*.

**Remark 3.** For an affine relation *R* to give rise to an affine succession rule

- Each string must have at most one successor, which only happens if  $\lambda_n$  is invertible; and
- The rule has to be a bijection, which only happens if  $\lambda_0$  is invertible.

**Example 1** (*Pure Cycling Register*)**.** An example of an affine succession rule is the one given by string rotation. For any string  $s = a_0 a_1 \cdots a_{n-1}$  we define

 $r_n(s) = a_1 \cdots a_{n-1} a_0$ 

so,  $a_n = a_0$ . For  $k = 1$ , the de Bruijn case,  $F_1(r_n)$  is the set of necklaces of length *n*. This is because necklaces of length *n* are by definition the set of equivalence classes of strings in  $\Gamma^n$  under rotations, and any two strings that are a rotation

of each other can be transformed into each other by successive application of  $\sigma$ , and therefore each cycle in the 2-factor  $F_1(r_n)$  is a set of strings under the same rotation equivalence class.

This succession rule is affine, because it is given by the affine relation

 $a_0a_1 \cdots a_n \in R \iff a_n - a_0 = 0$ 

**Example 2** (*Incremented Cycling Register*)**.** Another example of an affine succession rule is *incremented rotation*. For any string  $s = a_0 a_1 \cdots a_{n-1}$  we define

 $\iota_n(s) = a_1 \cdots a_{n-1}(a_0 + 1),$ 

so,  $a_n = a_0 + 1$ . An advantage of this particular succession rule *i* is that each cycle has an equal quantity of each symbol in  $\Gamma$ . This has applications in the construction of de Bruijn sequences with small discrepancy, see [\[10\]](#page-12-3) in contrast to [\[6\]](#page-11-5). Again, this is affine because it is given by the relation

 $a_0a_1 \cdots a_n \in R \iff a_n - a_0 = 1$ 

**Example 3** (*Xor Cycling Register*). The third example we consider is restricted to the special case  $|\Gamma| = 2$ , where the ring addition operation is the *xor*. In this ring, we define the succession rule for  $s = a_0a_1 \cdots a_{n-1}$  as

<span id="page-2-0"></span>
$$
x_n(s) = a_1 \cdots a_{n-1} a_n
$$

where  $a_n = a_1 + a_2 + a_3 + \cdots + a_{n-1}$ .

We now state the main result of this note:

**Theorem 1.** *Let n and k be positive integers such that k divides n or n divides k. The 2-factor Fk*(*rn*) *produced by the pure cycling register rule r<sup>n</sup> is extremal.*

When *k* does not divide *n* [Theorem](#page-2-0) [1](#page-2-0) is not necessarily true. For the case  $k = 2$ ,  $n = 3$  and  $\Gamma = \{0, 1\}$ , the successor rule *r*<sup>3</sup> produces a 2-factor of size 4 as shown in [Fig.](#page-3-0) [1\(a\),](#page-3-0) while the extremal factors have size 6. An example of such extremal factor is shown in [Fig.](#page-3-1)  $1(b)$  $1(b)$ . For such cases, where the hypothesis of [Theorem](#page-2-0) 1 do not hold, it remains an open problem to characterize the values of *k* and *n* where the conclusion does hold.

The second result in this note is a closed formula for the size of the 2-factors generated by affine rules. We use (*a* : *b*) for the greatest common divisor of the integers *a* and *b*. We write Γ [*X*] to denote the ring of polynomials in *X* with coefficients in  $\Gamma$ . We denote the ideal generated by  $P$  and  $Q$  in  $\Gamma[X]$  by  $(P, Q)$ .

<span id="page-2-1"></span>**Theorem 2.** *Let n and k be positive integers, and R be an affine rule given by*

$$
a_0a_1\cdots a_n\in R\iff c=\sum_{0\leq i\leq n}\lambda_ia_i,
$$

*for some coefficients* (λ*i*)*<sup>i</sup> and a constant term c. Then the size of the 2-factor associated succession rule* σ *is given by*

$$
|F_k(\sigma)| = \frac{k(s:\omega)}{s\omega} \sum_{(s:\omega)|d|\omega} \varphi(\omega/d) \left| \frac{\Gamma[X]}{(A, X^d - 1)} \right|
$$

*where*

- $\Lambda = \sum_{i=0}^n \lambda_i X^{n-i}$  is the characteristic polynomial of R,
- $\omega$  is any multiple of the multiplicative order of X in the group of units of the quotient ring of polynomials modulo  $\Lambda$ ,
- $\bullet$   $\varphi$  is Euler's totient function, and s is the length of the smallest cycle in the 2-factor. Equivalently, s can be defined as the *smallest multiple of k such that*

$$
c(1 + X + \dots + X^{s-1}) \in (\Lambda, X^s - 1).
$$

The next corollaries give the number of cycles in the 2-factors determined by two specific succession rules.

<span id="page-2-2"></span>**Corollary 1** (*2-Factor of Gn*,*<sup>k</sup> from Pure Cycling Register*)**.**

$$
|F_k(r_n)| = \frac{(n : k)}{n} \cdot \sum_{(n : k) \text{ } |d| \text{ } |n} \varphi(n/d)|\Gamma|^d.
$$

<span id="page-2-3"></span>**Remark 4.** When  $k = 1$  and  $| \Gamma | = 2$ ,  $|F_1(r_n)|$  is the number of binary irreducible polynomials whose degree divides *n*, see [[13](#page-12-4)].

<span id="page-3-0"></span>

(a) 2-Factor  $F_2(r_3)$  produced by Pure Cycling Register rule for  $\Gamma = \{0,1\}$ . The arcs of the cycles in the 2-factor are shown in magenta. There are 4 cycles in this 2-factor.





<span id="page-3-1"></span>(b) Extremal 2-factor of  $G_{3,2}$  for the alphabet  $\Gamma = \{0,1\}$ . The arcs of the cycles in the 2-factor are shown in blue. There are 6 cycles in this 2-factor, making it extremal.

Fig. 1. Pure Cycling Register induced 2-factors may not be extremal in astute graphs.

**Corollary 2** (*2-Factor of Gn*,*<sup>k</sup> from Incremented Cycling Register*)**.**

$$
|F_k(\iota_n)| = \frac{k(\operatorname{lcm}(k, bd_b(n)) : n)}{\operatorname{lcm}(k, bd_b(n))n} \sum_{(\operatorname{lcm}(k, bd_b(n)) : n) | d|n} \varphi(n/d) b^d
$$

*where b* =  $|\Gamma|$  *and d<sub>b</sub>*(*n*) *is the smallest divisor of n such that*  $n/d<sub>b</sub>(n)$  *<i>is coprime with b.* 

**Remark 5.** When  $k = 1$  and  $| \Gamma | = 2$ ,  $| F_1(t_n) |$  is the number of distinct output sequences from binary *n*-stage shift register which feeds back the complement of the last stage, see [[12\]](#page-12-5).

<span id="page-3-3"></span>**Corollary 3** (*2-Factor of Gn*,*<sup>k</sup> from Xor Cycling Register*)**.**

$$
|F_k(x_n)| = \frac{k}{2(n+1)} \sum_{d|n+1} \varphi(2d) 2^{(n+1)/d}.
$$

**Remark 6.** When  $k = 1$  and  $| \Gamma | = 2$ ,  $|F_1(x_n)|$  is the number of output sequences from  $(n - 1)$ -stage shift register which feeds back the mod 2 sum of the contents of the register, see [\[11](#page-12-6)].

# **2. Proof of [Theorem](#page-2-0) [1](#page-2-0)**

When *n* divides *k*, the result follows from the fact that the pure cycle register produces a 2-factor where each cycle has length exactly *k*, which is also the smallest possible length of a cycle in the graph *Gn*,*k*. Let us then consider the case that *k* divides *n*.

We use a basic tool from finite Fourier analysis [[3](#page-11-6)].

**Definition 8** (Discrete Fourier Transform)**.** Let  $\mu=e^{2\pi i/n}$  be a primitive root of unity of order *n*. Let us define C :  $\Gamma^n\to\mathbb{C}$ as

$$
C(a_0 \ldots a_{n-1}) = \sum_{i=0}^{n-1} a_i \mu^i.
$$

Here, we identify  $a_i \in \Gamma$  with the unique integer in the range [0, *b*) in its equivalence class.

<span id="page-3-2"></span>**Lemma 1.**  $C(r_n(s)) = \mu^{-1}C(s)$ .

**Proof.** Let  $s = a_0a_1 \cdots a_{n-1}$ . Then  $r(s) = a_1a_2 \cdots a_na_0$ . Then, we have:

$$
C(r(s)) = a_0 \mu^{n-1} + \sum_{i=1}^{n-1} a_i \mu^{i-1}.
$$

Since  $\mu^{n-1} = \mu^{-1}$ , we obtain

$$
C(r(s)) = \sum_{i=0}^{n-1} a_i \mu^{i-1} = \mu^{-1} \sum_{i=0}^{n-1} a_i \mu^i = \mu^{-1} C(s). \quad \Box
$$

<span id="page-4-0"></span>**Lemma 2.** Let  $(s_0, m_0), \ldots (s_{t-1}, m_{t-1})$  be the vertices of any cycle in the astute graph  $G_{n,k}$ . Then  $\sum C(s_i) = 0$ .

**Proof.** We have that

$$
\sum_{i=0}^{t-1} C(s_i) = \sum_{i=0}^{t-1} \sum_{j=0}^{n-1} (s_i)_j \mu^j.
$$

Since the strings *s<sup>i</sup>* form a cycle in the de Bruijn graph, we have that (*si*)*<sup>j</sup>* = (*si*+1)*j*−1, where the indices of *s* are taken modulo *t*. Then,  $(s_i)$  depends only of the sum  $i + j$  modulo *t*. Let  $w_{i+i} = (s_i)$  with the indices of w taken modulo *t*. We can rewrite the expression as

$$
\sum_{i=0}^{t-1} C(s_i) = \sum_{i=0}^{t-1} \sum_{j=0}^{n-1} w_{i+j} \mu^j.
$$

Swapping the inner and outer summations, we get

$$
\sum_{i=0}^{t-1} C(s_i) = \sum_{j=0}^{n-1} \sum_{i=0}^{t-1} w_{i+j} \mu^j = \sum_{j=0}^{n-1} \mu^j \sum_{i=0}^{t-1} w_{i+j}
$$

Notice that the sum  $\sum_{i=0}^{t-1} w_{i+j}$ ∑ Notice that the sum  $\sum_{i=0}^{t-1} w_{i+j}$  is constant on *j*, since  $i + j$  always takes every value modulo *t* exactly once. Therefore  $\sum_{i=0}^{t-1} C(s_i)$  is a multiple of  $\sum_{j=0}^{n-1} \mu^j$ , which is 0 because it is the sum

.

<span id="page-4-1"></span>**Lemma 3.** *Let s and t be two strings that are connected by an arc in the de Bruijn graph Gn*,1*. Then the complex numbers C*(*s*) and *C*( $r_n^{-1}(t)$ ) have the same imaginary part. Furthermore, they are equal exactly when  $s = r_n^{-1}(t)$ .

**Proof.** Since (*s*, *t*) is an arc in the de Bruijn graph, we can write

 $s = a_0 a_1 \cdots a_{n-1}$  $t = a_1 a_2 \cdots a_n$ .

Then,

 $r_n^{-1}(t) = a_n a_1 \cdots a_{n-1}.$ 

Expanding the definition of  $C(r^{-1}(t))$  we get

$$
C(r_n^{-1}(t)) = a_n \mu^0 + a_1 \mu^1 + a_2 \mu^2 + \cdots + a_{n-1} \mu^{n-1}.
$$

And for *C*(*s*) we get

$$
C(s) = a_0 \mu^0 + a_1 \mu^1 + a_2 \mu^2 + \cdots + a_{n-1} \mu^{n-1}.
$$

So, we have  $C(s) - C(r_n^{-1}(t)) = a_0 - a_n$ , which is a real number that is zero only when  $a_0 = a_n$ , which is precisely when  $s = r_n^{-1}(t)$ . □

<span id="page-4-2"></span>**Lemma 4.** *There exists a mapping from cycles in the 2-factor associated with the pure cycling register rule to vertices in the underlying graph such that every simple cycle in the graph contains at least one vertex that is associated with a pure cycling register cycle.*

**Proof.** Let  $(u_0, w_0), (u_1, w_1), \ldots, (u_{t-1}, w_{t-1})$  be any cycle in  $F_k(r_n)$ . There are two possibilities.

One possibility is that the transform *C*(*ui*) is real for all *u<sup>i</sup>* . In this case we take any arbitrary vertex in the 2-factor as the distinguished vertex.

The other possibility is that there exists some string  $u_i$  such that  $C(u_i)$  is not real. Let  $z = C(u_0)$ . Due to [Lemma](#page-3-2) [1,](#page-3-2) we have that  $C(u_i) = z\mu^{-i}$ . Since the length of any PCR cycle is a divisor of lcm(*n*, *k*) = *n* and the function  $i \mapsto z\mu^{-i}$  is



<span id="page-5-0"></span>**Fig. 2.** Transforms of the strings on the PCR cycle generated by  $s_0 = 123351$ . The large red point is the distinguished vertex for this PCR cycle.

periodic with a period equal to the order of  $\mu$  (which is *n*), the size of the 2-factor must be  $t = n$ , and the transforms of its strings form a regular *n*-sided polygon on the complex plane.

The distinguished vertex of the 2-factor will be the unique vertex  $(u_i, m_i)$  such that

 $\mathfrak{Im}(C(u_i)) < 0$  but  $\mathfrak{Im}(C(u_{i-1})) \geq 0$ ,

as exemplified in [Fig.](#page-5-0) [2](#page-5-0).

Now we have to prove that every cycle in the astute graph *Gn*,*<sup>k</sup>* contains at least one distinguished vertex. Let

 $(s_0, m_0), (s_1, m_1), \ldots, (s_{t-1}, m_{t-1})$ 

be any such cycle. We consider three cases:

*First case: There exists some string*  $s_i$  *<i>such that*  $C(s_i)$  *is not real.* The sum of  $C(s_i)$  over all *i* must be zero due to [Lemma](#page-4-0) [2.](#page-4-0) So, if the transforms are not always real, there must be a string where the imaginary part of the transform is positive and another where the imaginary part is negative. In particular, let  $s_i$  be any string such that  $\mathfrak{Im}(C(s_i)) < 0$  but  $\mathfrak{Im}(C(s_{i-1})) \geq 0$ . Since  $s_{i-1}$  and  $s_i$  are connected by an arc, [Lemma](#page-4-1) [3](#page-4-1) implies that  $C(s_{i-1})$  and  $C(r_n^{-1}(s_i))$  differ by a real number. And since  $\mathfrak{Im}(C(s_{i-1}))$  ≥ 0, then  $\mathfrak{Im}(C(r_n^{-1}(s_i)))$  ≥ 0 as well. This means that in the PCR cycle

$$
(s_i, m_i), (r_n(s_i), m_i + 1), (r_n^2(s_i), m_i + 2), \ldots, (r_n^{n-1}(s_i), m_i + n - 1) = (r_n^{-1}(s_i), m_i - 1)
$$

the distinguished vertex will be  $(s_i, m_i)$ , which is a vertex in the original cycle

 $(s_0, m_0), (s_1, m_1), \ldots, (s_{t-1}, m_{t-1}).$ 

Second case: The transform C(s<sub>i</sub>) is 0 for all i. Due to [Lemma](#page-3-2) [3,](#page-4-1)  $s_i = r_n^{-1}(s_{i+1})$  if and only if  $C(s_i) = C(r_n^{-1}(s_{i+1}))$ . By Lemma [1,](#page-3-2)  $C(r_n^{-1}(s_{i+1})) = \mu C(s_{i+1}) = \mu \cdot 0 = 0$ , so  $s_i = r_n^{-1}(s_{i+1})$  for all *i*. Hence, the cycle is actually a PCR cycle, so it must have at least one distinguished vertex.

Third case: The transform  $C(s_i)$  is always real, but not always 0. Let  $C(s_i) \in \mathbb{R} - \{0\}$ . Due to [Lemma](#page-4-1) [3](#page-4-1),  $C(r_n^{-1}(s_i)) = \mu C(s_i)$ differs from  $C(s_{i-1})$  by a real number. If  $n > 2$ ,  $\mu$  is a non-real complex number, and therefore  $\mu C(s_i)$  also has a non-zero imaginary component, which is equal to that of *C*(*si*−1). This is a contradiction, because we assumed *C*(*sj*) was real for all *j*. So we only have to analyze the cases  $n = 1$  $n = 1$  and  $n = 2$ . In both cases [Theorem](#page-2-0) 1 is implied by the fact that the PCR cycle of a vertex is the smallest possible cycle it belongs to in the astute graph.  $\Box$ 

**Proof** (*Proof of [Theorem](#page-2-0)* [1](#page-2-0))**.** Due to [Lemma](#page-4-2) [4,](#page-4-2) there exists a mapping from cycles in the 2-factor generated by the PCR to vertices in the underlying graph such that every simple cycle (not necessarily in the 2-factor) contains a distinguished vertex. To prove that the 2-factor produced by the Pure Cycling Register is extremal, we choose for each cycle in the 2-factor  $F_k(r_n)$  a distinguished vertex, and then we prove that any cycle in any 2-factor has at least one distinguished vertex, therefore the size of any 2-factor is at most the number of distinguished vertices.

We showed that every cycle in graph *Gn*,*<sup>k</sup>* contains at least one distinguished vertex of a cycle in the 2-factor determined by the Pure Cycling Register rule, so [Theorem](#page-2-0) [1](#page-2-0) is proved.

# **3. Proof of [Theorem](#page-2-1) [2](#page-2-1)**

#### *3.1. Burnside's lemma*

Our main tool for counting the number of cycles in a succession rule is the classical Burnside's Lemma [[5\]](#page-11-7). It states that for any finite group *G* acting on a set *S*, the following identity holds:

$$
|S/G| = \frac{1}{|G|} \sum_{g \in G} |S^g|,
$$

where *S*/*G* is the set of orbits of *S* under the action of *G*, and *S g* is the subset of *S* fixed by the action of *g*.

When considering succession rules, the set *S* is the set of vertices of an astute graph  $\overline{S} = V_{n,k}$  and  $\overline{G} = \langle A_k(\sigma) \rangle$  is the group generated by the action associated with a succession rule  $\sigma$ . In that case, the set of orbits  $S/G$  coincides with the 2-factor  $F_k(\sigma)$ . We thus have the following identity:

<span id="page-6-0"></span>
$$
|F_k(\sigma)| = \frac{1}{\omega} \sum_{i=0}^{\omega-1} |\operatorname{fix}(A_k(\sigma)^i)|,\tag{1}
$$

where  $\omega$  is the order of  $A_k(\sigma)$  and fix(f) is the set of fixed points of the function f.

Notice that the function  $i\mapsto |\operatorname{fix}(A_k(\sigma)^i)|$  is defined over all the integers, and it is periodic because  $A_k(\sigma)^i=A_k(\sigma)^{i+\omega}$ for all  $i \in \mathbb{Z}$ . Therefore, Eq. ([1\)](#page-6-0) asserts that the size of  $F_k(\sigma)$  is the average of the function  $i \mapsto |fix(A_k(\sigma)^i)|$  over one cycle. This average does not depend on which cycle is picked, because they all coincide with the average of the function  $i \mapsto |\operatorname{fix}(A_k(\sigma)^i)|$  in the range [0, *t*] when *t* tends to infinity. This gives rise to the following lemma.

<span id="page-6-1"></span>**Lemma 5.** Let k be a positive integer and  $\sigma$  :  $\Gamma^n \to \Gamma^n$  a succession rule. Let  $\omega$  be any positive integer such that

$$
|\operatorname{fix}(A_k(\sigma)^i)| = |\operatorname{fix}(A_k(\sigma)^{i+\omega})|, \quad \forall i \in \mathbb{Z}.
$$

*Then,*

$$
|F_k(\sigma)| = \frac{1}{\omega} \sum_{i=0}^{\omega-1} |\operatorname{fix}(A_k(\sigma)^i)|.
$$

For the de Bruijn case, since we identify  $V_{n,1}$  with  $\Gamma^n$ , we have that  $A_1(\sigma)=\sigma$ , and therefore  $|\operatorname{fix}(A_1(\sigma)^i)|=|\operatorname{fix}(\sigma^i)|.$ Let us analyze  $|\operatorname{fix}(A_k(\sigma)^i)|$  in the astute case.

Let  $(s, j) \in V_{n,k}$  be any vertex fixed by  $A_k(\sigma)^i$ . We have that

$$
(s, j) = A_k(\sigma)^{i}(s, j) = (\sigma^{i}(s), j + i).
$$

So  $(s,j)$  is a fixed point of  $A_k(\sigma)^i$  if and only if  $s$  is a fixed point of  $\sigma^i$  and  $k|i$ , which gives us the following identity:

$$
|\operatorname{fix}(A_k(\sigma)^i)| = k \mathbb{1}_{k|i} |\operatorname{fix}(\sigma^i)|,
$$

where  $\mathbb{1}_p$  is 1 when *p* is true, and 0 otherwise. Rewriting [Lemma](#page-6-1) [5](#page-6-1) with this identity we obtain the following.

<span id="page-6-3"></span>**Lemma 6.** Let k be a positive integer and  $\sigma$  :  $\Gamma^n \to \Gamma^n$  a succession rule. Let  $\omega$  be any positive integer such that

$$
k\mathbb{1}_{k|i}|\operatorname{fix}(\sigma^i)|=k\mathbb{1}_{k|i+\omega}|\operatorname{fix}(\sigma^{i+\omega})|,\quad\forall i\in\mathbb{Z}.
$$

*Then,*

$$
|F_k(\sigma)| = \frac{k}{\omega} \sum_{i=0}^{\omega-1} \mathbb{1}_{k|i} |\operatorname{fix}(\sigma^i)|.
$$

#### *3.2. GCDs of certain families of polynomials*

Here we deal with the families of polynomials  $X^n - 1$  and  $U_n = \frac{X^n - 1}{X - 1}$ . Although in an arbitrary ring not every pair of polynomials has a GCD, we show that any pair of polynomials in these classes has a GCD, and we compute it explicitly.

**Lemma 7.** 
$$
(U_n : U_m) = U_{(n:m)}
$$
.

**Proof.** Assuming that  $n < m$ , we have the following identity:

<span id="page-6-2"></span>
$$
U_m-X^{m-n}U_n=U_{m-n}.
$$

Then,  $(U_n: U_m) = (U_n: U_{m-n})$ . It is also true that  $(n:m) = (n, m-n)$ . This means that applying the euclidean algorithm over the polynomials  $U_n$  and  $U_m$  mirrors the steps taken during the application of the algorithm to the pair of integers *n* and *m*, which implies that the algorithm always terminates, and converges to  $U_{(n,m)}$   $\Box$ 

**Lemma 8.**  $(X^n - 1 : X^m - 1) = X^{(n:m)} - 1$ .

**Proof.** It follows directly from [Lemma](#page-6-2) [7](#page-6-2) by multiplying both sides by  $X - 1$ .  $\Box$ 

#### <span id="page-7-2"></span>**Lemma 9.** *If* Γ *is a field, then*

$$
(U_n: X^m - 1) = \begin{cases} X^{(n:m)} - 1 & \text{if } n/(n:m) \equiv 0 \mod |I| \\ U_{(n:m)} & \text{if } n/(n:m) \not\equiv 0 \mod |I|. \end{cases}
$$

**Proof.** Since  $\Gamma[X]$  is a principal ideal and  $(X^n - 1, X^m - 1) \subseteq (U_n, X^m - 1) \subseteq (U_n, U_m)$  we only need to decide whether  $(U_n, X^m-1)$  is generated by  $\hat{X}^{(n:m)}-1$  or by  $U_{(n:m)}$ . The former is true if and only if  $X^{(n:m)}-1 \mid U_n$ . Since

$$
U_n = \frac{X^{(n:m)} - 1}{X - 1} U_{n/(n:m)}(X^{(n:m)})
$$

we get  $X^{(n:m)}-1\mid U_n$  if and only if 1 is a root of  $U_{n/(n:m)}(X^{(n:m)})$ . The latter is equivalent to  $n/(n:m)=U_{n/(n:m)}(1)=0$ in  $\Gamma$ .  $\square$ 

#### *3.3. Affine succession rules*

In this section we compute the size of the set fix( $\sigma^k$ )  $\subseteq \Gamma^n$  for an affine succession rule  $\sigma$  :  $\Gamma^n \to \Gamma^n$  given by an affine relation *R*. For any  $s \in \mathrm{fix}(\sigma^k)$  we define its associated string  $w \in \varGamma^k$  by

$$
w_i=(\sigma^i(s))_0
$$

It is clear that w uniquely determines *s*. Indeed, due to the definition of succession rule,

$$
s_i = (\sigma^i(s))_0 = w_{(i \mod k)},
$$

this identity holds only because  $s\in\mathrm{fix}(\sigma^k)$ . Hence, the strings  $\sigma^i(s)$  repeat modulo *k*. Equivalently, if we write  $w^*=$  $www \cdots$  as the infinite concatenation of w with itself, the above claim states that  $s = w^*[0..n]$ . A similar argument shows that  $\sigma^i(s) = w^*[i..n+i)$ . Since  $\sigma$  is an affine succession rule we have that every  $n+1$  consecutive symbols in  $w^*$ satisfy the affine relation *R*. This condition can be encoded in the following polynomial series equation:

<span id="page-7-0"></span>
$$
\exists p : \deg(p) \le n \text{ and } \frac{w}{1 - X^k} \Lambda = p + \frac{c}{1 - X},\tag{2}
$$

where  $\Lambda$  is the characteristic polynomial of  $R$ ,  $c$  is its constant term, and we identify the string  $w$  with its generating polynomial. Under this identification,  $\frac{w}{1-X^k}$  is the generating function of the infinite string  $w^*$ .

The coefficient of degree *i* in  $\frac{w}{1-X^k}$   $\Lambda$  is the linear combination

$$
\sum_{j=0}^n \lambda_j w_{i-n+j}^*.
$$

Thus, when  $i \leq n$  we cannot guarantee that the result of this linear combination is  $c$ , and we need to introduce an ''error'' polynomial *p* of degree at most *n*.

If we multiply both sides of Eq. ([2\)](#page-7-0) by  $1 - X^k$  we get the equivalent expression:

<span id="page-7-1"></span>
$$
\exists p : \deg(p) \le n \text{ and } w \Lambda = p(1 - X^k) + cU_k
$$

which is, by definition, the same as

$$
w\Lambda \equiv cU_k \mod 1 - X^k. \tag{3}
$$

We claim that *any* string  $w \in \Gamma^k$  that satisfies Eq. ([3](#page-7-1)) is the associated string of some other string  $s \in \mathrm{fix}(\sigma^k)$ . Indeed, take any w that satisfies [\(3\)](#page-7-1). Since ([3\)](#page-7-1) is equivalent to ([2](#page-7-0)), every substring of length  $n + 1$  of  $w^*$  will satisfy the relation *R*. Then, if we take  $s = w^*[0..n)$  then  $\sigma^i(s) = w^*[i..n + i)$ , which will be cyclic modulo *k*, because  $w^*$  is cyclic modulo *k*.

The established bijection implies that  $|\operatorname{fix}(\sigma^k)|$  is the number of solutions to Eq. [\(3](#page-7-1)). Since  $w$  is always of length  $k$ , we can assume w is a polynomial in the quotient ring  $\Gamma[X]/(1-X^k)$  and the number of solutions of the equation will not change.

For an arbitrary polynomial  $P \in \Gamma[X]/(X^k-1)$ , the system of equations

$$
Aw \equiv P \qquad \text{mod } X^k - 1
$$

has a solution if and only if

 $u \cdot A - P = v \cdot (X^k - 1)$ 

for some polynomials *u*, v. This is equivalent to the condition

$$
P\in (A, X^k-1).
$$

 $\Lambda$ *w* 

Furthermore, if the system does have a solution, it has the same number of solutions as the associated linear system

$$
\equiv 0 \qquad \text{mod } X^k - 1.
$$

Since there are |Γ [*X*]/(*X <sup>k</sup>* − 1)| possibilities for w, and each possibility for *P* gets an equal number of solutions, each *P* gets exactly

$$
\frac{|\Gamma[X]/(X^k-1)|}{|(A, X^k-1)/(X^k-1)|} = |\Gamma[X]/(A, X^k-1)|
$$

solutions to the linear system, if there is at least one. This implies that

$$
|\operatorname{fix}(\sigma^k)| = \left|\frac{\Gamma[X]}{(A, X^k - 1)}\right| \mathbb{1}_{cU_k \in (A, X^k - 1)}.
$$

Let us further analyze the condition  $cU_k \in (\varLambda, X^k-1)$ . Let

$$
S = \{k \in \mathbb{N} : cU_k \in (\Lambda, X^k - 1)\}.
$$

We prove that *S* the set of all multiples of its least element  $\ell_{\sigma}$ . Let *d* be any multiple of  $\ell_{\sigma}$ . Since  $\ell_{\sigma} \in S$  we have that:

$$
cU_{\ell_{\sigma}}\in (\Lambda,X^{\ell_{\sigma}}-1)=(\Lambda,(X-1)U_{\ell_{\sigma}})
$$

The condition  $\ell_{\sigma}$  |*d* implies  $U_{\ell_{\sigma}}$  |*U<sub>d</sub>* and therefore it also holds that

$$
cU_d \in \frac{U_d}{U_{\ell_{\sigma}}}(\Lambda, (X-1)U_{\ell_{\sigma}}) \subseteq \left(\Lambda, \frac{U_d}{U_{\ell_{\sigma}}}(X-1)U_{\ell_{\sigma}}\right) = (\Lambda, (X-1)U_d) = (\Lambda, X^d - 1).
$$

Hence,  $d \in S$ .

Now take any 
$$
d \in S
$$
, and let  $g = \text{gcd}(d, \ell_{\sigma})$ . Then  $U_g \in (U_{\ell_{\sigma}}, U_d)$  due to Lemma 7. We also have that

$$
cU_{\ell_{\sigma}} \in (\Lambda, X^{\ell_{\sigma}} - 1) \quad \text{and} \quad cU_d \in (\Lambda, X^d - 1).
$$

Therefore,

$$
cU_g \in (A, X^{\ell_\sigma} - 1, X^d - 1) = (A, X^g - 1).
$$

Thus,  $g \in S$ . Since  $g|\ell_{\sigma}$  and  $\ell_{\sigma}$  is the least element in *S*, then  $g = \ell_{\sigma}$  and, as a result, *d* is a multiple of  $\ell_{\sigma}$ .

Notice that  $\ell_{\sigma}$  is the smallest-length cycle in the associated succession rule  $\sigma$ , since it is the first integer for which fix( $\sigma^{\ell_\sigma}$ ) is nonempty. Observe that when  $c=0$ , the zero string always has cycle length one, so  $\ell_\sigma=1$ .

We rewrite Eq. as

$$
\operatorname{fix}(\sigma^k) = \mathbb{1}_{\ell_{\sigma}|k} \cdot \left| \frac{\Gamma[X]}{(A, X^k - 1)} \right|.
$$

Now let us analyze  $\vert$ Γ [*X*]  $\frac{\Gamma[X]}{(\lambda X)^k-1}$ . Since the first coefficient of  $\Lambda$  is invertible, the polynomial *X* is invertible modulo  $\Lambda$ , so  $\frac{\Lambda}{\Lambda}$  is  $\frac{\$ there exists some  $\omega$  such that  $X^{\omega} \equiv 1 \mod \Lambda$ . Equivalently,  $\Lambda | X^{\omega} - 1$ .

Let *k* be any positive integer and  $g = (\omega : k)$ . We know that since  $(k : \omega) | k, X^{(k:\omega)} - 1 | X^k - 1$ . Therefore,

$$
(\Lambda, X^k - 1) \subseteq (\Lambda, X^{(k:\omega)} - 1).
$$

And also,

$$
(X^{(k:\omega)} - 1) = (X^k - 1, X^{\omega} - 1) \subseteq (A, X^k - 1).
$$

Consequently, the ideals (A,  $X^k-1$ ) and (A,  $X^{(k:\omega)}-1$ ) coincide. When we replace this into the formula for fix( $\sigma^k$ ), we get the following.

### <span id="page-8-0"></span>**Lemma 10.**

$$
fix(\sigma^k) = \mathbb{1}_{\ell_{\sigma}|k} \cdot \left| \frac{\Gamma[X]}{(A, X^{(k:\omega)} - 1)} \right|.
$$

# *3.4. Burnside's lemma for affine necklaces*

**Proof** (*Proof of [Theorem](#page-2-1)* [2\)](#page-2-1)**.** Let *k* be a positive integer and let σ be an affine succession rule with its characteristic polynomial  $\Lambda$ . Due to [Lemma](#page-6-3) [6,](#page-6-3) we have that

<span id="page-8-1"></span>
$$
|F_k(\sigma)| = \frac{k}{M} \sum_{i=0}^{M-1} \mathbb{1}_{k|i} |\operatorname{fix}(\sigma^i)|,\tag{4}
$$

for any  $M$  that is a cycle of the function  $i\mapsto 1_{k[i]}$  fix( $\sigma^i$ )|. Due to [Lemma](#page-8-0) [10](#page-8-0), we know that if  $\omega$  is the order of  $X$  modulo  $\Lambda$ , then

$$
fix(\sigma^k) = \mathbb{1}_{\ell_{\sigma} | k} \cdot \left| \frac{\Gamma[X]}{(A, X^{(k:\omega)} - 1)} \right|
$$

Therefore, *M* needs to be a cycle of

$$
i \mapsto \mathbb{1}_{k|i} \cdot \mathbb{1}_{\ell_{\sigma}|i} \cdot \left| \frac{\Gamma[X]}{(A, X^{(i\omega)} - 1)} \right|.
$$

To be a cycle of a product of functions, it suffices to be a multiple of the cycle length of each 2-factor. So, we have that one possible value of *M* is lcm( $k, \ell_{\sigma}, \omega$ ). Now rewriting Eq. [\(4](#page-8-1)) we get

$$
|F_k(\sigma)| = \frac{k}{M} \sum_{i=0}^{M-1} \mathbb{1}_{k|i} \cdot \mathbb{1}_{\ell_{\sigma}|k} \cdot \left| \frac{\Gamma[X]}{(\Lambda, X^{(i:\omega)} - 1)} \right|
$$
  
= 
$$
\frac{k}{M} \sum_{i=0}^{M-1} \mathbb{1}_{\text{lcm}(i,\ell_{\sigma})|i} \cdot \left| \frac{\Gamma[X]}{(\Lambda, X^{(i:\omega)} - 1)} \right|.
$$

.

Recall that each summand in the Burnside equation corresponds to fix( $A_k(\sigma)^i$ ), and the size of the ideal vector space  $\frac{\Gamma[X]}{\Gamma(A,X^{(i;\omega)}}$  $\left[\left(A, X^{(l, \omega)}-1\right)\right]$  smallest cycle in the 2-factor  $F_k(\sigma)$ . Let *S* be that cycle length. Rewriting the equation above we obtain the  $\frac{\Gamma[X]}{(\Lambda,X^{(i\omega)}-1)}$  is always positive. Hence, fix( $A_k(\sigma)^i$ ) is zero if and only if  $1_{\text{lcm}(i,\ell_{\sigma})|i}$  is zero. We conclude that  $\text{lcm}(i,\ell_{\sigma})$  is the  $\text{cm}$  is the smallest quality the smallest quality the smallest following:

$$
|F_k(\sigma)| = \frac{k}{M} \sum_{i=0}^{M-1} \mathbb{1}_{S|i} \cdot \left| \frac{\Gamma[X]}{(A, X^{(i:\omega)} - 1)} \right|
$$
  
= 
$$
\frac{k}{M} \sum_{i=0}^{M/S-1} \left| \frac{\Gamma[X]}{(A, X^{(i:S:\omega)} - 1)} \right|
$$
  
= 
$$
\frac{k}{M} \sum_{i=0}^{M/S-1} \left| \frac{\Gamma[X]}{(A, X^{(i:\omega/(S:\omega)) \cdot (S:\omega)} - 1)} \right|
$$
  
= 
$$
\frac{k}{S\omega/(S:\omega)} \sum_{i=0}^{\omega/(S:\omega)-1} \left| \frac{\Gamma[X]}{(A, X^{(i:\omega/(S:\omega)) \cdot (S:\omega)} - 1)} \right|
$$

The second equality uses that *S*|*M*, and the last equality uses that since  $M = \text{lcm}(S, \omega)$ , we have  $\omega/(S : \omega) = M/S$ . Since gcd(*i*,  $\omega/(S:\omega)$ ) iterates over all divisors of *d*, we can express that sum as follows:

.

$$
|F_k(\sigma)| = \frac{k(S:\omega)}{S\omega} \sum_{d| \frac{\omega}{(S:\omega)}} \varphi\left(\frac{\omega}{d(S:\omega)}\right) \left|\frac{\Gamma[X]}{(A, X^{d \cdot (S:\omega)} - 1)}\right|.
$$

This can be rewritten as

$$
|F_k(\sigma)| = \frac{k(S:\omega)}{S\omega} \sum_{(S:\omega)|d|\omega} \varphi(\omega/d) \left| \frac{\Gamma[X]}{(A, X^d - 1)} \right|.
$$

This completes the proof of [Theorem](#page-2-1) [2](#page-2-1).

## **4. Proof of the corollaries**

## *4.1. Proof of [Corollary](#page-2-2)* [1](#page-2-2)

**Proof of [Corollary](#page-2-2) [1.](#page-2-2)** The PCR rule for necklaces of order *n* is affine, and its associated affine relation is given by

$$
(a_i)_i \in R \iff 0 = a_0 - a_n.
$$

Then, its characteristic polynomial is  $A = X^n - 1$ . Using [Theorem](#page-2-1) [2](#page-2-1),

$$
|F_k(r_n)| = \frac{k(s: s\omega)}{\omega} \sum_{(s:\omega)|d|\omega} \varphi(\omega/d) \left| \frac{\Gamma[X]}{(A, X^d - 1)} \right|
$$

where

- $\omega$  is the order of *X* modulo  $\Lambda$ ,
- $\varphi$  is Euler's totient function, and
- *s* is the length of the smallest cycle in the 2-factor. Equivalently, *s* can be defined as the least multiple of *k* such that

$$
c(1 + X + \cdots + X^{s-1}) \in (\Lambda, X^s - 1).
$$

In this case the associated constant  $c$  is 0. That is, the rule is linear. Therefore  $s = k$  because the condition  $c(1 + X + \cdots + X^{s-1}) \in (A, X^s - 1)$  always holds.

The order of *X* modulo  $X^n - 1$  is *n*; therefore,  $g = (n : k)$ . Replacing these identities into the formula for  $|F_k(r_n)|$  we get

$$
|F_k(r_n)|=\frac{k(k:n)}{kn}\sum_{(k:n)|d|n}\varphi(n/d)\left|\frac{\Gamma[X]}{(A,X^d-1)}\right|.
$$

Observe that  $(A, X^d - 1) = (X^{(n:d)} - 1) = (X^d - 1)$ . Hence,

$$
\left|\frac{\Gamma[X]}{(A, X^d - 1)}\right| = \left|\frac{\Gamma[X]}{(X^d - 1)}\right| = b^d.
$$

Replacing this into the formula for  $|F_k(r_n)|$  we get

$$
|F_k(r_n)| = \frac{(k:n)}{n} \sum_{(k:n)|d|n} \varphi(n/d)b^d. \quad \Box
$$

*4.2. Proof of [Corollary](#page-2-3)* [2](#page-2-3)

**Proof of [Corollary](#page-2-3)** [2.](#page-2-3) For the incremented cycle register case, the associated characteristic polynomial is  $\Lambda = X^n - 1$  as in the PCR case, but the constant *c* is 1 instead of 0. To specialize [Theorem](#page-2-1) [2](#page-2-1), we have to find the smallest integer *s* that is a multiple of *k* and

$$
1 + X + \cdots + X^{s-1} \in (A, X^s - 1).
$$

Let  $d = (n : s)$ . Notice that  $\Lambda = X^n - 1$  and so  $(\Lambda, X^s - 1) = (X^d - 1)$ . Since the ideal is principal, the condition  $1+X+\cdots+X^{s-1}$  ∈  $(X^d-1)$  can be checked by reducing  $1+X+\cdots+X^{s-1}$  modulo the polynomial  $X^d-1$  and checking if the result is 0.

When we reduce a polynomial modulo  $X<sup>d</sup> − 1$ , the *i*th coefficient of the reduced polynomial is the sum of all the coefficients of the original polynomial that have degree congruent to *i* modulo *d*. The polynomial  $1 + X + \cdots + X^{s-1}$ has all coefficients equal to 1. So, when reduced modulo *X <sup>d</sup>* − 1, each resulting coefficient will be *s*/*d*, since for each *i* ∈ {0, . . . *d* − 1} there are *s*/*d* indices in {0, . . . *s*} congruent to *i* modulo *d*.

In order for  $1 + X + \cdots + X^{s-1}$  mod  $X^d - 1$  to be 0, we need  $s/d \equiv 0 \mod b$ . So *s* has to be a multiple of *b*, which we can write as  $s = bm$  for some *m*. Furthermore, we need the following condition to hold

 $s/d = bm/(bm:n) \equiv 0 \mod b$ .

This is equivalent to  $(bm:n)|m$ , which in turn is equivalent to

$$
\left(b\frac{m}{(m:n)}:\frac{n}{(m:n)}\right)\left|\frac{m}{(m:n)}\right|
$$

Notice that  $\frac{m}{(m:n)}$  is coprime with  $\frac{n}{(m:n)}$ , so

$$
\left(b\frac{m}{(m:n)}:\frac{n}{(m:n)}\right)=\left(b:\frac{n}{(m:n)}\right).
$$

And the only divisor of  $\frac{n}{(m.n)}$  that is also a divisor of  $\frac{m}{(m.n)}$  is 1, so the condition holds only when

$$
\left(b:\frac{n}{(m:n)}\right)=1,
$$

which is true precisely when  $d_b(n)|m$ . Since we also require that *s* be a multiple of *k*, the smallest possible value for *s* is:

 $s = \text{lcm}(k, bd_h(n)).$ 

We apply [Theorem](#page-2-1) [2](#page-2-1) as we did in the proof of [Corollary](#page-2-2) [1,](#page-2-2) with  $\omega = n$  and  $\vert$ Γ [*X*]  $\left| \frac{\Gamma[X]}{(\Lambda, X^d-1)} \right| = b^d$ . For any divisor *d* of the order  $\omega = n$ . Replacing this into the formula of [Theorem](#page-2-1) [2](#page-2-1) we get

$$
|F_k(\iota_n)| = \frac{k(\operatorname{lcm}(k, bd_b(n)) : n)}{\operatorname{lcm}(k, bd_b(n))n} \sum_{(\operatorname{lcm}(k, bd_b(n)) : n) | d | n} \varphi(n/d) b^d. \quad \Box
$$

#### *4.3. Proof of [Corollary](#page-3-3)* [3](#page-3-3)

**Proof of [Corollary](#page-3-3) [3.](#page-3-3)** The Xor rule for necklaces of order *n* is affine, and its associated affine relation is given by

$$
(a_i)_i \in R \iff a_n = a_0 + a_1 + \cdots + a_{n-1}.
$$

Consequently, its characteristic polynomial is  $\Lambda = 1 + X + \cdots + X^{n-1} - X^n$ , which is equal to  $U_{n+1}$  when  $|\Gamma| = 2$ . To specialize [Theorem](#page-2-1) [2](#page-2-1), we need to compute:

- The length of the smallest cycle, which is 1 since the rule is linear
- (A multiple of) the order  $\omega$  of *X* modulo  $\Lambda$
- The size of  $\vert$ Γ [*X*]  $\left| \frac{\Gamma[X]}{(A, X^d - 1)} \right|$  for all divisors  $d | \omega$ .

We know that  $\Lambda=U_{n+1}$  which divides  $X^{n+1}-1.$  Therefore,  $\omega=n+1$  is a multiple of the order of  $X$  modulo  $\Lambda.$  Due to [Lemma](#page-7-2) [9,](#page-7-2)

$$
(A: Xd - 1) = (U\omega : Xd - 1) = \begin{cases} Xd - 1 & \text{if } \omega/d \text{ is even} \\ Ud & \text{if } \omega/d \text{ is odd.} \end{cases}
$$

Hence,

$$
\left| \frac{\Gamma[X]}{(A, X^d - 1)} \right| = | \Gamma |^{d-1+1_{2}} \sqrt{\omega/d}.
$$

Replacing this in the statement of [Theorem](#page-2-1) [2](#page-2-1) we get

$$
|F_k(\sigma)| = \frac{k(s:\omega)}{s\omega} \sum_{(s:\omega)|d|\omega} \varphi(\omega/d) \left| \frac{\Gamma[X]}{(A, X^d - 1)} \right|
$$
  
= 
$$
\frac{k}{\omega} \sum_{d|\omega} \varphi(\omega/d) |\Gamma|^{d-1+1} \sqrt{2 \omega} d.
$$

If we do a change of variables  $d \mapsto \omega/d$ , and set  $|\Gamma| = 2$  we get

$$
= \frac{k}{\omega} \sum_{d|\omega} \varphi(d) 2^{\frac{1}{2}2|d} 2^{\omega/d-1}.
$$

Since  $\varphi(2d) = d$  when *d* is even and  $\varphi(2d) = 2d$  when *d* is odd,  $\varphi(d)2^{1/2d}$  reduces to  $\varphi(2d)$ 

$$
= \frac{k}{2\omega} \sum_{d|\omega} \varphi(2d) 2^{\omega/d}. \quad \Box
$$

# **Data availability**

No data was used for the research described in the article.

#### **Acknowledgments**

This project is supported by grant CONICET PIP 2022–2024 11220210100220CO and grant UBACyT 20020220100065BA Universidad de Buenos Aires.

# **References**

- <span id="page-11-1"></span>[1] N. Álvarez, V. Becher, P.A. Ferrari, S.A. Yuhjtman, Perfect necklaces, Adv. in Appl. Math. 80 (2016) 48–61, [http://dx.doi.org/10.1016/j.aam.2016.](http://dx.doi.org/10.1016/j.aam.2016.05.002) [05.002](http://dx.doi.org/10.1016/j.aam.2016.05.002).
- <span id="page-11-6"></span><span id="page-11-2"></span>[2] V. Becher, O. Carton, Normal numbers and nested perfect necklaces, J. Complexity 54 (2019) 101403, <http://dx.doi.org/10.1016/j.jco.2019.03.003>. [3] M. Beck, S. Robins, Computing the continuous discretely, second ed., in: Undergraduate Texts in Mathematics, Springer, New York, 2015,
- <span id="page-11-4"></span><http://dx.doi.org/10.1007/978-1-4939-2969-6>, integer-point enumeration in polyhedra, With illustrations by David Austin. [4] J. Berstel, D. Perrin, The origins of combinatorics on words, European J. Combin. 28 (3) (2007) 996–1022, [http://dx.doi.org/10.1016/j.ejc.2005.](http://dx.doi.org/10.1016/j.ejc.2005.07.019) [07.019](http://dx.doi.org/10.1016/j.ejc.2005.07.019), URL [https://www.sciencedirect.com/science/article/pii/S0195669805001629.](https://www.sciencedirect.com/science/article/pii/S0195669805001629)
- <span id="page-11-7"></span>[5] [W. Burnside, Theory of Groups of Finite Order, second ed., Dover Publications, Inc. New York, 1955.](http://refhub.elsevier.com/S0166-218X(24)00263-4/sb5)
- <span id="page-11-5"></span>[6] J. Cooper, C. Heitsch, The discrepancy of the lex-least de Bruijn sequence, Discrete Math. 310 (6–7) (2010) 1152–1159, [http://dx.doi.org/10.](http://dx.doi.org/10.1016/j.disc.2009.11.010) [1016/j.disc.2009.11.010](http://dx.doi.org/10.1016/j.disc.2009.11.010).
- <span id="page-11-0"></span>[7] [S.W. Golomb, Shift Register Sequences, Holden-Day, Inc., San Francisco, Calif.-Cambridge-Amsterdam, 1967, with portions co-authored by Lloyd](http://refhub.elsevier.com/S0166-218X(24)00263-4/sb7) [R. Welch, Richard M. Goldstein, and Alfred W. Hales.](http://refhub.elsevier.com/S0166-218X(24)00263-4/sb7)
- <span id="page-11-3"></span>[8] R. Hofer, G. Larcher, Discrepancy bounds for normal numbers generated by necklaces in arbitrary base, J. Complexity 78 (2023a) 101767, <http://dx.doi.org/10.1016/j.jco.2023.101767>, URL <https://www.sciencedirect.com/science/article/pii/S0885064X23000365>.
- <span id="page-12-2"></span>[9] R. Hofer, G. Larcher, The exact order of the discrepancy for Levin's normal number in base 2, J. Théor. Nombres Bordeaux 35 (3) (2023b) 999–1023, [http://dx.doi.org/10.5802/jtnb.1271.](http://dx.doi.org/10.5802/jtnb.1271)
- <span id="page-12-3"></span>[10] Y.J. Huang, A new algorithm for the generation of binary de Bruijn sequences, J. Algorithms 11 (1) (1990) 44–51, [http://dx.doi.org/10.1016/0196-](http://dx.doi.org/10.1016/0196-6774(90)90028-D) [6774\(90\)90028-D.](http://dx.doi.org/10.1016/0196-6774(90)90028-D)
- <span id="page-12-6"></span>[11] O.F. Inc, Entry A000013, The on-line encyclopedia of integer sequences, 2023, [https://oeis.org/A000013.](https://oeis.org/A000013)
- <span id="page-12-5"></span>[12] O.F. Inc, Entry A000016, The on-line encyclopedia of integer sequences, 2023, [https://oeis.org/A000016.](https://oeis.org/A000016)
- <span id="page-12-4"></span>[13] O.F. Inc, Entry A000031, The on-line encyclopedia of integer sequences, 2023, [https://oeis.org/A000031.](https://oeis.org/A000031)
- <span id="page-12-1"></span>[14] A. Lempel, On extremal factors of the de Bruijn graph, J. Combin. Theory Ser. B 11 (1971) 17–27, [http://dx.doi.org/10.1016/0095-8956\(71\)90009-](http://dx.doi.org/10.1016/0095-8956(71)90009-8) [8](http://dx.doi.org/10.1016/0095-8956(71)90009-8).
- <span id="page-12-0"></span>[15] J. Mykkeltveit, A proof of Golomb's conjecture for the de Bruijn graph, J. Combin. Theory Ser. B 13 (1972) 40–45, [http://dx.doi.org/10.1016/0095-](http://dx.doi.org/10.1016/0095-8956(72)90006-8) [8956\(72\)90006-8.](http://dx.doi.org/10.1016/0095-8956(72)90006-8)