On extremal factors of de Bruijn-like graphs

Nicolás Álvarez Verónica Becher Martín Mereb Ivo Pajor Carlos Miguel Soto

August 30, 2023

Abstract

In 1972 Mykkeltveit proved that the maximum number of vertex-disjoint cycles in the de Bruijn graphs of order n is attained by the pure cycling register rule, as conjectured by Golomb. We generalize this result to the tensor product of the de Bruijn graph of order n and a simple cycle of size k, when n divides k or vice versa. We also develop counting formulae for a large family of cycling register rules, including the linear register rules proposed by Golomb.

MSC Classification: 05C35; 05C45.
Keywords: de Bruijn graph, Astute graph, pure cycling register, perfect necklaces

1 Introduction and statement of results

In [5, Chapter VII Conjecture A] Golomb asked what is the maximal number of vertex-disjoint cycles in the de Bruijn graph of order n. He conjectured that this is exactly the number of cycles attained using the pure cycling register rule to partition the n-th de Bruijn graph. This conjecture was proved by Mykkeltveit [11], using Lempel's [10] 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* for a fixed alphabet Γ. These graphs are the tensor product of the de Bruijn graph of order 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].

The elements of Γ^n are referred to as *string*s 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 $a_ia_{i+1} \ldots a_{j-1}$.

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

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

such that $s[1..n) = t[0..n-1)$ and $j = i + 1$.

Remark 1. Notice that $G_{n,1}$ is the de Bruijn graph of order n. In this case, we identify the vertices $V_{n,1}$ with Γ^n.

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

To construct factors of de Bruijn and astute graphs, we consider *succession rules*. These are what Golomb in [5] calls Shift Registers.

Definition (Succession rule)**.** A succession rule is a bijective function $\sigma : \Gamma^n \to \Gamma^n$ such that for each string $s = a_0 a_1 \cdots a_{n-1}$, $\sigma(s) = a_1 a_2 \cdots a_{n-1} a_n$ 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 $\sigma(s)$, and in the astute graph there exists an arc from vertex (s, i) to $(\sigma(s), i + 1)$. This means that the succession rule σ can be thought to act on the vertices of the de Bruijn and astute graphs.

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

We interpret the alphabet Γ as the ring $\mathbb{Z}/b\mathbb{Z}$ where $b = |\Gamma|$, therefore we can do linear arithmetic on its symbols.

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

$$a_0 a_1 \cdots a_n \in R \iff c = \sum_{0 \le i \le n} \lambda_i a_i.$$

Definition (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 $a_0 a_1 \ldots a_{n-1}$, $\sigma(a_0 a_1 \ldots a_{n-1})$ is the unique string $a_1 \ldots a_n$ such that $a_0 a_1 \ldots a_n$ 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 λ_n is invertible; and

- The rule has to be bijective, which only happens if λ_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. Namely, the equivalence classes of Γ^n under string rotation.

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 ι is that each orbit has an equal quantity of each symbol in Γ. This has applications in the construction of de Bruijn sequences with small discrepancy, see [6] in contrast to [4].

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 $s = a_0 a_1 \cdots a_{n-1}$ as

$$\mathrm{x}_n(s) = a_1 \cdots a_{n-1} a_n$$

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 factor $F_k(r_n)$ produced by the pure cycling register rule r_n is extremal.*

When k does not divide n Theorem 1 is not necessarily true. For the case $k = 2$, $n = 3$ and $\Gamma = \{0, 1\}$, the successor rule r_3 produces a factor of size 4 as shown in Figure 1a, while the extremal factors have size 6. An example of such extremal factor is shown in Figure 1b. For such cases, where the hypothesis of Theorem 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 factors generated by affine rules.

We use $(a : b)$ for the greatest common divisor of the integers a and b. We write (P, Q) for the ideal generated by P and Q.

Theorem 2. *Let n and k be positive integers, and R be an affine rule given by*

$$a_0 a_1 \cdots a_n \in R \iff c = \sum_{0 \le i \le n} \lambda_i a_i,$$

for some coefficients $(\lambda_i)_i$ and a constant term c. Then the number of factors in the associated succession rule σ is given by

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

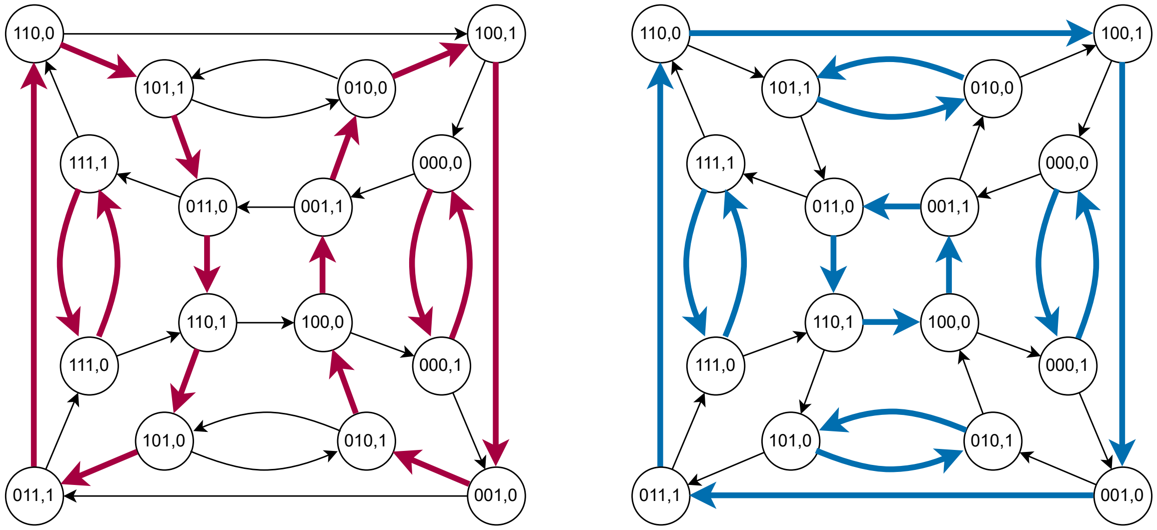
where

- $\Lambda = \displaystyle\sum_{i=0}^{n} \lambda_i X^{n-i}$ *is the characteristic polynomial of R,*
- ω *is any multiple of the order of X modulo Λ,*
- φ *is Euler's totient function, and*

  s *is the length of the smallest cycle in the factor. Equivalently, s can be defined as the smallest multiple of k such that*

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

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



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

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

Figure 1: Pure Cycling Register induced factors may not be extremal in astute graphs

Corollary 1 (Factor of $G_{n,k}$ from Pure Cycling Register)**.**

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

Remark 4. When $k = 1$ and $|\Gamma| = 2$, $|F_1(r_n)|$ is the number of binary irreducible polynomials whose degree divides n, see [9].

Corollary 2 (Factor of $G_{n,k}$ from Incremented Cycling Register)**.**

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

where $b = |\Gamma|$ and $d_b(n)$ is the smallest divisor of n such that $n/d_b(n)$ is coprime with b.

Remark 5. When $k = 1$ and $|\Gamma| = 2$, $|F_1(\iota_n)|$ is the number of distinct output sequences from binary n-stage shift register which feeds back the complement of the last stage, see [8].

Corollary 3 (Factor of $G_{n,k}$ from Xor Cycling Register)**.**

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

Remark 6. When $k = 1$ and $|\Gamma| = 2$, $|F_1(\mathrm{x}_n)|$ 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 [7].

2 Proof of Theorem 1

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

We use a basic tool from finite Fourier analysis [2].

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

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

Lemma 1. *$C(r_n(s)) = \mu^{-1} C(s)$.*

Proof. Let $s = a_0 a_1 \cdots a_{n-1}$. Then $r(s) = a_1 a_2 \cdots a_n a_0$. 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

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

□

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_i form a cycle in the de Bruijn graph, we have that $(s_i)_j = (s_{i+1})_{j-1}$, where the indices of s are taken modulo t. Then, $(s_i)_j$ depends only of the sum $i + j$ modulo t. Let $w_{i+j} = (s_i)_j$ with the indices of w taken modulo t. We can rewrite the expression as

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

With a change of variables $d = i + j$, we get

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

And the sum of the powers of a primitive root is 0, so we are done. □

Lemma 3. Let s and t be two strings that are connected by an arc in the de Bruijn graph $G_{n,1}$. Then $C(s) - C(r_n^{-1}(t)) \in \mathbb{R}$, and it is zero exactly when $s = r_n^{-1}(t)$.

Proof. Since s and t are connected in the de Buijn 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 $s = r_n^{-1}(t)$. □

To prove that the factor produced by the Pure Cycling Register is extremal, we choose for each cycle in the factor $F_k(r_n)$ a distinguished vertex, and then we prove that any cycle in any factor has at least one distinguished vertex, therefore the size of any factor is at most the number of distinguished vertices.

Let $(s_0, m_0), (s_1, m_1), \cdots, (s_{t-1}, m_{t-1})$ be any cycle in $F_k(r_n)$. There are two possibilities.

One possibility is that the transform $C(s_i)$ is real for all s_i. In this case we take any arbitrary vertex in the factor as the distinguished vertex.

The other possibility is that there exists some string s_i such that $C(s_i)$ is not real. Let $z = C(s_0)$. Due to Lemma 1, we have that $C(s_i) = z\mu^{-i}$. Since the length of any PCR cycle is a divisor of $\operatorname{lcm}(n, k) = n$ and $z\mu^{-i}$ has a cycle length equal to the order of μ (which is n), the size of the 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 factor will be the unique vertex (s_i, m_i) such that

$$\mathfrak{Im}(C(s_i)) < 0 \quad \text{but} \quad \mathfrak{Im}(C(s_{i-1})) \geq 0,$$

as exemplified in Figure 2. Now we have to prove that every cycle in the astute graph $G_{n,k}$ contains at least one distinguished vertex. Let

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

be any such cycle. We consider three cases:

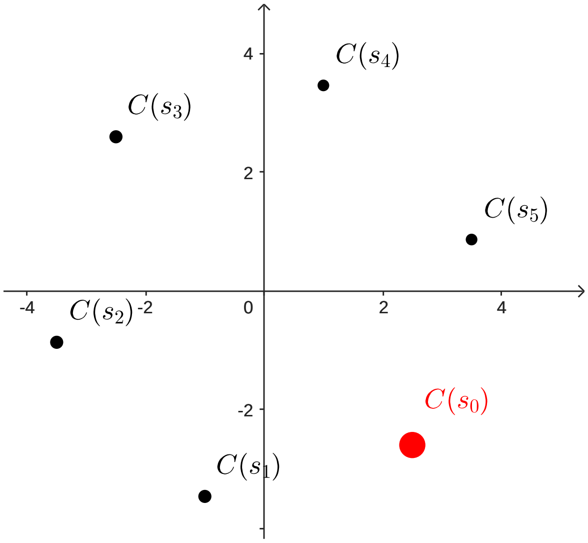


Figure 2: Transforms of the strings on the PCR cycle generated by $s_0 = 123351$. The large red point is the distinguished vertex for this PCR cycle.

First case: There exists some string s_i such that $C(s_i)$ is not real. The sum of $C(s_i)$ over all i must be zero due to Lemma 2. 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 3 implies that $C(s_{i-1})$ and $C(r_n^{-1}(s_i))$ differ by a real number. And since $\mathfrak{Im}(C(s_{i-1})) \geq 0$, then $\mathfrak{Im}(C(r_n^{-1}(s_i))) \geq 0$ as well. This means that in the PCR cycle

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

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

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

Second case: The transform $C(s_i)$ is 0 for all i. Due to Lemma 3, $s_i = r_n^{-1}(s_{i+1})$ if and only if $C(s_i) = C(r_n^{-1}(s_{i+1}))$. By Lemma 1, $C(r_n^{-1}(s_{i+1})) = \mu C(s_{i+1}) = \mu \cdot 0 = 0$, 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 3, $C(r_n^{-1}(s_i)) = \mu C(s_i)$ differs from $C(s_{i-1})$ by a real number. If $n > 2$, μ 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 $C(s_{i-1})$. This is a contradiction, because we assumed $C(s_j)$ was real for all j. So we only have to analyze the cases $n = 1$ and $n = 2$. In both cases Theorem 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.

We showed that every cycle in graph $G_{n,k}$ contains at least one distinguished vertex of a cycle in the factor determined by the Pure Cycling Register rule, Theorem 1 is proved.

3 Proof of Theorem 2

3.1 Burnside's Lemma

Our main tool for counting the number of cycles in a succession rule is the classical Burnside's Lemma [3]. 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 $S = V_{n,k}$ and $G = \langle A_k(\sigma) \rangle$ is the group generated by the action associated with a succession rule σ. In that case, the set of orbits S/G coincides with the factor $F_k(\sigma)$. We thus have the following identity:

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

where ω is the order of $A_k(\sigma)$ and $\operatorname{fix}(f)$ 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 cyclic because $A_k(\sigma)^i = A_k(\sigma)^{i+\omega}$ for all $i \in \mathbb{Z}$. Therefore, Equation (1) asserts that the size of $F_k(\sigma)$ is the average of the function $i \mapsto |\operatorname{fix}(A_k(\sigma)^i)|$ 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.

Lemma 4. Let k be a positive integer and $\sigma : \Gamma^n \to \Gamma^n$ a succession rule. Let ω 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 Γ^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 σ^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 4 with this identity we obtain the following.

Lemma 5. Let k be a positive integer and $\sigma : \Gamma^n \to \Gamma^n$ a succession rule. Let ω 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 $U_n = \dfrac{X^n-1}{X-1}$. 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 6. $(U_n : U_m) = U_{(n:m)}$.

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

$$U_m - X^{m-n}U_n = U_{m-n}.$$

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 $U_{(n:m)}$ □

Lemma 7. $(X^n - 1 : X^m - 1) = X^{(n:m)} - 1$.

Proof. It follows directly from Lemma 6 by multiplying both sides by $X - 1$. □

Lemma 8. If Γ is a field, then

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

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 $X^{(n:m)} - 1$ 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 Γ.

□

3.3 Affine Succession Rules

In this section we compute the size of the set $\mathrm{fix}(\sigma^k) \subseteq \Gamma^n$ 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 $w \in \Gamma^k$ 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 $w^* = wwwww\cdots$ 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 σ 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:

$$\exists p : \deg(p) \leq n \text{ and } \quad \frac{w}{1-X^k}\Lambda = p + \frac{c}{1-X}, \tag{2}$$

where Λ 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 Equation (2) by $1 - X^k$ we get the equivalent expression:

$$\exists p : \deg(p) \leq n \text{ and } \quad w\Lambda = p(1-X^k) + cU_k$$

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 Equation (3) is the associated string of some other string $s \in \mathrm{fix}(\sigma^k)$. Indeed, take any w that satisfies (3). Since (3) is equivalent to (2), 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 $|\mathrm{fix}(\sigma^k)|$ is the number of solutions to the Equation (3). Since w is always of length k, we can assume w is a polyomial 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

$$\Lambda w \equiv P \qquad \bmod X^k - 1$$

has a solution if and only if

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

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

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

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

$$\Lambda w \equiv 0 \qquad \bmod X^k - 1.$$

Since there are $|\Gamma[X]/(X^k - 1)|$ possibilities for w, and each possibility for P gets an equal number of solutions, each P gets exactly

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

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

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

Let us further analyze the condition $cU_k \in (\Lambda, X^k - 1)$. 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 ℓ_σ. Let d be any multiple of ℓ_σ. 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 $U_{\ell_\sigma} | U_d$ 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 $g = \gcd(d, \ell_\sigma)$. Then $U_g \in (U_{\ell_\sigma}, U_d)$ due to theorem 6. 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,

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

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

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

We rewrite Equation (4) as

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

Now let us analyze $\left| \frac{\Gamma[X]}{(\Lambda, X^k - 1)} \right|$. Since the first coefficient of Λ is invertible, the polynomial X is invertible modulo Λ, so there exists some ω 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,

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

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

Lemma 9.

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

3.4 Burnside's Lemma for affine necklaces

Let k be a positive integer and let σ be an affine succession rule with its characterisitc polynomial Λ. Due to Lemma 5, we have that

$$|F_k(\sigma)| = \frac{k}{M} \sum_{i=0}^{M-1} \mathbb{1}_{k|i} | \operatorname{fix}(\sigma^i)|, \tag{4}$$

for any M that is a cycle of the function $i \mapsto \mathbb{1}_{k|i} | \operatorname{fix}(\sigma^i)|$. Due to Lemma 9, we know that if ω is the order of X modulo Λ, then

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

Therefore, M needs to be a cycle of

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

To be a cycle of a product of functions, it suffices to be a multiple of the cycle length of each factor. So, we have that one possible value of M is $\operatorname{lcm}(k, \ell_\sigma, \omega)$. Now rewriting Equation (4) we get

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

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

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

The second equality uses that $S|M$, and the last equality uses that since $M = \mathrm{lcm}(S, \omega)$, we have $\omega/\left(S:\omega\right) = M/S$.

Since $\gcd(i, \omega/\left(S:\omega\right))$ iterates over all divisors of d, we can express that sum as follows:

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

This can be rewritten as

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

This completes the proof of Theorem 2.

4 Proof of the Corollaries

4.1 Proof of Corollary 1

Proof of Corollary 1. 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 $\Lambda = X^n - 1$. Using Theorem 2,

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

where

- ω is the order of X modulo Λ,
- φ is Euler's totient function, and

- s is the length of the smallest cycle in the 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 $c(1 + X + \cdots + X^{s-1}) \in (\Lambda, X^s - 1)$ 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

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

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

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

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

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

□

4.2 Proof of Corollary 2

Proof of Corollary 2. 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 2, we have to find the smallest integer s that is a multiple of k and

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

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 $1 + X + \cdots + X^{s-1} \in (X^d - 1)$ 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 $X^d - 1$, 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 $X^d - 1$, each resulting coefficient will be s/d, since for each $i \in \{0, \ldots d - 1\}$ there are s/d indices in $\{0, \ldots s\}$ 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

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

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

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

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:

$$s = \operatorname{lcm}(k, bd_b(n)).$$

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

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

□

4.3 Proof of Corollary 3

Proof of Corollary 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 2, we need to compute:

- The length of the smallest cycle, which is 1 since the rule is linear
- (A multiple of) the order ω of X modulo Λ
- The size of $\left|\frac{\Gamma[X]}{(\Lambda, X^d - 1)}\right|$ 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 Λ. Due to Lemma 8,

$$(\Lambda, X^d - 1) = ((U_\omega : X^d - 1)) = \begin{cases} X^d - 1 & \text{if } \omega/d \text{ is even} \\ U_d & \text{if } \omega/d \text{ is odd.} \end{cases}$$

Hence,

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

Replacing this in the statement of Theorem 2 we get

$$\begin{aligned} |F_k(\sigma)| &= \frac{k\,(s : \omega)}{s\omega} \sum_{(s:\omega)|d|\omega} \varphi\,(\omega/d) \left| \frac{\Gamma[X]}{(\Lambda, X^d - 1)} \right| \\ &= \frac{k}{\omega} \sum_{d|\omega} \varphi\,(\omega/d)\,|\Gamma|^{d-1+\mathbb{1}_{2 \,\mid\, \omega/d}}. \end{aligned}$$

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

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

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

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

□

References

[1] Nicolás Alvarez, Verónica Becher, Pablo A. Ferrari, and Sergio A. Yuhjtman. Perfect necklaces. *Advances in Applied Mathematics*, 80:48–61, 2016.

[2] Matthias Beck and Sinai Robins. *Computing the continuous discretely.* Undergraduate Texts in Mathematics. Springer, New York, second edition, 2015. Integer-point enumeration in polyhedra, With illustrations by David Austin.

[3] William Burnside. *Theory of groups of finite order.* Dover Publications, Inc., New York, 1955. 2d ed.

[4] Joshua Cooper and Christine Heitsch. The discrepancy of the lex-least de Bruijn sequence. *Discrete Mathematics*, 310(6-7):1152–1159, 2010.

[5] Solomon W. Golomb. *Shift register sequences.* Holden-Day, Inc., San Francisco, Calif.-Cambridge-Amsterdam, 1967. With portions co-authored by Lloyd R. Welch, Richard M. Goldstein, and Alfred W. Hales.

[6] Yue Jiang Huang. A new algorithm for the generation of binary de Bruijn sequences. *Journal of Algorithms*, 11(1):44–51, 1990.

[7] OEIS Foundation Inc. Entry A000013. The On-Line Encyclopedia of Integer Sequences https://oeis.org/A000013, 2023.

[8] OEIS Foundation Inc. Entry A000016. The On-Line Encyclopedia of Integer Sequences https://oeis.org/A000016, 2023.

[9] OEIS Foundation Inc. Entry A000031. The On-Line Encyclopedia of Integer Sequences https://oeis.org/A000031, 2023.

[10] Abraham Lempel. On extremal factors of the de Bruijn graph. *Journal of Combinatorial Theory. Series B*, 11:17–27, 1971.

[11] Johannes Mykkeltveit. A proof of Golomb's conjecture for the de Bruijn graph. *Journal of Combinatorial Theory. Series B*, 13:40–45, 1972.

Nicolás Álvarez
ICC CONICET Argentina - nico.alvarez@gmail.com

Verónica Becher
Departamento de Computación, Facultad de Ciencias Exactas y Naturales & ICC
Universidad de Buenos Aires & CONICET Argentina- vbecher@dc.uba.ar

Martín Mereb
Departamento de Matemática, Facultad de Ciencias Exactas y Naturales & IMAS
Universidad de Buenos Aires & CONICET Argentina- mmereb@gmail.com

Ivo Pajor
Departamento de Computación, Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires & Argentina- pajorivo@gmail.com

Carlos Miguel Soto
Departamento de Computación, Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires & Argentina- miguelsotocarlos@gmail.com