Lyndon pairs and the lexicographically greatest perfect necklace

Verónica Becher
vbecher@dc.uba.ar

Tomás Tropea
tomastropeaa@gmail.com

Departamento Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and ICC CONICET. Pabellón 0, Ciudad Universitaria, C1428EGA Buenos Aires, Argentina

Abstract. Fix a finite alphabet. A necklace is a circular word. For positive integers n and k, a necklace is (n,k)-perfect if all words of length n occur k times but at positions with different congruence modulo k, for any convention of the starting position. We define the notion of a Lyndon pair and we use it to construct the lexicographically greatest (n,k)-perfect necklace, for any n and k such that n divides k or k divides n. Our construction generalizes Fredricksen and Maiorana's construction of the lexicographically greatest de Bruijn sequence of order n, based on the concatenation of the Lyndon words whose length divide n.

1. Introduction

Let Σ be a finite alphabet with at least two symbols. A word on Σ is a finite sequence of symbols, and a necklace is the equivalence class of a word under rotations. Given two positive integers, n and k, a necklace is (n,k)-perfect if all words of length n occur k times but at positions with different congruence modulo k, for any convention of the starting position. The well known circular de Bruijn sequences of order n, see [3, 7, 8], that we call de Bruijn necklaces of order n, are exactly the (n,k)-perfect necklaces for $k=1$. For example, 11100100 is a $(2,2)$-perfect for $\Sigma=\{0,1\}$. The (n,k)-perfect necklaces correspond to Hamiltonian cycles in the tensor product of the de Bruijn graph with a simple cycle of length k.

A thorough presentation of perfect necklaces appears in [1]. With the purpose of constructing normal numbers with very fast convergence to normality M. Levin in [9] gives two constructions of perfect necklaces. One based on arithmetic progressions with difference coprime with the alphabet size which yields a (n,n)-perfect necklaces. The other based on Pascal triangle matrix which yields nested (n,n)-perfect necklaces when n is a power of 2. In [2] there is a method of constructing all nested (n,n)-perfect necklaces.

Assume the lexicographic order on words. A Lyndon word is a nonempty aperiodic word that is lexicographically greater than all of its rotations. For example, the Lyndon words over alphabet $\{0,1\}$ sorted by length and then in decreasing lexicographical order within each length are

$$1, 0, 10, 110, 100, 1110, 1100, 1000, 11110, 11100, 11010, 11000, 10100, 10000, ...$$

Lyndon words were introduced by Lyndon in mid 1950s [10, 11]. They provide a nice factorization of the free monoid: each word w of the free monoid Σ^* has a unique decomposition as a product $w=u_1 \dots u_n$ of a non-increasing sequence of Lyndon words $u_1, \dots u_n$

Date: May 28, 2024.

1991 *Mathematics Subject Classification.* Primary 68R15, 05A05; Secondary 11K16.

Key words and phrases. de Bruijn sequences, Lyndon words.

in the lexicographic order. The problem to compute the prime factorization of a given word has a solution in time linear to the length of the given word [4], see also [6]. Fredricksen and Maiorana [5] showed that the concatenation in decreasing lexicographic order of all the Lyndon words whose length divides a given positive integer number n, yields a de Bruijn necklace of order n. For example, the concatenation of the binary Lyndon words whose length divides four is (the spaces are for ease of reading),

$$1\ 1110\ 1100\ 10\ 1000\ 0.$$

This construction, together with the efficient generation of Lyndon words, provides an efficient method for constructing the lexicographically greatest de Bruijn necklace of each order n in linear time and logarithmic space.

In this note we present the notion of Lyndon pairs and we use it to generalize Fredricksen and Maiorana's algorithm to construct the lexicographically greatest (n, k)-perfect necklaces, for any n and k such that n divides k or k divides n.

2. Lyndon pairs in lexicographical order

2.1. **Lyndon pairs.** We assume a finite alphabet Σ with cardinality s, with $s \geq 2$. Without loss of generality we use $\Sigma = \{0, \ldots, s-1\}$. We use lowercase letters a, b, c possibly with subindices for alphabet symbols. Words are finite sequences of symbols that we write $a_1 a_2 .. a_n$ or with a capital letter A, B, C. We write a^ℓ to denote the word of length ℓ made just of a's. and we write A^ℓ to denote the word made of ℓ copies of A. The concatenation of two words A and B is written AB. The length of a word A is denoted with $|A|$. The positions of a word A are numbered from 1 to $|A|$. We use $>$ to denote the decreasing lexicographic order on words and we write $A \geq B$ with the when $A > B$ or $A = B$.

We use lowercase letters $h, \ldots, z$ to denote non-negative integers. We write $k|n$ to say that k divides n. We write $\mathbb{Z}_k$ for the set $\mathbb{Z}/k\mathbb{Z}$ of residues modulo k. We also use $<$ and $>$ for the natural orders on $\mathbb{Z}_k$ and $\mathcal{N}$ and, as usual, $u \leq v$ when $u < v$ or $u = v$; and $v \geq u$ when $v > u$ or $v = u$. When $u < v$ we may write $v > u$

We work with pairs in $\Sigma^n \times \mathbb{Z}_k$ when $k|n$ or $n|k$. This condition on n and k is assumed all along the sequel. We refer to pairs $\langle A, u\rangle$ with calligraphic letter $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ We consider the following order $\succ$ over $\Sigma^n \times \mathbb{Z}_k$.

Definition (order $\succ$ on $\Sigma^n \times \mathbb{Z}_k$)**.**

$$\langle A, u\rangle \succ \langle B, v\rangle \text{ exactly when either } (A > B \text{ and } u = v) \text{ or } (k-1-u) > (k-1-v).$$

The smallest the second component in $\mathbb{Z}_k$, the $\succ$-greater the pair. Among pairs with the same second component in $\mathbb{Z}_k$, the order $\succ$ is defined with the decreasing lexicographic order on Σ^n. Thus, $\langle (s-1)^n, 0\rangle$ is the $\succ$-greatest in $\Sigma^n \times \mathbb{Z}_k$. As usual, we write $\langle A, u\rangle \succeq \langle B, v\rangle$ exactly when $\langle A, u\rangle \succ \langle B, v\rangle$ or $\langle A, u\rangle = \langle B, v\rangle$.

It is possible to concatenate pairs in $\Sigma^n \times \mathbb{Z}_k$ having the same second component: the concatenation of $\langle A, u\rangle$ and $\langle B, u\rangle$ is $\langle AB, u\rangle$. We write $\langle A, u\rangle^t$ to denote the pair $\langle A^t, tu\rangle$.

Definition (rotation of a pair)**.** Given a pair $\langle a_1..a_n, u\rangle$ in $\Sigma^n \times \mathbb{Z}_k$ its left rotation is the pair $\langle a_n a_1..a_{n-1}, u-1\rangle$ and its right rotation is the pair $\langle a_2..a_n a_1, u+1\rangle$.

For $s = 3$, $n = k = 5$ the right rotation of pair $\langle 13212, 4\rangle$, is $\langle 21321, 0\rangle$, and its left rotation is $\langle 32121, 3\rangle$. The rotation function induces a relation between pairs: two pairs are related if successive rotations initially applied to the first yield the second. This relation is clearly reflexive and transitive. For pairs in $\Sigma^n \times \mathbb{Z}_k$, when $k|n$ or $n|k$, the rotation has an inverse, given by successive rotations. So the relation is also symmetric, hence, an equivalence relation.

Definition (maximal pair)**.** A necklace in $\Sigma^n \times \mathbb{Z}_k$ is a set of pairs that are equivalent under rotations,

$$\{\langle a_{i+1}..a_n a_1...a_i, u+i\rangle : 0 \le i < \max(n,k)\}$$

In each necklace we are interested in the pair that is maximal in the order $\succ$. We call it maximal .

For example, for $n=4$ and $k=2$ the pair $\langle 1110, 0\rangle$ is maximal among its rotations, hence $\langle 1011, 0\rangle$ is not maximal and $\langle 1101, 1\rangle$ is not maximal either. The pair $\langle 1010, 0\rangle$ is maximal among its rotations.

Observation 1. *When $n|k$ all pairs $\langle A, 0\rangle$ for $A \in \Sigma^n$ are maximal .*

Observation 2. *Let n and k positive integers such that $k|n$ or $n|k$. Then, $\mathcal{A} \in \Sigma^n \times \mathbb{Z}_k$ is maximal exactly when $\mathcal{A}^p$, for any $p \ge 1$, is maximal .*

Proof. ($\Longrightarrow$). In case $n|k$ all pairs in $\Sigma^n \times \mathbb{Z}_k$ with second component 0 are maximal . We prove it for $k|n$. Assume $\mathcal{A}^p$ is maximal rotation, but $\mathcal{A}$ is not. Then $\mathcal{A}$ has a rotation

$$\mathcal{R} = \langle (a_{j+1} \cdots a_n a_1 \cdots a_j), 0\rangle$$

where $k|j$ such that $\mathcal{R} \succ \mathcal{A}$. But this implies that

$$\langle a_{j+1} \cdots a_n a_1 \cdots a_j, 0\rangle^p \succ \langle a_1 \cdots a_n, 0\rangle^p,$$

contradicting that $\mathcal{A}^p$ is maximal .

($\Longleftarrow$). Assume $\mathcal{A}$ is maximal but $\mathcal{A}^p$ is not. Then, $\mathcal{A} = \langle A, 0\rangle$ and $\mathcal{A}^p$ has a rotation $\mathcal{R}^p = \langle R^p, 0\rangle$ that is $\succ$-greater than $\mathcal{A}^p$. Given that the second component of $\mathcal{A}$ and $\mathcal{R}$ is 0, necessarily $\mathcal{R} \succ \mathcal{A}$. But this contradicts that $\mathcal{A}$ was a maximal rotation. ☐

Observation 3. *If $\mathcal{A}$ is a maximal pair then none of its rotations are $\succ$-greater than $\mathcal{A}$, but there may be a rotation that is equal to $\mathcal{A}$.*

For a word $A = a_1..a_n$ we write A_i to denote its prefix of length i, that is, $a_1..a_i$.

Lemma 1. *Let n and k be positive integers such that $k|n$ or $n|k$. Let $\mathcal{A} = \langle A, 0\rangle$ in $\Sigma^n \times \mathbb{Z}_k$ be maximal and different from $\langle 0^n, 0\rangle$. Suppose $A = a_1...a_n$, let i be such that $a_i > 0$ and let*

$$\mathcal{B} = \langle A_{i-1}(a_i - 1)(s-1)^{j-i}, 0\rangle,$$

where if $n|k$ then $j = n$; and if $k|n$ then j is the smallest such that $k|j$ and $i \le j < n$. Then, $\mathcal{B}$ is maximal .

Proof. Assume $\mathcal{A}$ is a maximal pair and, by way of contradiction, assume that $\mathcal{B} = \langle A_{i-1}(a_i - 1)(s-1)^{j-i}, 0\rangle$ is not a maximal pair. Then there is some ℓ multiple of k such that :

$$\langle a_{\ell+1} \cdots a_{i-1}(a_i - 1)(s-1)^{j-i} A_\ell, 0 + \ell\rangle \succ \langle A_{i-1}(a_i - 1)(s-1)^{j-i}, 0\rangle$$

Necessarily

$$a_{\ell+1} \cdots a_{i-1} \ge a_1 \cdots a_{(i-1)-(\ell+1)+1}$$

Since $\mathcal{A}$ is a maximal pair,

$$a_1 \cdots a_{(i-1)-(\ell+1)+1} \ge a_{\ell+1} \cdots a_{i-1}.$$

Therefore, $a_{\ell+1} = a_1$, $a_{\ell+2} = a_2$, $\ldots$, $a_{i-1} = a_{(i-1)-(\ell+1)+1}$. This implies, $a_i - 1 \geq a_{(i-1)-(\ell+1)+2}$. Consequently,

$$a_{\ell+1} \cdots a_{i-1} a_i > a_1 \cdots a_{(i-1)-(\ell+1)+2},$$

but this contradicts that $\mathcal{A}$ es a maximal pair. $\square$

For example, let $s = 7$, $n = 6$, $k = 3$. Since $\mathcal{A} = \langle 456123, 0\rangle$ is maximal , then $\mathcal{B} = \langle 455, 0\rangle$ is maximal , defined by taking $i = 3$, which is multiple of k, and there is no filling with $s - 1 = 6$. Also $\mathcal{B} = \langle 366, 0\rangle$ is maximal , because we take $i = 1$ and we complete with $s - 1 = 6$ up to position $j = 3$.

We define the operator θ that given a pair in $\Sigma^n \times \mathbb{Z}_k$ with second component 0, but different from $\langle 0^n, 0\rangle$, it defines another pair in $\Sigma^n \times \mathbb{Z}_k$ with second component 0.

Definition (reduction)**.** *Let n and k be positive integers such that $k|n$. The reduction of a word $A = a_1 \cdots a_n$, denoted $\overline{A}$, is the word $a_1 \cdots a_p$ where p is the smallest such that $k \mid p$, $p \mid n$ and $a_1 \cdots a_n = (a_1 \cdots a_p)^{n/p}$. The reduction of a pair $\mathcal{A} = \langle A, u\rangle$, denoted with $\overline{\mathcal{A}}$, is the pair $\langle \overline{A}, u\rangle$.*

When $k|n$, the reduction always exists because one can take $p = n$. Notice that if a pair in $\Sigma^n \times \mathbb{Z}_k$ is non reduced with period p, such that $k|p$ and $p|n$ then it has n/p equal rotations. However, all the rotations of a reduced pair are pairwise different. For example, for $s = 8$, $n = 8$ and $k = 2$, $\overline{\langle 10101010, 0\rangle} = \langle 10, 0\rangle$; $\overline{\langle 01230123, 0\rangle} = \langle 0123, 0\rangle$; $\overline{\langle 01234567, 0\rangle} = \langle 01234567, 0\rangle$.

Definition (expansion)**.** *Let n and k be positive integers such that $n|k$. For $A \in \Sigma^n$, $\tilde{A} = A^{k/n} \in \Sigma^k$. For $\mathcal{A} = \langle A, u\rangle \in \Sigma^n \times \mathbb{Z}_k$, $\tilde{\mathcal{A}} = \langle \tilde{A}, u\rangle \in \Sigma^k \times \mathbb{Z}_k$.*

When $n|k$ the expansion always exists. For example for $s = 3$, $n = 2$ and $k = 8$, $\widetilde{\langle 12, 0\rangle} = \langle 12121212, 0\rangle$.

Definition (Lyndon pair)**.** *Let n and k be positive integers. When $k|n$, the Lyndon pairs are the reductions of the maximal pairs in $\Sigma^n \times \mathbb{Z}_k$. When $n|k$ with $n < k$, the Lyndon pairs are the expansions of the maximal pairs $\langle A, 0\rangle \in \Sigma^n \times \mathbb{Z}_k$.*

Thus, when $k|n$, the Lyndon pairs are elements in $Sigma^{\leq n} \times \mathbb{Z}_k$. But when $n|k$ the Lyndon pairs are elements in $\Sigma^n \times \mathbb{Z}_k$.

Observation 4. *Each Lyndon pair is strictly $\succ$-greater than all of its rotations.*

2.2. The operator θ.

Definition (operator θ)**.** *Let n and k be positive integers such that $k|n$ or $n|k$. For $\mathcal{A} = \langle A, 0\rangle = \langle a_1 \cdots a_n, 0\rangle$ in $\Sigma^n \times \mathbb{Z}_k$ such that $a_i > a_{i+1} = \cdots = a_n = 0$, we define the operator $\theta : \Sigma^n \times \mathbb{Z}_k \to \Sigma^n \times \mathbb{Z}_k$,*

$$\theta\langle A, 0\rangle = \langle [A_{i-1}(a_i - 1)(s-1)^{j-i}]^q A_{n-qj}, 0\rangle,$$

where,

if $n|k$ then $j = n$ and $q = 1$; so, $\theta\langle A, 0\rangle = \langle [A_{i-1}(a_i - 1)(s-1)^{n-i}, 0\rangle$;

If $k|n$ then j is the smallest integer such that $k|j$ and $j \geq i$, and q is the greatest integer such that $q \leq n/j$. Thus, in either case, $j = i + ((n - i) \mod k)$.

The operator θ is a function from $\Sigma^n \times \mathbb{Z}_k$ to $\Sigma^n \times \mathbb{Z}_k$. It is applicable on any pair with second component 0, except for $\langle 0^n, 0\rangle$. For example, for $s = 2$, $n = 6$ y $k = 2$, $\theta\langle 010000, 0\rangle = \langle 000000, 0\rangle$; $\theta\langle 011000, 0\rangle = \langle 010101, 0\rangle$; $\theta\langle 011101, 0\rangle = \langle 011100, 0\rangle$.

Definition (T yields the last)**.** *Let T be the integer such that $\theta^T\langle(s-1)^n,0\rangle = \langle(0^n,0\rangle)$.*

Lemma 2. *The list $(\theta^i\langle(s-1)^n,0\rangle)_{0\le i\le T}$ is strictly decreasing in $\succ$.*

Proof. Let's see that for every pair $\mathcal{A}=\langle A,0\rangle$, $\mathcal{A}\succ\theta\mathcal{A}$. Using the definition of θ,

$$\langle A,0\rangle \succ \langle [A_{i-1}(a_i-1)(s-1)^{j-i}]^q A_{n-qj},0\rangle.$$

Since both pairs have second component 0, there is some i such that

$$a_1\cdots a_i = A_{i-1}a_i > A_{i-1}(a_i-1).$$

We conclude that $(\theta^i\langle(s-1)^n,0\rangle)_{0\le i\le T}$ is strictly decreasing in $\succ$. □

Observation 5. *When $n|k$ the operator θ yields a bijection between maximal pairs.*

Proof. Every pair of the form $\langle A,0\rangle$ is maximal because, when k is a multiple of n, there is just this unique rotation. The definition of θ ensures that θ goes through all the pairs of the form $\langle A,0\rangle$. in $\succ$-decreasing order. The operator θ can be used forward for every pair $\langle A,0\rangle$ except for $\langle 0^n,0\rangle$, and it can be used backwards for every pair $\langle A,0\rangle$ except for $\langle 0^n,0\rangle$. Thus, except for the extremes, we can obtain the successor and the predecessor of a maximal pair in the order $\succ$. □

When $k|n$ the operator θ is not injective nor surjective over pairs with second component 0. For example, for $s=2$, $n=4$ and $k=2$, we see θ is not injective because $\theta\langle 0100,0\rangle = \langle 0000,0\rangle$ and also $\theta\langle 0001,0\rangle = \langle 0000,0\rangle$. To see that θ is not surjective observe that the pair $\mathcal{B}=\langle 1011,0\rangle$ is not in the image of θ, because there is no pair $\mathcal{A}$ such that $\theta\mathcal{A}=\mathcal{B}$.

It is possible to construct the reverse of list $(\theta^i\langle(s-1)^n,0\rangle)_{0\le i\le T}$. In case $n|k$ divides k the operator θ yields a bijection between maximal pairs. In case $k|n$ and $k<n$, θ is not injective, there are pairs that have more than one preimage by θ. However, except for $\langle(s-1)^n,0\rangle$ every element has one predecessor in the list $(\theta^i(\langle s-1)^n,0\rangle)_{0\le i\le T}$, which is just one of the possible preimages by θ.

Lemma 3. *Let n and k positive integers such that $k|n$. Every element $\mathcal{A}=\langle A,0\rangle$ in $(\theta^i\langle(s-1)^n,0\rangle)_{0\le i\le T}$, except $\langle(s-1)^n,0\rangle$, has a predecessor which is the preimage of $\mathcal{A}$ by θ given by*

$$\langle A_{u-1}(a_u+1))0^{n-u},0\rangle$$

where

$$A=\left(A_u(s-1)^{r-u}\right)^w A_v$$

$a_u<(s-1)$ and r is the smallest multiple of k such that $v<r$ and $r-k\le u\le r$.

Proof. First notice that this factorization always exists

$$A=(A_r)^w A_v = (A_u(s-1)^{r-u})^w A_v.$$

If $w=1$ the $r=n$, $v=0$ and $A=A_n=A_r=A_u(s-1)^{r-u}$. Let $\mathcal{B}=\langle B,0\rangle$ be the pair obtained by undoing the transformation done by the θ operator, knowing that $\mathcal{A}$ and $\mathcal{B}$ are in $(\theta^i\langle(s-1)^n,0\rangle)_{0\le i\le T}$,

$$\theta\mathcal{B}=\langle [B_{i-1}(b_i-1)(s-1)^{j-i}]^q B_{n-qj},0\rangle = \mathcal{A}$$

where i satisfies $b_{i+1}=\ldots b_n=0$, and $i\le j$ and $k|j$. Thus,

$$[B_{i-1}(b_i-1)(s-1)^{j-i}]^q B_{n-qj} = (A_u(s-1)^{r-u})^w A_v.$$

The word B is determined by $r=j, u=i, w=q, v=n-qj$ and $A_{i-1}(a_i+1)=B_i$. Then,

$$\langle B_{i-1}b_i0^{n-j},0\rangle = \langle A_{u-1}(a_u+1))0^{n-u},0\rangle,$$
$$\langle [B_{i-1}(b_i-1)(s-1)^{j-i}]^q B_{n-qj},0\rangle = [A_{u-1}a_u(s-1)^{r-u}]^w A_v.$$

□

Lemma 4. *If $\mathcal{A} \succ \mathcal{B} \succ \mathcal{C}$, with $\mathcal{C} = \mathcal{A}\theta$, then $\mathcal{B}$ is not maximal .*

Proof. Let $\mathcal{A} = \langle A, 0\rangle$, $\mathcal{B} = \langle B, 0\rangle$, $\mathcal{C} = \mathcal{A}\theta = \langle C, 0\rangle$ and indices i, j such that

$$\begin{aligned} A &= a_1\cdots a_i0^{n-i}, \text{with } a_i>0,\\ B &= b_1\cdots b_n,\\ C &= c_1\cdots c_n = [a_1\cdots a_{i-1}(a_i-1)(s-1)^{j-i}]^q a_1\cdots a_{n-qj}.\end{aligned}$$

Assume $\mathcal{A} \succ \mathcal{B} \succ \mathcal{C}$ and, by way of contradiction suppose $\mathcal{B}$ is a maximal pair. Since $\mathcal{A} \succ \mathcal{B}$,

$$b_1 = a_1, \quad b_2 = a_2, \quad \ldots, \quad b_{i-1} = a_{i-1},$$

and it should be $a_i > b_i$. Since $\mathcal{B} \succ \mathcal{C}$ then $b_i \geq c_i = a_i - 1$; hence, $b_i = a_i - 1$. Then, we have $B_i = C_i$. And for $\ell = 1, \ldots j-i$ we have $b_{i+\ell} = s-1$. This is because $\mathcal{B} \succ \mathcal{C}$ and $c_{i+1}\cdots c_{i+\ell} = (s-1)^{j-i}$, then $b_{i+1}\cdots b_{i+\ell} = (s-1)^{j-i}$, given that $s-1$ is the lexicographically greatest symbol.

We now show that the other symbols in $\mathcal{B}$ and $\mathcal{C}$ also coincide. Since $\mathcal{B}$ is a maximal pair, $B \geq b_{j+1}\cdots b_n b_1\cdots b_j$ and we know that $a_1 = b_1 \geq b_{j+1}$. Since $\mathcal{B} \succ \mathcal{C}$ we have $b_{j+1} \geq c_{j+1} = a_1$, hence $b_{j+1} = a_1$. Repeating this argument we obtain for $m = 1, .., q$ and $p = 1, .., i-1$ we have:

$b_{mj+p} = a_p$, for $mj + p \leq n$,
$b_{mj+i} = a_i - 1$ and
$b_{mj+i+\ell} = s-1$ for $\ell = 1, .., j-i$.

This would imply $\mathcal{B} = \mathcal{C} = \theta\mathcal{A}$. We conclude that there is no maximal pair $\mathcal{B}$ such that $\mathcal{A} \succ \mathcal{B} \succ \mathcal{C}$. □

Definition (list $\mathcal{M}$ of maximal pairs of $\Sigma^n \times \mathbb{Z}_k$)**.** *Let n and k be positive integers such that $n|k$ or $k|n$. Define $\mathcal{M}$ by removing from $(\theta^i\langle (s-1)^n, 0\rangle)_{0\leq i\leq T}$ the pairs that are not maximal. In case $n|k$ all the elements all maximal, none is removed. In case $k|n$ with $k<n$, for each pair $\mathcal{A}$ in $(\theta^i\langle (s-1)^n, 0\rangle)_{0\leq i\leq T}$, except $\mathcal{A} = \langle 0^n, 0\rangle$, remove $\mathcal{A}\theta$, $\mathcal{A}\theta^2, \ldots \mathcal{A}\theta^{h-1}$ where h is the least such that $\theta^h(\mathcal{A})$ is maximal . Let M be the number of elements of the list $\mathcal{M}$.*

Lemma 5. *The list $\mathcal{M}$ starts with the $\succ$-greatest pair $\langle (s-1)^n, 0\rangle$, ends with the $\succ$-smallest pair $\langle 0^n, 0\rangle$, and contains all maximal pairs in strictly decreasing $\succ$-order.*

Proof of Lemma 5. Case $k \mid n$.

(1) By definition, the list $\mathcal{M}$ starts with $\langle (s-1)^n, 0\rangle$.
(2) The maximal pairs are in $\succ$-decreasing order: This is because the list $\mathcal{M}$ is constructed by successive applications of the operator θ and by Lemma 2 the list is strictly $\succ$-decreasing.
(3) No maximal pair is missing: Suppose $\mathcal{A}$ is in $\mathcal{M}$ and let h be the least such that $\theta^h(\mathcal{A})$ is maximal . To argue by contradiction, suppose there is a maximal pair $\mathcal{B}$ such that $\mathcal{A} \succ \mathcal{B} \succ \theta^h(\mathcal{A})$. Then, there is i, $0 \leq i < h$, such that

$$\mathcal{A} \succ \theta\mathcal{A} \succ \ldots \succ \theta^i\mathcal{A} \succ \mathcal{B} \succ \theta^{i+1}\mathcal{A} \succ \ldots \succ \theta^{h-1}\mathcal{A} \succ \theta^h\mathcal{A}.$$

Thus, $\mathcal{B}$ appears in between some $\mathcal{D}$ and $\theta\mathcal{D}$. By Lemma 4 this is impossible.

(4) The list $\mathcal{M}$ ends with $\langle 0^n, 0\rangle$: There is no $\mathcal{A}$ such that $\langle 0^n, 0\rangle \succ \mathcal{A}$, and the list $\mathcal{L}$ is strictly $\succ$-decreasing. By Lemma 2, the operator θ applies to any pair except $\langle 0^n, 0\rangle$.

Case $n \mid k$, $n < k$: It is the same proof as in the previous case, but simpler because θ yields exactly all the maximal pairs in $\succ$-decreasing order. $\square$

Example 1. *Let $s = 2$, $n = 6$ y $k = 2$. The list $\mathcal{M}$ of maximal pairs is:*
$\langle 111111, 0\rangle, \langle 111110, 0\rangle, \langle 111101, 0\rangle, \langle 111100, 0\rangle, \langle 111010, 0\rangle, \langle 111001, 0\rangle, \langle 111000, 0\rangle, \langle 110110, 0\rangle,$
$\langle 110101, 0\rangle, \langle 110100, 0\rangle, \langle 110010, 0\rangle, \langle 110001, 0\rangle, \langle 110000, 0\rangle, \langle 101010, 0\rangle, \langle 101001, 0\rangle, \langle 101000, 0\rangle,$
$\langle 100101, 0\rangle, \langle 100100, 0\rangle, \langle 100001, 0\rangle, \langle 100000, 0\rangle, \langle 010101, 0\rangle, \langle 010100, 0\rangle, \langle 010000, 0\rangle, \langle 000000, 0\rangle.$

The next is the key lemma.

Lemma 6. *If $\mathcal{A} = \langle A, 0\rangle$ is followed by $\mathcal{B} = \langle B, 0\rangle$ in the list of maximal pairs $\mathcal{M}$ then A is a prefix of $\overline{AB}$.*

Proof. We can write $\mathcal{A}$ as
$$\overline{\mathcal{A}}^q = \langle (A_i 0^{p-i})^q, 0\rangle,$$
where $q = n/p$ and $a_i > 0$. If $q = 1$ then $\overline{\mathcal{A}} = \mathcal{A} = \langle A, 0\rangle$ and $\overline{\mathcal{A}\mathcal{B}} = \langle A, 0\rangle\langle \overline{B}, 0\rangle = \langle A\overline{B}, 0\rangle$. Otherwise, $q > 1$ and, since $\mathcal{B} = \theta\mathcal{A}^h$ for the smallest h such that it is maximal , the shape of $\mathcal{B}$ is
$$\mathcal{B} = \langle \overline{A}^{q-1} A_{i-1}(a_i - 1)(s-1)^{j-i} C, 0\rangle$$
for some word C and for j the smallest such that $i \le j \le p$ and $k \mid j$. Since $\mathcal{B} = \langle B, 0\rangle$ and B starts with $\overline{A}^{q-1} A_{i-1}(a_i - 1)$ we have $\langle \overline{B}, 0\rangle = \langle B, 0\rangle$, hence $\overline{\mathcal{B}} = \mathcal{B}$. Then,
$$\overline{\mathcal{A}\mathcal{B}} = \langle \overline{A}, 0\rangle\langle \overline{B}, 0\rangle = \langle \overline{A}, 0\rangle\langle B, 0\rangle = \langle \overline{A}B, 0\rangle = \langle \overline{A}\overline{A}^{q-1} A_{i-1}(a_i - 1)(s-1)^{j-i} C, 0\rangle.$$
In both cases A is prefix of $\overline{AB}$. $\square$

3. Statement and proof of Theorem 1

Fix a finite alphabet Σ with cardinality $s \ge 2$. Recall that a necklace is a circular word and a necklace is (n, k)-perfect if all words of length n occur k times but at positions with different congruence modulo k, for any convention of the starting position.

Theorem 1. *Let n and k be positive integers such that $k|n$ or $n|k$. The concatenation of the words in Lyndon pairs ordered lexicographically yields the lexicographically greatest (n, k)-perfect necklace.*

Here is an example for $s = 2$, $n = 6$ y $k = 2$. The lexicographically greatest (n, k)-perfect necklace is obtained by concatenating the following words (the symbol | is just used here for ease of reading):

11 | 111110 | 111101 | 111100 | 111010 | 111001 | 111000 | 110110 | 110101 | 110100 | 110010 | 110001 | 110000 | 10 | 101001 | 101000 | 100101 | 100100 | 100001 | 100000 | 01 | 010100 | 010000 | 00

We nowgive the proof of Theorem 1. We start with a definition.

Definition (list $\mathcal{L}$ of Lyndon pairs)**.** *Let n and k be positive integers such that $k|n$ or $n|k$. We define the list $\mathcal{L}$ as the list of Lyndon pairs of the elements in the list $\mathcal{M}$. Since $\mathcal{M}$ has M elements, $\mathcal{L}$ has M elements as well.*

Assume $n|k$ or $k|n$. Let X be the concatenation of the words in Lyndon pairs in the order given by $\mathcal{L}$.

3.1. **Necklace X has length $s^n k$.** The length of X is

$$\sum_{\overline{\mathcal{A}} \text{ is Lyndon}} |\overline{\mathcal{A}}|.$$

Each Lyndon pair is the reduction of a maximal pair. The length of a Lyndon pair $\mathcal{A} = \langle \overline{A}, 0\rangle$ is the length of the word $\overline{A}$, which is the number of different rotations of $\mathcal{A}$, see Observation 4. The necklace X is defined by concatenating the words of all the reduced maximal pairs of $\Sigma^n \times \mathbb{Z}_k$ in $\succ$-decreasing order. Suppose $\overline{\mathcal{A}} = \langle \overline{A}, 0\rangle = \langle a_1 \cdots a_p, 0\rangle$. Then $\mathcal{A}$ has p different rotations : $\langle a_1 \cdots a_p, 0\rangle$, $\langle a_2 \cdots a_n a_1, 1\rangle$,..., $\langle a_p a_1 \cdots a_{n-1}, p-1\rangle$. So, each of the p symbols of $\overline{A}$ is the start of a different rotation of $\mathcal{A}$.

By Lemma 6 we know that if $\mathcal{A} = \langle A, 0\rangle$ is a maximal pair followed by the maximal pair $\mathcal{B} = \langle B, 0\rangle$ then in X we have $\overline{AB}$. and A is a prefix of $\overline{AB}$. Since we concatenate in X all the words of the reduced maximal pairs exactly once, we conclude that there is exactly one position for each of the different pairs in $\Sigma^n \times \mathbb{Z}_k$. Thus, the length of X is $s^n k$.

3.2. **Necklace X is (n,k)-perfect, with $k|n$.** To prove that X is a (n,k)-perfect necklace we need to show that each word of length n occurs exactly k times, at positions with different congruence modulo k. In this proof we number the positions of X starting at 0; this is convenient for the presentation because the positions with congruence 0 are multiple of k.

We say that we find a pair $\langle A, u\rangle$ in $X = x_0 x_1 ... x_{s^n k-1}$ when there is a position p in X such that $0 \leq p < s^n k$ and $x_p .. x_{p+|A|-1} = A$. To prove that X is a (n,k)-perfect necklace we need to find all the rotations of all the maximal pairs in $\Sigma^n \times \mathbb{Z}_k$ in the necklace X. Each maximal pair $\mathcal{A} = \langle A, 0\rangle$ has eaxactly $|\overline{A}|$ many rotations.

Case $A = (s-1)^n$. Since $k|n$, $\overline{\mathcal{A}} = \langle \overline{A}, 0\rangle = \langle (s-1)^k, 0\rangle$. Application of θ on $\mathcal{A}$ yields the maximal pair

$$\theta\mathcal{A} = \mathcal{B} = \langle B, 0\rangle = \langle (s-1)^{n-1}(s-2), 0\rangle.$$

Since $\overline{\mathcal{A}}$ and $\overline{\mathcal{B}}$ are consecutive Lyndon pairs in $\mathcal{L}$, the construction of X puts $\overline{A}$ followed by $\overline{B}$.

$$\overline{\mathcal{A}\mathcal{B}} = \langle (s-1)^k, 0\rangle \langle (s-1)^{n-1}(s-2), 0\rangle = \langle (s-1)^{n+k-1}(s-2), 0\rangle.$$

This yields all rotations of $\mathcal{A}$: $\langle (s-1)^n, 0\rangle$, .., $\langle (s-1)^n, k-1\rangle$.

Case $A = 0^n$. Since $k|n$, $\overline{\mathcal{A}} = \langle \overline{A}, 0\rangle = \langle 0^k, 0\rangle$. Consider the maximal pairs for $i = 0, 1, ..k-1$.

$$\mathcal{B} = \langle B, 0\rangle = \langle 0^i 1 0^{k-1-i} 0^{n-k}, 0\rangle.$$

For each of these $\mathcal{B}$ the next maximal pair is

$$\theta\mathcal{B} = \mathcal{C} = \langle C, 0\rangle = \langle (0^{i+1}(s-1)^{k-1-i})^{n/k}, 0\rangle.$$

Since $\overline{\mathcal{C}} = \langle 0^{i+1}(s-1)^{k-1-i}, 0\rangle$, it is clear that $\mathcal{C}$ is maximal because all the rotations of $\mathcal{C}$ that have second component 0 are identical to $\mathcal{C}$. Since $\overline{\mathcal{B}}$ and $\overline{\mathcal{C}}$ are consecutive Lyndon pairs in $\mathcal{L}$, the construction of X puts $\overline{B}$ followed by $\overline{C}$. Notice that for each $i = 0, 1, .., k-1$

$$\overline{\mathcal{B}\mathcal{C}} = \langle 0^i 1 0^{k-1-i} 0^{n-k} 0^{i+1} (s-1)^{k-i-1}, 0\rangle = \langle 0^i 1 0^n (s-1)^{k-i-1}, 0\rangle$$

gives rise to the pair $\langle 0^n, i+1\rangle$. Thus, we have all the rotations $\mathcal{A} = \langle 0^n, 0\rangle$, which are $\langle 0^n, 0\rangle$, .., $\langle 0^n, k-1\rangle$.

Case $(s-1)^n > A > 0^n$**.** Every maximal pair $\mathcal{A} = \langle A, 0\rangle$ different from $\langle 0^n, 0\rangle$ has the form

$$\mathcal{A} = \langle (A_p)^{q+1}, 0\rangle,$$

where A_p is $\overline{A}$ with p the smallest integer such that $A = (A_p)^{q+1}$ and $q+1 = (n/p)$.

Subcase $q > 0$. The pair $\mathcal{A}$, different from $\langle 0^n, 0\rangle$, always has a successor maximal pair in the list $\mathcal{M}$:

$$\mathcal{B} = \theta\mathcal{A} = \langle (A_p)^q A_{i-1}(a_i - 1)(s-1)^{j-i} b_{i+1} \cdots b_p, 0\rangle,$$

where i is the largest with $a_i > 0$, j is the smallest multiple of k with $j \geq i$. Notice that $A_p \neq 0^p$. Since $\overline{\mathcal{B}} = \mathcal{B}$,

$$\overline{\mathcal{A}\mathcal{B}} = \langle (A_p)^{q+1} A_{i-1}(a_i - 1)(s-1)^{j-i} b_{i+1} \cdots b_p, 0\rangle,$$

and it also yields the first i left rotations of $\mathcal{A}$, which are $\langle a_{r+1} \cdots a_n A_r, r\rangle$, with $0 \leq r < i$.

It remains to identify in our constructed necklace X the $p - i$ right rotations of $\mathcal{A} = \langle (A_p)^{q+1}, 0\rangle$. These are of the form

$$\langle a_{p-i+1} ... a_p A_p^q A_{p-i}, i\rangle,$$

To see this consider $\mathcal{Q}$ the predecessor of $\mathcal{A}$ by θ,

$$\theta\mathcal{Q} = \mathcal{A} = \langle (A_p)^{q+1}, 0\rangle$$

where

$$\mathcal{Q} = \langle A_{p-1}(a_p + 1)0^{pq}, 0\rangle.$$

To see that $\mathcal{Q}$ is a maximal pair consider first $\mathcal{R} = \langle A_p, 0\rangle$ which, by Observation 2, is maximal. We now argue that A_p has no proper suffix $a_i \cdots a_p$ which coincides with A_{p-i+1}. If there were such a prefix we could construct the rotation of $\mathcal{R}$ given by the pair $\mathcal{S} = \langle a_i \cdots a_p A_{i-1}, 0\rangle$ and one of the following would be true:

- $\mathcal{S} = \mathcal{R}$: but this is impossible by Observation 4.
- $\mathcal{R} \succ \mathcal{S}$: Since $\mathcal{R} = \langle A_{p-i+1} a_{p-i+2} \cdots a_p, 0\rangle$, $\mathcal{S} = \langle a_i ... a_p A_{i-1}, 0\rangle$ and we assumed $a_i \cdots a_p = A_{p-i+1}$, necessarily $a_{p-i+2} \cdots a_p > A_{i-1}$. But this contradicts that $\mathcal{R}$ is a maximal pair.
- $\mathcal{S} \succ \mathcal{R}$: There is a rotation of $\mathcal{R}$ which is $\succ$-greater than $\mathcal{R}$, contradicting that $\mathcal{R}$ is a maximal pair.

We conclude that all suffixes $a_i \cdots a_p$ of R are lexicographically smaller than A_{p-i}. Therefore, all suffixes of $a_i \cdots a_{p-1}(a_p + 1)$ of Q are lexicographically smaller than or equal to A_{p-i}. We already argue that $\mathcal{Q}$ is indeed maximal . For any rotation of $\mathcal{Q}$ of the form $\langle a_i \cdots a_{p-1}(a_p + 1)0^{pq} a_1 \cdots a_{i-1}, 0\rangle$, for $i > 1$, we argued that $a_i \cdots a_{p-1}(a_p + 1)$ is lexicographically smaller than or equal to A_{p-i}. Moreover, $a_i \cdots a_{p-1}(a_p + 1)0^{pq}$ is lexicographically smaller than $Q = A_{p-1}(a_p + 1)0^{pq}$. For any rotation of $\mathcal{Q}$ that starts with a suffix of 0^{pq} it can not be maximal , because for any m, $Q_m > 0^m$, otherwise $\mathcal{A}$ would not be maximal .

We have the maximal successive pairs $\mathcal{Q} = \langle Q, 0\rangle$, $\theta\mathcal{Q} = \mathcal{A} = \langle A, 0\rangle$ and $\mathcal{B} = \theta\mathcal{A} = \langle B, 0\rangle$. From the arguments above, $\overline{Q} = Q$ and $\overline{B} = B$. Since $\mathcal{A} = \langle (A_p)^{q+1}, 0\rangle$ and we assumed $q > 1$, A is not equal to $\overline{A}$. Then, $\overline{QAB} = Q\overline{A}B$. Observe that the last pq symbols of Q followed by $\overline{A}$ followed by the first pq symbols of B give rise to

$$\langle 0^{pq}(A_p)^{q+1}, 0\rangle,$$

where 0 is because p is a multiple of k. We now identify in this pair the $p-i$ rotations of $\mathcal{A}$ to the right. Since $\mathcal{A} = \langle A, 0\rangle$ where $A = (A_i 0^{p-i})^{q+1}$, after r rotations to the right of $\mathcal{A}$, for $r = 0, 1, .., p-i$, we obtain

$$\langle 0^r (A_p)^q A_i 0^{p-i-r}, -r\rangle.$$

Subcase $q = 0$. We need to see that for the maximal pairs $\mathcal{A} = \langle A, 0\rangle$ where A is reduced, that is $A = \overline{A}$, we can find all rotations of $\mathcal{A}$ in the constructed X.

If $A = 0^{k-1}10^{n-k}$ then $\theta\mathcal{A} = \mathcal{B} = \langle B, 0\rangle = \langle 0^n, 0\rangle$, which is the last maximal pair in the list $\mathcal{M}$ and $\overline{B} = 0^k$. Then,

$$\overline{AB} = 0^{k-1}10^{n-k}0^k = 0^{k-1}10^n,$$

and we can find k left rotations of $\mathcal{A}$ which are of the form

$$\langle 0^{k-1-r}10^{n-k+r}, r\rangle \text{ for } r = 0, 1, ..k-1.$$

Now assume $A \neq 0^{k-1}10^{n-k}$. Let's write

$$\mathcal{A} = \langle A_i 0^{n-i}, 0\rangle,$$

with i such that $a_i > a_{i+1} = \ldots = a_n = 0$, $\mathcal{B} = \langle B, 0\rangle = \theta\mathcal{A}$ and $\mathcal{C} = \langle C, 0\rangle = \theta\mathcal{B} = \theta^2\mathcal{A}$. Then $\mathcal{B}$ is of the form

$$\mathcal{B} = \langle A_{i-1}(a_i - 1)(s-1)^{j-i} b_{j+1} \cdots b_p, 0\rangle,$$

with j is the least multiple of k with $j \geq i$. Since $A = \overline{A}$, and by Lemma 6 B is a prefix of $\overline{BC}$, then AB is a prefix of $\overline{ABC}$, and we can find the first i left rotations of $\mathcal{A}$ in it, which are of the form

$$\langle a_{r+1} \cdots a_n A_r, r\rangle, \text{ for } r = 0, 1, .., i-1.$$

It remains to find $n-i$ right rotations of $\mathcal{A}$, which are of the form

$$\langle 0^r A_i 0^{n-i-r}, -r\rangle$$

for $r = 1, \ldots, n-i$. Equivalently we can write it as

$$\langle 0^{n-i-h} A_i 0^h, (n-i-h)\rangle$$

for $h = 0, 1, \ldots, n-i-1$.

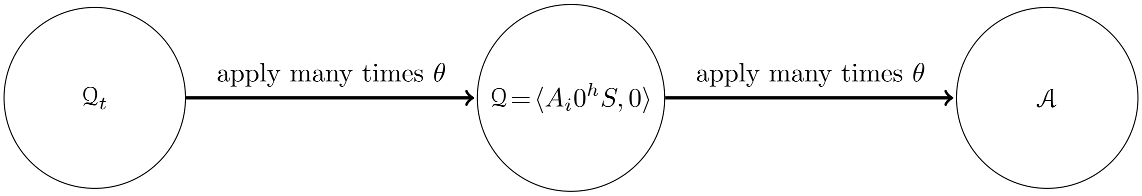
If $h = 0$ and $A_i = (s-1)^i$ then the pair is $\langle 0^{n-i}(s-1)^i, n-i\rangle$. We find it in the pair $\langle 0^n(s-1)^n, 0\rangle$, which results from the concatenation of the words in the last two Lyndon pairs in $\mathcal{L}$ and the first two, that is $\mathcal{L}_{M-1}$ and $\mathcal{L}_M$, followed by $\mathcal{L}_1$ and $\mathcal{L}_2$.

If $h \neq 0$ or $A_i \neq (s-1)^i$ then we find the right rotations of $\mathcal{A}$ in the concatenation of the words of three Lyndon pairs, that we call $\overline{\mathcal{Q}_t\mathcal{P}\mathcal{R}}$ and we define below. Recall that $\mathcal{A} = \langle A, 0\rangle = \langle A_i 0^{n-i}, 0\rangle$, with i such that $a_i > a_{i+1} = \ldots = a_n = 0$, is a maximal pair. We claim there is a unique pair $\mathcal{Q}$ of the form

$$\mathcal{Q} = \langle A_i 0^h S, 0\rangle,$$

where $|S| = n - (i+h)$ is not zero. Notice that $\mathcal{Q}$ may not be maximal.

Since $\mathcal{Q}$ is a pair $\succ$-greater than $\mathcal{A}$, we know $\mathcal{Q}$ is a predecessor of $\mathcal{A}$ by θ. The pair $\mathcal{Q}_t$ is the closest predecessor of $\mathcal{Q}$ by θ that is a maximal pair. For this, consider the succesive application of the operator θ, which allows us to traverse the list $(\theta^l\langle(s-1)^n, 0\rangle)_{1\leq l\leq T}$. Our interest is to traverse it backwards. Here is a diagram:



For $l = 1, 2, 3, \ldots$ let $\mathcal{Q}_l = \langle Q_l, 0\rangle$ be the pair such that $\theta^l \mathcal{Q}_l = \mathcal{Q}$. For $l = 1$,

$$\mathcal{Q}_1 = \langle Q_1, 0\rangle = \langle A_{u_1-1}(a_{u_1}+1)0^{n-u_1}, 0\rangle,$$

$$\theta\mathcal{Q}_1 = \langle (A_{u_1}(s-1)^{r_1-u_1})^{w_1} A_{v_1} S, 0\rangle = \langle (A_{r_1})^{w_1} A_{v_1} S, 0\rangle = \langle A_i 0^h S, 0\rangle = \mathcal{Q},$$

where r_1 is the least multiple of k such that $v_1 < r_1$, $a_{u_1} < (s-1)$, $A_{r_1} = A_{u_1}(s-1)^{r_1-u_1}$ and $r_1 - k \le u_1 \le r_1$. If $\mathcal{Q}_1$ is a maximal pair then fix $t = 1$ and we have finished the search. Otherwise we consider the predecessor of $\mathcal{Q}_1$ by θ,

$$\mathcal{Q}_2 = \langle Q_2, 0\rangle = \langle A_{u_2-1}(a_{u_2}+1)0^{n-u_2}, 0\rangle$$

such that

$$\theta\mathcal{Q}_2 = \langle (A_{u_2}(s-1)^{r_2-u_2})^{w_2} A_{v_2}, 0\rangle = \langle (A_{r_2})^{w_2} A_{v_2}, 0\rangle = \mathcal{Q}_1,$$

where r_2 is the least multiple of k, $v_2 < r_2$, and $a_{u_2} < (s-1)$ such that $A_{r_2} = A_{u_2}(s-1)^{r_2-u_2}$. We know that $a_{u_2} < (s-1)$ exists because $\mathcal{Q}_1 \ne \langle (s-1)^n, 0\rangle$ (otherwise $\mathcal{Q}_1$ would be a maximal pair), and then there is $a_i < s-1$ for some $1 \le i \le r_2$. Notice that $u_2 < u_1$. If $\mathcal{Q}_2$ is a maximal pair then we fix $t = 2$ and we have finished the search. Otherwise, we repeat this procedure. In this way the predecessor of $\mathcal{Q}_{l-1}$ by θ is

$$\mathcal{Q}_l = \langle Q_l, 0\rangle = \langle A_{u_l-1}(a_{u_l}+1)0^{n-u_l}, 0\rangle,$$

such that

$$\theta\mathcal{Q}_l = \langle (A_{u_l}(s-1)^{r_l-u_l})^{w_l} A_{v_l}, 0\rangle = \langle (A_{r_l})^{w_l} A_{v_l}, 0\rangle = \mathcal{Q}_{l-1},$$

where r_l is multiple of k, $v_l < r_l$, $u_l < u_{l-1}$ and $a_{u_l} < (s-1)$ such that

$$A_{r_l} = A_{u_l}(s-1)^{r_l-u_l}.$$

Eventually we find t such that $\mathcal{Q}_t$ is a maximal pair.

Consider now the three consecutive maximal pairs in the list $\mathcal{M}$,

$$\mathcal{Q}_t = \langle Q_t, 0\rangle, \mathcal{P} = \langle P, 0\rangle, \text{ and } \mathcal{R} = \langle R, 0\rangle.$$

Let $\overline{\mathcal{Q}_t}, \overline{\mathcal{P}}$ and $\overline{\mathcal{R}}$ be the corresponding Lyndon pairs. $\mathcal{P}$ always exists, because it's either $\mathcal{A}$ or a maximal pair before it, and $\mathcal{R}$ is $\langle 0^n, 0\rangle$ in the worst case. Notice that $\overline{Q_t}$ ends with 0^{n-u_t} and, by Lemma 6, $\overline{PR}$ starts with P. Therefore, $\overline{\mathcal{Q}_t \mathcal{P}\mathcal{R}}$ contains $0^{n-u_t}P = 0^{n-u_t}A_i 0^h C$, for some C. Since $\mathcal{P}$ is a maximal pair $\succ$-greater than or equal to $\mathcal{A}$, we can assert that its prefix is $A_i 0^h$. Finally, we have

$$u_t \le u_1 \le i + h,$$

because $u_t < u_{t-1} < \ldots < u_1$ and $u_1 \le i + h$ because u_1 was the position of a symbol in $A_i 0^h$, which has length $i + h$. Then,

$$\begin{aligned} u_t \le u_1 \le i + h &\iff -u_t \ge -u_1 \ge -i - h \\ &\iff n - u_t \ge n - u_1 \ge n - i - h \\ &\iff n - u_t \ge n - i - h. \end{aligned}$$

We conclude that $\overline{\mathcal{Q}_t\mathcal{P}\mathcal{R}}$ contains

$$\langle 0^{n-i-h}A_i0^h, (|\overline{Q_h}| - (n-i-h))\rangle.$$

This can be rewritten as the pair

$$\langle 0^{n-i-h}A_i0^h, -(n-i-h)\rangle,$$

because $|\overline{Q_t}|$ is a multiple of k, hence $\mathcal{Q}_t$ has second component 0.

3.3. **Necklace X is (n,k)-perfect, with $n|k$, $n<k$.** This proof is similar to the the proof for the case $k|n$, but it uses the corresponding definitions of θ, the list $\mathcal{M}$ and the notion of expansion for the list $\mathcal{L}$. The proof becomes simpler because it requires fewer cases and each case is easier. To find the rotations of $\langle 0^n, 0\rangle$ is similar to the case $k|n$. To find the rotations to the left for pairs $\langle A, 0\rangle$ such that $(s-1)^n > A > 0^n$ is also similar to the case $k|n$, because the number of positions i to be considered in the two definitions of θ coincide. However, to find rotations to the right now is much simpler than in the case $k|n$ because the operator θ is now a bijection. Given that all the pairs obtained by application of θ are maximal pairs, for each pair we just need to find the predecessor by θ. In contrast, when $k|n$, the operator θ is not bijective and we had to justify how we find the maximal pair $\mathcal{Q}_t$.

3.4. **Necklace X is the lexicographically greatest (n,k)-perfect necklace.** Given an (n,k)-perfect necklace fix the starting position 0 in the first symbol of the occurrence of $(s-1)^n$ that yields the lexicographically maximal rotation $a_0 \dots a_{s^nk-1}$.. For each position i such that $0 \le i \le ks^n - 1$ define the pair $\langle a_i...a_{i+n-1}, i \mod k\rangle$. This of course includes the pairs starting at positions s^nk-n+1, ..., s^nk-1 because we are dealing with a necklace. Notice that the symbol a_{i+n-1} is not just in the pair starting at position i, but also also in the pairs starting at positions $i+1, ..., i+n-1$. The definition of (n,k)-perfect necklace ensures that this set of pairs is exactly $\Sigma^n \times \mathbb{Z}_k$. In the lexicographically greatest (n,k)-perfect necklace, as i increases from 0 to s^nk-1, at each position i we find the $\succ$-greatest remaining pair $\langle a_i...a_{i+n-1}, i \mod k\rangle$ given that the symbols $a_i, .., a_{i+n-2}$ have already been determined. Thus, the choices for the maximal pair starting at position i are just the choices for the symbol a_{i+n-1}. Let's see how our construction achieves this.

Recall that M is the number of maximal pairs in $\Sigma^n \times \mathbb{Z}_k$ with second component 0. This is exactly the length of the list $\mathcal{M}$ and the length of the list $\mathcal{L}$. Our construction concatenates all the words in Lyndon pairs of $\Sigma^n \times \mathbb{Z}_k$ in $\succ$-decreasing order, given in $\mathcal{L} = (\langle w_i, 0\rangle)_{1\le i\le M}$. So, $\mathcal{L}$ is the list of all the reduced maximal pairs of $\Sigma^n \times \mathbb{Z}_k$ with second component 0 in $\succ$-decreasing order. Let $\ell_0 = 0$ and $(\ell)_{1\le i\le M}$ be sum of the lengths of the first i Lyndon pairs in the list $\mathcal{L}$,

$$\ell_i = \sum_{j=1}^{i} |\mathcal{L}_j|.$$

As we already showed, $\ell_M = s^nk$. Our construction yields the word $X = a_0, a_1, \dots a_{s^nk-1}$ where $\mathcal{L}_1 = \langle a_0, \dots a_{\ell_1-1}, 0\rangle$, $\mathcal{L}_2 = \langle a_{\ell_1}, \dots a_{\ell_2-1}, 0\rangle$, $\dots$, $\mathcal{L}_M = \langle a_{\ell_{M-1}}, \dots, \dots a_{\ell_M-1}, 0\rangle$. That is, for each i such that $0 \le i \le s^nk-1$, $\mathcal{L}_i = \langle w_i, 0\rangle$, where $w_i = a_{\ell_{i-1}}, \dots a_{\ell_i-1}$, w_i is lexicographically greater than all of its rotations, and w_i is lexicographically greater than all the rotations of all the w_j, for $j > i$. Our construction greedily picks, one after the other, the $\succ$-greatest remaining reduced maximal pair with second component 0. By Lemma 6 we know that the concatenation of the reduced maximal pairs in decreasing $\succ$-order gives rise to the same pairs without reduction.

Implicitly our construction defines a function $f : \Sigma^n \to \Sigma$, such that

$$a_{i+n} = f(a_i, a_{i+1}, ..., a_{i+n-1}).$$

Since the $\succ$-order among the pairs with the same second component is just the lexicographical order, $a_0, a_2, \ldots a_{s^n k-1}$ is the lexicographically greatest among all (n, k)-perfect necklaces because in position $i+n$ the symbol a_{i+n} is the lexicographically greatest among all the symbols that could have been allocated to define a perfect necklace. Consequently, if $b_0, b_1, \ldots, b_{s^n k-1}$ is another (n, k)-perfect necklace then, necessarily, there is a position p such that $b_0 = a_0, b_1 = a_1, \ldots, b_{p-1} = a_{p-1}$ but $a_p > b_p$. This completes the proof of Theorem 1.

Acknowledgements

This work is supported by the projects CONICET PIP 11220210100220CO and UBACyT 20020220100065BA.

References

[1] N. Álvarez, V. Becher, P. Ferrari, and S. Yuhjtman. Perfect necklaces. *Advances in Applied Mathematics*, 80:48 – 61, 2016.

[2] V. Becher and O. Carton. Normal numbers and perfect necklaces. *Journal of Complexity*, 54(101403), 2019.

[3] N. G. de Bruijn. A combinatorial problem. *Koninklijke Nederlandse Akademie v. Wetenschappen*, 49:758–764, 1946. Indagationes Mathematicae 8 (1946) 461-467.

[4] J.-P. Duval. Mots de Lyndon et périodicité. *RAIRO Informatique Théorique*, 14(2):181–191, 1980.

[5] H. Fredricksen and J. Maiorana. Necklaces of beads in k colors and k-ary de Bruijn sequences. *Discrete Mathematics*, 23(3):207–210, 1978.

[6] D. Knuth. *The art of computer programming. Volume 3.* Addison-Wesley Series in Computer Science and Information Processing. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. Sorting and searching.

[7] N. Korobov. Normal periodic systems and their applications to the estimation of sums of fractional parts. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 15(1):17–46, 1951.

[8] N. Korobov. On normal periodic systems. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 16(3):211–216, 1952.

[9] M. B. Levin. On the discrepancy estimate of normal numbers. *Acta Arithmetica*, 88(2):99–111, 1999.

[10] R. C. Lyndon. On Burnside's problem. *Transactions of the American Mathematical Society*, 77:202–215, 1954.

[11] R. C. Lyndon. On Burnside's problem. II. *Transactions of the American Mathematical Society*, 78:329–332, 1955.