# Gray code orders for $q$-ary words avoiding a given factor

A. Bernini* S. Bilotta* R. Pinzani* A. Sabri† V. Vajnovszki†

April 17, 2018



**Abstract**

Based on BRGC inspired order relations we define Gray codes and give a generating algorithm for $q$-ary words avoiding a prescribed factor. These generalize an early 2001 result and a very recent one published by some of the present authors, and can be seen as an alternative to those of Squire published in 1996. Among the involved tools, we make use of generalized BRGC order relations, ultimate periodicity of infinite words, and word matching techniques.

## 1 Introduction

A very special way for listing a class of combinatorial objects is the so called *combinatorial Gray code*, where two consecutive objects differ 'in some pre-specified small way' [7]. In [17] a general definition is given, where a Gray code is defined as 'an infinite set of word-lists with unbounded word-length such that the Hamming distance between any two adjacent words is bounded independently of the word-length' (the Hamming distance is the number of positions in which the words differ).

In [6] Guibas and Odlyzko enumerated the set of length $n$ words avoiding an arbitrary factor, and a systematic construction and enumeration results for particular factor avoidance in binary case are considered in [2, 3]. In Gray code context, Squire in his early paper [12] explores the possibility of listing factor avoiding words such that consecutive words differ in only one position, and by 1 or $-1$ in this position, and in [14] is given a Gray code and a generating algorithm for binary words avoiding $\ell$ consecutive 1s. The result in [14] was recently generalized in [1] where two Gray codes (one prefix partitioned and the other trace partitioned) for $q$-ary words avoiding a factor constituted by $\ell$ consecutive equal symbols are given.

Here, we adopt a different approach by relaxing Squire's 'one position constraint' and give Gray codes for length $n$ words avoiding any given factor, where consecutive words differ in at most three positions. Our definitions for these Gray codes are based on two order relations inspired from the original *Binary Reflected Gray Code* [5]; similar techniques were used previously (less or more explicitly) for other combinatorial classes, see for example [18, 11, 15] and the references therein. More precisely, we characterize forbidden factors inducing *zero periodicity* (defined later), which is a crucial notion for our construction; and we show that the zero periodicity property of a forbidden factor is a sufficient condition for the set of words avoiding this

*Dipartimento di Matematica e Informatica "Ulisse Dini", Università degli Studi di Firenze, Viale G.B. Morgagni 65, 50134 Firenze, Italy. {bernini}{bilotta}{pinzani}@dsi.unifi.it

†LE2I, Université de Bourgogne, BP 47 870, 21078 Dijon Cedex, France. {ahmad.sabri}{vvajnov}@u-bourgogne.fr

factor when listed in the appropriate order to be a Gray code. However, this is not a necessary condition and we show that there are forbidden factors with no zero periodicity property, and the set of words avoiding one of them when listed in the appropriate order yields a Gray code. Also, among all $q^\ell$ forbidden factors of length $\ell$ on a $q$-ary alphabet, all but $\ell+1$ of them induce zero periodicity; and when a Gray code is prohibited by lack of zero periodicity property of the forbidden factor, we give a simple transformation of this factor which allows to eventually obtain the desired Gray code. Finally, we give a constant average time generating algorithm for these Gray codes using $\ell \cdot q$ extra space and a Knuth-Morris-Pratt word matching technique [8]. A `C` implementation of the obtained algorithm is on the web site of the last author [16].

Although in [9] it is proved that the set of words avoiding a given factor is 'reflectable' under some conditions on the alphabet cardinality and the forbidden factor, our construction yields Gray codes for any alphabet and forbidden factor, and has a natural algorithmic implementation.

## 2 Notations and definitions

### Words over a finite alphabet

An *alphabet* $A$ is simply a set of symbols, and a length $n$ *word* is a function $\{1, 2, \dots, n\} \to A$, and $\epsilon$ is the empty (i.e., length zero) word. We adopt the convention that lower case bold letters represent words, for example $\boldsymbol{a} = a_1 a_2 \dots a_n$; and $A^n$ denotes the set of words of length $n$ over $A$, $A^* = \cup_{n \geq 0} A^n$ and $A^+ = \cup_{n \geq 1} A^n$. For $\boldsymbol{a} \in A^*$, $|\boldsymbol{a}|$ denotes the length of $\boldsymbol{a}$, or equivalently, the number of symbols in $\boldsymbol{a}$, and $|\boldsymbol{a}|_{\neq 0}$ the number of non-zero symbols in $\boldsymbol{a}$. An infinite word is a function $\mathbb{N} \to A$, and $A^\infty$ is the set of infinite words over $A$. For $\boldsymbol{a} \in A^*$, $i \geq 0$, $\boldsymbol{a}^i$ is the word obtained by $i$ repetitions of $\boldsymbol{a}$ ($\boldsymbol{a}^0$ being the empty word $\epsilon$) and $\boldsymbol{a}^\infty$ is the infinite *periodic* word $\boldsymbol{aaa}\dots$. The word $\boldsymbol{a} \in A^\infty$ is *ultimately periodic* if there are $\boldsymbol{b} \in A^*$ and $\boldsymbol{c} \in A^+$ such that $\boldsymbol{a} = \boldsymbol{b}\boldsymbol{c}^\infty$, and we say that $\boldsymbol{a}$ has *ultimate period* $\boldsymbol{c}$. Incidentally, we will make use of *left infinite words*, which are infinite words $\boldsymbol{a}$ of the form $\boldsymbol{a} = \dots a_{-3} a_{-2} a_{-1}$. Left infinite words are reverse of infinite words, and formally a left infinite word is a function $\{\dots, -3, -2, -1\} \to A$, and for $\boldsymbol{a} \in A^*$, $\boldsymbol{a}^{-\infty}$ is the left infinite word $\dots \boldsymbol{aaa}$.

The word $\boldsymbol{f}$ is a *factor* of the word $\boldsymbol{a}$ if there are words $\boldsymbol{b}$ and $\boldsymbol{c}$ such that $\boldsymbol{a} = \boldsymbol{bfc}$; when $\boldsymbol{b} = \epsilon$ (resp. $\boldsymbol{c} = \epsilon$), then $\boldsymbol{f}$ is a *prefix* (resp. *suffix*) of $\boldsymbol{a}$; and in this case, the prefix or the suffix is *proper* if $\boldsymbol{f} \neq \boldsymbol{a}$ and $\boldsymbol{f} \neq \epsilon$.

For a word $\boldsymbol{a}$ and a set of words $X$ we denote by $\boldsymbol{a}|X$ the set of words in $X$ having the prefix $\boldsymbol{a}$, and by $X(\boldsymbol{a})$ those avoiding $\boldsymbol{a}$, i.e, the words in $X$ which do not contain $\boldsymbol{a}$ as a factor. Thus, for example, $\boldsymbol{p}|X(\boldsymbol{f})$ is the set of words in $X$ having prefix $\boldsymbol{p}$ and avoiding $\boldsymbol{f}$. Clearly $A^*(\boldsymbol{a}) = \cup_{n \geq 0} A^n(\boldsymbol{a})$.

Through this paper we consider the alphabet $A_q = \{0, 1, \dots, q-1\}$ with $q \geq 2$.

### Gray codes

We will adopt the following definition: a list of same length words is a Gray code if there is a $d$ such that the Hamming distance between any consecutive words in the list is bounded from above by $d$; and often we refer to this list as a $d$-Gray code, and so, for example, a 3-Gray code is also a 4-Gray code. In addition, if for any two consecutive words in the list the leftmost and the rightmost positions where they differ are separated by at most $e-1$ symbols, then the Gray code is called *$e$-close*.

## Order relations

Our constructions of Gray codes for factor avoiding words are based on two order relations on $A_q^n$ we will define below. The first one captures the order induced by *$q$-ary Reflected Gray Code* [4], which is the natural extension of *Binary Reflected Gray Code* introduced by Frank Gray [5]; and the second one is a variation of the previous one.

**Definition 1** ([14, 15, 11])**.** Let $\boldsymbol{s} = s_1 s_2 \dots s_n$ and $\boldsymbol{t} = t_1 t_2 \dots t_n$ be two words in $A_q^n$, $k$ be the leftmost position where they differ, and $u = \sum_{i=1}^{k-1} s_i = \sum_{i=1}^{k-1} t_i$. We say that $\boldsymbol{s}$ is less than $\boldsymbol{t}$ in *Reflected Gray Code order*, denoted by $\boldsymbol{s} \prec \boldsymbol{t}$, if either

- $u$ is even and $s_k < t_k$, or
- $u$ is odd and $s_k > t_k$.

It follows that the set $A_q^n$ listed in $\prec$ order yields precisely the $q$-ary Reflected Gray Code (see [4, 19]) where two consecutive words differ in one position and by 1 or $-1$ in this position.

For a set of same length words $X$, we refer to *$\prec$-first* (resp.*$\prec$-last*) word in $X$ for the first (resp. last) word in $X$ with respect to $\prec$ order.

In the binary case, Definition 1 can be re-expressed as following. For $\boldsymbol{s}$, $\boldsymbol{t}$ and $k$ as in Definition 1, let $v$ be the number of non-zero symbols in the length $k-1$ prefix of $\boldsymbol{s}$ and of $\boldsymbol{t}$. Then $\boldsymbol{s} \prec \boldsymbol{t}$ if either

- $v$ is even and $s_k < t_k$, or
- $v$ is odd and $s_k > t_k$.

By 'adding' $u$ in Definition 1 and $v$ defined above we obtain a new order relation.

**Definition 2.** Let $\boldsymbol{s} = s_1 s_2 \dots s_n$ and $\boldsymbol{t} = t_1 t_2 \dots t_n$ be two words in $A_q^n$, $k$ be the leftmost position where they differ, $u = \sum_{i=1}^{k-1} s_i = \sum_{i=1}^{k-1} t_i$, and $v$ be the number of non-zero symbols in the length $k-1$ prefix of $\boldsymbol{s}$. We say that $\boldsymbol{s}$ is less than $\boldsymbol{t}$ in *Dual Reflected Gray Code order*, denoted by $\boldsymbol{s} \lhd \boldsymbol{t}$, if either

- $u + v$ is even and $s_k < t_k$, or
- $u + v$ is odd and $s_k > t_k$.

Clearly, listing a set of words in $\prec$ or in $\lhd$ order gives a prefix partitioned list, in the sense that words with the same prefix are consecutive. See Table 2 in Appendix for the set $A_3^4$ listed in $\lhd$ order.

For a set of same length words $X$, we refer to *$\lhd$-first* (resp.*$\lhd$-last*) word in $X$ for the first (resp. last) word in $X$ with respect to $\lhd$ order. In the following, without explicitly precise otherwise we will consider $\prec$ order on $A_q^*$ when $q$ is even and $\lhd$ order when $q$ is odd.

The next remark says that $\lhd$ order yields a Gray code when $q$ is odd, but generally for $q$ even, $A_q^n$ listed in $\lhd$ order is not a Gray code.

**Remark 1.** For any odd $q \geq 3$ and $n \geq 1$, the set $A_q^n$ listed in $\lhd$ order is a Gray code where two consecutive words differ in at most two adjacent positions. In addition, if $\boldsymbol{s}$ and $\boldsymbol{t}$ are two consecutive words in this list and $k$ is the leftmost position where they differ, then

- $s_k = t_k \pm 1$, and
- if $s_{k+1} \neq t_{k+1}$, then $\{s_{k+1}, t_{k+1}\} = \{0, q-1\}$.

We define the *parity* of the word $\boldsymbol{a} = a_1 a_2 \dots a_n \in A_q^n$ according to two cases:

- when $q$ is even, the parity of $\boldsymbol{a}$ is the parity of the integer $\sum_{i=1}^{n} a_i$, and
- when $q$ is odd, the parity of $\boldsymbol{a}$ is the parity of the integer $\sum_{i=1}^{n} a_i + |\boldsymbol{a}|_{\neq 0}$.

For example, if $\boldsymbol{a} = 0222$, then

- considering $\boldsymbol{a} \in A_4^*$, the parity of $\boldsymbol{a}$ is even, and given by $0+2+2+2=6$, and
- considering $\boldsymbol{a} \in A_5^*$, the parity of $\boldsymbol{a}$ is odd, and given by $0+2+2+2+3=9$.

Now we introduce a critical concept for our purposes: we say that the forbidden factor $\boldsymbol{f} \in A_q^*$ *induces zero periodicity* on $A_q^\infty$ if for any $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$ the first and the last (as mentioned previously, with respect to $\prec$ order for even $q$, or $\lhd$ order for odd $q$) words in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ both have ultimate period 0. Consequently, if $\boldsymbol{f} \in A_q^*$ *does not* induce zero periodicity on $A_q^\infty$, it follows that there exists a $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$ such that the first and/or the last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ do not have ultimate period 0.

Whether or not $\boldsymbol{f}$ induces zero periodicity on $A_q^\infty$ depends on $q$; and $q$ will often be understood from the context. For example $\boldsymbol{f} = 3130$ induces zero periodicity on $A_6^\infty$ but not on $A_4^\infty$. Indeed, the first word in $\prec$ order in $A_6^\infty(\boldsymbol{f})$ and having prefix 313 is $3135000\dots$, and the last one is $3131500\dots$; whereas the last word in $\prec$ order in $A_4^\infty(\boldsymbol{f})$ and with the same prefix is the periodic word $31313\dots$.

In Section 4 it is shown that the *Graycodeness* of the set $A_q^n(\boldsymbol{f})$ listed in the appropriate order is intimately related to that $\boldsymbol{f}$ induces zero periodicity.

## Outline of the paper

Avoiding a factor of length one is equivalent to shrink the underlying alphabet, but for the sake of generality, we will consider most of the time forbidden factors of any positive length.

In the next section we will characterize the forbidden factors $\boldsymbol{f}$ inducing zero periodicity on $A_q^\infty$. By means of three sets $U_q, V, W_q \subset A_q^+$, these factors are characterized in Corollary 1. The non-zero ultimate periods produced by forbidden factors that do not induce zero periodicity have the form $1(q-1)0^m$, $10^m$ or $(q-2)$, see Propositions 1, 3 and 5.

Theorems 4 to 6 and Proposition 12 in Section 4 prove that the property of $\boldsymbol{f}$ to induce zero periodicity guarantees the set $A_q^n(\boldsymbol{f})$ listed in the appropriate order to be a Gray code. However, this property of $\boldsymbol{f}$ is not a necessary condition: there are two 'special' forbidden factors, namely $\boldsymbol{f} = 0^\ell$ and $\boldsymbol{f} = (q-1)0^\ell$ belonging to $U_q$ (and so, which do not induce zero periodicity) but $A_q^n(\boldsymbol{f})$ listed in $\prec$ order is still a Gray code. These two cases are discussed in Section 4.2. Table 1 summarizes the Graycodeness for the set $A_q^n(\boldsymbol{f})$ if $\boldsymbol{f}$ does not induce zero periodicity. Section 4 ends by showing that simple transformations of forbidden factors $\boldsymbol{f}$ which do not induce zero periodicity allow to obtain Gray code for the set $A_q^n(\boldsymbol{f})$.

Finally, we present in Section 5 an efficient generating algorithm for the obtained Gray codes.

# 3 Periodicity

As stated above, without restriction, the set $A_q^n$ listed in $\prec$ order is a 1-Gray code for any $q \geq 2$, and listed in $\lhd$ order a 2-Gray code for odd $q \geq 3$. Roughly, it is due to that for any $\boldsymbol{p} \in A_q^*$ the first and the last words—with respect to $\prec$ order for any $q \geq 2$, or $\lhd$ order for odd $q \geq 3$—in the set $\boldsymbol{p}|A_q^\infty$ are among the three infinite words: $\boldsymbol{p}0^\infty$, $\boldsymbol{p}(q-1)0^\infty$, and $\boldsymbol{p}(q-1)^\infty$. This phenomenon is no longer true if an arbitrary factor $\boldsymbol{f}$ is forbidden. For example, if $\boldsymbol{f} = 130$, then the $\prec$-last word in $03|A_4^\infty(\boldsymbol{f})$ is $0300000\ldots$, and the $\prec$-first one in $13|A_4^\infty(\boldsymbol{f})$ is $1313131\ldots$; and 0300000 and 1313131 are consecutive words in $A_4^7(\boldsymbol{f})$, in $\prec$ order. Or, for $\boldsymbol{f} = 223$, the $\prec$-last word in $123|A_4^\infty(\boldsymbol{f})$ is $123300\ldots$ and the $\prec$-first one in $122|A_4^\infty(\boldsymbol{f})$ is $122222\ldots$; and 123300 and 122222 are consecutive words in $A_4^6(\boldsymbol{f})$, in $\prec$ order.

However, it is easy to understand the next remark.

**Remark 2.** If $\boldsymbol{f} \in A_q^*$ is a forbidden factor ending by a symbol other than 0 or $q-1$, then for any $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$, $q \geq 2$ and even (resp. $q \geq 3$ and odd), the set formed by the $\prec$-first and the $\prec$-last (resp. the $\lhd$-first and the $\lhd$-last) words in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ is $\{\boldsymbol{p}0^\infty, \boldsymbol{p}(q-1)0^\infty\}$.

In other words, the previous remark says that any factor ending by a symbol other than 0 or $q-1$ induces zero periodicity. However, there exist forbidden factors ending by 0 or $q-1$ that do induce zero periodicity. For example, with $\boldsymbol{f} = 120$, the $\prec$-first and $\prec$-last words in $12|A_4^\infty(\boldsymbol{f})$ are $1230000\ldots$ and $1213000\ldots$.

In the following we will use (often implicitly) the next straightforward remark which provides the form of the words on the right of a fixed prefix $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$, with respect to the appropriate order. It is obtained by the following observation: the first/last word in $\boldsymbol{p}|A_q^n(\boldsymbol{f})$ is the appropriate prefix of the first/last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$.

**Remark 3.** Let $q$ be even, $\boldsymbol{f} \in A_q^*$ be a forbidden factor, $\boldsymbol{p}, \boldsymbol{r} \in A_q^*(\boldsymbol{f})$, and let $\boldsymbol{p}$ have even (resp. odd) parity such that $\boldsymbol{pr} \in A_q^*(\boldsymbol{f})$. Then:

- If $\boldsymbol{pr}$ is a prefix of the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, then $\boldsymbol{r}$ is the smallest word, in $\prec$ order, with this property; that is, if $\boldsymbol{s} \in A_q^*$ with $|\boldsymbol{s}| = |\boldsymbol{r}|$ and $\boldsymbol{s} \neq \boldsymbol{r}$ is such that $\boldsymbol{ps}$ is the prefix of some word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, then $\boldsymbol{r} \prec \boldsymbol{s}$.
- If $\boldsymbol{pr}$ is a prefix of the $\prec$-last (resp. $\prec$-first) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, then $\boldsymbol{r}$ is the largest word, in $\prec$ order, with this property; that is, if $\boldsymbol{s} \in A_q^*$ with $|\boldsymbol{s}| = |\boldsymbol{r}|$ and $\boldsymbol{s} \neq \boldsymbol{r}$ is such that $\boldsymbol{ps}$ is the prefix of some word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, then $\boldsymbol{s} \prec \boldsymbol{r}$.

Similar results hold when $q$ is odd by replacing $\prec$ by $\lhd$ and considering the words parity as in the definition given after Remark 1.

Remark 4 below specifies the form of the first and last words in $A_q^n$, subject to the additional constraint that they do not begin by 0 or $q-1$. Later on we will see that when $\boldsymbol{f}$ does not induce zero periodicity, then the possible non-zero periods of the first or the last word in $\boldsymbol{p}|A_q^*(\boldsymbol{f})$ are related to those words. This remark will be used in the proofs of Propositions 1, 3 and 5.

**Remark 4.**

- For $q$ even, the first word in $A_q^n$, with respect to $\prec$ order, which does not begin by 0 is $1(q-1)0^{n-2}$.

- For $q$ odd, the first word in $A_q^n$, with respect to $\lhd$ order, which does not begin by 0 is $10^{n-1}$.
- For $q \geq 2$ (even or odd), the last word in $A_q^n$, with respect to the appropriate order, which does not begin by a $q-1$ is $(q-2)(q-1)0^{n-2}$.

In the following we need the technical lemma below.

**Lemma 1.** *Let $\boldsymbol{u}, \boldsymbol{g} \in A_q^*$ and $\boldsymbol{v} \in A_q^+$ be such that $\boldsymbol{g}$ is a suffix of both $\boldsymbol{u}$ and $\boldsymbol{uv}$. Then there exist a $j \geq 0$ and a (possibly empty) suffix $\boldsymbol{w}$ of $\boldsymbol{v}$ such that $\boldsymbol{g} = \boldsymbol{w}\boldsymbol{v}^j$ (or equivalently, $\boldsymbol{g}$ is a suffix of the left infinite word $\boldsymbol{v}^{-\infty}$).*

*Proof.* We prove the statement by induction on $k = \lfloor \frac{|\boldsymbol{g}|}{|\boldsymbol{v}|} \rfloor$. When $k = 0$, since the length of $\boldsymbol{g}$ is less than that of $\boldsymbol{v}$, and $\boldsymbol{g}$ is a suffix of $\boldsymbol{uv}$ the statement follows by considering $j = 0$ and $\boldsymbol{w} = \boldsymbol{g}$.

Let now $k = \lfloor \frac{|\boldsymbol{g}|}{|\boldsymbol{v}|} \rfloor > 0$. In this case the length of $\boldsymbol{g}$ is greater than that of $\boldsymbol{v}$, and it follows that $\boldsymbol{v}$ is a suffix of both $\boldsymbol{g}$ and $\boldsymbol{u}$. By considering $\boldsymbol{u}'$ and $\boldsymbol{g}'$ such that

- $\boldsymbol{u} = \boldsymbol{u}'\boldsymbol{v}$
- $\boldsymbol{g} = \boldsymbol{g}'\boldsymbol{v}$

we have that $\boldsymbol{g}'$ is a suffix of both $\boldsymbol{u}'$ and $\boldsymbol{u} = \boldsymbol{u}'\boldsymbol{v}$. Since $|\boldsymbol{g}'| = |\boldsymbol{g}| - |\boldsymbol{v}|$ we have that $\lfloor \frac{|\boldsymbol{g}'|}{|\boldsymbol{v}|} \rfloor = k-1$ and the statement follows by induction on $k$. $\square$

### 3.1 Forbidden factor ending by 0 and not inducing zero periodicity

We will determine, according to the parity of $q$, the form of the first and the last words in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ having no ultimate period 0 for $\boldsymbol{f}$ ending by 0, and consequently the form of the forbidden factors $\boldsymbol{f}$ that do not induce zero periodicity.

#### $q$ even

**Proposition 1.** *Let $q \geq 2$ be even and $\boldsymbol{f} \in A_q^+$ be a forbidden factor ending by 0 and not inducing zero periodicity. Let also $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$ be such that one of the $\prec$-first or the $\prec$-last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ does not have ultimate period 0, and let $\boldsymbol{a}$ be this word. Then $\boldsymbol{a}$ is ultimately periodic, more precisely there is an $m \geq 0$ such that either*

*(i)* $\boldsymbol{a} = \boldsymbol{p}0^i(1(q-1)0^m)^\infty$, *for some $i \leq m$, or*

*(ii)* $\boldsymbol{a} = \boldsymbol{p}((q-1)0^m1)^\infty$.

*Proof.* We will show that when $\boldsymbol{p}$ has even (resp. odd) parity, then either

1. $\boldsymbol{a}$ is the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, and in this case $\boldsymbol{a}$ has the form given in point $(i)$ above, or
2. $\boldsymbol{a}$ is the $\prec$-last (resp. $\prec$-first) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, and in this case $\boldsymbol{a}$ has the form given in point $(ii)$ above.

For the point 1, considering the parity of $\boldsymbol{p}$ and since $\boldsymbol{a}$ does not have ultimate period 0, there is an $i \geq 0$ such that $\boldsymbol{p}0^{i+1}$ contains the factor $\boldsymbol{f}$, but $\boldsymbol{p}0^i$ does not. Now, by Remark 4, since $\boldsymbol{f}$ ends by a 0, it follows that there is an $m \geq 0$ such that $\boldsymbol{p}0^i\,1(q-1)0^m$ is a prefix of $\boldsymbol{a}$, but $\boldsymbol{p}0^i 1(q-1)0^{m+1}$ is not. Thus, the length maximal 0 suffix of $\boldsymbol{f}$ is $0^{m+1}$, and reasoning in the same way, it follows that there is an $m' \geq 0$ such that $\boldsymbol{p}0^i\,1(q-1)0^m\,1(q-1)0^{m'}$ is a prefix of $\boldsymbol{a}$, but $\boldsymbol{p}0^i\,1(q-1)0^m\,1(q-1)0^{m'+1}$ is not. Since $0^{m+1}$ is the length maximal 0 suffix of $\boldsymbol{f}$, necessarily $m' = m$, and the statement holds by iterating this construction.

Similarly, point 2 holds considering that $\boldsymbol{p}(q-1)$ is a prefix of $\boldsymbol{a}$ and there is an $m \geq 0$ such that $\boldsymbol{p}(q-1)0^{m+1}$ contains the factor $\boldsymbol{f}$. □

Now we characterize the forbidden factors $\boldsymbol{f} \in A_q^+$ ending by 0, for even $q \geq 2$, and not inducing zero periodicity.

For even $q \geq 2$, let define the set $U_q \subset A_q^+$ as

$$U_q = \bigcup_{m \geq 0} \{\boldsymbol{b}0 \,|\, \boldsymbol{b} \text{ a suffix of } (1(q-1)0^m)^{-\infty}\}. \tag{1}$$

Alternatively, $U_q$ is the set of words of the form $\boldsymbol{b}0$, where $\boldsymbol{b}$ is either empty, or for some $m \geq 0$, a factor of $(1(q-1)0^m)^\infty$ ending by $0^m$ if $m > 0$ and ending by $q-1$ elsewhere. Clearly, $U_q$ contains exactly $n$ words of length $n$, for example, $U_4 \cap A_4^5 = \{00000, 30000, 13000, 01300, 13130\}$.

**Proposition 2.** *For even $q \geq 2$, if a forbidden factor $\boldsymbol{f} \in A_q^+$ ending by 0 does not induce zero periodicity, then $\boldsymbol{f} \in U_q$.*

*Proof.* If $\boldsymbol{f}$ is such a factor, then by Proposition 1 there is a $\boldsymbol{p} \in A^*(\boldsymbol{f})$ and an $m \geq 0$ such that, $\boldsymbol{a}$, the $\prec$-first or the $\prec$-last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ is

- $\boldsymbol{a} = \boldsymbol{p}0^i(1(q-1)0^m)^\infty$, for some $i \leq m$, or
- $\boldsymbol{a} = \boldsymbol{p}((q-1)0^m1)^\infty$.

Let $\boldsymbol{g}$ be the word obtained from $\boldsymbol{f}$ after erasing its last 0. In the first case it follows that $\boldsymbol{g}$ is a suffix of both $\boldsymbol{p}0^i\,1(q-1)0^m$ and $\boldsymbol{p}0^i\,1(q-1)0^m\,1(q-1)0^m$, and by Lemma 1 the statement holds. The proof is similar for the second case. □

**Remark 5.** If $\boldsymbol{f} \in U_q$ and $q$ is even, then $\boldsymbol{f}$ does not induce zero periodicity. Indeed, let for example $\boldsymbol{f} = \boldsymbol{b}0$ with $\boldsymbol{b}$ a suffix of $(1(q-1)0^m)^{-\infty}$ be as in relation (1). Then either the first word in $\boldsymbol{b}|A_q^\infty(\boldsymbol{f})$ when $\boldsymbol{b}$ has even parity, or the last word in $\boldsymbol{b}|A_q^\infty(\boldsymbol{f})$ when $\boldsymbol{b}$ has odd parity, has ultimate period $1(q-1)0^m$.

**Example 1.** Let $\boldsymbol{f} = 301300 \in U_4$ be a forbidden factor and let consider the prefix $\boldsymbol{p} = 0021301 \in A_4^*$. The $\prec$-first word in $\boldsymbol{p}|A_4^\infty(\boldsymbol{f})$ is $\boldsymbol{p}30(130)^\infty$.

Combining Proposition 2 and Remark 5 we have the following theorem.

**Theorem 1.** *For even $q \geq 2$, the forbidden factor $\boldsymbol{f} \in A_q^+$ ending by 0 does not induce zero periodicity if and only if $\boldsymbol{f} \in U_q$.*

### $q$ odd

Now we give the odd $q$ counterpart of the previous results.

**Proposition 3.** *Let $q \geq 3$ be odd and $\boldsymbol{f} \in A_q^+$ be a forbidden factor ending by $0$ and not inducing zero periodicity. Let also $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$ be such that one of the $\lhd$-first or the $\lhd$-last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ does not have ultimate period $0$, and let $\boldsymbol{a}$ be this word. Then $\boldsymbol{a}$ is ultimately periodic, more precisely there is an $m \geq 0$ such that either*

- $\boldsymbol{a} = \boldsymbol{p}0^i(10^m)^\infty$*, for some $i \leq m$, or*
- $\boldsymbol{a} = \boldsymbol{p}(q-1)(0^m1)^\infty$.

*Proof.* The proof is similar to that of Proposition 1 and considering the second point of Remark 4. □

Now we characterize the forbidden factor $\boldsymbol{f} \in A_q^+$ ending by 0, for odd $q \geq 3$, and not inducing zero periodicity.

For $q \geq 3$, let define the set $V \subset A_q^+$ as

$$V = \bigcup_{m \geq 0} \{\boldsymbol{b}0 \,|\, \boldsymbol{b} \text{ a suffix of } (10^m)^{-\infty}\}. \tag{2}$$

Alternatively, $V$ is the set of words of the form $\boldsymbol{b}0$, where $\boldsymbol{b}$ is either empty, or for some $m \geq 0$, a factor of $(10^m)^\infty$ ending by $0^m$ if $m > 0$ (and ending by 1 elsewhere). Notice that $V$ does not depend on $q$, i.e. $V \subset A_q^+$ for any $q \geq 2$. Clearly, $V$ contains exactly $n$ words of length $n$, for example, $V \cap A_q^5 = \{00000, 10000, 01000, 10100, 11110\}$, for any $q \geq 2$.

Considering Proposition 3 and the definition of $\lhd$ order relation, with the same arguments as in the proof of Proposition 2 we have the next result.

**Proposition 4.** *For odd $q \geq 3$, if a forbidden factor $\boldsymbol{f} \in A_q^+$ ending by $0$ does not induce zero periodicity, then $\boldsymbol{f} \in V$.*

**Remark 6.** If $\boldsymbol{f} \in V$ and $q \geq 3$ is odd, then $\boldsymbol{f}$ does not induce zero periodicity on $A_q^\infty(\boldsymbol{f})$. Indeed, let for example $\boldsymbol{f} = \boldsymbol{b}0$ with $\boldsymbol{b}$ a suffix of $(10^m)^{-\infty}$ be as in relation (2). Then either the first word in $\boldsymbol{b}|A_q^\infty(\boldsymbol{f})$ when $\boldsymbol{b}$ has even parity, or the last word in $\boldsymbol{b}|A_q^\infty(\boldsymbol{f})$ when $\boldsymbol{b}$ has odd parity, has ultimate period $10^m$.

**Example 2.** Let $\boldsymbol{f} = 0100010000 \in V$ a the forbidden factor and let consider the prefix $\boldsymbol{p} = 430100010 \in A_5^*$. The $\lhd$-last word in $\boldsymbol{p}|A_5^\infty(\boldsymbol{f})$ is $\boldsymbol{p}00(1000)^\infty$.

Combining Proposition 4 and Remark 6, we have the following theorem.

**Theorem 2.** *For odd $q \geq 3$, the forbidden factor $\boldsymbol{f} \in A_q^+$ ending by $0$ does not induce zero periodicity if and only if $\boldsymbol{f} \in V$.*

### 3.2 Forbidden factor ending by $q-1$ and not inducing zero periodicity

The next proposition holds for $q \geq 3$ (even or odd), and the case for $q = 2$ is stated in the remark that follows it.

**Proposition 5.** *Let $q \geq 3$ (even or odd) and $\boldsymbol{f} \in A_q^+$ be a forbidden factor ending by $q-1$ and not inducing zero periodicity. Let also $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$ be such that one of the first or the last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, with respect to the appropriate order, does not have ultimate period $0$, and let $\boldsymbol{a}$ be this word. Then $\boldsymbol{a} = \boldsymbol{p}(q-2)^\infty$.*

*Proof.* Nor $\boldsymbol{p}0$ neither $\boldsymbol{p}(q-1)$ can not be a prefix of $\boldsymbol{a}$; otherwise, in the first case $\boldsymbol{a} = \boldsymbol{p}0^\infty$ and in the second one $\boldsymbol{a} = \boldsymbol{p}(q-1)0^\infty$. By the third point of Remark 4 and since $\boldsymbol{f}$ ends by a $q-1$, it follows that $\boldsymbol{p}(q-2)$ is a prefix of $\boldsymbol{a}$, but $\boldsymbol{p}(q-2)(q-1)$ is not (otherwise $\boldsymbol{a} = \boldsymbol{p}(q-2)(q-1)0^\infty$). Again, $\boldsymbol{p}(q-2)(q-2)$ is a prefix of $\boldsymbol{a}$, but $\boldsymbol{p}(q-2)(q-2)(q-1)$ is not; and finally $\boldsymbol{a} = \boldsymbol{p}(q-2)^\infty$. □

When $q = 2$, the ultimate $(q-2)$ period of $\boldsymbol{a}$ in Proposition 5 becomes 0 period, and so, for $q = 2$ any forbidden factor $\boldsymbol{f} \in A_q^+$ ending by $q-1 = 1$ induces zero periodicity. Thus, below we will consider only factors ending by $q-1$ and not inducing zero periodicity only for $q \geq 3$ (even or odd).

For $q \geq 3$, let define the set $W_q$ as

$$W_q = \bigcup_{\ell \geq 0} \{(q-2)^\ell (q-1)\}. \tag{3}$$

With the previous terminology, $W_q$ is the set of words of the form $\boldsymbol{b}(q-1)$ with $\boldsymbol{b}$ a suffix of $(q-2)^{-\infty}$. Clearly, $W_q$ contains exactly one word of each length, and for example, $W_4 = \{3, 23, 223, 2223, 22223, \ldots\}$.

**Proposition 6.** *For $q \geq 3$ (even or odd), if the forbidden factor $\boldsymbol{f} \in A_q^+$ ending by $q-1$ does not induce zero periodicity, then $\boldsymbol{f} \in W_q$.*

*Proof.* Let $\boldsymbol{f}$ be such a factor, and $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$ such that, with respect to the appropriate order, the first word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ has not ultimate period 0 (the case of the first word being similar). Let also $\boldsymbol{g}$ be the (possibly empty) word obtained from $\boldsymbol{f}$ after erasing its last symbol $q-1$. By Proposition 5, $\boldsymbol{g}$ is a suffix of both $\boldsymbol{p}(q-2)$ and $\boldsymbol{p}(q-2)(q-2)$, and by Lemma 1 the statement holds. □

**Remark 7.** If $\boldsymbol{f} \in W_q$, then $\boldsymbol{f}$ does not induce zero periodicity. Indeed, let for example $\boldsymbol{b} = (q-2)^\ell$, for some $\ell \geq 0$, and $\boldsymbol{f} = \boldsymbol{b}(q-1)$ be as in relation (3). Then the last word in $\boldsymbol{b}|A_q^\infty(\boldsymbol{f})$ has ultimate period $(q-2)$.

**Example 3.** Let $\boldsymbol{f} = 223 \in W_4$ be a forbidden factor and let consider the prefix $\boldsymbol{p} = 2322 \in A_4^*$. The $\prec$-first word in $\boldsymbol{p}|A_4^\infty(\boldsymbol{f})$ is $\boldsymbol{p}2^\infty$. And when $\boldsymbol{f} = 12 \in W_3$ and $\boldsymbol{p} = 01 \in A_3^*$, the $\triangleleft$-last word in $\boldsymbol{p}|A_3^\infty(\boldsymbol{f}) = \boldsymbol{p}1^\infty$.

Combining Proposition 6 and Remark 7, we have the following theorem.

**Theorem 3.** *For $q \geq 3$ (even or odd), the forbidden factor $\boldsymbol{f} \in A_q^+$ ending by $q-1$ does not induce zero periodicity if and only if $\boldsymbol{f} \in W_q$.*

Even we will not make use later, it is worth to mention the following remark.

**Remark 8.** For $q$ even (resp. odd), if $\boldsymbol{f}$, $|\boldsymbol{f}| \geq 2$, does not have the form $0^\ell$ nor $(q-1)0^\ell$ (resp. the form $0^\ell$) for some $\ell \geq 1$, then for any $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$, at least one among the $\prec$-first and the $\prec$-last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ (resp. the $\triangleleft$-first and the $\triangleleft$-last word in $\boldsymbol{p}|A_q^\infty$) has ultimate period 0.

### 3.3 Forbidden factor inducing zero periodicity

Here we characterize the first and the last words in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ when the forbidden factor $\boldsymbol{f}$ induces zero periodicity; the resulting ultimate 0 periodic words will be used in the next section.

**Proposition 7.** *Let $q \geq 2$ be even, $\boldsymbol{f} \in A_q^+ \setminus U_q$ be a forbidden factor ending by 0, $\ell \geq 1$ be the length of the maximal 0 suffix of $\boldsymbol{f}$, and $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$. If $\boldsymbol{a}$ is the $\prec$-first or the $\prec$-last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, then $\boldsymbol{a}$ has the form*

$$\boldsymbol{p}\boldsymbol{r}0^\infty,$$

*where*

1. *$\boldsymbol{r} = \epsilon$ or $\boldsymbol{r} = 0^i1(q-1)$ for some $i$, $0 \leq i \leq \ell - 1$, if $\boldsymbol{a}$ is the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ and $\boldsymbol{p}$ has even (resp. odd) parity, or*
2. *$\boldsymbol{r} = q-1$ or $(q-1)0^{\ell-1}1(q-1)$ if $\boldsymbol{a}$ is the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ and $\boldsymbol{p}$ has odd (resp. even) parity.*

*Proof.* We prove the first point, the second one being similar. Let $\boldsymbol{a}$ be the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ with $\boldsymbol{p}$ having even (resp. odd) parity. Let also suppose that $\boldsymbol{r}$ has not the form prescribed in point 1. Reasoning as in the proof of Proposition 1 it follows that $0^i1(q-1)0^{\ell-1}1(q-1)0^{\ell-1}$ is a prefix of $\boldsymbol{r}$, for some $i$, $0 \leq i \leq \ell - 1$, and finally, by Lemma 1 that $\boldsymbol{f} \in U_q$, which leads to a contradiction. $\square$

The proof of the next proposition is similar to that of Proposition 7.

**Proposition 8.** *Let $q \geq 3$ be odd, $\boldsymbol{f} \in A_q^+ \setminus V$ be a forbidden factor ending by 0, $\ell \geq 1$ be the length of the maximal 0 suffix of $\boldsymbol{f}$, and $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$. If $\boldsymbol{a}$ is the $\triangleleft$-first or the $\triangleleft$-last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$, then $\boldsymbol{a}$ has the form*

$$\boldsymbol{p}\boldsymbol{r}0^\infty,$$

*where*

1. *$\boldsymbol{r} = \epsilon$ or $\boldsymbol{r} = 0^i1$ for some $i$, $0 \leq i \leq \ell - 1$, if $\boldsymbol{a}$ is the $\triangleleft$-first (resp. $\triangleleft$-last) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ and $\boldsymbol{p}$ has even (resp. odd) parity, or*
2. *$\boldsymbol{r} = q-1$ or $(q-1)0^{\ell-1}1$ if $\boldsymbol{a}$ is the $\triangleleft$-first (resp. $\triangleleft$-last) word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ and $\boldsymbol{p}$ has odd (resp. even) parity.*

It is routine to check the next two propositions.

**Proposition 9.** *Let $q \geq 3$, $\boldsymbol{f} \in A_q^+ \setminus W_q$ be a forbidden factor ending by $q-1$, and $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$. If $\boldsymbol{a}$ is the first or the last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ with respect to the appropriate order, then $\boldsymbol{a}$ has the form*

$$\boldsymbol{p}\boldsymbol{r}0^\infty,$$

*where $\boldsymbol{r}$ is either $\epsilon$, or $q-1$, or $(q-2)(q-1)$.*

As mentioned in Remark 2, forbidden factors ending by other symbol than 0 or $q-1$ induce zero periodicity, and we have the following proposition.

**Proposition 10.** *If $\boldsymbol{f} \in A_q^*$ is a forbidden factor that does not end by* 0 *nor by $q-1$, then for any $\boldsymbol{p} \in A_q^*(\boldsymbol{f})$, with respect to the appropriate order, both the first and the last word in $\boldsymbol{p}|A_q^\infty(\boldsymbol{f})$ have the form:*

$$\boldsymbol{pr}0^\infty,$$

*where $\boldsymbol{r}$ is either $\epsilon$ or $q-1$.*

We will see later that Propositions 7 to 10 above describe sufficient (but not a necessary) conditions for the Graycodeness of $A_q^n(\boldsymbol{f})$.

We conclude this section by the next corollary which summarizes the results in Remark 2 and Theorems 1, 2 and 3, and we will refer it later.

**Corollary 1.** *The forbidden factor $\boldsymbol{f} \in A_q^*$ induces zero periodicity if and only if either:*

- *$\boldsymbol{f}$ does not end by* 0 *nor by $q-1$, or*
- *$q = 2$ and $\boldsymbol{f} \notin U_2$, or*
- *$q \geq 4$ is even and $\boldsymbol{f} \notin U_q \cup W_q$, or*
- *$q \geq 3$ is odd and $\boldsymbol{f} \notin V \cup W_q$.*

# 4 Gray codes

In this section we show that for forbidden factors $\boldsymbol{f}$ inducing zero periodicity on $A_q^\infty$ (as stated in Corollary 1) consecutive words—in $\prec$ order for $q$ even, or $\lhd$ order for $q$ odd—in $A_q^n(\boldsymbol{f})$, beyond the common prefix, have all symbols 0, except the first few of them; and this ensures that the set $A_q^n(\boldsymbol{f})$ listed in an appropriate order is a Gray code.

Nevertheless, the property of $\boldsymbol{f}$ to induce zero periodicity is not a necessary condition. Indeed, listing the set $A_q^n(\boldsymbol{f})$ in $\prec$ order with:

- $\boldsymbol{f} = 0^\ell$ for any $q$ (not necessarily even), or
- $\boldsymbol{f} = (q-1)0^\ell$ for $q$ even,

where $\ell \geq 1$, yields a 1-Gray code, despite $\boldsymbol{f} \in U_q$ (and so, $\boldsymbol{f}$ does not induce zero periodicity for $q$ even). This particular cases are discussed in Section 4.2, and we show that such factors $\boldsymbol{f}$, $\boldsymbol{f} \geq 2$, are the only ones giving Gray codes for forbidden factors not inducing zero periodicity. In particular, the Gray code obtained for $A_q^n(0^\ell)$ is one of those defined in [1] as a generalization of a Gray code in [14]. Finally, for forbidden factors $\boldsymbol{f}$ for which $\prec$ nor $\lhd$ does not produce Gray codes on $A_q^n(\boldsymbol{f})$, we give simple transformations of $\boldsymbol{f}$, and eventually obtain Gray codes for $A_q^n(\boldsymbol{f})$ (in order other than $\prec$ or $\lhd$).

We will make use later of the following property of forbidden factors ending by 0 or $q-1$: for *any* $q \geq 2$, if $\boldsymbol{f}$ ends by 0 or $q-1$, then any two consecutive words in $A_q^n(\boldsymbol{f})$, in both $\prec$ and $\lhd$ order, differ by 1 or $-1$ in the leftmost position where they differ.

**Proposition 11.** *Let $q \geq 2$ and $\boldsymbol{f} \in A_q^+$ be a forbidden factor ending by 0 or $q-1$, and $\boldsymbol{a} = a_1a_2\dots a_n$ and $\boldsymbol{b} = b_1b_2\dots b_n$ be two words in $A_q^n(\boldsymbol{f})$, consecutive with respect to $\prec$ or $\lhd$ order. If $k$ is the leftmost position where $\boldsymbol{a}$ and $\boldsymbol{b}$ differ, then $b_k = a_k + 1$ or $b_k = a_k - 1$.*

*Proof.* If $\boldsymbol{f}$ ends by 0 let us suppose that $b_k < a_k - 1$. It follows that $\boldsymbol{f}$ is a suffix of $a_1a_2\dots(a_k-1)$, so $a_k - 1 = 0$ and thus $b_k < 0$, which is a contradiction. The proof when $b_k > a_k + 1$ or when $\boldsymbol{f}$ ends by $q-1$ is similar. □

## 4.1 Factors inducing zero periodicity

We show that for factors $\boldsymbol{f}$ as in Corollary 1 the set $A_q^n(\boldsymbol{f})$ listed in $\prec$ or $\lhd$ order is a Gray code.

**Proposition 12.** *If $q$ is even (resp. odd) and $\boldsymbol{f} \in A_q^+$ does not end by 0 nor $q-1$, then $A_q^n(\boldsymbol{f})$, $n \geq 1$, listed in $\prec$ (resp. $\lhd$) order is a 2-adjacent Gray code.*

*Proof.* Let $\boldsymbol{a}, \boldsymbol{b} \in A_q^n(\boldsymbol{f})$, $\boldsymbol{a} = a_1a_2\dots a_n$ and $\boldsymbol{b} = b_1b_2\dots b_n$ be two consecutive words with respect to the appropriate order, and $k$ be the leftmost position where they differ. Since $\boldsymbol{f}$ does not end by 0 nor by $q-1$, it follows that $q \geq 3$, and considering the definitions of $\prec$ and $\lhd$ order, we have in both cases (see Remark 1) $\{a_{k+1}, b_{k+1}\} \subset \{0, q-1\}$ and $a_{k+2}\dots a_n = b_{k+2}\dots b_n = 0^{n-k-1}$. In any case, $\boldsymbol{a}$ and $\boldsymbol{b}$ differ in position $k$ and possibly in position $k+1$. □

**Theorem 4.** *If $q \geq 2$ is even, $\boldsymbol{f} \in A_q^+ \setminus U_q$ ends by 0, and $\ell$ is the length of the maximal 0 suffix of $\boldsymbol{f}$, then $A_q^n(\boldsymbol{f})$, $n \geq 1$, listed in $\prec$ order is an at most $(\ell+2)$-close 3-Gray code.*

*Proof.* Let $\boldsymbol{a}, \boldsymbol{b} \in A_q^n(\boldsymbol{f})$, $\boldsymbol{a} = a_1a_2\dots a_n$ and $\boldsymbol{b} = b_1b_2\dots b_n$ be two consecutive words with respect to the appropriate order, and $k$ be the leftmost position where $\boldsymbol{a}$ and $\boldsymbol{b}$ differ. By Proposition 11, $b_k = a_k + 1$ or $b_k = a_k - 1$ and so the prefixes $\boldsymbol{a}' = a_1a_2\dots a_k$ and $\boldsymbol{b}' = b_1b_2\dots b_k$ have different parity. Two cases arise according to the parity of $\boldsymbol{a}'$.

• $\boldsymbol{a}'$ has even parity, and so $\boldsymbol{b}'$ has odd parity. By point 2 of Proposition 7

$$\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{x},$$

and

$$\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{y},$$

with $\boldsymbol{x}$ and $\boldsymbol{y}$ being the $n-k$ prefixes of $\boldsymbol{r}0^\infty$ and of $\boldsymbol{r}'0^\infty$, where $\{\boldsymbol{r}, \boldsymbol{r}'\} \subset \{(q-1), (q-1)0^{\ell-1}1(q-1)\}$. Thus $\boldsymbol{a}$ and $\boldsymbol{b}$ differ in position $k$ and possibly in positions $k+\ell+1$ and $k+\ell+2$ if $k+\ell+1 \geq n$.

• $\boldsymbol{a}'$ has odd parity, and so $\boldsymbol{b}'$ has even parity. By point 1 of Proposition 7 either

(i) $a_{k+1}a_{k+2}\dots a_n = b_{k+1}b_{k+2}\dots b_n = 0^{n-k}$, or

(ii) at least one of $a_{k+1}a_{k+2}\dots a_n$ or $b_{k+1}b_{k+2}\dots b_n$ is the $n-k$ prefix of a word of the form $0^i1(q-1)0^\infty$.

In case (i) $\boldsymbol{a}$ and $\boldsymbol{b}$ differ only in position $k$. And in case (ii) let us suppose that $a_{k+1}a_{k+2}\dots a_n$ is the length $n-k$ prefix of $0^i1(q-1)0^\infty$ (the corresponding case for $b_{k+1}b_{k+2}\dots b_n$ being similar). Considering that $b_k = a_k + 1$ or $b_k = a_k - 1$ it follows that $b_{k+1}b_{k+2}\dots b_n$ is the length $n-k$ prefix of $0^\infty$ and so $\boldsymbol{a}$ and $\boldsymbol{b}$ differ in positions $k$, and (possibly) $k+i+1$ and $k+i+2$.

In any case, $\boldsymbol{a}$ and $\boldsymbol{b}$ differ in at most three positions which are at most $\ell+2$ apart from each other. □

**Example 4.** The words 00230130 and 00330000 are consecutive in $A_4^8(2300)$ listed in $\prec$ order. They differ in 3 positions which are 4-close, and are in the worst case since the list is a 4-close 3-Gray code.

Considering the possible values of $\boldsymbol{r}$ in Proposition 8 it is easy to see that for $\boldsymbol{f} \notin V$ ending by 0 and $q$ odd, the set $A_q^n(\boldsymbol{f})$ listed in $\lhd$ order is a 4-Gray code. The next theorem gives a more restrictive result.

**Theorem 5.** *If $q \geq 3$ is odd, $\boldsymbol{f} \in A_q^+ \setminus V$ ends by 0, and $\ell$ is the length of the maximal 0 suffix of $\boldsymbol{f}$, then $A_q^n(\boldsymbol{f})$, $n \geq 1$, listed in $\lhd$ order is an at most $(\ell+1)$-close 3-Gray code.*

*Proof.* Let $\boldsymbol{a}, \boldsymbol{b} \in A_q^n(\boldsymbol{f})$, $\boldsymbol{a} = a_1a_2\dots a_n$ and $\boldsymbol{b} = b_1b_2\dots b_n$ be two consecutive words, in $\lhd$ order, and $k$ be the leftmost position where $\boldsymbol{a}$ and $\boldsymbol{b}$ differ. If $\boldsymbol{a}'$ and $\boldsymbol{b}'$ are the length $k$ prefix of $\boldsymbol{a}$ and $\boldsymbol{b}$, by Proposition 8

$$\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{x},$$

and

$$\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{y},$$

with $\boldsymbol{x}$ and $\boldsymbol{y}$ being the $n-k$ prefixes of $\boldsymbol{r}0^\infty$ and of $\boldsymbol{r}'0^\infty$, where $\{\boldsymbol{r}, \boldsymbol{r}'\} \subset \{\epsilon, 0^i1, (q-1), (q-1)0^{\ell-1}1\}$, for some $i$, $0 \leq i \leq \ell-1$. The statement holds by showing that $\{\boldsymbol{r}, \boldsymbol{r}'\} \subset \{0^i1, (q-1)0^{\ell-1}1\}$ is not possible. Indeed, let us suppose that $\boldsymbol{r} = 0^i1$ for some $i$, $0 \leq i \leq \ell-1$, and $\boldsymbol{r}' = (q-1)0^{\ell-1}1$ (the case $\boldsymbol{r} = (q-1)0^{\ell-1}1$ and $\boldsymbol{r}' = 0^i1$ being similar). This happens when both $\boldsymbol{a}'$ and $\boldsymbol{b}'$ have both odd parity. By Proposition 11, $b_k = a_k + 1$ or $b_k = a_k - 1$, and since $a_1a_2\dots a_{k-1} = b_1b_2\dots b_{k-1}$ it follows that $a_k = 1$ and $b_k = 0$. Since $\boldsymbol{r}' = (q-1)0^{\ell-1}1$, the factor $\boldsymbol{f}$ must end by $(q-1)0^\ell$ and since $a_k = 1$ it follows that $\boldsymbol{r} = \epsilon$, which leads to a contradiction. $\square$

**Example 5.** By Theorem 5, the sets $A_5^9(31000)$ and $A_5^9(24000)$ listed in $\lhd$ order are 4-close 3-Gray codes. However, it is easy to check that in particular, $A_5^9(31000)$ is a 3-close 3-Gray code, and $A_5^9(24000)$ is 4-close 2-Gray code. For example:

- the words 001304000 and 001310010 are consecutive in $A_5^9(31000)$ when listed in $\lhd$ order; they differ in 3 positions which are 3-close; and
- the words 001140000 and 001240010 are consecutive in $A_5^9(24000)$ when listed in $\lhd$ order; they differ in 2 positions which are 4-close.

**Theorem 6.** *If $q$ is even (resp. odd) and $\boldsymbol{f} \in A_q^+ \setminus W_q$ ends by $q-1$, then $A_q^n(\boldsymbol{f})$ listed in $\prec$ order (resp. $\lhd$ order) is a 2-close 3-Gray code (that is, a 3-adjacent Gray code).*

*Proof.* Let $k$ be the leftmost position where two consecutive words $\boldsymbol{a} = a_1a_2\dots a_n$ and $\boldsymbol{b} = b_1b_2\dots b_n$, in $A_q^n(\boldsymbol{f})$ differ. Exhausting the possible values of $a_k$ and $b_k$, and since $\boldsymbol{f} \notin W_q$ ends by $q-1$ it follows that $a_i = b_i = 0$ for all $i > k+2$ (see also Proposition 9). $\square$

### 4.2 Particular cases

As mentioned before, there are two cases when $\boldsymbol{f} \in U_q$, $q \geq 2$ and even, but $A_q^n(\boldsymbol{f})$ listed in $\prec$ order is a Gray code; these are $\boldsymbol{f} = 0^\ell$ and $\boldsymbol{f} = (q-1)0^\ell$, $\ell \geq 1$. Moreover, it turns out that $A_q^n(0^\ell)$, $q \geq 3$ and odd, also gives Gray code if listed in $\prec$ order. Similar phenomenon does not occur for $\boldsymbol{f} \in V$, i.e., the set $A_q^n(\boldsymbol{f})$ listed in $\lhd$ order is not a Gray code for any $\boldsymbol{f} \in V$, $|\boldsymbol{f}| \geq 2$ and $q \geq 3$ odd, see for example Remark 9.

Before discussing these particular forbidden factors we introduce some notations.
Let $q \geq 2$, $\ell \geq 1$, and let define the infinite words:

$$\begin{aligned} \boldsymbol{u} &= (0^{\ell-1}1(q-1))^\infty, \\ \boldsymbol{v} &= ((q-1)0^{\ell-1}1)^\infty. \end{aligned} \tag{4}$$

Notice that $\boldsymbol{u}$ and $\boldsymbol{v}$ are suffixes to each other, and they are related with the infinite words occurring in Proposition 1. It is easy to see that $\boldsymbol{u}$ and $\boldsymbol{v}$ are, respectively, the $\prec$-first and $\prec$-last word in $A_q^\infty(0^\ell)$ for even $q$; and thus the length $n$ prefix of $\boldsymbol{u}$ and $\boldsymbol{v}$ are, respectively, the $\prec$-first and $\prec$-last word in $A_q^n(0^\ell)$.

Moreover, for any $\boldsymbol{p} \in A_q^k(0^\ell)$ with $1 \leq k \leq n$ and $q$ even

- the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^n(0^\ell)$ is $\boldsymbol{p}\boldsymbol{v}'$ if $\boldsymbol{p}$ has an odd (resp. even) parity, where $\boldsymbol{v}'$ is the length $n-k$ prefix of $\boldsymbol{v}$;
- if $\boldsymbol{p}$ does not end by 0, then the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^n(0^\ell)$ is $\boldsymbol{p}\boldsymbol{u}'$ if $\boldsymbol{p}$ has an even (resp. odd) parity, where $\boldsymbol{u}'$ is the length $n-k$ prefix of $\boldsymbol{u}$.

Now let $q \geq 3$ and odd, $\ell \geq 1$, and let define the infinite words:

$$\begin{aligned} \boldsymbol{s} &= 0^{\ell-1}1(q-1)^\infty, \\ \boldsymbol{t} &= (q-1)^\infty, \end{aligned} \tag{5}$$

and $\boldsymbol{s}$ and $\boldsymbol{t}$ have similar property as $\boldsymbol{u}$ and $\boldsymbol{v}$ for $q$ odd and with same $\prec$ order.

**The case $\boldsymbol{f} = 0^\ell$**

**Proposition 13.** *For $q \geq 2$ (even or odd), and $\ell, n \geq 1$, the set $A_q^n(0^\ell)$ listed in $\prec$ order is a Gray code where two consecutive words differ in one position and by 1 or $-1$ in this position.*

*Proof.* Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two consecutive words, in $\prec$ order, in $A_q^n(0^\ell)$, $\boldsymbol{a}' = a_1a_2\dots a_k$ and $\boldsymbol{b}' = b_1b_2\dots b_k$ be the length $k$ prefix of $\boldsymbol{a}$ and $\boldsymbol{b}$, with $k$ the leftmost position where $\boldsymbol{a}$ and $\boldsymbol{b}$ differ.

When $q$ is even, with $\boldsymbol{u}'$ and $\boldsymbol{v}'$ the length $(n-k)$ prefix of $\boldsymbol{u}$ and $\boldsymbol{v}$ defined in relation (4), we have

- $\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{v}'$ and $\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{v}'$ if $\boldsymbol{a}'$ has an even parity (and so, by Proposition 11, $\boldsymbol{b}'$ has odd parity);
- $\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{u}'$ and $\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{u}'$, elsewhere, since $a_k \neq 0$ and $b_k \neq 0$ by considering the parity of the common length $k-1$ prefix of $\boldsymbol{a}$ and $\boldsymbol{b}$.

Similarly, when $q$ is odd, with $\boldsymbol{s}'$ and $\boldsymbol{t}'$ the length $(n-k)$ prefix of $\boldsymbol{s}$ and $\boldsymbol{t}$ defined in relation (5), we have

- $\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{t}'$ and $\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{t}'$ if $\boldsymbol{a}'$ has an even parity (given by $\sum_{i=1}^{k} a_i$);
- $\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{s}'$ and $\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{s}'$ (since, $a_k \neq 0$ and $b_k \neq 0$ ), elsewhere.

In both cases $\boldsymbol{a}$ and $\boldsymbol{b}$ differ only in position $k$. $\square$

**The case $\boldsymbol{f} = (q-1)0^\ell$ for $q$ even**

Let $q \geq 2$ be even, $1 \leq k \leq n$, and $\boldsymbol{v}'$ be the $n-k$ prefix of $\boldsymbol{v}$ defined in relation (4). For any $\boldsymbol{p} \in A_q^k((q-1)0^\ell)$ with $\ell \geq 1$ and $1 \leq k \leq n$

- the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^n((q-1)0^\ell)$ is $\boldsymbol{p}\boldsymbol{v}'$ if $\boldsymbol{p}$ has odd (resp. even) parity;
- if $\boldsymbol{p}$ does not end by 0 nor by $q-1$, then the $\prec$-first (resp. $\prec$-last) word in $\boldsymbol{p}|A_q^n((q-1)0^\ell)$ is $\boldsymbol{p}0^{n-k}$ if $\boldsymbol{p}$ has even (resp. odd) parity.

**Proposition 14.** *For $q \geq 2$ even, and $\ell, n \geq 1$, the set $A_q^n((q-1)0^\ell)$ listed in $\prec$ order is a Gray code where two consecutive words differ in one position and by $1$ or $-1$ in this position.*

*Proof.* Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two consecutive words, in $\prec$ order, in $A_q^n((q-1)0^\ell)$, $\boldsymbol{a}' = a_1a_2\ldots a_k$ and $\boldsymbol{b}' = b_1b_2\ldots b_k$ be the length $k$ prefix of $\boldsymbol{a}$ and $\boldsymbol{b}$ with $k$ the leftmost position where $\boldsymbol{a}$ and $\boldsymbol{b}$ differ.

If $\boldsymbol{a}'$ has even parity (and so, by Proposition 11, $\boldsymbol{b}'$ has odd parity), then by the above considerations $\boldsymbol{a} = \boldsymbol{a}'\boldsymbol{v}'$ and $\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{v}'$.

If $\boldsymbol{a}'$ has odd parity, by considering the parity of the common length $k-1$ prefix of $\boldsymbol{a}$ and $\boldsymbol{b}$ it follows that $a_k \neq q-1$ and $b_k \neq q-1$, and again, by the above considerations we have $\boldsymbol{a} = \boldsymbol{a}'0^{n-k}$ and $\boldsymbol{b} = \boldsymbol{b}'0^{n-k}$.

In both cases $\boldsymbol{a}$ and $\boldsymbol{b}$ differ only in position $k$. $\square$

### 4.3 Factors preventing Graycodeness

A consequence of the next remark and proposition, is Corollary 2 below. Proposition 15 sounds like Remark 8 and says that if $\boldsymbol{f}$, $|\boldsymbol{f}| \geq 2$ ($|\boldsymbol{f}| = 1$ being trivial), does not induce zero periodicity (see Corollary 1), and it is not in one of the two particular cases above, then consecutive words, with respect to the appropriate order, in $A_q^n(\boldsymbol{f})$ can differ in arbitrary many positions for enough large $n$. One of these particular cases is explained below.

**Remark 9.** For $q \geq 3$ and odd, $\ell \geq 2$ and $\boldsymbol{f} = 0^\ell$, the set $A_q^n(\boldsymbol{f})$ listed in $\lhd$-order is not a Gray code. Indeed, for example, the words $02\boldsymbol{z}'$ and $1\boldsymbol{z}''$ are consecutive in $\lhd$-order in $A_q^n(\boldsymbol{f})$, where $\boldsymbol{z}'$ and $\boldsymbol{z}''$ are appropriate length prefixes of $(0^{\ell-1}1)^\infty$, and they differ in arbitrary many positions for enough large $n$.

**Proposition 15.** *Let $\boldsymbol{f} \in A_q^+$, $q \geq 2$ and $|\boldsymbol{f}| \geq 2$, be a forbidden factor not inducing zero periodicity, other than $0^\ell$ or $(q-1)0^\ell$, $\ell \geq 1$. Let also $\boldsymbol{a}$ and $\boldsymbol{b}$ be two consecutive words, in appropriate order, in $A_q^n(\boldsymbol{f})$, $n \geq 1$, and $k$ the leftmost position where $\boldsymbol{a}$ and $\boldsymbol{b}$ differ. If*

- *$\boldsymbol{a}' = a_1a_2\ldots a_k$ and $\boldsymbol{b}' = b_1b_2\ldots b_k$ are, respectively, the length $k$ prefix of $\boldsymbol{a}$ and $\boldsymbol{b}$, and*
- *$\boldsymbol{a}''$ and $\boldsymbol{b}''$ are, respectively, the last word in $\boldsymbol{a}'|A_q^\infty(\boldsymbol{f})$ and the first word in $\boldsymbol{b}'|A_q^\infty(\boldsymbol{f})$, in appropriate order,*

*then at most one among $\boldsymbol{a}''$ and $\boldsymbol{b}''$ does not have ultimate period* 0.

*Proof.* Since $\boldsymbol{f}$ does not induce zero periodicity, we prove the statement according to $\boldsymbol{f}$ belongs to $U_q$, $V$ or $W_q$ (see Corollary 1), and supposing that $\boldsymbol{a}''$ does not have ultimate period 0 (the corresponding case for $\boldsymbol{b}''$ being similar).
If $\boldsymbol{f} \in U_q$, $q \geq 2$ and $\boldsymbol{f}$ does not have the form $0^\ell$ nor $(q-1)0^\ell$:

- When $\boldsymbol{a}'$ has odd parity, since $a_k$ must be a symbol of $\boldsymbol{f}$, it follows that $a_k \in \{0, 1, q-1\}$. From the parity of $\boldsymbol{a}'$, it follows that $a_k = 0$ implies that $b_k = a_k - 1$, and $a_k = q-1$ that $b_k = a_k + 1$, which are not possible, and necessarily $a_k = 1$. Thus, either $\boldsymbol{a}'' = \boldsymbol{a}'0^\infty$ (which is a contradiction with the non-zero periodicity of $\boldsymbol{a}$) or $\boldsymbol{b}'' = \boldsymbol{b}'0^\infty$.
- When $\boldsymbol{a}'$ has even parity, then $\boldsymbol{a}'' = \boldsymbol{a}'\boldsymbol{v}$ and since $b_k \neq a_k$, $\boldsymbol{b}'' = \boldsymbol{b}'(q-1)0^\infty$, with $\boldsymbol{v}$ defined in relation (4).

If $\boldsymbol{f} \in V$, $q \geq 3$ and odd, and $\boldsymbol{f}$ does not have the form $0^\ell$:

- $\boldsymbol{a}'$ can not have even parity, otherwise $\boldsymbol{a}'' = \boldsymbol{a}'(q-1)0^\infty$, which is a contradiction with the non-zero periodicity of $\boldsymbol{a}$;
- When $\boldsymbol{a}'$ has odd parity, the symbol $a_k$ must be one of the forbidden factor, so $a_k \in \{0, 1\}$. But $a_k = 0$, implies $b_k = a_k - 1$, which again is not possible; and $a_k = 1$ implies $b_k = 0$, and so $\boldsymbol{b}'' = (q-1)0^\infty$, which does not contain the factor $\boldsymbol{f}$ if it is different from $0^\ell$.

Finally, when $\boldsymbol{f} \in W_q$, $q \geq 3$ (even or odd) then $\boldsymbol{a}'' = \boldsymbol{a}'(q-2)^\infty$ and $\boldsymbol{b}''$ is either $\boldsymbol{b}'0^\infty$ (this can occur if $q$ is odd) or $\boldsymbol{b}'(q-1)0^\infty$. □

Table 1 summarizes the cases occurring in Proposition 15.

A consequence of Remark 9, Propositions 13 to 15 and Corollary 1, is the corollary below.

**Corollary 2.**

- *For even $q \geq 2$ and $|\boldsymbol{f}| \geq 2$, the set $A_q^n(\boldsymbol{f})$ listed in $\prec$ order is a Gray code for any $n \geq 1$ if and only if $\boldsymbol{f} \in \{0^\ell, (q-1)0^\ell\}_{\ell \geq 1} \cup W_2 \cup (A_q^* \setminus (U_q \cup W_q))$.*
- *For odd $q \geq 3$ and $|\boldsymbol{f}| \geq 2$, the set $A_q^n(\boldsymbol{f})$ listed in $\lhd$ order is a Gray code for any $n \geq 1$ if and only if $\boldsymbol{f} \in A_q^* \setminus (V \cup W_q)$.*

## 4.4 Obtaining Gray code if $\boldsymbol{f}$ does not induce zero periodicity and beyond the particular cases

According to the previous results, if the forbidden factor $\boldsymbol{f}$ does not induce zero periodicity, then the set $A_q^n(\boldsymbol{f})$ listed in $\prec$ or $\lhd$ order is not a Gray code, except for the two particular cases in Section 4.2. Now we show how a simple transformation allows to define Gray codes, with the same Hamming distance and closeness properties as for factors that induce zero periodicity, when $\boldsymbol{f}$ does not have this property. By Theorems 1, 2 and 3, the last symbol of a factor that does not induce zero periodicity is either 0, or $q-1$ when $q \geq 3$.

Let define the transformation $\phi : A_q \to A_q$ depending on $\boldsymbol{f}$ as

- when the last symbol of $\boldsymbol{f}$ is 0, then $\phi(0) = 1$, $\phi(1) = 0$, and $\phi(x) = x$ if $x \notin \{0, 1\}$; and

| $q$ | Order relation | The set for forbidden factor $\boldsymbol{f}$ | The set of ultimate periods of the last word in $\boldsymbol{a}\vert A_q^\infty(\boldsymbol{f})$ and the first one in $\boldsymbol{b}\vert A_q^\infty(\boldsymbol{f})$ | Graycodeness of $A_q^n(\boldsymbol{f})$ |
|---|---|---|---|---|
| even | $\prec$ | $U_q \setminus \{0^\ell, (q-1)0^\ell\}_{\ell \geq 1}$ | $\{1(q-1)0^{\ell-1}, 0\}$ | Not Gray code |
| even | $\prec$ | $\{0^\ell, (q-1)0^\ell\}_{\ell \geq 1}$ | $\{1(q-1)0^{\ell-1}\}$ | 1-Gray code |
| odd | $\prec$ | $\{0^\ell\}_{\ell \geq 1}$ | $\{(q-1)\}$ | 1-Gray code |
| odd | $\lhd$ | $V \setminus \{0^\ell\}_{\ell \geq 1}$ | $\{10^{\ell-1}, 0\}$ | Not Gray code |
| odd | $\lhd$ | $\{0^\ell\}_{\ell \geq 2}$ | $\{10^{\ell-1}\}$ | Not Gray code |
| $q \geq 3$ even (resp. odd) | $\prec$ (resp. $\lhd$) | $W_q \cap A_q^{\geq 2}$ | $\{(q-2), 0\}$ | Not Gray code |

Table 1: The Graycodeness of $A_q^n(\boldsymbol{f})$ listed in appropriate order together with the ultimate periods of the last word in $\boldsymbol{a}|A_q^\infty(\boldsymbol{f})$ and the first word in $\boldsymbol{b}|A_q^\infty(\boldsymbol{f})$, when at least one of them does not have ultimate period 0, and $\boldsymbol{a}$ and $\boldsymbol{b}$ are consecutive words; and $A_q^{\geq 2}$ is the set of words on $A_q$ of length at least two. These summarize Propositions 13 to 15, and Corollary 2.

- when the last symbol of $\boldsymbol{f}$ is $q-1$ (and so, $q \geq 3$), then $\phi(q-2) = q-1$, $\phi(q-1) = q-2$, and $\phi(x) = x$ if $x \notin \{q-2, q-1\}$.

In both cases, $\phi$ is an involution, that is, $\phi^{-1} = \phi$. By abuse of notation, for $\boldsymbol{w} \in A_q^*$, $\phi(\boldsymbol{w})$ is the word obtained from $\boldsymbol{w}$ by replacing each of its symbols $x$ by $\phi(x)$, and for a list $\mathcal{L}$ of words, $\phi(\mathcal{L})$ is the list obtained from $\mathcal{L}$ by replacing each word $\boldsymbol{w}$ in $\mathcal{L}$ by $\phi(\boldsymbol{w})$.

If $\boldsymbol{f}$ is a forbidden factor that does not induce zero periodicity, then $\phi(\boldsymbol{f})$ does not end by 0 nor by $q-1$, and so it induces zero periodicity, see Remark 2. In this case $\phi(\mathcal{L})$ is a Gray code for the set $A_q^n(\boldsymbol{f})$, where $\mathcal{L}$ is the set $A_q^n(\phi(\boldsymbol{f}))$ listed in $\prec$ order for $q$ even, and in $\lhd$ order for $q$ odd.

## 5 Algorithm considerations

Here we give a generating algorithm for the set $A_q^n(\boldsymbol{f})$, $n \geq 1$, for any forbidden factor $\boldsymbol{f} \in A_q^\ell$, $\ell \geq 2$ (the case $\ell = 1$ being trivial). This generating algorithm produces recursively prefixes of words in $A_q^n(\boldsymbol{f})$, in $\prec$ order if $q$ is even, or in $\lhd$ order if $q$ is odd, and in particular, it generates the previously discussed Gray codes for $A_q^n(\boldsymbol{f})$. We will show that this algorithm is efficient, except for the trivial factors of the form $00\ldots01$ or $11\ldots10$, for which a simple transformation of them makes the generating algorithm efficient.

The generating procedure GenAvoid in Figure 1 expands recursively a current generated prefix $w_1w_2\ldots w_{k-1}$ ($k$ being the first parameter of GenAvoid) to $w_1w_2\ldots w_{k-1}j$, with $j$ covering the alphabet $A_q$ in increasing or decreasing order, according to the value of $dir \in \{0,1\}$, the second parameter of the procedure, which is the parity of the word $w_1w_2\ldots w_{k-1}$. Moreover, when the length $(\ell-1)$ prefix of $\boldsymbol{f} = f_1f_2\ldots f_\ell$ is a suffix of $w_1w_2\ldots w_{k-1}$, the value $f_\ell$ is skipped for $j$ in order not to produce the forbidden factor. To do this efficiently, the third parameter, $i$, of procedure GenAvoid is the length of the maximal prefix of the forbidden factor $\boldsymbol{f}$ which is also a suffix of the current generated word $w_1w_2\dots w_{k-1}$; and $h$ in this procedure is the length of the maximal suffix of $w_1w_2\dots w_{k-1}j$ which is also a prefix of $\boldsymbol{f}$, and it is given by $M_{i,j}$. So, when $h$ is equal to $\ell$ (the length of the forbidden factor), the current value of $j$ is skipped for the prefix expansion.

Now we explain in more details the array $M$ used by algorithm GenAvoid. For a forbidden factor $\boldsymbol{f} = f_1f_2\dots f_\ell \in A_q^\ell$, the $\ell \cdot q$ size two dimensional array $M$ is defined as: for $i \in \{0, 1, \dots, \ell-1\}$ and $j \in \{0, 1, \dots, q-1\} = A_q$, $M_{i,j}$ is the length of the maximal suffix of $f_1f_2\dots f_ij$ which is also a prefix $\boldsymbol{f}$. For instance, for $q = 4$ and $\boldsymbol{f} = 012011 \in A_4^6$ we have

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ \mathbf{1} & \mathbf{0} & \mathbf{3} & 0 \\ 4 & 0 & 0 & 0 \\ 1 & 5 & 0 & 0 \\ 1 & 6 & 3 & 0 \end{bmatrix},$$

and, for example (see the entries in boldface in $M$)

- $M_{2,0} = \mathbf{1}$, since the length of the longest suffix of $f_1f_20 = 010$ which is a prefix of $\boldsymbol{f}$ is 1,
- $M_{2,1} = \mathbf{0}$, since there is no suffix of $f_1f_21 = 011$ which is a prefix of $\boldsymbol{f}$,
- $M_{2,2} = \mathbf{3}$, since $f_1f_22 = 012$ (of length 3) is a prefix of $\boldsymbol{f}$.

```
procedure GenAvoid(k, dir, i)
if k = n + 1 then type;
else if dir = 0 then S := ⟨0, ..., q − 1⟩; else S := ⟨q − 1, ..., 0⟩;
     for j in S
         h := M[i, j];
         if h ≠ ℓ then
                 w[k] := j; m := (dir + j)  mod 2;
                 if q is odd and j ≠ 0 then m := (m + 1)  mod 2;
                 GenAvoid(k + 1, m, h);
```

Figure 1: Algorithm producing the set $A_q^n(\boldsymbol{f})$, listed in $\prec$ order if $q$ is even or in $\lhd$ order if $q$ is odd. The initial call is GenAvoid$(1, 0, 0)$, and it uses array $M$, initialized in a preprocessing step by MakeArray; and $\mathcal{S}$ is the list of symbols in the alphabet $A_q$ in increasing or decreasing order.

The array $M$ is initialized, in an $O(\ell \cdot q)$ time preprocessing step, by procedure MakeArray in Figure 2, which in turn uses array $b = b_0b_1b_2\dots b_\ell$, the *border* array of $\boldsymbol{f}$ defined as (see for instance [10]): $b_i$, $0 \le i \le \ell$, is the length of the border of $f_1f_2\dots f_i$, that is, the length of the longest factor which is both a proper prefix and a proper suffix of $f_1f_2\dots f_i$; and by convenience $b_0 = -1$. For example if $\ell = 8$ and $\boldsymbol{f} = 01001010$, then $b_0b_1\dots b_8 = -100112323$; and for instance, $b_5 = 2$ since 01 is the longest proper prefix which is also a suffix of $f_1f_2\dots f_5 = 01001$. Actually, the border array $b$ is a main ingredient for Knuth-Morris-Pratt word matching algorithm in [8] and it is initialized by an $O(\ell)$ time complexity preprocessing step by procedure MakeBorder in Figure 3, see again [10].

```
procedure MakeArray()
for j := 0 to q − 1
    for i := 0 to ℓ − 1
        if f[i + 1] = j then M[i, j] := i + 1;
        else if i > 0 then M[i, j] := M[b[i], j];
            else M[i, j] := 0;
```

Figure 2: Algorithm initializing the array $M$.

```
procedure MakeBorder()
b[0] := −1;
i := 0;
for j := 1 to (ℓ − 1)
    b[j] := i;
    while (i ≥ 0 and f[j + 1] ≠ f[i + 1])
        i := b[i];
    i := i + 1;
b[ℓ] := i;
```

Figure 3: Procedure computing the border array $b$ of the length $\ell$ forbidden factor $\boldsymbol{f}$, and used by MakeArray.

Before analyzing the time complexity of the generating algorithm GenAvoid we show that, if in the underlying tree induced by recursive calls of GenAvoid there are degree-one successive calls, then $q = 2$ and the forbidden factor has the form $00\ldots01$ or $11\ldots10$. See Figure 4 for words in $A_2^n(001)$ produced by degree-one consecutive calls of GenAvoid.

For a length $\ell \geq 2$ forbidden factor $\boldsymbol{f}$ let $\boldsymbol{w} \in A_q^*(\boldsymbol{f})$ and $i, j \in A_q$ such that $\boldsymbol{w}ij \in A_q^*(\boldsymbol{f})$ and:

- $\boldsymbol{w}k$ ends by $\boldsymbol{f}$ for any $k \in A_q$, $k \neq i$, and
- $\boldsymbol{w}ik$ ends by $\boldsymbol{f}$ for any $k \in A_q$, $k \neq j$.

In other words, when the current word is $\boldsymbol{w}$ as above, then the call of GenAvoid is a degree-one call (producing $\boldsymbol{w}i$) which in turn produces a degree-one call (producing $\boldsymbol{w}ij$). By the two conditions above, it follows that $q = 2$ and $i = j$. When $i = j = 0$, the length $\ell - 1$ suffix of $\boldsymbol{w}$ is

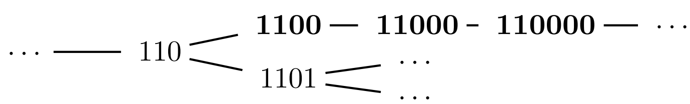


Figure 4: In boldface a ‘branch’ of words produced by consecutive degree-one calls in the generating tree of $A_2^n(001)$.

equal to the $\ell-1$ suffix of $\boldsymbol{w}0$, which in this case must be $0^{\ell-1}$, and finally $\boldsymbol{f}=0^{\ell-1}1$. Similarly, when $i=j=1$, it follows that $\boldsymbol{f}=1^{\ell-1}0$.

Let now $\boldsymbol{f}$ be a length $\ell\geq 2$ forbidden factor, and either $q\geq 3$ or $q=2$ and $\boldsymbol{f}$ is not $0^{\ell-1}1$ nor $1^{\ell-1}0$. In this case, by the previous considerations, each recursive call of GenAvoid is either:

- a terminal call, or
- a call producing at least two recursive calls, or
- a call producing one recursive call, which in turn is in one of the two cases above,

and by Ruskey's CAT principle in [13], it follows that, with the previous restrictions on $q$ and $\boldsymbol{f}$, GenAvoid runs in constant amortized time, and so is an efficient generating algorithm.

Nevertheless, for the particular factors above, when $\ell=2$, $A_2^n(1^{\ell-1}0)$ is trivially the set $\{0^n,0^{n-1}1,0^{n-2}11,\ldots,1^n\}$, and $A_2^n(0^{\ell-1}1)$ the set $\{0^n,10^{n-1},110^{n-2},\ldots,1^n\}$. And for $\ell\geq 3$, both sets $A_2^n(1^{\ell-1}0)$ and $A_2^n(0^{\ell-1}1)$ can be generated efficiently in Gray code order. Indeed, for $A_2^n(1^{\ell-1}0)$ with $\ell\geq 3$ it is enough to generate (efficiently) the Gray code for $A_2^n(01^{\ell-1})$ (see Theorem 6) and then reverse each generated word; and for $A_2^n(0^{\ell-1}1)$ it is enough to generate the Gray code for $A_2^n(1^{\ell-1}0)$ as previously, then complement each symbol in each word. The following scheme describes this method (see the example in Table 3):

$$A_2^n(01^{\ell-1})\xrightarrow{\text{Reverse}}A_2^n(1^{\ell-1}0)\xrightarrow{\text{Complement}}A_2^n(0^{\ell-1}1).$$

Finally, notice that the generating order ($\prec$ or $\lhd$ in our case) does not affect the efficiency of the generating algorithm, which can obviously be modified to produce same set of factor avoiding words in lexicographical order. A `C` implementation of our generating algorithm is on the web site of the last author [16].

## 6 Conclusions

We introduce two order relations on the set of length $n$ $q$-ary words, and show that the set of words avoiding any from among the $q^\ell$ factors of length $\ell\geq 2$, except $\ell-1$ or $\ell$ of them according to the parity of $q$, when listed in the appropriate order is an (at most) 3-Gray code. For each of the excepted factors we give a simple transformation which allows to eventually obtain similar Gray codes. Finally, an efficient generating algorithm for the derived Gray codes is given.

# Appendix

| | | | | | |
|---|---|---|---|---|---|
| 0 0 0 0 | 0 1 **2** 2 | 1 0 **1** **0** | 1 **2** 2 0 | 2 2 **2** 2 | 2 0 **1** 2 |
| 0 0 0 **1** | 0 1 2 **1** | 1 0 1 **1** | 1 2 2 **1** | 2 2 2 **1** | 2 0 1 **1** |
| 0 0 0 **2** | 0 1 2 **0** | 1 0 1 **2** | 1 2 2 **2** | 2 2 2 **0** | 2 0 1 **0** |
| 0 0 **1** **0** | 0 **2** 2 0 | 1 0 **2** 2 | 1 2 **1** 2 | 2 **1** 2 0 | 2 0 **0** **2** |
| 0 0 1 **1** | 0 2 2 **1** | 1 0 2 **1** | 1 2 1 **1** | 2 1 2 **1** | 2 0 0 **1** |
| 0 0 1 **2** | 0 2 2 **2** | 1 0 2 **0** | 1 2 1 **0** | 2 1 2 **2** | 2 0 0 **0** |
| 0 0 **2** 2 | 0 2 **1** 2 | 1 **1** **0** 0 | 1 2 **0** **2** | 2 1 **1** 2 | |
| 0 0 2 **1** | 0 2 1 **1** | 1 1 0 **1** | 1 2 0 **1** | 2 1 1 **1** | |
| 0 0 2 **0** | 0 2 1 **0** | 1 1 0 **2** | 1 2 0 **0** | 2 1 1 **0** | |
| 0 **1** **0** 0 | 0 2 **0** **2** | 1 1 **1** **0** | **2** 2 0 0 | 2 1 **0** **2** | |
| 0 1 0 **1** | 0 2 0 **1** | 1 1 1 **1** | 2 2 0 **1** | 2 1 0 **1** | |
| 0 1 0 **2** | 0 2 0 **0** | 1 1 1 **2** | 2 2 0 **2** | 2 1 0 **0** | |
| 0 1 **1** **0** | **1** **0** 0 0 | 1 1 **2** 2 | 2 2 **1** **0** | 2 **0** **2** 0 | |
| 0 1 1 **1** | 1 0 0 **1** | 1 1 2 **1** | 2 2 1 **1** | 2 0 2 **1** | |
| 0 1 1 **2** | 1 0 0 **2** | 1 1 2 **0** | 2 2 1 **2** | 2 0 2 **2** | |

Table 2: The set $A_3^4$ listed in $\lhd$ order, inducing a 2-Gray code. The list is columnwise and the changed symbols are in bold.

| $A_2^4(011)$ (a) | $A_2^4(110)$ (b) | $A_2^4(001)$ (c) |
|---|---|---|
| 0 0 0 0 | 0 0 0 0 | 1 1 1 1 |
| 0 0 0 **1** | **1** 0 0 0 | **0** 1 1 1 |
| 0 0 **1** **0** | **0** **1** 0 0 | **1** **0** 1 1 |
| 0 **1** **0** **1** | **1** **0** **1** 0 | **0** **1** **0** 1 |
| 0 1 0 **0** | **0** 0 1 0 | **1** 1 0 1 |
| **1** 1 0 0 | 0 0 1 **1** | 1 1 0 **0** |
| 1 1 0 **1** | **1** 0 1 1 | **0** 1 0 0 |
| 1 1 **1** 1 | 1 **1** 1 1 | 0 **0** 0 0 |
| 1 1 1 **0** | **0** 1 1 1 | **1** 0 0 0 |
| 1 **0** 1 0 | 0 1 **0** 1 | 1 0 **1** 0 |
| 1 0 **0** **1** | **1** **0** 0 1 | **0** **1** 1 0 |
| 1 0 0 **0** | **0** 0 0 1 | **1** 1 1 0 |

Table 3: (a) The set $A_2^4(011)$ listed in $\prec$ order, inducing 3-adjacent Gray code; (b) the reverse of the list in (a), giving Gray code for $A_2^4(110)$; (c) the complement of the list in (b), giving Gray code for $A_2^4(001)$. The changed symbols are in bold

# References


[1] A. Bernini, S. Bilotta, R. Pinzani and V. Vajnovszki, Two Gray codes for $q$-ary $k$-generalized Fibonacci strings. *ICTCS13*, Palermo-Italy, September 9–11, 2013.

[2] S. Bilotta, E. Pergola and R. Pinzani, A construction for a class of binary words avoiding $1^j0^i$. *PU.M.A.*, **23**(2), 81–102, 2012.

[3] S. Bilotta, D. Merlini, E. Pergola and R. Pinzani, Pattern $1^{j+1}0^j$ avoiding binary words. *Fund. Inform.*, **117**, 35–55, 2012.

[4] M.C. Er, On generating the $N$-ary reflected Gray code. *IEEE Transaction on Computers*, **33**(8), 739–741, 1984.

[5] F. Gray, *Pulse code communication*, U.S. Patent 2632058, 1953.

[6] L.J. Guibas and A.M. Odlyzko, Periods in strings. *J. Combin. Theory Ser. A*, **30**(1), 19–42, 1981.

[7] J. Joichi, D.E. White and S.G. Williamson, Combinatorial Gray codes. *Siam J. on Computing*, **9**, 130–141, 1980.

[8] D.E. Knuth, J.H. Morris and V.R. Pratt, Fast pattern matching in strings. *SIAM J. on Computing*, **6**:323–350, 1977.

[9] Y. Li and J. Sawada, Gray codes for reflectable languages. *Inf. Process. Lett.*, **109**(5) 296-300, 2009.

[10] M. Lothaire, *Applied Combinatorics on Words*. Cambridge University Press, New York, 2005.

[11] A. Sabri and V. Vajnovszki, Reflected Gray code based orders on some restricted growth sequences. To appear, *The Computer Journal*.

[12] M. Squire, Gray codes for $A$-free strings. *Electr. J. Combinatorics*, **3**, paper R17, 1996.

[13] F. Ruskey, *Combinatorial Generation*, book in preparation.

[14] V. Vajnovszki, A loopless generation of bitstrings without $p$ consecutive ones. *DMTCS–Springer*, 227–240, 2001.

[15] V. Vajnovszki and R. Vernay, Restricted compositions and permutations: from old to new Gray codes. *Inf. Process. Lett.*, **111**(13), 650–655, 2011.

[16] V. Vajnovszki, `v.vincent.u-bourgogne.fr/0ABS/publi.html`.

[17] T. Walsh, Generating Gray Codes in $O(1)$ worst-case time per word. *4th Discrete Mathematics and Theoretical Computer Science Conference*, Dijon-France, 7–12 July 2003 (LNCS, **2731**, 73–88).

[18] T. Walsh, Loop-free sequencing of bounded integer compositions. *Journal of Combinatorial Mathematics and Combinatorial Computing*, **33**, 323–345, 2000.

[19] S.G. Williamson, *Combinatorics for Computer Science*. Computer Science Press, Rockville, Maryland, 1985.