ELSEVIER

Theoretical Computer Science 182 (1997) 245–256

Theoretical Computer Science

Note

Hajós factorizations and completion of codes

Nguyen Huong Lam*

Hanoi Institute of Mathematics, P.O. Box 631, Bo Ho, 10000 Hanoi, Viet Nam

Received July 1996; revised December 1996
Communicated by D. Perrin

Abstract

Given a Hajós factorization of $\mathbf{Z}_n$, De Felice has shown that $a^n + a^P b + ba^Q$ is finitely completable to a d-code, $d \leqslant 3$. In this article, we prove that the code $a^n + a^P ba^Q$ is finitely completable but not always to a d-code with $d \leqslant 3$.

1. Introduction

Let $A = \{a, b\}$ be a binary alphabet and A^* the free monoid of words on A. The symbol ε stands for the empty word and $A^+ = A^* - \{\varepsilon\}$. A subset C of A^* is a *code* if the equality

$$c_1, \ldots, c_i = d_1, \ldots, d_j$$

with $c_1, \ldots, c_i, d_1, \ldots, d_j \in C$ implies $i = j$ and $c_1 = d_1, \ldots, c_i = d_i$. A code is said to be *maximal* (on A) if every strictly larger subset (of A^*) is not a code. A standard application of Zorn lemma shows that every code is included in, or is *completed to* a maximal code, frequently called a *completion* of it. Despite its elegance and importance, such a universal theorem provides no means to do the completion and gives no insight into structure of maximal codes. Although the completion of codes is a central theme, a systematic treatment has genuinely emerged in the mid-1980s when Ehrenfeutch and Rozenberg proved constructibly that every recognizable code can be completed to a maximal code which is also recognizable. Then follow a number of interesting results conjectured by Berstel and Perrin. It is worth mentioning first, the work of Bruyère, Wang and Zhang on the possibility of completing recognizable codes of a given finite deciphering delay within this class (1990). Recently, Shen and Zhang

* E-mail: nhlam@thevinh.ac.vn.

0304-3975/97/$17.00 © 1997 — Elsevier Science B.V. All rights reserved
PII S0304-3975(97)00032-7

have proved that every recognizable bifix code is completed to a recognizable bifix-maximal codes of degree not surpassing a given computable constant (1995).

In the field of finite codes, the scene is somewhat austere. When a code has a finite code as a completion, we say that it has a *finite completion*, or it is *finitely completable*. Markov (1969) and Restivo [9] discovered that not every finite code has a finite completion and they first produced explicit examples. Then arose a fundamental question: when does a finite code have a finite completion? Besides some special results scattered in the literature, no systematic study on codes having finite completions is known. Recently, De Felice, based on a characterization of finite maximal 1-, 2- and 3-codes, has made investigations on a class of finite codes commonly encountered in the field. The result is directly related to Hajós factorizations: De Felice has shown, among other things, that the code $a^n + ba^P + a^Qb$, where (P,Q) is a Hajós factorization of the additive group $\mathbf{Z}_n$ of residues modulo n, has a finite completion which is, moreover, a d-code with $d \leqslant 3$. In this paper, we, in one respect, strengthen this result by proving that the code $a^n + a^Pba^Q$ is completable to a finite code, and in the other, we indicate those (P,Q) for which $a^n + a^Pba^Q$ cannot have a d-code, $d \leqslant 3$, as completion.

We devote the next section to the theory of Hajós factorization, as it has an underlaying meaning to this article. One may skip the section, considering it as an appendix, only take Theorem 2.8 on trust.

2. Hajós factorization

We give in this section a concise exposition of Hajós factorization construction. One may see De Felice [5, 6] for historical background or Hajós [8] and Sands [11] for original settings.

2.1. General theory

Let G be an (additive) abelian group. A *factorization* of G is a pair of subsets (*summands*) (K,L) such that each element of G is represented uniquely as a sum of elements in K and L, that is $K+L=\{k+l\colon\ k\in K,\ l\in L\}=G$ and $k_1+l_1=k_2+l_2$, $k_1,k_2\in K$, $l_1,l_2\in L$ imply $k_1=k_2$, $l_1=l_2$. A subset of G is called *periodic* if there is a nonzero element $g\in G$, *period* of S, satisfying $g+S=S$. A factorization is *periodic* if a summand of it is a periodic subset.

Finite abelian groups fall into two classes: ones, which are called *good* groups, for which every factorization is periodic and the other, called *bad* ones, which admit a nonperiodic factorization. The class of good groups has a remarkable feature that it is closed under formation of subgroups and factor groups [2].

Let $S=\{s_1,\dots,s_q\}$ be a subset of G, for a subset T we define $S\circ T$ as the collection of subsets of the form $R=\{s_1+t_1,\dots,s_q+t_q\}$ where $t_1,\dots,t_q$ are (not necessarily different) elements in T. The operations $+$, defined already for subsets, and $\circ$ now extend in a natural way to collections of subsets of G.

The construction goes as follows. Let $G = G_0 > G_1 > \cdots > G_r > G_{r+1} = \{0\}$ be a strictly descending chain of subgroups of G and let H_i be a system of coset representatives of G_{i-1} by G_i, $G_{i-1} = H_i + G_i$ for each $i = 1,2,\ldots,r+1$. Consider the following collections $\Delta_0 = \{0\}$, $\Theta_0 = \{0\};\ldots;\Delta_1 = \{0\}+H_1$, $\Theta_1 = \{0\}\circ H_1;\ldots;\Theta_{i+1} = \Delta_i \circ H_{i+1}$, $\Delta_{i+1} = \Theta_i + H_{i+1};\ldots;\Delta = \Delta_{r+1} = \Theta_r + H_{r+1}$, $\Theta = \Theta_{r+1} = \Delta_r \circ H_{r+1}$. If needed to be specific, we indicate the parameters: $\Delta_i(H_1,\ldots,H_i)$, $\Theta_i(H_1,\ldots,H_i)$. The following result was first proved by Hajós [8] and Sands [11].

Theorem 2.1 (Hajós [8]; Sands [11]). *If H_i is a system of coset representatives of G_{i-1} by G_i, for every $i = 1,\ldots,r+1$, then every pair (T,R) of (Θ,Δ) is a periodic factorization of G.*

Definition 2.2. Any pair of subsets obtained by a construction based on each H_i, $1 \leqslant i \leqslant r+1$, being a system of coset representatives of G_{i-1} by G_i, which is actually a factorization, is called a *Hajós factorization.*

Remark 2.3. If we remove the condition on H_i being a system of coset representatives G_{i-1} by G_i and require solely $G_{i-1} = H_i + G_i$ to be held, the construction leads sometimes to a factorization and sometimes not. Example, consider a particular case $r = 1$, $G = H_1 + G_1$, $G_1 = H_2 + \{0\}$. We have $H_1 = \{0\} + H_1 \in \Delta_1$, $\{h\} \in \{0\} \circ H_1 \in \Theta_1$ for any $h \in H_1$ that leads to $G_1 \in \Delta$, $H_1 \in \Theta$. But (G_1, H_1) is a factorization of G iff H_1 is a system of coset representatives of G/G_1. Nevertheless, it can be shown that if the process results in a factorization, then this one can equally be obtained by a construction based on any fixed system of coset representatives H_i' of G_{i-1}/G_i for each i, that is, in regard to Hajós factorization, the generating capacity of unrestricted construction remains the same.

The next result shows the recursive character of Hajós factorization construction for which we provide a sample proof. We say that a subset S of a group G is *reduced* modulo a subgroup G' if no two elements of S are in the same coset class by G'.

Theorem 2.4. *(R,T) is a Hajós factorization of G, if and only if there is a subgroup G', a pair (K,L) of subsets reduced modulo G' such that $(\rho(K),\rho(L))$ ($\rho : G \to G/G'$,* canonical homomorphism) *is a Hajós factorization of the factor group G/G' and $R = K + G'$, $T \in L \circ G'$.*

Proof. Explicitly, the theorem claims that (T,R) is a Hajós factorization iff there exist (K,L) reduced modulo G' such that $T \in L \circ G'$, $R = K + G'$, a chain of subgroups $G_0^* = G/G' > G_1^* > \cdots > G_{r-1}^* > G_r^* = \{0\}$ of G/G', subsets $H_1^*,\ldots,H_r^*$ of G/G' such that $G_{i-1}^* = H_i^* + G_i^*$, H_i^* reduced modulo G_i^*, $i = 1,\ldots,r$ and

(a) $\rho(K) \in \Delta(H_1^*,\ldots,H_r^*)$, $\rho(L) \in \Theta(H_1^*,\ldots,H_r^*)$ or

(b) $\rho(K) \in \Theta(H_1^*,\ldots,H_r^*)$, $\rho(L) \in \Delta(H_1^*,\ldots,H_r^*)$.

We take, for each $i = 1,\ldots,r$, a system H_i of coset representatives of H_i^* that is a subset H_i reduced modulo G' and $\rho(H_i) = H_i^*$. Clearly, putting $G_r = G'$, $G_{i-1} = H_i +$

G_i, $i=r,\dots,1$, we have $G_0=G>G_1>\cdots>G_{r-1}>G_r=G'$ a descending chain of subgroups of G with H_i reduced modulo G_i.

Now (a) is

$$\rho(K)\in\Delta_r(\rho(H_1),\dots,\rho(H_r)),\rho(L)\in\Theta_r(\rho(H_1),\dots,\rho(H_r)).$$

Since for any subsets $S,T\subseteq G$: $\rho(S)+\rho(T)=\rho(S+T)$, and if S reduced modulo G', $\rho(S)\circ\rho(T)=\rho(S\circ T)$, then (a) is equivalent to

$$\rho(K)\in\rho(\Delta_r(H_1,\dots,H_r)),\qquad \rho(L)\in\rho(\Theta_r(H_1,\dots,H_r))$$

that in turn is equivalent to

$$K\in K'\circ G',\ L\in L'\circ G'\quad\text{for some } K'\in\Delta_r(H_1,\dots,H_r),\ L'\in\Theta_r(H_1,\dots,H_r)$$

or

$$\begin{aligned}R\in(\Delta_r(H_1,\dots,H_r)\circ G')+G'&=\Delta_r(H_1,\dots,H_r)+G'\\&=(\Theta_{r-1}(H_1,\dots,H_{r-1})+H_r)+G_r\\&=\Theta_{r-1}(H_1,\dots,H_{r-1})+G_{r-1}\end{aligned}$$

and

$$\begin{aligned}T\in(\Theta_r(H_1,\dots,H_r)\circ G')\circ G'&=\Theta_r(H_1,\dots,H_r)\circ G'\\&=(\Delta_{r-1}(H_1,\dots,H_{r-1})\circ H_r)\circ G_r\\&=\Delta_{r-1}(H_1,\dots,H_{r-1})\circ G_{r-1}.\end{aligned}$$

Now set $H_1'=H_1,\dots,H_{r-1}'=H_{r-1}$, $H_r'=G_{r-1}$; $G_0'=G_0,\dots,G_{r-1}'=G_{r-1}$, $G_r'=\{0\}$. We have that $G_{i-1}'=G_i'+H_i'$, H_i' reduced modulo G_i' and, thus, (a) is equivalent to

$$R\in\Delta_r(H_1',\dots,H_r'),\qquad T\in\Theta_r(H_1',\dots,H_r')$$

that is when (R,T) is a Hajós factorization.

For the case (b), the argument goes similarly. The proof is complete. □

Last, but not least, the assertion below is originally the motivation of the Hajós construction.

Theorem 2.5 (Hajós [8]). *Every factorization of good group is a Hajós factorization.*

2.2. Hajós factorization of cyclic groups

Further on, we deal exclusively with factorizations of the cyclic group, that is, the additive group $\mathbf{Z}_n$ of residues modulo an integer n. Hajós factorization invariably means that of $\mathbf{Z}_n$ for some n. Usually, one represents an element of $\mathbf{Z}_n$ by its representative modulo n in a complete set of residues modulo n frequently which is $\{0,1,\dots,n-1\}$.

We find it convenient to give the following formulation of the concept of factorization of cyclic group in terms of integers. A *factorization* of $\mathbf{Z}_n$ is a pair of subsets (*summands*) (P,Q) of nonnegative integers such that for every $i=0,1,\ldots,n-1$ there exists one and only one pair $(p,q)\in(P,Q)$ satisfying $i\equiv p+q \bmod n$. We single out such factorizations that every number $0,1,\ldots,n-1$ is a (unique) sum,*without modulo n*,of two elements, each from one summand and that we call *Krasner factorizations.* The Krasner factorization is a Hajós one by the following characterization given by De Felice.

Theorem 2.6 (De Felice [6]). *A factorization (P,Q) of $\mathbf{Z}_n$ is a Hajós factorization if and only if there exists a Krasner factorization (I,J) such that both (I,Q) and (P,J) are factorizations of $\mathbf{Z}_n$.*

De Felice gives also the "polynomial" form of Hajós factorization. Let a be a variable, $\mathbb{N}$ be the set of natural numbers (including zero), $\mathbf{Z}$ the ring of integers, $\mathbf{Z}[a]$ the polynomial ring over $\mathbf{Z}$. For a subset T of $\mathbb{N}$, denote $a^T=\sum_{t\in T}a^t$ with $a^{\emptyset}=1$.

Theorem 2.7 (De Felice [6]). *A pair (P,Q) is a Hajós factorization of $\mathbf{Z}_n$ if and only if there exists a Krasner factorization (I,J) and subsets M,L of $\mathbb{N}$ such that $x^P=x^I(x^M(x-1)+1)$ and $x^Q=x^J(x^L(x-1)+1)$.*

We conclude this section with paraphrasing Theorem 2.4 into the language of $\mathbf{Z}_n$.

Theorem 2.8. *A factorization (P,Q) is a Hajós factorization of $\mathbf{Z}_n$ if and only if there are integers m and s such that $s>1$, $n=ms$ and a Hajós factorization (P_1,Q_1) of $\mathbf{Z}_m$ such that $P=P_1+\{0,\ldots,(s-1)m\}$ and $Q\in Q_1\circ\{0,\ldots,(s-1)m\}$, i.e. $Q=\{q_1+i_1m,\ldots,q_t+i_tm\}$ for some nonnegative integers $i_1,\ldots,i_t<s$, where $Q_1=\{q_1,\ldots,q_t\}$.*

Proof. Indeed, we choose the set $\{0,1,\ldots,n-1\}$ for representing $\mathbf{Z}_n$ without affecting the generality of consideration. Then a unique subgroup G' of order s of $\mathbf{Z}_n$, where $n=ms$, is $G'=\{0,m,\ldots,(s-1)m\}$. Putting $Q_1=\{q\colon\ 0\leqslant q<m,\ l=q+im,\ l\in L,\ i\geqslant 0\}$, $P_1=\{p\colon\ 0\leqslant p<m,\ k=p+jm,\ k\in K,\ j\geqslant 0\}$, we see that P_1,Q_1 are reduced modulo G' and $\rho(P_1)=\rho(K)$, $\rho(Q_1)=\rho(L)$. Hence, for a factorization (P,Q), the condition that $P=K+G'$, $Q\in L\circ G'$ and $(\rho(K),\rho(L))$ is a Hajós factorization of $\mathbf{Z}_n/G'=\mathbf{Z}_m$ holds if and only if $P=P_1+G'$, $Q\in Q_1\circ G'$ and (P_1,Q_1) is a Hajós factorization of $\mathbf{Z}_m$ that is the claim of Theorem 2.8.

3. Completion

Given a subset C, we define two subsets of integers:

$$P=\{i\colon\ a^ib^+\cap C\neq\emptyset\},\qquad Q=\{j\colon\ b^+a^j\cap C\neq\emptyset\}.$$

Finite maximal codes have an interesting property in connection with factorizations that when C is a code containing a^n and b, (P,Q) forms a factorization of $\mathbf{Z}_n$ (see [10]).

Thus, it is necessary for a code containing a^n and b to have finite completions that the corresponding pair (P,Q) be itself a factorization or be completable to a factorization of $\mathbf{Z}_n$. De Felice [6] showed that if (P,Q) is a Hajós factorization of $\mathbf{Z}_n$, the code $C=\{a^n\}\cup a^P b \cup ba^Q$, where given $T\subseteq \mathbb{N}$ a^T stands for the set $\{a^t\colon\ t\in T\}$, has a finite completion which is a d-code with $d\leqslant 3$. In this section, we make a stronger statement. But first, we begin by a recourse to the notion of composition of codes [1]. Let Z be a code on an alphabet A and Y be a code on an alphabet $B=\operatorname{alph} Y$, that is, B is the smallest alphabet containing all letters occuring in words of Y. We say that the codes Y and Z are *composable* (through β) if there is a bijection β from B onto Z. As β defines an injective morphism $B^*\to A^*$, the set $X=\beta(Y)\subseteq Z^*\subseteq A^*$ is a code which we denote by $X=Y\circ_\beta Z$ and we call a *composition* of Y and Z; we say also that X has a *decomposition* $X=Y\circ_\beta Z$.

The following proposition is instrumental.

Proposition 3.1 (Restivo et al. [10]). *If Y and Z have finite completions on A and B, respectively, then X has finite completions on A.*

As far as the Hajós factorization (P,Q) of $\mathbf{Z}_n$ is concerned, by Theorem 2.8, we can choose nonnegative integers $s,m,\ i_1,i_2,\dots,i_t$ so that $s>1$, $n=sm$ and

$$P=P_1+\{0,m,\dots,(s-1)m\},$$

$$Q=\{q_1+i_1m,q_2+i_2m,\dots,q_t+i_tm\},$$

where $Q_1=\{q_1,q_2,\dots,q_t\}$ and (P_1,Q_1) is, again, a Hajós factorization of $\mathbf{Z}_m$. The result of this section is the following:

Theorem 3.2. *The code $C=\{a^n\}\cup a^P ba^Q$, where (P,Q) a Hajós factorization of $\mathbf{Z}_n$, for any $n\geqslant 1$, has a finite completion.*

In order to prove Theorem 3.2 we need two further simple lemmas.

Lemma 3.3. *Let $C=\{a^m\}\cup a^{P_1}ba^{Q_1}$ be a code having finite completions. Then $C_1=\{a^m\}\cup a^{P_1}ba^{Q}$ is also a code having finite completions.*

Proof of Lemma 3.3. The fact that C_1 is a code, requires only a straightforward check. Consider now the mapping β from the alphabet $B=\{x_a\}\cup\{x_{pq}\colon\ p\in P_1,\ q\in Q_1\}$ into $A^*=\{a,b\}^*$ defined as $\beta(x_a)=a^m$, $\beta(x_{pq})=a^pba^q$, $p\in P_1$, $q\in Q_1$. The set $Y=\{x_a\}\cup\{x_{pq_t}x_a^{i_t}\colon\ p\in P_1,\ q_t\in Q_1\}$ is clearly a prefix code on the alphabet B. Since we have the decomposition $C_1=Y\circ_\beta C$, and C has a finite completion by assumption, Proposition 3.1 guarantees C_1 to have finite completions. □

Lemma 3.4. *If a code C containing a^m is finite maximal then $C_1=\{a^{sm}\}\cup\{\varepsilon,a^{2m},\dots,a^{(s-1)m}\}(C\backslash\{a^m\})$ is also a finite maximal code for $s\geqslant 1$.*

Proof of Lemma 3.4. First, that C_1 is a code, is verified by a direct verification. Further, it is maximal since, under the Bernoulli distribution $\mu(a)=\frac{1}{2}$, $\mu(b)=\frac{1}{2}$ the value $\mu(C)=1$ and

$$\mu(C_1)=\left(1+\cdots+\frac{1}{2^{(s-1)m}}\right)\mu(C\backslash\{a^m\})+\frac{1}{2^{sm}}$$
$$=\frac{1-1/2^{sm}}{1-1/2^m}\left(1-\frac{1}{2^m}\right)+\frac{1}{2^{sm}}=1. \qquad \square$$

Proof of Theorem 3.2. The proof now is done by induction on n. The conclusion is trivial when $n=1$: $P=\{p\}$, $Q=\{q\}$ and $C=\{a,a^pba^q\}$ is two-element code which has finite completions in virtue of [10, Corollary 2]. Suppose that it holds for all integers $<n$. By induction hypothesis, the code $C_2=\{a^m\}\cup a^{P_1}ba^{Q_1}$ has a finite completion. By Lemma 3.3, the code $C_1=\{a^m\}\cup a^{P_1}ba^{Q}$ has finite completions and by Lemma 3.4 the code $C=\{a^{sm}\}\cup\{\varepsilon,\dots,a^{(s-1)m}\}(C_1\backslash\{a^m\})=\{a^n\}\cup\{\varepsilon,\dots,a^{(s-1)m}\}a^{P_1}ba^{Q}=\{a^n\}\cup a^{P_1+\{0,\dots,(s-1)m\}}ba^{Q}=\{a^n\}\cup a^{P}ba^{Q}$ has finite completions. The proof is completed. $\square$

We should remark that the foregoing theorem implicitly provides a constructive procedure for completion. Consider a small illustrative example.

Example 3.5. Let

$$C=\{a^6,b,ba,ba^5,a^3b,a^3ba,a^3ba^5\}$$

with $n=6$, $(P,Q)=(\{0,3\},\ \{0,1,5\})$ a Hajós factorization of $\mathbf{Z}_6$. We have $(P_1,Q_1)=(\{0\},\{0,1,2\})$, a factorization of $\mathbf{Z}_3$.

Put $C_0=\{a^3\}\cup a^{P_1}ba^{Q_1}=\{a^3,b,ba,ba^2\}$ which is a maximal code.

Put $C_1=\{a^3\}\cup a^{P_1}ba^{Q}$. Let $B=\{x_a,x_{00},x_{01},x_{02}\}$. Put $Y=\{x_a,x_{00},x_{01},x_{02}x_a\}$ which has, as a prefix code, a (prefix-maximal) completion $\bar{Y}=\{x_a,x_{00},x_{01},x_{02}x_a;\ x_{02}x_{00},x_{02}x_{01},x_{02}x_{02}\}$. Put $\beta(x_a)=a^3$, $\beta(x_{00})=b$, $\beta(x_{01})=ba$ and $\beta(x_{02})=ba^2$. Then C_1 has a completion $\bar{C}_1=\bar{Y}\circ_\beta\bar{C}_0=\{a^3,b,ba,ba^5;ba^2b,ba^2ba,ba^2ba^2\}$.

Finally C has a finite completion $\bar{C}=\{a^6\}\cup\{\varepsilon,a^3\}(\bar{C}_1\backslash\{a^3\})=\{a^6,b,ba,ba^5;\ ba^2b,ba^2ba,ba^2ba^2;a^3b,a^3ba,a^3ba^5;a^3ba^2b,a^3ba^2ba,a^3ba^2ba^2\}$.

4. Embedding in d-codes

In this section, we inquire into when a code, of the form under consideration is included in a d-code, i.e. a code, each word of which contain no more than d occurences of b, for some $d\leqslant 3$.

For a finite alphabet, let denote $\mathbf{Z}\langle A\rangle$ the ring of polynomials with integer coefficients in noncommuting variables from A. It is common to identify a finite subset X of A^* with its characteristic polynomial $\chi(X)=\sum_{w\in X}w$. Characteristic polynomial has all nonzero coefficients equal to 1 and, conversely, a polynomial having 0 or 1 as coefficients is indeed a characteristic one of a unique subset of A^*.

Now let (P,Q) be a Hajós factorization of $\mathbf{Z}_n$ and (I,J) a Krasner one. Let us say that the Krasner factorization (I,J) is a companion factorization of (P,Q) if (I,Q) and (P,J) are both factorizations of $\mathbf{Z}_n$ (see Theorem 2.6). Notice that given (P,Q), (I,J) the subsets M,L as in Theorem 2.7 are uniquely determined.

In the following assertion we locate a property that the code $C=\{a^n\}\cup a^Pba^Q$ having a d-code, $d\leqslant 3$, as a completion must possess.

Proposition 4.1. *The code $C=\{a^n\}\cup a^Pba^Q$, where (P,Q) is a factorization $\mathbf{Z}_n$, has a finite maximal d-code, $d\leqslant 3$, as a completion if and only if there exists a Krasner companion factorization of (P,Q) for which at least one of the polynomials $a^M(a-1)a^L+a^L$ and $a^M(a-1)a^L+a^M$ has nonnegative coefficients.*

Proof. Let X be a code, we denote by X_1 the subset of words of X containing exactly one occurence of b. Concerning X_1, when X is a maximal d-code, $d\leqslant 3$, the results from Restivo [9, Theorem 3.5], De Felice [4, Theorem 5.1] and [3, Theorems 4.3 and 4.6] can be summed up in the following formulation:

Proposition 4.2. *Let X be a finite maximal d-code, $d\leqslant 3$ with $a^n\in X$. Then, there exists a Krasner factorization (I,J) of $\mathbf{Z}_n$ so that X_1, or its mirror image, has the form*

$$X_1=a^Iba^J+\sum_{t\in T}a^I(a-1)a^{M_t}ba^t+\sum_{i\in I}a^iba^{L_i}(a-1)a^J,$$

where for every $i\in I$, L_i,R_i, T are subsets of $\mathbb{N}$ related by $a^{R_i}=a^J(a^{L_i}(a-1)+1)$, $T\subseteq\bigcup_{i\in I}R_i$; for every $t\in T$, M_t is a subset of $\mathbb{N}$ such that the polynomial $a^I(a^{M_t}(a-1)+1)$ is of nonnegative coefficients and if $t\notin J$ the polynomial $a^{M_t}(a-1)a^{L_i}+a^{M_t}$ is of nonnegative coefficients for all $i\in I$.

Suppose that $C=\{a^n\}\cup a^Pba^Q$ has a finite completion X, which is a d-code X, $d\leqslant 3$. Then X_1 has the form described in the Proposition 4.2. Indeed, we have the inclusion $a^Pba^Q\subseteq X_1$. Since $a^n\in X$ and a^Pba^Q makes up n words in all, we conclude then that $X_1=a^Pba^Q$ [1]; [7, Proposition 2.1]. It is known that once $a^I(a^S(a-1)+1)$ is of nonnegative coefficients, for S,I subsets of nonnegative integers, it is of coefficients 0 or 1 [3, Remark 6.7]. Therefore, we can put T_t the subset of $\mathbb{N}$ such that

$$a^{T_t}=a^I(a^{M_t}(a-1)+1). \tag{1}$$

We rewrite X_1 as

$$X_1=\sum_{i\in I}a^iba^{R_i}+\sum_{t\in T}a^{T_t}ba^t-a^Iba^T=a^Pba^Q. \tag{2}$$

For each $t\in T$, set $K_t=\{i\in I:R_i\ni t\}$. Since $X_1\cap a^*ba^t=a^{K_t}ba^t+a^{T_t\setminus I}ba^t-a^{I\setminus T_t}ba^t$ has coefficients $0,1$, we must have $I\setminus T_t\subseteq K_t$ and

$$X_1\cap a^*ba^t=a^{T_t\setminus I}ba^t+a^{K_t\setminus(I\setminus T_t)}ba^t=a^{T_t\setminus I\cup K_t\setminus(I\setminus T_t)}ba^t. \tag{3}$$

If there is $t \in T \backslash Q$ then the left-hand side of (3) is an emptyset which forces the right-hand one to vanish, namely, $T_t \backslash I = \emptyset$, or $T_t \subseteq I$ and $K_t \backslash (I \backslash T_t) = \emptyset$, or $K_t = I \backslash T_t$. Since $|T_t| = |I|$ we have $T_t = I$ and thus $K_t = I \backslash T_t = \emptyset$. This means that $t \notin R_i$ for all $i \in I$: a contradiction. So $T \subseteq Q$. Also it is evident by (2) that $R_i \subseteq Q$ for every $i \in I$.

Now we prove that $R_i = Q$ for all $i \in I$ and $T_t = P$ for all $t \in T$. Put $H_i = \{t \in T : T_t \ni i\}$. Then

$$X_1 \cap a^i ba^* = a^i ba^{R_i} + a^i b^{H_i} - a^i ba^T = a^i b^{R_i} - a^i ba^{T \backslash H_i}. \tag{4}$$

If $P \cap I \neq \emptyset$, fix any $p \in P \cap I$. Then $X_1 \cap a^p ba^* = a^p ba^Q$ and, by (4), $Q = R_p \backslash (T \backslash H_p)$. But $R_p \subseteq Q$ $(p \in I)$ implies $R_p = Q$ which, in turn, together with the fact that $|R_i| = |R_p| = |Q|$ and $R_i \subseteq Q$ implies for all i that $R_i = Q$. Next, for every $t \in T$, by (3) and by $K_t \subseteq I$, we have $P = T_t \backslash I \cup K_t \backslash (I \backslash T_t) \subseteq T_t \backslash I \cup I \backslash (I \backslash T_t) = T_t \backslash I \cup (I \cap T_t) = T_t$. Now the fact that $|Q| = |R_i| = |J|$ implies $|P| = |I|$ and take into account $P \subseteq T_t$ and $|I| = |T_t|$, we get $P = T_t$ for all $t \in T$.

If, on the contrary, $P \cap I = \emptyset$ then for every i the left-hand side of (4) is empty, hence $R_i = T \backslash H_i$, in particular $R_i \subseteq T$. So $T = \bigcup_{i \in I} R_i$ in view of the assumption. From (2) it follows immediately that $Q \subseteq T$. In this instance, together with the fact that $T \subseteq Q$, we get $T = Q$. The relation $P = T_t \backslash I \cup K_t \backslash (I \backslash T_t)$, since $K_t \subseteq I$ and $P \cap I = \emptyset$, gives $K_t \backslash (I \backslash T_t) = \emptyset$ and $P = T_t \backslash I$. Since $T_t \subseteq P \cup I$ we get $T_t = P \cup (T_t \cap I)$ for every t. But now this fact implies T_t are the same for all t by the following observation. For any $s, t \in T$, we have by (1)

$$a^P + a^{T_s \cap I} = a^I (a^{M_s}(a-1) + 1),$$

$$a^P + a^{T_t \cap I} = a^I (a^{M_t}(a-1) + 1).$$

Subtracting the second equation from the first one we come to

$$a^{T_s \cap I} - a^{T_t \cap I} = a^I (a^{M_s} - a^{M_t})(a-1).$$

If not identically zero, the left-hand side is a polynomial of a degree lesser than that in the right-hand one, hence, $a^{T_s \cap I}$ must equal to $a^{T_t \cap I}$ that leads to $T_s = T_t$. As $i \in T_t$ if and only if $t \in H_i$ it follows that H_i are all the same (notice that H_i are of the same cardinality) and, thus, $R_i = T \backslash H_i$ are all the same. Since $Q = \bigcup_{i \in I} R_i$, we have $R_i = Q$ for all i. Finally, as $T = Q$, we have $H_i = \emptyset$ for all i, so $T_t \cap I = \emptyset$ that means $T_t = P$ for all t.

The equalities $R_i = Q$ and $T_t = P$ for all $i \in I, t \in T$ show that the Krasner factorization (I, J) is a companion factorization for (P, Q) and all M_t are the same, $M_t = M$ all L_i are the same, $L_i = L$ and indeed for them the polynomial $a^M(a-1)a^L + a^M$ must have been in nonnegative coefficients.

Conversely, given a Krasner companion factorization (I, J) of (P, Q) for them $a^M (a-1)a^L + a^M$ holds. We explicitly give one maximal d-code, $d \leqslant 3$ that contains C. Consider the following polynomial

$$\chi = (a^I + a^I ba^L)(a + b - 1)(a^J a^M ba^Q) + 1.$$

Expanding and summing up the terms of the same occurence of b we have

$$\chi = a^n + a^P ba^Q + a^I b(a^M(a-1)a^L + a^M) + a^I ba^L ba^J + a^I ba^L ba^M ba^Q$$

which virtually has nonnegative coefficients by virtue of the assumption that all coefficients of $a^M(a-1)a^L + a^M$ are nonnegative. By a result of Reutemauer (see [4, Remark 2.1]), χ has coefficients 0,1 and the corresponding subset of A^* of which χ is the characteristic polynomial, is a finite maximal d-code containing $C = a^P ba^Q$ with $d \leqslant 3$. The proof is complete. □

The completion procedure in Section 3 does not lead to 3-codes in all instances, unlike De Felice's result. This is not a defect to the method, since we can indicate, in the next last section, a Hajós factorization that do not satisfy the Proposition 4.1, hence, the corresponding code is not completable to a finite maximal d-code with $d \leqslant 3$. We do not know if there is a universal constant d such that $a^n + a^P ba^Q$ has a finite d-code as a completion for all n and all Hajós factorizations (P, Q) of $\mathbf{Z}_n$. However, if d is supposed to depend on n, such a constant exists. Analyzing the proof of Theorem 3.2 (Lemma 3.3) one can find that one possible value is $d(n) = n$.

5. Example

Here we construct Hajós factorizations with the property that for each of their Krasner companion factorizations the corresponding polynomials $a^M(a-1)a^L + a^L$ and $a^M(a-1)a^L + a^M$ both have some negative coefficients. More specifically, the Hajós factorizations of our considerations will possess only one Krasner companion factorization and this one shall possess the indicated property. They are as follows.

Let $n = 8p$, p a positive integer and

$$P = \{0, 4, 6, \ldots, 2p-2, 2p+2; 4p, 4p+4, 4p+6, \ldots, 6p-2, 6p+2\},$$

$$Q = \{0, 3, 2p, 6p+3\},$$

$$I = \{0, 2, \ldots, 2p-2; 4p, 4p+2, \ldots, 6p-2\}, \qquad J = \{0, 1, 2p, 2p+1\}.$$

First, we see that (I, J) is a Krasner factorization, (P, J) and (I, Q) are factorizations. Second, (P, Q) is a Hajós factorization of $\mathbf{Z}_n$: $a^P = a^I((a-1)a^M + 1)$ and $a^Q = a^J(a^L(a-1)+1)$ with $M = \{0, 2p+2, 2p+4, \ldots, 4p\} + 1$ and $L = \{2, 3\}$.

We can verify that

$$a^M(a-1)a^L + a^L$$

has the coefficient of a^3 equal -1 and the polynomial

$$a^M(a-1)a^L + a^M$$

has the coefficient of a^{2p+5} equal -1.

We show that the above Krasner factorization is the only companion factorization of (P,Q) if p is a prime number not 3. We employ always the fact that Hajós factorizations are periodic.

Let (I,J) be a Krasner companion factorization of (P,Q) of $\mathbf{Z}_{8p}$. Note that (I,Q) and (P,J) by definition are also Hajós factorizations (with the same companion (I,J)). Since Q is nonperiodic modulo $8p$, it follows that I periodic modulo $8p$ of possible periods $d=2,4,8,p,2p$ or $4p$.

If the period is 2 then $|I|\geqslant 4p$ which is impossible. If the period is 4 then $I=\{0,4,8,\ldots,8p-4\}$ which implies $J=\{0,1,2,3\}$ contradicting (P,J) a factorization of $\mathbf{Z}_{8p}$. If the period is 8 then $I=\{0,8,\ldots,8(p-1);i,i+8,\ldots,i+8(p-1)\}$ for some i, $0<i<8$ which implies $\{0,1,\ldots,i-1\}\subseteq J$, hence $i\leqslant 4$. If $i\geqslant 3$ then $2\in J$ and if $i<3$ then $4\in J$ as $2i-1<5$ and (I,J) a Krasner factorization. But this is in contradiction with (P,J) a factorization by the fact that P includes 4, 6 when $p>3$. If $p=2$ then $3,4\in Q$ and (I,Q) cannot be a factorization of $\mathbf{Z}_{16}$ by $1+3=4$, $0+3=3+0$ or $4+0=4$ when $i=1,3$ or 4. So remains the case $p=2, i=2$ which leads to $I=\{0,2,8,10\}$, $J=\{0,1,4,5\}$.

Next, if the period is p or $2p$, we have $2p\in I$, but then the equality $2p+6p=8p=0$ modulo $8p$ shows a contradiction. It remains to consider I to be periodic with period $4p$.

We write then $I=I_1+\{0,4p\}$, $P=P_1+\{0,4p\}$ with $P_1=\{0,4,\ldots,2p-2,2p+2\}$ and put $Q_1=\{0,3,2p,2p+3\}$. We see that (I_1,J) is a Krasner factorization, (P_1,J), (I_1,Q_1) and (P_1,Q_1) are (Hajós) factorizations of $\mathbf{Z}_{4p}$.

Clearly, P_1 is nonperiodic modulo $4p$, so J must be periodic. If J is periodic with period p, then $J=\{0,p,2p,3p\}$, hence $I_1=\{0,1,\ldots,p-1\}$. But then (I_1,Q_1) is not a factorization of $\mathbf{Z}_{4p}$ by $(2p+3)+0=2p+3$ when $p>3$ and $1+3=4=2p$ when $p=2$. Therefore, J is periodic with period $2p$. Put $J=J_1+\{0,2p\}$, $P_{11}=\{0,2,4,\ldots,2(p-1)\}$ and put $Q_{11}=\{0,3\}$. We have then (P_{11},J_1), (I_1,Q_{11}), (P_{11},Q_{11}) are Hajós factorizations and (I_1,J_1) is a Krasner factorization of $\mathbf{Z}_{2p}$. Now it is straightforward to see that either $J_1=\{0,p\}$, $I_1=\{0,1,\ldots,p-1\}$ which is impossible since $3\in I_1, 3\in Q_{11}$ when $p>3$ and $1\in I_1, 3\in Q_{11}$ when $p=2$ are in contradiction with (I_1,Q_{11}) a factorization; or $J_1=\{0,1\}$, $I_1=\{0,2,\ldots,2(p-1)\}$. This case leads to $J=\{0,1,2p,2p+1\}$ and $I=\{0,2,\ldots,2p-2;4p,4p+2,\ldots,6p-2\}$. Note that when $p=2$, $I=\{0,2,8,10\}$ and $J=\{0,1,4,5\}$ what we worked out before. The uniqueness is proved.

Remark 4.3. A Hajós factorization may have multiple companions. This is the case for $p=3$ in the above example. Let $P=\{0,4,8;12,16,20\}$ and $Q=\{0,3,6,21\}$ form a Hajós factorization of $\mathbf{Z}_{24}$. Beside the Krasner factorization above with $p=3$

$$I=\{0,2,4;12,14,16\},\qquad J=\{0,1,6,7\}$$

it has another companion

$$I'=\{0,4,8,12,16,20\},\qquad J'=\{0,1,2,3\}.$$

Moreover, this companion has $M=\emptyset$ and, thus, the corresponding polynomials both have nonnegative coefficients.

We should note that Theorem 5.1 of [6] also provides a proof of the uniqueness of Krasner companion, p is prime not 3, in a more systematic way.

Acknowledgements

The author is greatly indebted to D. Perrin for his willing assitance. Thanks also go to the referee for his sharp and constructive criticism, in particular, for pointing out the Proposition 4.2 and Remark 4.3.

References

[1] J. Berstel, D. Perrin, Theory of Codes, Academic Press, New York, 1985.

[2] N.G. de Bruijn, On the Factorization of Finite Abelian groups, Indag. Math. 15 (1953) 258–264.

[3] C. De Felice, Construction of a family of finite maximal codes, Theoret. Comput. Sci. 63 (1989) 157–184.

[4] C. De Felice, On the Factorization Conjecture, Lecture Notes in Computer Science, vol. 577 (1992) Springer Berlin, pp. 545–556.

[5] C. De Felice, Completing codes by Hajós factorization of groups, in: J. Almeida, G.M.S. Gomes, P.V. Silva (Eds.), Semigroups, Automata and Languages, World Scientific, Porto, 1996, pp. 59–66.

[6] C. De Felice, An Application of Hajós factorizations to variable-length codes, Theoret. Comput. Sci. 164 (1996) 223–252.

[7] C. De Felice, A. Restivo, Some results on finite maximal codes, RAIRO Inform. Théor. Appl. 19 (1985) 383–403.

[8] G. Hajós, Sur la factorisation des groupes abéliens, Časopis Pest, Mat. Fys. 74 (1950) 157–162.

[9] A. Restivo, On codes having no finite completions, Discrete Math. 17 (1977) 309–316.

[10] A. Restivo, S. Salemi, T. Sportelli, Completing codes, RAIRO Inform. Théor. Appl. 23 (1989) 135–147.

[11] A.D. Sands, On the factorisation of finite abelian groups, Acta Math. Acad. Sci. Hungar. 8 (1957) 65–86.