Article

DNA Code Constructions from Some Codes over a Family of Finite Rings

Abdullah Dertli*[1] & Yasemin Çengellenmiş[2]

[1]Dept. of Math., Ondokuz Mayıs University, Samsun, Turkey
2Dept. of Math., Trakya University, Edirne, Turkey

Abstract

Firstly, by using a non-trivial automorphism θ_i over A_i, where $i = 2,3, \dots, k$, we define the skew cyclic codes over a family of the rings A_k. By using them, we obtain the reversible DNA codes. In the second method, the necessary and sufficient condition of cyclic code over A_i to be reversible and reversible complement are given, where $i = 2,3, \dots, k$. By introducing a map, the DNA codes are obtained from them. In the third method, the special linear codes over A_i are established, where $i = 2,3, \dots, k$. By using them, the reversible and reversible complement DNA codes are obtained.

Keywords: DNA codes, skew codes, reversibility.

1. Introduction

The transmission and storage of information take place in digital platforms and the coding theory is necessary to correct and detect errors in the platform. There is another platform. In the platform, correcting and detecting errors is necessary but it does not take place digital. It is DNA.

It is well known that DNA contains a genetic program for the biological development of life and has two strands which are linked by Watson-Crick pairing so that every A is linked with a T and every C with a G, and vice versa, where A, T, C, G are the four bases of a DNA sequence. The idea of computing with DNA was given by T. Head in [8]. L. Adleman performed the computation using DNA strands in [1].

A specific set of DNA sequences are required to perform computation using DNA strands with particular properties. This paper aims to obtain the set of DNA strands satisfying various constraints, by using some error-correcting codes over a family of finite rings which enjoy DNA properties. One of the constraints is reverse constraints. This leads to reversible codes. The other one is the reverse complement constraint. This leads to reversible complement codes.

To obtain reversible DNA codes, some authors considered skew cyclic codes. The reversibility problem for DNA 8-bases and DNA $2^{s+1}k$-bases is solved in [5] and [6] respectively by using

*Correspondence Author: Abdullah Dertli, Department of Mathematics, Ondokuz Mayıs University, Samsun, Turkey.
Email: abdullah.dertli@gmail.com

skew cyclic codes over the finite rings $F_{16} + uF_{16} + vF_{16} + uvF_{16}$ where $u^2 = u, v^2 = v, uv = vu$ and $F_{4^{2k}}[u_1, \dots, u_s]/\langle u_1{}^2 - u_1, \dots, u_s^2 - u_s \rangle$ where $k, s \geqslant 1, u_i u_j = u_j u_i$. The r eversibility problem arises from the fact that the pairing of nucleotides in two different strands of a DNA sequence is done in the opposite direction and in reverse order. For example, let us consider the codeword (DNA string) GTTAGGCA which corresponds to a codeword (a_1, a_2). The reverse of (a_1, a_2) is (a_2, a_1). However, the vector (a_2, a_1) corresponds to GGCAGTTA which is not the reverse of GTTAGGCA. The reverse of GTTAGGCA is $ACGGATTG$.

To obtain the DNA codes, some authors used cyclic DNA codes of length n that enjoy some of the properties of DNA. In [10], by introducing a map, a family of cyclic codes over the ring $F_2[u]/< u^4 - 1 >$ is mapped to DNA codes.

In [11], the design of linear codes over $F_2 + uF_2 + vF_2 + uvF_2, u^2 = 0, v^2 = v, uv = vu, F_2 = \{0,1\}$ is presented by using σ-set, where σ is a nontrivial automorphism on this finite ring. By using these linear codes, the authors obtained DNA codes with the other method.

In this paper, firstly, a non-trivial automorphism θ_i over $A_i = A_{i-1} + u_i A_{i-1}$, where $i = 2,3, \dots, k, u_i^2 = u_i, A_0 = F_2$ is defined. By introducing skew cyclic codes over a family of the finite rings $A_k = F_2[u_1, u_2, \dots, u_k]/< u_i^2 - u_i, u_i u_j - u_j\, u_i >$, where $1 \leq i, j \leq k$, the reversible DNA codes are obtained from them. With the other method, the necessary and sufficient conditions of cyclic codes over A_i, where $i = 2, \dots, k$ to be reversible and reversible complement is given. By introducing a map, the DNA codes are obtained from these types of codes. As a last, the linear codes over A_i are designed, by using θ_i-set for $i = 2,3, \dots, k$. By using these type codes, the reversible or reversible complement DNA codes are obtained.

2. Preliminaries

In [4], a family of the finite rings $A_k = F_2[u_1, u_2, \dots, u_k]/\langle u_i^2 - u_i, u_i u_j - u_j u_i \rangle$, where $1 \leq i, j \leq k, A_0 = F_2$ was introduced. It contains the commutative finite rings with characteristic 2 and cardinality $4^{2^{i-1}}$ for $i = 1,2, \dots, k$.

The finite rings of the family are written as recursively

$$A_i = A_{i-1} + u_i A_{i-1}$$

where $u_i^2 = u_i, i = 1,2,3, \dots, k, A_0 = F_2$. For example $A_2 = A_1 + u_2 A_1 = (F_2 + u_1 F_2) + u_2(F_2 + u_1 F_2), u_1^2 = u_1, u_2^2 = u_2, u_1 u_2 = u_2 u_1, F_2 = \{0,1\}$.

In [4], the map on A_i where $i = 1, , \dots, k$ was defined as follows

$$\begin{aligned} \varphi_i \quad &: A_i \longrightarrow A_{i-1}^2 \\ x_{i-1} + y_{i-1} u_i \quad &\mapsto (x_{i-1}, x_{i-1} + y_{i-1}) \end{aligned}$$

where $x_{i-1}, y_{i-1} \in A_{i-1}, u_i^2 = u_i$, for $i = 1,2,3, \dots, k$ and $A_0 = F_2$.

By using a correspondence ξ_1 between the elements A_1 and the set $\{A,T,C,G\}$ such as $\xi_1(0)=A,\xi_1(1)=T,\xi_1(u_1)=G,\xi_1(1+u_1)=C$, we define a correspondence ξ_2 between the elements of the finite ring $A_2=F_2+u_1F_2+u_2F_2+u_1u_2F_2$, where $u_1^2=u_1,u_2^2=u_2,u_1u_2=u_2u_1$ and

DNA double pairs as follows

Elements α	Gray images	DNA double pairs $\xi_2(\alpha)$
0	$(0,0)$	AA
1	$(1,1)$	TT
u_1	(u_1,u_1)	GG
u_2	$(0,1)$	AT
u_1u_2	$(0,u_1)$	AG
$1+u_1$	$(1+u_1,1+u_1)$	CC
$1+u_2$	$(1,0)$	TA
u_1+u_2	$(u_1,1+u_1)$	GC
$u_1+u_1u_2$	$(u_1,0)$	GA
$u_2+u_1u_2$	$(0,1+u_1)$	AC
$1+u_1u_2$	$(1,1+u_1)$	TC
$1+u_1+u_1u_2$	$(1+u_1,1)$	CT
$1+u_1+u_2$	$(1+u_1,u_1)$	CG
$1+u_2+u_1u_2$	$(1,u_1)$	TG
$u_1+u_2+u_1u_2$	$(u_1,1)$	GT
$1+u_1+u_2+u_1u_2$	$(1+u_1,0)$	CA

By using the map φ_3 and ξ_2, we established ξ_3 correspondence between the element of A_3 and DNA 4-bases $\beta=x_2+y_2u_3\longmapsto\big(\xi_2(x_2),\xi_2(x_2+y_2)\big)$ as follows

Elements β	DNA 4-bases $\xi_3(\beta)$
0	$AAAA$
1	$TTTT$
u_1	$GGGG$
u_2	$ATAT$
u_3	$AATT$

$$\vdots$$

By using the matching and the elements of A_2 and $S_{D_{16}}=\{AA,TT,\dots,GG\}$ and by using the Gray map from A_i to A_{i-1}^2, we can define ξ_i correspondence between the elements of the finite ring A_i and DNA 2^{i-1}-bases for $i=2,\dots,k$ as follows

$$\begin{aligned}\xi_i\quad &: A_i\longrightarrow A_{i-1}^2\longrightarrow\{A,T,G,C\}^{2^{i-1}}\\ x_{i-1}+y_{i-1}u_i\quad &\longmapsto\ (x_{i-1},x_{i-1}+y_{i-1})\longmapsto l\end{aligned}$$

where $l=\big(\xi_{i-1}(x_{i-1}),\xi_{i-1}(x_{i-1}+y_{i-1})\big)$.

It can be written that $\xi_i = \gamma_i \varphi_i$, where a map γ_i is defined from A_{i-1}^2 to 2^{i-1}-bases as follows,

$$\gamma_i(s_{i-1}, t_{i-1}) = \left(\xi_{i-1}(s_{i-1}), \xi_{i-1}(t_{i-1})\right)$$

where $s_{i-1}, t_{i-1} \in A_{i-1}$ for $i = 2, \dots, k$.

In [2], a nontrivial automorphism was defined on A_2 as follows

$$\begin{array}{rcl} \theta_2 & : & A_2 \longrightarrow A_2 \\ x_1 + y_1 u_2 & \longmapsto & x_1 + (1 + u_2) y_1 \end{array}$$

where $x_1, y_1 \in A_1 = F_2 + u_1 F_2, u_1^2 = u_1$.

By defining a nontrivial automorphism on A_i as follows, for $i = 3, \dots, k$, we can define the skew cyclic codes over A_i, for $i = 2, \dots, k$.

$$\begin{array}{rl} \theta_i & : A_i \longrightarrow A_i \\ x_{i-1} + y_{i-1} u_i & \longmapsto \theta_i(x_{i-1} + y_{i-1} u_i) = q \end{array}$$

where $q = \theta_{i-1}(x_{i-1}) + (1 + u_i)\theta_{i-1}(y_{i-1}), x_{i-1}, y_{i-1} \in A_{i-1}$, for $i = 3, \dots, k$. The order of θ_i, for $i = 2, \dots, k$ is 2.

The rings

$$A_i[x, \theta_i] = \left\{b_0^i + b_1^i x + \cdots + b_{n-1}^i x^{n-1} : b_j^i \in A_i, n \in N, = 2, \dots, k, j = 0,1, \dots, n-1\right\}$$

are skew polynomial rings with the usual polynomial addition and the multiplication as follows

$$(a_i x^s)\left(b_i x^j\right) = a_i \theta_i^s(b_i) x^{s+j}$$

where $a_i, b_i \in A_i$, for $i = 2, \dots, k$. They are non-commutative rings.

Definition 1: *A subset C_i of A_i^n, where $i = 2, \dots, k$ is called a skew cyclic code of length n if C_i satisfies the following conditions,*

1. *C_i is a submodule of A_i^n*
2. *If $\boldsymbol{c}_i = \left(c_0^i, c_1^i, \dots, c_{n-1}^i\right) \in C_i$, then $\theta_i(\boldsymbol{c}_i) = \left(\theta_i\left(c_{n-1}^i\right), \theta_i\left(c_0^i\right), \dots, \theta_i\left(c_{n-2}^i\right)\right) \in C_i$,*

where θ_i is the skew cyclic shift operator.

In polynomial representation, a skew cyclic code of length n over A_i is defined as a left ideal of the quotient ring $A_{i_{\theta_i,n}} = A_i[x, \theta_i]/\langle x^n - 1\rangle$, if the order of θ_i divides n, that is, if n is even. If the order of θ_i does not divide n, a skew cyclic code of length n over A_i is defined as a left $A_i[x, \theta_i]$-submodule of $A_{i_{\theta_i,n}}$, since the set $A_{i_{\theta_i,n}} = A_i[x, \theta_i]/\langle x^n - 1\rangle = \{f_i(x) + \langle x^n - 1\rangle : f_i(x) \in A_i[x, \theta_i]\}$ is a left $A_i[x, \theta_i]$-module with the multiplication from left defined by

$$r_i(x)(f_i(x) + \langle x^n - 1\rangle) = r_i(x) f_i(x) + \langle x^n - 1\rangle$$

for any $r_i(x) \in A_i[x, \theta_i]$.

In either case, the following holds.

Theorem 2: *Let C_i be a skew cyclic code over A_i and let $f_i(x)$ be a polynomial in C_i of minimal degree, where $i = 2, \dots, k$. If the leading coefficient of $f_i(x)$ is a unit in A_i, then $C_i = \langle f_i(x) \rangle$, where $f_i(x)$ is a right divisor of $x^n - 1$.*

3. Reversible DNA codes

In this section, the reversible DNA codes are obtained by using the skew cyclic codes over A_i for $i = 2, \dots, k$.

Definition 3: *For $\boldsymbol{x}_i = (x_0^i, x_1^i, \dots, x_{n-1}^i) \in A_i^n$, the vector $(x_{n-1}^i, x_{n-2}^i, \dots, x_1^i, x_0^i)$ is called the reverse of $\boldsymbol{x}_i$ and is denoted by $\boldsymbol{x}_i^r$. A linear code C_i of length n over A_i is said to be reversible if $\boldsymbol{x}_i^r \in C_i$ for every $\boldsymbol{x}_i \in C_i$, where $i = 2, \dots, k$.*

We can express the matching the elements of A_2 and $S_{D_{16}} = \{AA, TT, \dots, GG\}$ by means of the automorphism θ_2 as follows.

Each element $\alpha_2 = x + yu_2 \in A_2$, where $x, y \in A_1 = F_2 + u_1 F_2, u_1^2 = u_1$ and $\theta_2(\alpha_2)$ are mapped to DNA double pairs which are reverse of each other. Since a correspondence the elements of the finite ring A_2 and DNA double pairs is ξ_2, so we have $\xi_2(u_2) = AT$, while $\xi_2(\theta_2(u_2)) = TA$.

By using a map $\xi_i = \gamma_i \circ \varphi_i$, where the map γ_i is from A_{i-1}^2 to 2^{i-1}-bases as follows,

$$\gamma_i(s_{i-1}, t_{i-1}) = (\xi_{i-1}(s_{i-1}), \xi_{i-1}(t_{i-1}))$$

where $s_{i-1}, t_{i-1} \in A_{i-1}$ for $i = 2, \dots, k$, we can explain a relationship between skew cyclic codes and DNA codes. $\xi_i(s_i)$ and $\xi_i(\theta_i(s_i))$ are DNA reverse of each other $s_i = a_{i-1} + u_i b_{i-1}$, where $a_{i-1}, b_{i-1} \in A_{i-1}, i = 2, \dots, k$.

For $s_i = a_{i-1} + u_i b_{i-1} \in A_i, i = 2, \dots, k$, we have

$$\begin{aligned} \xi_i(s_i) &= \gamma_i(\varphi_i(a_{i-1} + u_i b_{i-1})) \\ &= \gamma_i(a_{i-1}, a_{i-1} + b_{i-1}) \\ &= (\xi_{i-1}(a_{i-1}), \xi_{i-1}(a_{i-1} + b_{i-1})) \end{aligned}$$

On the other hand,

$$\begin{aligned}\xi_i(\theta_i(s_i)) &= \xi_i(\theta_{i-1}(a_{i-1}) + (1+u_i)\theta_{i-1}(b_{i-1})) \\ &= \xi_i(\theta_{i-1}(a_{i-1}+b_{i-1}) + u_i\theta_{i-1}(b_{i-1})) \\ &= \gamma_i\left(\varphi_i(\theta_{i-1}(a_{i-1}+b_{i-1}) + u_i\theta_{i-1}(b_{i-1}))\right) \\ &= \gamma_i(\theta_{i-1}(a_{i-1}+b_{i-1}), \theta_{i-1}(a_{i-1})) \\ &= \left(\xi_{i-1}(\theta_{i-1}(a_{i-1}+b_{i-1})), \xi_{i-1}(\theta_{i-1}(a_{i-1}))\right)\end{aligned}$$

where $i = 2, \ldots, k$.

This map can be extended as follows. For any $\mathbf{d}_i = (d_0^i, \ldots, d_{n-1}^i) \in A_i^n$, where $i = 2, \ldots, k$

$$\left(\xi_i(d_0^i), \xi_i(d_1^i), \ldots, \xi_i(d_{n-1}^i)\right)^r = w$$

where $w = \left(\xi_i\left(\theta_i(d_{n-1}^i)\right), \ldots, \xi_i\left(\theta_i(d_1^i)\right), \xi_i\left(\theta_i(d_0^i)\right)\right)$.

Example 4: Let $i = 3$. If $d_3 = (1 + u_1u_2) + u_3(1 + u_1 + u_2) \in A_3$, then we get

$$\begin{aligned}\xi_3(d_3) &= \gamma_3(\varphi_3(d_3)) = \gamma_3(1 + u_1u_2, u_1 + u_2 + u_1u_2) \\ &= (\xi_2(1 + u_1u_2), \xi_2(u_1 + u_2 + u_1u_2)) = TCGT\end{aligned}$$

On the other hand,

$$\begin{aligned}\xi_3(\theta_3(d_3)) &= \xi_3(\theta_2(1 + u_1u_2) + (1+u_3)\theta_2(1 + u_1 + u_2)) \\ &= \xi_3(\theta_2(u_1 + u_2 + u_1u_2) + u_3\theta_2(1 + u_1 + u_2)) \\ &= \gamma_3\left(\varphi_3(\theta_2(u_1 + u_2 + u_1u_2) + u_3\theta_2(1 + u_1 + u_2))\right) \\ &= \gamma_3(\theta_2(u_1 + u_2 + u_1u_2), \theta_2(1 + u_1u_2)) \\ &= \left(\xi_2(\theta_2(u_1 + u_2 + u_1u_2)), \xi_2(\theta_2(1 + u_1u_2))\right) \\ &= TGCT\end{aligned}$$

Definition 5: *Let C_i be a code of length n over A_i, for $i = 2, \ldots, k$. If $\xi_i(\boldsymbol{d}_i)^r \in \xi_i(C_i)$ for all $\boldsymbol{d}_i \in C_i$, then C_i or equivalently $\xi_i(C_i)$ is called a reversible DNA code.*

Definition 6: *Let $g_i(x) = a_0^i + a_1^i x + a_2^i x^2 + \cdots + a_s^i x^s$ be a polynomial of degree s over A_i. $g_i(x)$ is called a palindromic polynomial if $a_t^i = a_{s-t}^i$ for all $t \in \{0,1, \ldots, s\}$. $g_i(x)$ is called a θ_i-palindromic polynomial if $a_t^i = \theta_i(a_{s-t}^i)$ for all $t \in \{0,1, \ldots, s\}$, for $i = 2, \ldots, k$.*

As the order of θ_i is 2 , a skew cyclic code of odd length n over A_i with respect to θ_i is an ordinary cyclic code. So we will take the length n to be even.

The next two theorems show that palindromic and θ_i-palindromic polynomials generate reversible DNA codes.

Theorem 7: *Let $C_i = \langle f_i(x)\rangle$ be a skew cyclic code of length n over A_i, where $f_i(x)$ is a right divisor of $x^n - 1$ and $deg(f_i(x))$ is odd. If $f_i(x)$ is a θ_i-palindromic polynomial, then $\xi_i(C_i)$ is a reversible DNA code, for $i = 2, \dots, k$.*

Proof. Let $f_i(x)$ be a θ_i-palindromic polynomial and $f_i(x) = a_0^i + a_1^i x + \dots + a_{2s-1}^i x^{2s-1}$. So $a_t^i = \theta_i(a_{2s-1-t}^i)$, for all $t = 0,1,\dots,s-1$. Let $h_i(x) = h_0^i + h_1^i x + \dots + h_{2k-1}^i x^{2k-1}$. Let b_l^i be the coefficient of x^l in $h_i(x)f_i(x)$, where $l = 0,1,\dots,n-1$. For any $p < n/2$, the coefficient of x^p in $h_i(x)f_i(x)$ is

$$b_p^i = \sum_{j=0}^{p} h_j^i \theta_i^j(a_{p-j}^i)$$

and the coefficient of x^{n-p} is $b_{n-p}^i = \sum_{j=0}^{p} h_{2k-1-j}^i \theta_i^{2k-1-j}(a_{2s-1-(p-j)}^i)$.

The polynomial $h_i(x)f_i(x) = \sum_{d=0}^{2k-1} h_d^i x^d f_i(x)$ corresponds to a vector $\mathbf{b}_i = (b_0^i, b_1^i, \dots, b_{n-1}^i) \in C_i$.

The vector $\xi_i(\mathbf{b}_i)^r = \left(\left(\xi_i(b_0^i), \dots, \xi_i(b_{n-1}^i)\right)\right)^r$ is equal to the vector $\xi_i(\mathbf{z}_i)$, where the vector $\mathbf{z}_i$ corresponds to the polynomial $\sum_{d=0}^{2k-1} \theta_i(h_d^i) x^{2k-1-d} f_i(x)$. So $\xi_i(C_i)$ is a reversible DNA code.

Theorem 8: *Let $C_i = \langle f_i(x)\rangle$ be a skew cyclic code of length n over A_i, where $f_i(x)$ is a right divisor of $x^n - 1$ and $deg(f_i(x))$ is even. If $f_i(x)$ is a palindromic polynomial, then $\xi_i(C_i)$ is a reversible DNA code, for $i = 2, \dots, k$.*

Proof. Let $f_i(x)$ be a palindromic polynomial with even degree so that $f_i(x) = a_0^i + a_1^i x + \dots + a_{2s}^i x^{2s}$ and $a_t^i = a_{2s-t}^i$, for all $t = 0,1,\dots,s$. Let $h_i(x) = h_0^i + h_1^i x + \dots + h_{2k}^i x^{2k}$. Let b_l^i be the coefficient of x^l in $h_i(x)f_i(x)$, where $l = 0,1,\dots,n-1$. For any $p < n/2$, the coefficient of x^p in $h_i(x)f_i(x)$ is

$$b_p^i = \sum_{j=0}^{p} h_j^i \theta_i^j(a_{p-j}^i)$$

and the coefficient of x^{n-p} is $b_{n-p}^i = \sum_{j=0}^{t} h_{(2k)-j}^i \theta_i^{(2k)-j}(a_{2s-(p-j)}^i)$.

The polynomial $h_i(x)f_i(x) = \sum_{d=0}^{2k} h_d^i x^d f_i(x)$ corresponds to a vector $\mathbf{b}_i = (b_0^i, b_1^i, \dots, b_{n-1}^i) \in C_i$.

The vector $\xi_i(\mathbf{b}_i)^r = \left(\left(\xi_i(b_0^i), \dots, \xi_i(b_{n-1}^i)\right)\right)^r$ is equal to the vector $\xi_i(\mathbf{z}_i)$, where the vector $\mathbf{z}_i$ corresponds to the polynomial $\sum_{d=0}^{2k} \theta_i(h_d^i) x^{2k-d} f_i(x)$. So $\xi_i(C_i)$ is a reversible DNA code.

Theorem 9: *Let $x^n - 1 = h_i(x)f_i(x) \in A_i[x, \theta_i]$, where the degree of $f_i(x)$ is odd. If $f_i(x)$ is a θ_i-palindromic polynomial, then $h_i(x)$ is a palindromic polynomial.*

Proof. Let $f_i(x) = a_0^i + a_1^i x + \cdots + a_{2s-1}^i x^{2s-1}$. As the length n is even, then $h_i(x) = h_0^i + h_1^i x + \cdots + h_{2k-1}^i x^{2k-1}$. Since $f_i(x)$ is a θ_i-palindromic polynomial, then $a_t^i = \theta_i\left(a_{2s-1-t}^i\right)$ for all $t = 0,1, \dots, s-1$. Let b_l^i be the coefficient of x^l in $h_i(x)f_i(x)$, where $l = 0,1, \dots, n-1$. For any $p < n/2$, the coefficient of x^p in $h_i(x)f_i(x)$ is

$$b_p^i = \sum_{j=0}^{p} h_j^i \theta_i^j\left(a_{p-j}^i\right)$$

and the coefficient of x^{n-p} is $b_{n-p}^i = \sum_{j=0}^{p} h_{2k-1-j}^i \theta_i^{2k-1-j}\left(a_{2s-1-(p-j)}^i\right)$. By using the fact that $b_0^i = b_n^i = 0$ and $b_t^i = 0$ for all $t = 1,2, \dots, n-1$, it can be shown that $h_t^i = h_{2k-1-t}^i$ for all $t = 0,1,..,k-1$ by induction, as in [7].

4. Reversible and reversible complement codes over A_k

In this section, the necessary and sufficient conditions of cyclic codes over A_i to be reversible and reversible complement are given. By using the map, the DNA codes are obtained from these codes.

Firstly, we characterize the reversible codes over A_i, where $i = 1,2, \dots, k$.

In [9], it is proved that the cyclic code over $GF(q)$ generated by the monic polynomial $g(x)$ is reversible if and only if $g(x)$ is self reciprocal.

Theorem 10: *Let $C_1 = u_1 C_1^1 \oplus (1+u_1)C_1^2$ be a cyclic code of arbitrary length n over A_1. Then C_1 is reversible if and only if C_1^1 and C_1^2 are reversible codes over F_2 and both of them are cyclic codes over F_2.*

Theorem 11: *Let $C_i = u_i C_{i-1}^1 \oplus (1+u_i)C_{i-1}^2$ be a cyclic code of arbitrary length n over A_i, where $i = 2,3, \dots, k$. Then C_i is reversible if and only if C_{i-1}^1 and C_{i-1}^2 are reversible codes over A_{i-1}, where $i = 2,3, \dots, k$ and both of them are cyclic codes over A_{i-1}, where $i = 2,3, \dots, k$.*

Proof. Let C_{i-1}^1, C_{i-1}^2 be reversible codes. For any $\mathbf{b}_i \in C_i, \mathbf{b}_i = u_i \mathbf{b}_{i-1}^1 + (1+u_i)\mathbf{b}_{i-1}^2$, where $\mathbf{b}_{i-1}^1 \in C_{i-1}^1, \mathbf{b}_{i-1}^2 \in C_{i-1}^2$. As C_{i-1}^1, C_{i-1}^2 are reversible codes, $(\mathbf{b}_{i-1}^1)^r \in C_{i-1}^1, (\mathbf{b}_{i-1}^2)^r \in C_{i-1}^2$, so $\mathbf{b}_i^r = u_i(\mathbf{b}_{i-1}^1)^r + (1+u_i)(\mathbf{b}_{i-1}^2)^r \in C_i$. Hence C_i is reversible codes.

On the other hand, let C_i be a reversible code over A_i. So for any $\mathbf{b}_i = u_i \mathbf{b}_{i-1}^1 + (1+u_i)\mathbf{b}_{i-1}^2 \in C_i$, where $\mathbf{b}_i^1 \in C_{i-1}^1, \mathbf{b}_{i-1}^2 \in C_{i-1}^2$, we get $\mathbf{b}_i^r = u_i(\mathbf{b}_{i-1}^1)^r + (1+u_i)(\mathbf{b}_{i-1}^2)^r \in C_i$. Let $\mathbf{b}_i^r = u_i(\mathbf{b}_{i-1}^1)^r + (1+u_i)(\mathbf{b}_{i-1}^2)^r = u_i \mathbf{s}_{i-1}^1 + (1+u_i)\mathbf{s}_{i-1}^2$, where $\mathbf{s}_{i-1}^1 \in C_{i-1}^1, \mathbf{s}_{i-1}^2 \in C_{i-1}^2$. Therefore C_{i-1}^1 and C_{i-1}^2 are reversible codes over A_{i-1}.

Secondly, we characterize the reversible complement codes over A_i, where $i = 1,2,3, \dots, k$.

Definition 12: *For $\boldsymbol{x}_i = \left(x_0^i, x_1^i, \dots, x_{n-1}^i\right) \in A_i^n$, the vector $\left(\overline{x_{n-1}^i}, \overline{x_{n-2}^i}, \dots, \overline{x_1^i}, \overline{x_0^i}\right)$ is called the reversible complement of $\boldsymbol{x}_i$ and is denoted by $\boldsymbol{x}_i^{rc}$, where $\overline{x_j^i}$ represents the complement of the* *elements* $x_j^i, j = 0,1, \dots, n-1$. *A linear code* C_i *of length* n *over* A_i *is said to be reversible complement if* $\boldsymbol{x}_i^{rc} \in C_i$, *for every* $\boldsymbol{x}_i \in C_i$.

Lemma 13: *For any* $c_i \in A_i$, *where* $i = 1, \dots, k$ *we have* $c_i + \overline{c_i} = 1$.

Lemma 14: *Let* $a_i, b_i \in A_i$, *where* $i = 1, \dots, k$, *then* $\overline{a_i + b_i} = \overline{a_i} + \overline{b_i} + 1$.

Theorem 15: *Let* $C_i = u_i C_{i-1}^1 \oplus (1 + u_i) C_{i-1}^2$ *be a cyclic code of arbitrary length* n *over* A_i, *where* $i = 1,2,3, \dots, k$. *Then* C_i *is reversible complement if and only if* C_i *is reversible and* $(\overline{0}, \overline{0}, \dots, \overline{0}) \in C_i$, *where* C_{i-1}^1, C_{i-1}^2 *are both cyclic codes over* $A_{i-1}, i = 1,2,3, \dots, k$.

Proof. Since C_i is reversible complement, for any $\mathbf{d}_i = \left(d_0^i, \dots d_{n-1}^i\right) \in C_i, \mathbf{d}_i^{rc} = \left(\bar{d^\iota}_{n-1}, \dots, \bar{d}^i\right) \in C_i$. Since C_i is a linear code, so $(0,0, \dots, 0) \in C_i$. By using Lemma 13, we get

$$\mathbf{d}_i^r = \left(d_{n-1}^i, \dots, d_0^i\right) = \left(\overline{d_{n-1}^\iota}, \dots, \overline{d_0^\iota}\right) + (1,1,1, \dots, 1) \in C_i$$

Hence for any $\mathbf{d}_i \in C_i$, we have $\mathbf{d}_i^r \in C_i$.

On the other hand, let C_i be reversible code over A_i. So, for any $\mathbf{d}_i = \left(d_0^i, \dots d_{n-1}^i\right) \in C_i$, then $\mathbf{d}_i^r = \left(d_{n-1}^i, \dots, d_0^i\right) \in C_i$. For any $\mathbf{d}_i \in C_i$,

$$\mathbf{d}_i^{rc} = \left(\overline{d_{n-1}^\iota}, \dots, \overline{d_0^\iota}\right) = \left(d_{n-1}^i, \dots . d_0^i\right) + (1, \dots, 1) \in C_i$$

So, C_i is reversible complement code over A_i. By a cyclic DNA code over A_i of length n, we mean a cyclic code that has the reverse complement property, where $i = 1,2, .., k$.

Corollary 16: *Let* C_i *be a cyclic DNA code of length* n *over* A_i *and minimum Hamming distance* d, *where* $i = 2, \dots, k$. *Then* $\xi_i(C_i)$ *is a DNA code of length* $2^{i-1}n$ *over the alphabet* $\{A, T, C, G\}$ *with minimum Hamming distance at least* d.

5. Reversible and Reversible Complement DNA Codes

In this section, we will design linear codes over A_i, where $i = 2, \dots, k$, by using θ_i-set, where θ_i is a non trivial automorphism for $i = 2, .., k$ in order to obtain DNA codes.

Definition 17: *Let* $f_{0,1}, \dots, f_{0,2^i}$ *be polynomials dividing* $x^n - 1$ *over* F_2 *and let* $f_{i-1,1}, f_{i-1,2}$ *be polynomials with* $deg\, f_{i-1,1} = t_{i-1,1}$, $deg\, f_{i-1,2} = t_{i-1,2}$ *and both are over* A_{i-1} *for* $i = 2, \dots, k$. *Let* $f_i = u_i f_{i-1,1} + (1 + u_i) f_{i-1,2} \in A_i[x]$ *and*

$$\begin{aligned}
f_{i-1,1} &= u_{i-1}f_{i-2,1} + (1+u_{i-1})f_{i-2,2}\\
f_{i-1,2} &= u_{i-1}f_{i-2,3} + (1+u_{i-1})f_{i-2,4}\\
f_{i-2,1} &= u_{i-2}f_{i-3,1} + (1+u_{i-2})f_{i-3,2}\\
f_{i-2,2} &= u_{i-2}f_{i-3,3} + (1+u_{i-2})f_{i-3,4}\\
f_{i-2,3} &= u_{i-2}f_{i-3,5} + (1+u_{i-2})f_{i-3,6}\\
f_{i-2,4} &= u_{i-2}f_{i-3,7} + (1+u_{i-2})f_{i-3,8}\\
&\vdots\\
f_{1,1} &= u_1 f_{0,1} + (1+u_1)f_{0,2}\\
f_{1,2} &= u_1 f_{0,3} + (1+u_1)f_{0,4}\\
&\vdots\\
f_{1,2^{i-1}} &= u_1 f_{0,2^{i}-1} + (1+u_1)f_{0,2^{i}}
\end{aligned}$$

Let $m_i = min\{n - t_{i-1,1}, n - t_{i-1,2}\}$*. The set* $L(f_i)$ *is called a* θ_i*-set and is defined as* $L(f_i) = \{E_0, E_1, \ldots, E_{m_i-1}, F_0, F_1, \ldots, F_{m_i-1}\}$*, where* $E_j = x^j f_i, F_j = x^j \theta_i(h_i), 0 \leq j \leq m_i - 1, i = 2, \ldots, k.$
If $\deg f_{0,2s} \geq \deg f_{0,2s-1}$

$$h_{i,1,s} = u_1 x^{\deg f_{0,2s} - \deg f_{0,2s-1}} f_{0,2s-1} + (1+u_1) f_{0,2s}$$

otherwise,

$$h_{i,1,s} = u_1 f_{0,2s-1} + (1+u_1) x^{\deg f_{0,2s-1} - \deg f_{0,2s}} f_{0,2s}$$

where $s = 1,2, \ldots, 2^{i-1}$ and
If $\mathrm{degh}_{i,1,2t} \geq \mathrm{degh}_{i,1,2t-1}$

$$h_{i,2,t} = u_2 x^{\mathrm{degh}_{i,1,2t} - \deg h_{i,1,2t-1}} h_{i,1,2t-1} + (1+u_2) h_{i,1,2t}$$

otherwise,

$$h_{i,2,t} = u_2 h_{i,1,2t-1} + (1+u_2) x^{\deg h_{i,1,2t-1} - \deg h_{i,1,2t}} h_{i,1,2t}$$

where $t = 1,2, \ldots, 2^{i-2}$ and

If $\mathrm{degh}_{i,i-2,2v} \geq \deg_{i,i-2,2v-1}$,

$$h_{i,i-1,v} = u_{i-1} x^{\deg h_{i,i-2,,2v} - \deg h_{i,i-2,,2v-1}} h_{i,i-2,2v-1} + (1+u_{i-1}) h_{i,i-2,2v}$$

otherwise,

$$h_{i,i-1,v} = u_{i-1} h_{i,i-2,2v-1} + (1+u_{i-1}) x^{\mathrm{degh}\, h_{i,i-2,2v-1} - deg\, h_{i,i-2,2v}} h_{i,i-2,2v}$$

where $v = 1,2$ and

If $\mathrm{degh}_{i,i-1,2} \geq \mathrm{degh}_{i,i-1,1}$

$$h_i = u_i x^{\mathrm{degh}_{i,i-1,2} - \mathrm{degh}_{i,i-1,1}} h_{i,i-1,1} + (1+u_i) h_{i,i-1,2}$$

otherwise,

$$h_i = u_i h_{i,i-1,1} + (1+u_i) x^{\deg h_{i,i-1,1} - \deg h_{i,i-1,2}} h_{i,i-1,2}.$$

$L(f_i)$ generates a linear code C_i over A_i, where $i = 2, \dots, k$. It will be denoted by $C_i = \langle f_i \rangle_{\theta_i}$ or $C_i = \langle L(f_i) \rangle$. It means that it is A_i-submodule generated by the set $L(f_i)$, where $i = 2, \dots, k$.

Let $f_i = a_0^i + a_1^i x + \cdots + a_t^i x^t \in A_i[x], \theta_i(h_i) = b_0^i + b_1^i x + \cdots + b_s^i x^s$, where $i = 2, \dots, k$. The A_i-submodule can be considered to be generated by the rows of the following matrix

$$L(f_i) = \begin{bmatrix} E_0 \\ F_0 \\ E_1 \\ F_1 \\ E_2 \\ F_2 \\ \dots \end{bmatrix} = \begin{bmatrix} a_0^i & a_1^i & a_2 & \dots & a_t^i & 0 & \dots & \dots & & \cdots & 0 \\ b_0^i & b_1^i & \cdots & \cdots & b_t^i & b_{t+1}^i & \dots & b_s^i & 0 & \cdots & 0 \\ 0 & a_0^i & a_1^i & a_2^i & \cdots & a_t^i & 0 & 0 & & \cdots & 0 \\ 0 & b_0^i & b_1^i & b_2^i & \cdots & \cdots & \cdots & \cdots & b_s^i & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots & \cdots & \dots & \dots & & \cdots & \vdots \end{bmatrix}$$

Theorem 18: *Let $f_{0,1}, \dots, f_{0,2^i}$ be self reciprocal polynomials dividing $x^n - 1$ over F_2, where $i = 2, \dots, k$. So $C_i = \langle L(f_i) \rangle$ is a linear code over A_i and $\xi_i(C_i)$ is a reversible DNA code, where ξ_i is from C_i to $S^{2^{i-1}}_{D_4^{2^{i-1}}}$, for $i = 2, \dots, k$.*

Proof. It is proved as in the proof of the Theorem 4.3 in [11].

Corollary 19: *Let $f_{0,1}, \dots, f_{0,2^i}$ be self reciprocal polynomials dividing $x^n - 1$ over F_2 and $C_i = \langle L(f_i) \rangle$ be a cyclic code over A_i. If $\frac{x^n - 1}{x - 1} \in C_i$, then $\xi_i(C_i)$ is a reversible complement DNA code.*

Example 20:

$$f_{0,1}(x) = x + 1$$
$$f_{0,2}(x) = x^2 + x + 1$$
$$f_{0,3}(x) = x^6 + x^3 + 1$$
$$f_{0,4}(x) = x + 1$$

where all of them divide $x^9 - 1$ over F_2. Hence

$$f_2 = u_2\big(u_1 f_{0,1} + (1 + u_1) f_{0,2}\big) + (1 + u_2)\big(u_1 f_{0,3} + (1 + u_1) f_{0,4}\big)$$

over A_2. That is

$$f_2 = u_1(1 + u_2)x^6 + u_1(1 + u_2)x^3 + u_2(1 + u_1)x^2 + (u_1(1 + u_2) + 1)x + 1.$$

Since $h_{2,1,1} = u_1 x f_{0,1} + (1 + u_1) f_{0,2}$ and $h_{2,1,2} = u_1 f_{0,3} + x^5 (1 + u_1) f_{0,4}$, we get $h_2 = u_2 x^4 h_{2,1,1} + (1 + u_2) h_{2,1,2} = x^6 + (1 + u_1 + u_1 u_2)x^5 + x^4(1 + u_1)u_2 + x^3(1 + u_2)u_1 + u_1(1 + u_2)$. So $\theta_2(h_2) = x^6 + (1 + u_1 u_2)x^5 + (1 + u_1 + u_2 + u_1 u_2)x^4 + u_1 u_2 x^3 + u_1 u_2$.
Since $m_2 = 3$, we consider the generator matrix of C,

$$\begin{bmatrix} E_0 \\ F_0 \\ E_1 \\ F_1 \\ E_2 \\ F_2 \end{bmatrix}$$

where $E_0 = f_2, E_1 = xf_2, E_2 = x^2 f_2, F_0 = \theta_2(h_2), F_1 = x\theta_2(h_2), F_2 = x^2\theta_2(h_2)$. If we take $\alpha_0 = 0, \alpha_1 = 1, \alpha_2 = 1, \beta_0 = u_2, \beta_1 = 0, \beta_2 = 0$ then $\alpha_0 E_0 + \alpha_1 E_1 + \alpha_2 E_2 + \beta_0 F_0 + \beta_1 F_1 + \beta_2 F_2 == (u_1 + u_1u_2)x^8 + (u_1 + u_1u_2)x^7 + u_2x^6 + (u_1 + u_2)x^5 + x^4(u_1 + u_2) + (1 + u_1 + u_2 + u_1u_2)x^3 + (u_1 + u_1u_2)x^2 + x + u_1u_2$.

It is correspondence to the codeword

$$\mathbf{d}_1 = (u_1u_2, 1, u_1 + u_1u_2, 1 + u_1 + u_2 + u_1u_2, u_1 + u_2, u_1 + u_2, u_2, u_1 + u_1u_2, u_1 + u_1u_2)$$

Hence $\xi_2(\mathbf{d}_1) = AGTTGACAGCGCATGAGA$.
Moreover, $\theta_2(\alpha_0)F_2 + \theta_2(\alpha_1)F_1 + \theta_2(\alpha_2)F_0 + \theta_2(\beta_0)E_2 + \theta_2(\beta_1)E_1 + \theta_2(\beta_2)E_0 = (u_1 + u_1u_2)x^8 + x^7 + u_1u_2x^6 + (u_2 + u_1u_2)x^5 + (1 + u_1 + u_2)x^4 + (1 + u_1 + u_2)x^3 + (1 + u_2)x^2 + u_1u_2x + u_1u_2$ correspondences to the codeword

$$\mathbf{d}_2 = (u_1u_2, u_1u_2, 1 + u_2, 1 + u_1 + u_2, 1 + u_1 + u_2, u_2 + u_1u_2, u_1u_2, 1, u_1 + u_1u_2)$$

Hence $\xi_2(\mathbf{d}_2) = AGAGTACGCGACAGTTGA$. So $\left(\xi_2(\mathbf{d}_2)\right)^r = \xi_2(\mathbf{d}_1)$.

Received April 11, 2023; Accepted August 13, 2023

References

[1] Adleman L., Molecular computation of the solutions to combinatorial problems, Science 266 (1994), 1021-1024.

[2] Barra A., Dougherty S. T., Muchlis A., Muchtadi-Alamsyah I., Solé P., Suprijanto D., Yemen O., $\Theta_S -$ cyclic codes over A_k, arXiv preprint arXiv:1707.04681, 2017.

[3] Cengellenmis Y., Dertli A., On the cyclic DNA codes over the finite ring, Acta Universitatis Apulensis 58 (2019), 1-11.

[4] Cengellenmis Y., Dertli A., Dougherty S., Codes over an infinite family of rings with a Gray map, Des. Codes Cryptogr. 72 (2014), 559-580.

[5] Gursoy F., Oztas S. E., Siap I., Reversible DNA codes over $F_{16} + uF_{16} + vF_{16} + uvF_{16}$, arXiv: 1703.10189v1, 2017.

[6] Gursoy F., Oztas S. E., Ozkan A., Reversible DNA codes over a family of non-chain ring, arXiv:1711.02385v1, 2017.

[7] Gursoy F., Oztas S. E., Siap I., Reversible DNA codes using skew polynomial rings, Applicable Algebra in Engineering, Communication and Computing 28(2017), 311-320.

[8] Head T., Formal language theory and DNA: an analysis of the generative capacity of specific recombinant behaviors, Bulletin of mathematical biology 49(1987), 737-759.

[9] Massey J. L., Reversible codes, Information and Control 7 (1964), 369-380.

[10] Yildiz B., Siap I., Cyclic codes over $F_2[u]/< u^4 - 1 >$ and applications to DNA codes, Computers Mathematics with Applications 63 (2012), 11691176.

[11] Zhu S., Chen X., Cyclic DNA codes over $F_2 + uF_2 + vF_2 + uvF_2$ and their applications, J. Appl. Math. Comput. 55 (2017), 479-493.