MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(04)**(2023), 332–343.
http://doi.org/10.26637/mjm1104/001

On the DNA codes

ABDULLAH DERTLI *[1], **YASEMIN CENGELLENMIS** [2]

[1] *Department of Mathematics, Ondokuz Mayıs University Samsun, Turkey.*
[2] *Department of Mathematics, Trakya University, Edirne, Turkey.*

***R**eceived 17 February 2021; **A**ccepted 12 August 2023*

Abstract. In this paper, by using three different methods, the DNA codes are obtained from some codes over a family of the rings $D_i = D_1 \left[w_2, ..., w_i\right] / \left\langle w_i^2 - w_i, w_i w_j - w_j w_i \right\rangle$, where $i = 2, ..., r, j = 1, 2, ...r$ and $D_1 = F_2 + uF_2 + w_1(F_2 + uF_2), u^2 = 0, w_1^2 = w_1, uw_1 = w_1 u, F_2 = \{0, 1\}$.

AMS Subject Classifications: 94B05, 95B60.

Keywords: DNA codes, skew codes, reversibilty.

Contents

1. Introduction

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

It is well known that DNA contains 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 [7]. L. Adleman performed the computation using DNA strands in [1].

To perform computation using DNA strands, a specific set of DNA sequences are required with particular properties. The aim of this paper is to obtain the set of DNA strands satisfying various constraints, by using the some error correcting codes over a family of finite rings which enjoy DNA properties. One of the constraints is

*Corresponding author. Email address: **abdullah.dertli@gmail.com** (Abdullah DERTLI), **ycengellenmis@gmail.com** (Yasemin CENGELLENMIS)

reverse constaint. This leads to reversible codes.The other one is reverse complement constraint. This leads to reversible complement codes.

In order 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 [4] and [5] respectively by using 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 $F_{4^{2k}}[u_1, ..., u_s]/\langle {u_1}^2 - u_1, ..., {u_s}^2 - u_s\rangle$ where $k, s \geqslant 1, u_i u_j = u_j u_i$. Reversibility problem arises from the fact that the pairing of nucleotides in two different strands of a DNA sequence is done in opposite direction and 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 $ACGGATTG$.

In order to obtain the DNA codes, some authors used cyclic DNA codes of length n that enjoy some of the properties of DNA. In [9], by introducing a map, a family of cyclic codes over the ring $F_2[u]/ < u^4 - 1 >$ is mapped to DNA codes.

In [10], the design of linear codes over $D_1 = F_2 + uF_2 + vF_2 + uvF_2, u^2 = 0, v^2 = v, uv - vu, F_2 = \{0, 1\}$ is presented by using σ-set, where σ is a nontrivial automorphism on the finite ring D_1. 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 $D_i = D_{i-1} + w_i D_{i-1}$, where $i = 2, 3, ..., r, w_i^2 = w_i, D_1 = F_2 + uF_2 + w_1(F_2 + uF_2), u^2 = 0, w_1^2 = w_1, uw_1 = w_1 u, F_2 = \{0, 1\}$ is defined. By introducing skew cyclic codes over a family of the finite rings $D_i = D_1[w_2, ..., w_i] / \langle w_i^2 - w_i, w_i w_j - w_j w_i\rangle$, where $i = 2, ..., r, j = 1, 2, ..., r$, the reversible DNA codes are obtained from them. With the other method, the necessary and sufficient conditions of cyclic codes over D_i, where $i = 1, ..., r$ to be reversible and reversible complement are given. By introducing a map, the DNA codes are obtained from these type codes. As a last, the linear codes over D_i are designed, by using θ_i-set for $i = 2, 3, ..., r$. By using these type codes, the reversible or reversible complement DNA codes are obtained.

2. Preliminaries

A family of the finite rings $D_i = D_{i-1} + w_i D_{i-1}$, where $i = 2, 3, ..., r, w_i^2 = w_i, D_1 = F_2 + uF_2 + w_1(F_2 + uF_2), u^2 = 0, w_1^2 = w_1, uw_1 = w_1 u$ contains the commutative finite rings with characteristic 2 and cardinality 4^{2^i} for $i = 1, 2, ..., r$.

The finite rings of the family are written as recursively

$$D_i = D_{i-1} + w_i D_{i-1}$$

where $i = 2, 3, ..., r, w_i^2 = w_i, D_1 = F_2 + uF_2 + w_1(F_2 + uF_2), u^2 = 0, w_1^2 = w_1, uw_1 = w_1 u, F_2 = \{0, 1\}$.

In [10], the map φ_1 was defined as follows

$$\begin{aligned} \varphi_1 \quad : \quad & D_1 \longrightarrow (F_2 + uF_2)^2 \\ a + bw_1 \longmapsto & (a, a + b) \end{aligned}$$

where $a, b \in F_2 + uF_2, u^2 = 0, w_1^2 = w_1$.

We define the map on D_i where $i = 2, ..., r$ as follows

$$\begin{aligned} \varphi_i \quad : \quad & D_i \longrightarrow D_{i-1}^2 \\ x_{i-1} + y_{i-1} w_i \longmapsto & (x_{i-1}, x_{i-1} + y_{i-1}) \end{aligned}$$

where $x_{i-1}, y_{i-1} \in D_{i-1}, w_i^2 = w_i$ for $i = 2, 3, ..., r$.

In [10], they defined a ξ_1 correspondence between the elements of the finite ring $D_1 = F_2 + uF_2 + w_1F_2 + uw_1F_2$, where $u^2 = 0, w_1^2 = w_1, uw_1 = w_1u$ and DNA double pairs as follows

Elements α	Gray images	DNA double pairs $\xi_1(\alpha)$
0	$(0,0)$	AA
1	$(1,1)$	GG
u	(u,u)	TT
w_1	$(0,1)$	AG
uw_1	$(0,u)$	AT
$1+u$	$(1+u,1+u)$	CC
$1+w_1$	$(1,0)$	GA
$u+w_1$	$(u,1+u)$	TC
$u+uw_1$	$(u,0)$	TA
w_1+uw_1	$(0,1+u)$	AC
$1+uw_1$	$(1,1+u)$	GC
$1+u+uw_1$	$(1+u,1)$	CG
$1+u+w_1$	$(1+u,u)$	CT
$1+w_1+uw_1$	$(1,u)$	GT
$u+w_1+uw_1$	$(u,1)$	TG
$1+u+w_1+uw_1$	$(1+u,0)$	CA

By using the map φ_2 and ξ_1, we established ξ_2 correspondence between the element of D_2 and DNA 4-bases $x_1 + y_1w_2 \longmapsto (\xi_1(x_1), \xi_1(x_1+y_1))$ as follows

Elements β	DNA 4-bases $\xi_2(\beta)$
0	$AAAA$
1	$GGGG$
u	$TTTT$
w_1	$AGAG$
w_2	$AAGG$
...	

By using the matching and the elements of D_1 and $S_{D_{16}} = \{AA, TT, ..., GG\}$ and by using the Gray map from D_i to D_{i-1}^2, we can define ξ_i correspondence between the elements of the finite ring D_i and DNA 2^i-bases for $i = 2, ..., r$ as follows

$$\begin{aligned} \xi_i \quad &: \quad D_i \longrightarrow D_{i-1}^2 \longrightarrow \{A,T,G,C\}^{2^i} \\ x_{i-1} + y_{i-1}w_i &\longmapsto (x_{i-1}, x_{i-1}+y_{i-1}) \longmapsto q \end{aligned}$$

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

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

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

where $s_{i-1}, t_{i-1} \in D_{i-1}$ for $i = 2, ..., r$.

In [10], a nontrivial automorphism was defined on D_1 as follows

$$\begin{aligned} \theta_1 \quad &: \quad D_1 \longrightarrow D_1 \\ x_0 + y_0w_1 &\longmapsto x_0 + (1+w_1)y_0 \end{aligned}$$

where $x_0, y_0 \in F_2 + uF_2, u^2 = 0$.

By defining a nontrivial automorphism on D_i as follows, for $i = 2, ..., r$, we can define the skew cyclic codes over D_i, for $i = 2, ..., r$.

$$\begin{aligned} \theta_i \quad &: \quad D_i \longrightarrow D_i \\ x_{i-1} + y_{i-1}w_i &\longmapsto \theta_i(x_{i-1} + y_{i-1}w_i) = l \end{aligned}$$

where $l = \theta_{i-1}(x_{i-1}) + (1 + w_i)\theta_{i-1}(y_{i-1})$ and $x_{i-1}, y_{i-1} \in D_{i-1}$, for $i = 2, ..., r$.

The order of θ_i, for $i = 1, 2, ..., r$ is 2.

The rings

$$D_i[x, \theta_i] = \{b_0^i + b_1^i x + ... + b_{n-1}^i x^{n-1} : b_j^i \in D_i, n \in N, i = 2, ..., k, j = 0, 1, ..., n-1\}$$

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

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

where $a_i, b_i \in D_i$, for $i = 1, ..., r$. They are non-commutative rings.

Definition 2.1. *A subset C_i of D_i^n, where $i = 1, ..., r$ is called a skew cyclic code of length n if C_i satisfies the following conditions,*

1. *C_i is a submodule of D_i^n*

2. *If $\mathbf{c}_i = (c_0^i, c_1^i, ..., c_{n-1}^i) \in C_i$, then $\theta_i(\mathbf{c}_i) = \left(\theta_i(c_{n-1}^i), \theta_i(c_0^i), ..., \theta_i(c_{n-2}^i)\right) \in C_i$, where θ_i is the skew cyclic shift operator.*

In polynomial representation, a skew cyclic code of length n over D_i is defined as a left ideal of the quotient ring $D_{i_{\theta_i,n}} = D_i[x, \theta_i]/\langle x^n - 1\rangle$, 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 D_i is defined as a left $D_i[x, \theta_i]$-submodule of $D_{i_{\theta_i,n}}$, since the set $D_{i_{\theta_i,n}} = D_i[x, \theta_i]/\langle x^n - 1\rangle = \{f_i(x) + \langle x^n - 1\rangle : f_i(x) \in D_i[x, \theta_i]\}$ is a left $D_i[x, \theta_i]$-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 $r_i(x) \in D_i[x, \theta_i]$.

In either case, the following holds.

Theorem 2.2. *Let C_i be a skew cyclic code over D_i and let $f_i(x)$ be a polynomial in C_i of minimal degree. If the leading coefficient of $f_i(x)$ is a unit in D_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 D_i for $i = 1, 2, ..., r$.

Definition 3.1. *For $\mathbf{x}_i = \left(x_0^i, x_1^i, ..., x_{n-1}^i\right) \in D_i^n$, the vector $\left(x_{n-1}^i, x_{n-2}^i, ..., x_1^i, x_0^i\right)$ is called the reverse of $\mathbf{x}_i$ and is denoted by $\mathbf{x}_i^r$. A linear code C_i of length n over D_i is said to be reversible if $\mathbf{x}_i^r \in C_i$ for every $\mathbf{x}_i \in C_i$, where $i = 1, 2, ..., r$.*

We can express the matching the elements of D_1 and $S_{D_{16}} = \{AA, TT, ..., GG\}$ by means of the automorphism θ_1 as follows.

Each element $\alpha_1 = x + yw_1 \in D_1$, where $x, y \in F_2 + uF_2, u^2 = 0$ and $\theta_1(\alpha_1)$ are mapped to DNA double pairs which are reverse of each other. Since a correspondence the elements of the finite ring D_1 and DNA double pairs is ξ_1, so we have $\xi_1(w_1) = AG$, while $\xi_1(\theta_1(w_1)) = GA$.

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

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

where $s_{i-1}, t_{i-1} \in D_{i-1}$ for $i = 2, ..., r$, 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 $s_i = a_{i-1} + w_i b_{i-1}, a_{i-1}, b_{i-1} \in D_{i-1}$, where $a_{i-1}, b_{i-1} \in D_{i-1}, i = 2, ..., r$.

For $s_i = a_{i-1} + w_i b_{i-1} \in D_i, i = 2, ..., r$, we have

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

On the other hand,

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

where $i = 2, ..., r$.

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

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

Example 3.2. *Let $i = 2$. If $d_2 = (1 + uw_1) + w_2(1 + u + w_1) \in D_2$, then we get*

$$\begin{aligned}\xi_2(d_2) &= \gamma_2(\varphi_2(d_2)) = \gamma_2(1 + uw_1, u + w_1 + uw_1) \\ &= (\xi_1(1 + uw_1), \xi_1(u + w_1 + uw_1)) = GCTG.\end{aligned}$$

On the other hand,

$$\begin{aligned}\xi_2(\theta_2(d_2)) &= \xi_2(\theta_1(1 + uw_1) + (1 + w_2)\theta_1(1 + u + w_1)) \\ &= \xi_2(\theta_1(u + w_1 + uw_1) + w_2\theta_1(1 + u + w_1)) \\ &= \gamma_2(\varphi_2(\theta_1(u + w_1 + uw_1) + w_2\theta_1(1 + u + w_1))) \\ &= \gamma_2(\theta_1(u + w_1 + uw_1), \theta_1(1 + uw_1)) \\ &= (\xi_1(\theta_1(u + w_1 + uw_1)), \xi_1(\theta_1(1 + uw_1))) \\ &= GTCG.\end{aligned}$$

Example 3.3. *Let $i = 3$. If $d_3 = [(1 + uw_1) + w_2(1 + u + w_1)] + w_3(1 + w_2) \in D_3$, then we get*

$$\begin{aligned}\xi_3(d_3) &= \gamma_3(\varphi_3(d_3)) = \gamma_3((1 + uw_1) + w_2(1 + u + w_1), uw_1 + w_2(u + w_1)) \\ &= (\xi_2((1 + uw_1) + w_2(1 + u + w_1)), \xi_2(uw_1 + w_2(u + w_1))) \\ &= GCTGATTG.\end{aligned}$$

On the other hand,

$$\begin{aligned}\xi_3\left(\theta_3(d_3)\right) &= \xi_3\left(\theta_2((1+uw_1)+w_2(1+u+w_1))+(1+w_3)\theta_2(1+w_2)\right)\\ &= \xi_3\left(\theta_2(uw_1+w_2(u+w_1))+w_3\theta_2(1+w_2)\right)\\ &= \gamma_3(\varphi_3\left(\theta_2(uw_1+w_2(u+w_1))+w_3\theta_2(1+w_2)\right))\\ &= \gamma_3\left(\theta_2(uw_1+w_2(u+w_1)),\theta_2(1+uw_1+w_2(1+u+w_1))\right)\\ &= \left(\xi_2\left(\theta_2(uw_1+w_2(u+w_1))\right),\xi_2\left(\theta_2(1+uw_1+w_2(1+u+w_1))\right)\right)\\ &= GTTAGTCG.\end{aligned}$$

Definition 3.4. *Let C_i be a code of length n over D_i, for $i=1,...,r$. If $\xi_i(\mathbf{d}_i)^r \in \xi_i(C_i)$ for all $\mathbf{d}_i \in C_i$, then C_i or equivalently $\xi_i(C_i)$ is called a reversible DNA code.*

Definition 3.5. *Let $g_i(x)=a_0^i+a_1^ix+a_2^ix^2+...+a_s^ix^s$ be a polynomial of degree s over D_i. $g_i(x)$ is called a palindromic polynomial if $a_t^i=a_{s-t}^i$ for all $t\in\{0,1,...,s\}$. $g_i(x)$ is called a θ_i-palindromic polynomial if $a_t^i=\theta_i(a_{s-t}^i)$ for all $t\in\{0,1,...,s\}$, for $i=1,...,r$.*

As the order of θ_i is 2, a skew cyclic code of odd length n over D_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 3.6. *Let $C_i=\langle f_i(x)\rangle$ be a skew cyclic code of length n over D_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 $i=1,...,r$.*

Proof. Let $f_i(x)$ be a θ_i-palindromic polynomial and $f_i(x)=a_0^i+a_1^ix+...+a_{2s-1}^ix^{2s-1}$. So $a_t^i=\theta_i(a_{2s-1-t}^i)$, for all $t=0,1,...,s-1$. Let $h_i(x)=h_0^i+h_1^ix+\cdots+h_{2k-1}^ix^{2k-1}$. Let b_l^i be the coefficient of x^l in $h_i(x)f_i(x)$, where $l=0,1,\ldots,n-1$. 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 $h_i(x)f_i(x)=\sum_{d=0}^{2k-1}h_d^ix^df_i(x)$ corresponds to a vector $\mathbf{b}_i=(b_0^i,b_1^i,...,b_{n-1}^i)\in C_i$.

The vector $\xi_i(\mathbf{b}_i)^r=((\xi_i(b_0^i),...,\xi_i(b_{n-1}^i)))^r$ 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 3.7. *Let $C_i=\langle f_i(x)\rangle$ be a skew cyclic code of length n over D_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 $i=1,...,r$.*

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

$$b^i_p = \sum_{j=0}^{p} h^i_j \theta^j_i (a^i_{p-j})$$

and the coefficient of x^{n-p} is

$$b^i_{n-p} = \sum_{j=0}^{t} h^i_{(2k)-j} \theta_i^{(2k)-j} (a^i_{2s-(p-j)}).$$

The polynomial $h_i(x)f_i(x) = \sum_{d=0}^{2k} h^i_d x^d f_i(x)$ corresponds to a vector $\mathbf{b}_i = (b^i_0, b^i_1, ..., b^i_{n-1}) \in C_i$.

The vector $\xi_i(\mathbf{b}_i)^r = ((\xi_i(b^i_0), ..., \xi_i(b^i_{n-1})))^r$ is equal to the vector $\xi_i(\mathbf{z}_i)$, where the vector $\mathbf{z}_i$ corresponds to the polynomial $\sum_{d=0}^{2k} \theta_i(h^i_d) x^{2k-d} f_i(x)$. So $\xi_i(C_i)$ is a reversible DNA code. ■

Theorem 3.8. *Let $x^n - 1 = h_i(x)f_i(x) \in D_i[x, \theta_i]$, 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 $f_i(x) = a^i_0 + a^i_1 x + ... + a^i_{2s-1} x^{2s-1}$. As the length n is even, then $h_i(x) = h^i_0 + h^i_1 x + \cdots + h^i_{2k-1} x^{2k-1}$. Since $f_i(x)$ is a θ_i-palindromic polynomial, then $a^i_t = \theta_i(a^i_{2s-1-t})$ for all $t = 0, 1, ..., s-1$. Let b^i_l be the coefficient of x^l in $h_i(x)f_i(x)$, where $l = 0, 1, \ldots, n-1$. For any $p < n/2$, the coefficient of x^p in $h_i(x)f_i(x)$ is

$$b^i_p = \sum_{j=0}^{p} h^i_j \theta^j_i (a^i_{p-j})$$

and the coefficient of x^{n-p} is $b^i_{n-p} = \sum_{j=0}^{p} h^i_{2k-1-j} \theta_i^{2k-1-j} (a^i_{2s-1-(p-j)})$. By using the fact that $b^i_0 = b^i_n = 0$ and $b^i_t = 0$ for all $t = 1, 2, ..., n-1$, it can be shown that $h^i_t = h^i_{2k-1-t}$ for all $t = 0, 1, .., k-1$ by induction, as in [6]. ■

4. Reversible and reversible complement codes over D_r

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

In [10], they characterized the reversible codes over D_1 as follows.

Theorem 4.1. *[10] Let $C_1 = w_1 C_0^{1} \oplus (1 + w_1) C_0^{2}$ be a cyclic code of arbitrary length n over D_1. Then C_1 is reversible if and only if C_0^{1} and C_0^{2} are reversible codes over $F_2 + uF_2, u^2 = 0$ and both of them are cyclic codes over $F_2 + uF_2, u^2 = 0$.*

In [3] and [8], the necessary and sufficient conditions of cyclic codes over the ring $F_2 + uF_2, u^2 = 0$ to be reversible were given in case of the length n is odd or even, respectively.

In [2], the reversible codes over D_2 were characterized as follows;

Theorem 4.2. *[2] Let $C_2 = w_2 C_1^{1} \oplus (1 + w_2) C_1^{2}$ be a cyclic code of arbitrary length n over D_2. Then C_2 is reversible if and only if C_1^{1} and C_1^{2} are reversible codes over D_1 and both of them are cyclic codes over D_1.*

Firstly, we characterize the reversible codes over D_i, where $i = 3, ..., r$.

Theorem 4.3. *Let $C_i = w_i C_{i-1}^{1} \oplus (1+w_i) C_{i-1}^{2}$ be a cyclic code of arbitrary length n over D_i, where $i = 3, ..., r$. Then C_i is reversible if and only if C_{i-1}^{1} and C_{i-1}^{2} are reversible codes over D_{i-1}, where $i = 3, ..., r$ and both of them are cyclic codes over D_{i-1}, where $i = 3, ..., r$.*

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

On the other hand, let C_i be a reversible code over D_i. So for any $\mathbf{b}_i = w_i\mathbf{b}^1_{i-1} + (1+w_i)\mathbf{b}^2_{i-1} \in C_i$, where $\mathbf{b}^1_i \in C^1_{i-1}$,$\mathbf{b}^2_{i-1} \in C^{2}_{i-1}$, we get $\mathbf{b}^r_i = w_i\left(\mathbf{b}^1_{i-1}\right)^r + (1+w_i)\left(\mathbf{b}^2_{i-1}\right)^r \in C_i$. Let $\mathbf{b}^r_i = w_i\left(\mathbf{b}^1_{i-1}\right)^r + (1+w_i)\left(\mathbf{b}^2_{i-1}\right)^r = w_i\mathbf{s}^1_{i-1} + (1+w_i)\mathbf{s}^2_{i-1}$,where $\mathbf{s}^1_{i-1} \in C^1_{i-1}$,$\mathbf{s}^2_{i-1} \in C^{2}_{i-1}$. Therefore C^{1}_{i-1} and C^{2}_{i-1} are reversible codes over D_{i-1}. ■

In [10] and [2], they characterized the reversible complement codes over D_1 and D_2, respectively. Secondly, we characterize the reversible complement codes over D_i, where $i = 3, ..., r$.

Definition 4.4. *For $\mathbf{x}_i = \left(x^i_0, x^i_1, ..., x^i_{n-1}\right) \in D^n_i$, the vector $\left(\overline{x^i_{n-1}}, \overline{x^i_{n-2}}, ..., \overline{x^i_1}, \overline{x^i_0}\right)$ is called the reversible complement of $\mathbf{x}_i$ and is denoted by $\mathbf{x}^{rc}_i$, where $\overline{x^i_j}$ represents the complement of the elements x^i_j, $0 \leq j \leq n-1$. A linear code C_i of length n over D_i is said to be reversible complement if $\mathbf{x}^{rc}_i \in C_i$, for every $\mathbf{x}_i \in C_i$.*

Lemma 4.5. *For any $c_i \in D_i$,where $i = 1, ..., r$ we have $c_i + \overline{c_i} = u$.*

Lemma 4.6. *Let $a_i, b_i \in D_i$, where $i = 1, ..., r$, then $\overline{a_i + b_i} = \overline{a_i} + \overline{b_i} + u$.*

Theorem 4.7. *[10] Let $C_1 = w_1C^{1}_0 \oplus (1+w_1)C^{2}_0$ be a cyclic code of arbitrary length n over D_1. Then C_1 is reversible complement if and only if C_1 is reversible and $(\overline{0}, \overline{0}, ..., \overline{0}) \in C_1$, where C^{1}_0, C^{2}_0 are both cyclic codes over $F_2 + uF_2, u^2 = 0$.*

Theorem 4.8. *[2] Let $C_2 = w_2C^{1}_1 \oplus (1+w_2)C^{2}_1$ be a cyclic code of arbitrary length n over D_2. Then C_2 is reversible complement if and only if C_2 is reversible and $(\overline{0}, \overline{0}, ..., \overline{0}) \in C_2$, where C^{1}_1, C^{2}_1 are both cyclic codes over D_1.*

Theorem 4.9. *Let $C_i = w_iC^{1}_{i-1} \oplus (1+w_i)C^{2}_{i-1}$ be a cyclic code of arbitrary length n over D_i, where $i = 3, ..., r$. Then C_i is reversible complement if and only if C_i is reversible and $(\overline{0}, \overline{0}, ..., \overline{0}) \in C_i$, where C^{1}_{i-1}, C^{2}_{i-1} are both cyclic codes over $D_{i-1}, i = 3, ..., r$.*

Proof. Since C_i is reversible complement, for any $\mathbf{d}_i = (d^i_0, ...d^i_{n-1}) \in C_i, \mathbf{d}^{rc}_i = (\overline{d^i}_{n-1}, ..., \overline{d^i}_0) \in C_i$. Since C_i is a linear code, so $(0, 0, ..., 0) \in C_i$. By using Lemma 4.5, we get

$$\mathbf{d}^r_i = (d^i_{n-1}, ..., d^i_0) = (\overline{d^i}_{n-1}, ..., \overline{d^i}_0) + (u, u, u, ..., u) \in C_i.$$

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

On the other hand, let C_i be reversible code over D_i. So, for any $\mathbf{d}_i = (d^i_0, ...d^i_{n-1}) \in C_i$, then $\mathbf{d}^r_i = (d^i_{n-1}, ..., d^i_0) \in C_i$. For any $\mathbf{d}_i \in C_i$,

$$\mathbf{d}^{rc}_i = (\overline{d^i}_{n-1}, ..., \overline{d^i}_0) = (d^i_{n-1}, ...d^i_0) + (u, ..., u) \in C_i.$$

So, C_i is reversible complement code over D_i. ■

By a cyclic DNA code over D_i of length n, we mean a cyclic code that has the reverse complement property, where $i = 1, 2, .., r$.

Corollary 4.10. *Let C_i be a cyclic DNA code of length n over D_i and minimum Hamming distance d, where $i = 1, 2, .., r$. Then $\xi_i(C_i)$ is a DNA code of length 2^in over the alphabet $\{A, T, C, G\}$ with minimum Hamming distance at least d.*

5. Reversible and Reversible Complement DNA Codes

In [10], the design of linear codes over D_1 was presented. It was obtained DNA codes from them.

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

Definition 5.1. *Let* $f_{0,1}, ..., f_{0,2^i}$ *be polynomials dividing* $x^n - 1$ *over* $F_2 + uF_2, u^2 = 0$ *and* $f_{i-1,1}, f_{i-1,2}$ *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* D_{i-1}, *for* $i = 2, ..., r$. *Let* $f_i = w_i f_{i-1,1} + (1 + w_i) f_{i-1,2} \in D_i[x]$ *and*

$$\begin{aligned} f_{i-1,1} &= w_{i-1} f_{i-2,1} + (1 + w_{i-1}) f_{i-2,2}, \\ f_{i-1,2} &= w_{i-1} f_{i-2,3} + (1 + w_{i-1}) f_{i-2,4}, \\ f_{i-2,1} &= w_{i-2} f_{i-3,1} + (1 + w_{i-2}) f_{i-3,2}, \\ f_{i-2,2} &= w_{i-2} f_{i-3,3} + (1 + w_{i-2}) f_{i-3,4}, \\ f_{i-2,3} &= w_{i-2} f_{i-3,5} + (1 + w_{i-2}) f_{i-3,6}, \\ f_{i-2,4} &= w_{i-2} f_{i-3,7} + (1 + w_{i-2}) f_{i-3,8}, \\ &\vdots \\ f_{1,1} &= w_1 f_{0,1} + (1 + w_1) f_{0,2}, \\ f_{1,2} &= w_1 f_{0,3} + (1 + w_1) f_{0,4}, \\ &\vdots \\ f_{1,2^{i-1}} &= w_1 f_{0,2^i-1} + (1 + w_1) f_{0,2^i}. \end{aligned}$$

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*

$$L(f_i) = \{E_0, E_1, ..., E_{m_i-1}, F_0, F_1, ..., F_{m_i-1}\},$$

where $E_j = x^j f_i, F_j = x^j \theta_i(h_i), 0 \leq j \leq m_i - 1, i = 2, ..., r$.

If deg $f_{0,2s} \geq deg f_{0,2s-1}$,

$$h_{i,1,s} = w_1 x^{deg f_{0,2s} - deg f_{0,2s-1}} f_{0,2s-1} + (1 + w_1) f_{0,2s}$$

otherwise,

$$h_{i,1,s} = w_1 f_{0,2s-1} + (1 + w_1) x^{deg f_{0,2s-1} - deg f_{0,2s}} f_{0,2s}$$

where $s = 1, 2, ..., 2^{i-1}$ *and*

if $deg h_{i,1,2t} \geq deg h_{i,1,2t-1}$

$$h_{i,2,t} = w_2 x^{deg h_{i,1,2t} - deg h_{i,1,2t-1}} h_{i,1,2t-1} + (1 + w_2) h_{i,1,2t}$$

otherwise,

$$h_{i,2,t} = w_2 h_{i,1,2t-1} + (1 + w_2) x^{deg h_{i,1,2t-1} - deg h_{i,1,2t}} h_{i,1,2t}$$

where $t = 1, 2, ..., 2^{i-2}$ *and*

$$\vdots$$

if $deg h_{i,i-2,2v} \geq deg h_{i,i-2,2v-1}$

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

otherwise,

$$h_{i,i-1,v} = w_{i-1}h_{i,i-2,2v-1} + (1+w_{i-1})x^{degh_{i,i-2,2v-1}-degh_{i,i-1,2v}}h_{i,i-2,2v}$$

where $v = 1,2$ *and*

If $degh_{i,i-1,2} \geq degh_{i,i-1,1}$,

$$h_i = w_i x^{degh_{i,i-1,2}-degh_{i,i-1,1}}h_{i,i-1,1} + (1+w_i)h_{i,i-1,2}$$

otherwise,

$$h_i = w_i h_{i,i-1,1} + (1+w_i)x^{degh_{i,i-1,1}-degh_{i,i-1,2}}h_{i,i-1,2}.$$

$L(f_i)$ *generates a linear code* C_i *over* D_i, *where* $i = 2, ..., r$. *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* D_i*-submodule generated by the set* $L(f_i)$, *where* $i = 2, ..., r$.

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

$$L(f_i) = \begin{bmatrix} E_0 \\ F_0 \\ E_1 \\ F_1 \\ E_2 \\ F_2 \\ \cdots \end{bmatrix}$$

$$= \begin{bmatrix} a_0^i & a_1^i & a_2 & \cdots & a_t^i & 0 & \cdots & \cdots & & \cdots & 0 \\ b_0^i & b_1^i & \cdots & \cdots & b_t^i & b_{t+1}^i & \cdots & b_s^i & 0 & \cdots & 0 \\ 0 & a_0^i & a_1^i & a_2^i & \cdots & a_t^i & 0 & 0 & & \cdots & 0 \\ 0 & b_0^i & b_1^i & b_2^i & \cdots & \cdots & \cdots & \cdots & b_s^i & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & & \cdots & \vdots \end{bmatrix}.$$

Theorem 5.2. *Let* $f_{0,1}, ..., f_{0,2^i}$ *be self reciprocal polynomials dividing* $x^n - 1$ *over* $F_2 + uF_2, u^2 = 0$. *So* $C_i = \langle L(f_i) \rangle$ *is a linear code over* D_i *and* $\xi_i(C_i)$ *is a reversible DNA code, where* ξ_i *is from* C_i *to* $S_{D_4}^{2^i n}$, *for* $i = 2, ..., r$.

Proof. It is proved as in the proof of the Theorem 4.3 in [10]. ■

Corollary 5.3. *Let* $f_{0,1}, ..., f_{0,2^i}$ *be self reciprocal polynomials dividing* $x^n - 1$ *over* $F_2 + uF_2, u^2 = 0$ *and* $C_i = \langle L(f_i) \rangle$ *be a cyclic code over* D_i. *If* $u\frac{x^n-1}{x-1} \in C_i$, *then* $\xi_i(C_i)$ *is a reversible complement DNA code.*

Example 5.4.

$$\begin{aligned} f_{0,1}(x) &= x + 1, \\ f_{0,2}(x) &= x^2 + x + 1, \\ f_{0,3}(x) &= x^6 + x^3 + 1, \\ f_{0,4}(x) &= x + 1, \end{aligned}$$

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

$$f_2 = w_2\left(w_1 f_{0,1} + (1+w_1) f_{0,2}\right) + (1+w_2)\left(w_1 f_{0,3} + (1+w_1) f_{0,4}\right)$$

over D_2. *That is*

$$f_2 = w_1(1+w_2)x^6 + w_1(1+w_2)x^3 + w_2(1+w_1)x^2 + (w_1(1+w_2)+1)x + 1.$$

Since $h_{2,1,1} = w_1 x f_{0,1} + (1+w_1) f_{0,2}$ *and* $h_{2,1,2} = w_1 f_{0,3} + x^5(1+w_1) f_{0,4}$, *we get* $h_2 = w_2 h_{2,1,1} + (1+w_2) h_{2,1,2}$ *Then we have* $h_2 = x^6 + (1+w_1+w_1w_2)\, x^5 + x^4\,(1+w_1)\, w_2 + (1+w_2)\, w_1 x^3 + w_1\,(1+w_2)$. *So* $\theta_2\,(h_2) = x^6 + (1+w_2\,(1+w_1))\, x^5 + w_1\,(1+w_2)\, x^4 + (1+w_1)\, w_2 x^3 + w_2\,(1+w_1)$.
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 $E_0 = F_2, E_1 = x f_2, E_2 = x^2 f_2, F_0 = \theta_2(h_2), F_1 = x\theta_2(h_2), F_2 = x^2\theta_2(h_2)$. *If we take* $\alpha_0 = 0, \alpha_1 = 1, \alpha_2 = 1, \beta_0 = 0, \beta_1 = 0, \beta_2 = 1$, *then* $\alpha_0 E_0 + \alpha_1 E_1 + \alpha_2 E_2 + \beta_0 F_0 + \beta_1 F_1 + \beta_2 F_2 = x + x^2\,(w_1+w_2) + w_2\,(1+w_1)\, x^3 + (w_1+w_2)\, x^4 + (w_1+w_2)\, x^5 + w_1\,(1+w_2)\, x^6 + (1+w_1+w_2)\, x^7 + (1+w_1+w_1w_2)\, x^8$.
It is correspondence to the codeword

$$\mathbf{d}_1 = \begin{pmatrix} 0, 1, w_1+w_2, w_2\,(1+w_1)\,, w_1+w_2, w_1+w_2, \\ w_1\,(1+w_2)\,, 1+w_1+w_2, 1+w_1+w_1w_2 \end{pmatrix}.$$

Hence $\xi_2(\mathbf{d}_1) = AAAAGGGGAGGAAAGAAGGAAGGAAGAAGAAGGAGG$.

Moreover $\theta_2\,(\alpha_0)\, F_2 + \theta_2\,(\alpha_1)\, F_1 + \theta_2\,(\alpha_2)\, F_0 + \theta_2\,(\beta_0)\, E_2 + \theta_2\,(\beta_1)\, E_1 + \theta_2\,(\beta_2)\, E_0 = 1 + w_2\,(1+w_1) + x\,(1+w_1+w_2) + x^2\,(w_2\,(1+w_1)) + x^3\,(w_1+w_2) + x^4\,(w_1+w_2) + x^5\,(1+w_1+w_2) + x^6\,(1+w_1+w_2) + x^7$ *correspondences to the codeword*

$$\mathbf{d}_2 = \begin{pmatrix} 1+w_2\,(1+w_1)\,, 1+w_1+w_2, w_2\,(1+w_1)\,, w_1+w_2, \\ w_1+w_2, 1+w_1+w_2, 1+w_1+w_2, 1, 0 \end{pmatrix}.$$

Hence $\xi_2(\mathbf{d}_2) = GGAGGAAGAAGAAGGAAGGAAGAAAGGAGGGGAAAA$.

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

6. Conclusion

By using three different methods, the DNA codes are obtained from the some error correcting codes over the family of finite rings.

References

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

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

[3] K. Guenda and T.A. Gulliver, Construction of cyclic codes over $F_2 + uF_2$ for DNA computing, *Applicable Algebra in Engineering, Communication and Computing*, **24**(2013), 445-459.

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

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

[6] F. GURSOY, E. S. OZTAS AND I. SIAP, Reversible DNA codes using skew polynomial rings, *Applicable Algebra in Engineering, Communication and Computing*, **28**(2017), 311–320.

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

[8] J. LIANG AND W. LIQI, On cyclic DNA codes over $F_2 + uF_2$, *Journal of Applied Mathematics and Computing*, **51**(2016), 81–91.

[9] B. YILDIZ AND I. SIAP, Cyclic codes over $F_2[u]/ < u^4 - 1 >$ and applications to DNA codes., *Computers Mathematics with Applications*, **63**(2012), 1169–1176.

[10] S. ZHU AND X. CHEN, Cyclic DNA codes over $F_2 + uF_2 + vF_2 + uvF_2$ and their applications, *Journal of Applied Mathematics and Computation*, **55**(2017), 479–493.