J. Algebra Comb. Discrete Appl.
4(1) • 93–102

Received: 28 July 2016
Accepted: 25 October 2016

Journal of Algebra Combinatorics Discrete Structures and Applications

On DNA codes from a family of chain rings[*]

Research Article

Elif Segah Oztas, Bahattin Yildiz, Irfan Siap

Abstract: In this work, we focus on reversible cyclic codes which correspond to reversible DNA codes or reversible-complement DNA codes over a family of finite chain rings, in an effort to extend what was done by Yildiz and Siap in [20]. The ring family that we have considered are of size 2^{2^k}, $k = 1, 2, \cdots$ and we match each ring element with a DNA 2^{k-1}-mer. We use the so-called u^2-adic digit system to solve the reversibility problem and we characterize cyclic codes that correspond to reversible-complement DNA-codes. We then conclude our study with some examples.

2010 MSC: 94B15, 92D10

Keywords: Cyclic codes, Chain rings, Reversible codes, DNA codes

1. Introduction

In [3], Adleman proposed a solution to an instance of the NP-complete problem of the directed Hamiltonian Path problem, using DNA molecules. This ground breaking approach of using the DNA as a computational tool has led to many similar studies since its appearance. In [6] and [4], the advantages of these studies were demonstrated by a molecular program that led to breaking the Data Encryption System. In [11], Mansuripur et al. show that DNA molecules can be used as a storage medium.

The DNA sequences consist of four bases, namely adenine (A), thymine (T), guanine (G), cytosine (C). The governing principle in its duplication is the well known Watson-Crick property (WCC). According to WCC A and T bound to each other and G and C bound to each other on the opposite strands. A and G are called the complements of T and C respectively or vice versa.

Acting like an error-correcting code in nature, the DNA, has naturally attracted the attention of coding theorists in their researches. Consequently the concept of a DNA-code, that is a code that has

[*] *A part of this study is presented in The International Conference on Coding theory and Cryptography (ICCC2015, Algeria).*

Elif Segah Oztas (Corresponding Author), Irfan Siap; Department of Mathematics, Yildiz Technical University, Istanbul, Turkey (email: elifsegahoztas@gmail.com, irfan.siap@gmail.com).
Bahattin Yildiz; Istanbul, Turkey (email: bahattinyildiz@gmail.com).

http://dx.doi.org/10.13069/jacodesmath.96056

ISSN 2148-838X

special properties (i.e. reversible and complement) akin to the DNA, was introduced and has been a focal point of research in recent years. In many other studies, some constraints such as the Hamming distance constraint, the reverse constraint, the reverse-complement constraint and the fixed GC-content constraint are also considered [7, 9, 10, 12]. In [1], Aboluion et al. develop a new approach and extend the studies of [8] and [9]. They added the reverse-complement constraint to further prevent the unwanted hybridizations. In such works as [2, 7–9], the focus was on constructing large sets of DNA codewords by using minimum Hamming distance.

With the four letter ambient alphabet of the DNA, many of the early works on DNA codes, use algebraic structures of size four to construct DNA codes. To this end, we can cite [2], in which DNA codes are studied over the Galois field of size 4. As another possible such alphabet, Siap et al. studied DNA codes over the finite ring $\mathbb{F}_2[u]/(u^2-1)$ in [19]. In [15], DNA-double pairs are used with the field $\mathbb{F}_{16}$ and optimal codes are obtained. In [17], a generalization of [15] was done to include DNA $2k$-bases (DNA pairs) over a suitable ring.

In this paper, we extend on the work done in [20], in which DNA pairs were matched with the elements of the ring $\mathbb{F}_2[u]/(u^4-1)$ of size 16. Here, we match DNA 2^{k-1}-bases (mers)

$$\{\underbrace{AA...AA}_{2^{k-1}}, AT...TT, ...\}$$

with elements of the ring $\mathcal{R}_{2^k} = \mathbb{F}_2[u]/(u^{2^k}-1)$. The main idea of the paper is to lay out the theory behind generating reversible and reversible-complement DNA codes over this ring. In doing so, we first tackle the issue of ring-reversibility versus the DNA reversibility. A problem that arises in all the cases where DNA k-mers ($k \geq 2$) are matched with ring elements, it can be best explained in the following example: Let $(u_1, u_2, u_3) \in \mathcal{R}_4$ be a codeword that corresponds to the DNA string ATAGCC. The reverse of (u_1, u_2, u_3) is (u_3, u_2, u_1) and (u_3, u_2, u_1) corresponds to CCAGAT. But CCAGAT is not the reverse of ATAGCC. This problem is solved by using the so-called u^2-adic system. After giving the theoretical results concerning the constructions of DNA codes over the ring, we give examples of DNA-codes thus obtained.

The rest of the work is organized as follows. In section 2, we give some brief background on the ring $\mathcal{R}_{2^k}$. In section 3, we introduce the u^2-adic digit system for the ring $\mathcal{R}_{2^k}$ and demonstrate how it can be used to solve the reversibility problem. In section 4, we give the main results about cyclic DNA codes over $\mathcal{R}_{2^k}$ together with many examples. We finish the paper in section 5 with some concluding remarks and directions for possible future research on the related topics.

2. Preliminaries

We consider the finite chain ring $\mathcal{R}_{2^k} = \mathbb{F}_2[u]/(u^{2^k}-1) = \{a_0 + a_1u + ... + a_{2^k-1}u^{2^k-1} | a_i \in \mathbb{F}_2, 1 \leq i \leq 2^k-1, u^{2^k} = 1\}$ in this work. As can easily be observed from its structure, $\mathcal{R}_{2^k}$ is a commutative, characteristic 2 ring of size 2^{2^k}. The chain of ideals can be listed as

$$\{0\} = ((1+u)^{2^k}) \subset ((1+u)^{2^k-1}) \subset \cdots \subset ((1+u)^3) \subset ((1+u)^2) \subset (1+u) \subset (1) = \mathcal{R}_{2^k}.$$

A linear code over $\mathcal{R}_{2^k}$ of length n is defined naturally as an $\mathcal{R}_{2^k}$-submodule of $(\mathcal{R}_{2^k})^n$. By using certain elementary row and column operations, it can be shown that any linear code C over $\mathcal{R}_{2^k}$ of length n is equivalent to a code whose generating matrix can be put into the following standard form:

$$\begin{bmatrix} I_{t_0} & M_1^1 & M_2^1 & \dots & M_{2^k}^1 \\ 0 & (1+u)I_{t_1} & (1+u)M_1^2 & \dots & (1+u)M_{2^k-1}^2 \\ 0 & 0 & (1+u)^2I_{t_2} & \dots & (1+u)^2M_{2^k-2}^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \dots & \dots & (1+u)^{2^k-1}I_{t_{2^k-1}} & (1+u)^{2^k-1}M_1^{2^k-1} \end{bmatrix}, \quad (1)$$

where M_i^j's are matrices over $\mathcal{R}_{2^k}$. As is the case with codes over finite chain rings, instead of the dimension, we can talk about the type of a code. A code with the above generator matrix is said to have type $(t_0, t_1, t_2, \cdots, t_{2^k-1})$ and the size of the code is given by

$$|C| = (2^{2^k})^{t_0}(2^{2^k-1})^{t_1}(2^{2^k-2})^{t_2} \cdots (2)^{t_{2^k-1}} = 2^{2^k t_0 + (2^k-1)t_1 + \cdots + t_{2^k-1}}.$$

A subset C of $\mathcal{R}_{2^k}^n$ is called a cyclic code of length n if C is a submodule of $\mathcal{R}_{2^k}^n$ and if C is invariant under σ, namely $\sigma(C) = C$, where σ is the right cyclic shift on $\mathcal{R}_{2^k}^n$. σ acts on $\mathcal{R}_{2^k}^n$ as

$$\sigma(c_0, c_1, \cdots, c_{n-1}) = (c_{n-1}, c_0, c_1, \cdots, c_{n-2}). \tag{2}$$

In studying cyclic codes, it is essential to make use of their algebraic structures by matching each codeword $c = (c_0, c_1, \cdots, c_{n-1}) \in C$ to a polynomial $c_0 + c_1x + \cdots + c_{n-1}x^{n-1} \in \mathcal{R}_{2^k}[x]$. Thus cyclic codes of length n in that case correspond to the ideals of the quotient ring $\mathcal{R}_{2^k}[x]/(x^n - 1)$. As will be needed later, we include in this section, the definition of the reciprocal polynomial. For $g(x) \in \mathcal{R}_{2^k}[x]$, the reciprocal of $g(x)$ is denoted by $g(x)^* = x^{deg(g)}g(1/x)$.

A code C is said to be reversible if $c^r \in C$ for all $c \in C$, where c^r corresponds to the reverse of c. More precisely if $c = (c_0, c_1, \cdots, c_{n-1})$, then $c^r = (c_{n-1}, c_{n-2}, \cdots, c_0)$. A code is said to be complement if $c^c \in C$ for each $c \in C$, where c^c denotes a suitably defined complement of c. In general, algebraically the complement corresponds to adding a fixed constant to each of the coordinates. A code is said to be a reversible-complement code if it is both reversible and complement. DNA codes are defined as codes that satisfy both the reversible and complement property.

3. u^2-adic digit system for DNA 2^{k-1}-mers

In [16], the u^2-adic digit system was introduced for DNA k-mers. We will make use of the function η from [16], which operates on DNA single bases in the following form:

$$\eta(A) = 0, \eta(T) = 1 + u, \eta(G) = 1, \text{ and } \eta(C) = u. \tag{3}$$

The u^2-adic digit system can be used for all the rings of the form $\mathcal{R}_{2k} = \mathbb{F}_2[u]/(u^{2k} - 1)$, in particular for the ring $\mathcal{R}_{2^k}$, that is in question in our work.

The following definitions are obtained from [16] by modifying them to go along with the ring $\mathcal{R}_{2^k}$:

Definition 3.1. *Every element of the ring $\mathcal{R}_{2^k}$ can be expressed as a linear combination of $1, u^2, u^4, \ldots, u^{2^k-2}$, with coefficients from $\{0, 1, u, 1+u\} = \mathbb{F}_2[u]/(u^2 - 1)$. We can view this as a number-base system with the digits 1s (units), u^2s, u^4s, u^6s, u^8s, ... , u^{2^k-2}s. We call this system the u^2-adic system.*

Definition 3.2. *Let $b_1b_2...b_{2^{k-1}}$ be a DNA 2^{k-1}-bases (2^{k-1}-mers) where $b_i \in \{A, T, G, C\}$. This corresponds to an element in the ring $\mathcal{R}_{2^k}$, where this correspondence is given by the following:*

$$\zeta(b_1b_2...b_{2^{k-1}}) = \alpha \in \mathcal{R}_{2^k}, \tag{4}$$

where

$$\alpha = \eta(b_{2^{k-1}})1 + \eta(b_{2^{k-1}-1})u^2 + \eta(b_{2^{k-1}-2})u^4 + \cdots + \eta(b_1)u^{(2^k-2)} = \sum_{t=1}^{2^{k-1}} \eta(b_t)u^{(2^k-2t)}.$$

The following lemma expresses the effect of multiplying by u^2:

Lemma 3.3. *If $\alpha \in \mathcal{R}_{2^k}$ and $\zeta(b_1b_2 \cdots b_{2^{k-1}}) = \alpha$, then $\zeta^{-1}(u^2\alpha) = b_2 \cdots b_{2^{k-1}}b_1$.*

Proof. According to the structure of u^2-adic digits, multiplying by u^2 shifts the u^2-adic digits cyclically to the left. The result then follows from the definition of ζ. □

For $p(u) \in \mathcal{R}_{2^k}$, let $p(u)_{u^2}$ or p_{u^2} represent the u^2-adic digits of $p(u)$ as follows: If $p(u) = a_0 1 + a_1 u^2 + a_2 u^4 + ... + a_{(2^{k-1}-1)} u^{(2^k-2)}$, then let

$$p_{u^2} = a_{(2^{k-1}-1)} ... a_2 a_1 a_0.$$

Denote by ε, the function $\varepsilon(p_{u^2}) = p(u)$ and by $[p_{u^2}]^r$, the reverse vector, namely $[p_{u^2}]^r = a_0 a_1 a_2 ... a_{(2^{k-1}-1)}$. Then by Lemma 3.3 we have

$$u^2 p_{u^2} = a_{(2^{k-1}-2)} ... a_2 a_1 a_0 a_{(2^{k-1}-1)},$$

where $a_i \in \mathbb{F}_2[u]/(u^2-1)$.

Example 3.4. *Let us consider the ring* $\mathcal{R}_{16} = \mathbb{F}_2[u]/(u^{16}-1)$.

$$AAATGACC \to \zeta(AAATGACC) = u + u^3 + u^6 + u^8 + u^9 \to 000\bar{u}10uu,$$

where $\bar{u} = 1 + u$. *Multiplying by* u^2, *we get* $u^2(u + u^3 + u^6 + u^8 + u^9) = u^3 + u^5 + u^8 + u^{10} + u^{11} \to 00\bar{u}10uu0$ *(*u^2*-adic digit system in* $\mathcal{R}_2$*).*

Digits:	u^{14}	u^{12}	u^{10}	u^8	u^6	u^4	u^2	*1*
Base in $\mathcal{R}_2$*:*	*0*	*0*	$\bar{u}$	1	0	u	u	0
DNA:	*A*	*A*	*T*	*G*	*A*	*C*	*C*	*A*

The following theorem provides us with reverse-complement sets of DNA 2^{k-1}-mers.

Theorem 3.5. *Every ideal in* $\mathcal{R}_{2^k}$ *corresponds to a reversible-complement set of DNA* 2^{k-1}*-mers.*

Proof. We know, from the ideal structure of $\mathcal{R}_{2^k}$, that every ideal is principally generated by $(1+u)^m$ for some $m = 0, 1, 2, \dots 2^k - 1$. So let $\langle (1+u)^m \rangle$ be any ideal. We will consider two cases separately.
Case 1: m is even. If $p = p(u) = (1+u)^m$, then $p_{u^2} = 0^\ell c_1 c_2 \cdots c_{2^{k-1}-\ell}$ in the u^2-adic system, where 0^ℓ corresponds to a vector of zeros of length ℓ. We assume that $c_1 \neq 0$. Now, since $(1+u)^m = ((1+u)^2)^{m/2} = (1+u^2)^{m/2}$, and $\binom{m/2}{i} = \binom{m/2}{m/2-i}$, we see that $(c_1 c_2 \cdots c_{2^{k-1}-\ell})$ is self-reversible, i.e., $(c_1 c_2 \cdots c_{2^{k-1}-\ell})^r = (c_1 c_2 \cdots c_{2^{k-1}-\ell})$.

Then, for $0 \le i \le \ell$ we have (in u^2-adic system)

$$[\left(u^2\right)^i \cdot p_{u^2}]^r = \left(u^2\right)^{\ell-i} \cdot p_{u^2}$$

and for $\ell + 1 \le i \le 2^{k-1} - 1$ we have

$$[\left(u^2\right)^i \cdot p_{u^2}]^r = \left(u^2\right)^{\ell+2^{k-1}-i} \cdot p_{u^2}.$$

Thus u^2-multiples of $(1+u)^m$ are self-reversible. And since $[(u(1+u)^m)_{u^2}]^r = u[((1+u)^m)_{u^2}]^r$, we see that $u(1+u)^m$ is also self-reversible. But then this implies, by linearity of self-reversability, that any element in the ideal is self-reversible.
Case 2: m is odd. Then we can write

$$(1+u)^m = (1+u)(1+u)^{m-1} = (1+u)^{m-1} + u(1+u)^{m-1}.$$

We can then use Case 1 for $(1+u)^{m-1}$. If we then take any polynomial $q(u) = \sum_{i=0}^{2^{k-1}-1} q_i u^{2i}$, we have

$$\varepsilon([q(u)(1+u)^m]^r) \in ((1+u)^m),$$

since the reversibility is linear and closed under multiplying by $a+ub$ and u^2. This proves that every ideal corresponds to a reversible set of DNA 2^{k-1}-mers.

According to the Watson-Crick complement and the correspondence we have defined the complement of an element $\alpha \in \mathcal{R}_{2^k}$ is just given by $\alpha+1+u+u^2+\cdots+u^{2^k-1}$. Because of the ideal structure, we know that each ideal contains the generator of the minimal ideal, which is given by $(1+u)^{2^k-1}$. So if $\alpha \in I$ for some ideal I of $\mathcal{R}_{2^k}$, then $\alpha+(1+u)^{2^k-1} \in I$ as well. But since $(1+u)^{2^k-1} = 1+u+u^2+\cdots+u^{2^k-1}$ in $\mathcal{R}_{2^k}$, we see that the complement of each element in I is also in I.

This proves that each ideal corresponds to a reversible-complement set of DNA 2^{k-1}-mers. □

Let us give an illustrative example.

Example 3.6. *Consider $p(u)=(1+u)^2$ and $J=(p(u))$ in $\mathcal{R}_8=\mathbb{F}_2[u]/(u^8-1)$. Take $q(u)=u^6+u+1$. And let us compute $[(q(u)p(u))_{u^2}]^r$. First, we see that*

$$q(u)p(u)=(u^6+u+1)(u^2+1)=u^8+u^6+(u+1)u^2+u+1=u^6+(u+1)u^2+u \to (q(u)p(u))_{u^2}=10\bar{u}u.$$

Notice that $p(u)=u^2+1$ and so,

$$u^2+1=0\cdot u^6+0\cdot u^4+1\cdot u^2+1 \to 0011.$$

This means $\ell=2$. Thus we have

$$[u^6p(u)_{u^2}]^r=[(u^2)^3p(u)_{u^2}]^r=(u^2)^{\ell+2^{k-1}-3}p(u)=(u^2)^3p(u)=u^6p(u)=u^6+1$$

$$[up(u)_{u^2}]^r=u[(u^2)^0p(u)_{u^2}]^r=u(u^2)^{\ell-0}p(u)=uu^4p(u)=u^5p(u)=u^5+u^7$$

$$[p(u)_{u^2}]^r=[(u^2)^0p(u)_{u^2}]^r=(u^2)^{\ell-0}p(u)=(u^2)^2p(u)=u^4p(u)=u^6+u^4.$$

Hence,

$$[(q(u)p(u))_{u^2}]^r=[u^6p(u)_{u^2}]^r+[up(u)_{u^2}]^r+[p(u)_{u^2}]^r=u^6+1+u^5+u^7+u^6+u^4$$

$$[(q(u)p(u))_{u^2}]^r=u\cdot u^6+(u+1)\cdot u^4+1=u\bar{u}01.$$

Now let us verify that indeed we get the reverse DNA 4-mer:

$$[(q(u)p(u))_{u^2}]=u^6+0\cdot u^4+(u+1)u^2+u=10\bar{u}u \to \zeta^{-1}(q(u)p(u))=GATC.$$

$$[q(u)p(u)_{u^2}]^r=u\cdot u^6+(u+1)\cdot u^4+0\cdot u^2+1=u\bar{u}01 \to \zeta^{-1}([q(u)p(u)]^r)=CTAG.$$

4. Cyclic codes for DNA codes over $\mathcal{R}_{2^k}$

Cyclic codes, have been one of the most common methods by which reversible-complement codes over different alphabets have been obtained. With the vast literature on the structural properties of cyclic codes over finite chain rings especially, this proves to be a fruitful direction to take. We will follow the footsteps of similar works to this effect. The following two theorems can easily be proven by following the similar steps as in [14] and [5].

Theorem 4.1. *Assume that*

$$C = (f_0, (1+u)f_1, (1+u)^2 f_2, ..., (1+u)^{2^k-1} f_{2^k-1})$$

generates a cyclic code of length n over $\mathcal{R}_{2^k}$ with $f_{2^k-1}|...f_3|f_2|f_1|f_0|(x^n-1)$ over $\mathbb{F}_2$. Then the dual of the code is given by

$$C^{\perp} = (\bar{f}^*_{2^k-1}, ..., (1+u)^{2^k-2}\bar{f_1}^*, (1+u)^{2^k-1}\bar{f_0}^*)$$

where $\bar{f_i} = (x^n-1)/f_i$ and if $f_i = 0$ then $f_i = f_{i-1}$ is considered $(f_{-1} = x^n - 1)$ $(0 \leq i \leq 2^k - 1)$.

Further, if C is cyclic code of even length and $f_{2^k-1-i} = \bar{f}^*_i$ for $0 \leq i \leq 2^k - 1$, then C is a cyclic self dual code.

The following is a classical result about reversible codes:

Theorem 4.2. *[13] The cyclic code generated by a monic polynomial $g(x)$ is reversible if and only if $g(x)$ is self-reciprocal where $g(x)|(x^n-1)$.*

The following lemma, taken from [2] explains the basic properties of reciprocals.

Lemma 4.3. *Let $f(x)$ and $g(x)$ be polynomials in $\mathcal{R}_{2^k}$ with $\deg(f) \geq \deg(g)$. Then*

1. *$[f(x)g(x)]^\star = f(x)^\star g(x)^\star$ and*
2. *$[f(x)+g(x)]^\star = f(x)^\star + x^{\deg(f)-\deg(g)}g(x)^\star$.*

We are now ready to state the main result of this section:

Theorem 4.4. *Let $C = (f_0, (1+u)f_1, (1+u)^2 f_2, ..., (1+u)^{2^k-1} f_{2^k-1})$ be a cyclic code over $\mathcal{R}_{2^k}$ with $f_{2^k-1}|...f_3|f_2|f_1|f_0|(x^n-1)$ over $\mathbb{F}_2$. Then the dual of the code is given by*

$$C^{\perp} = (\bar{f}^*_{2^k-1}, ..., (1+u)^{2^k-2}\bar{f_1}^*, (1+u)^{2^k-1}\bar{f_0}^*)$$

with $\bar{f_i} = (x^n-1)/f_i$. Moreover,

- *If f_i's ($\bar{f_i}$'s) are self reciprocal polynomials then C ($C^\perp$) is a reversible cyclic code and $\zeta^{-1}(C)$ ($\zeta^{-1}(C^\perp)$) is a reversible DNA code.*
- *If f_i's ($\bar{f_i}$'s) are self reciprocal polynomials and f_0($\bar{f}^*_{2^k-1}$) divides $(x^n-1)/(x+1)$ over $\mathbb{F}_2$ then C ($C^\perp$) is a reversible cyclic code with complement property and so $\zeta^{-1}(C)$ ($\zeta^{-1}(C^\perp)$) is a reversible complement DNA code.*

Proof. If $c(x) \in C$, then $c(x) = f_0 a_0 + (1+u)f_1 a_1 + ... + (1+u)^{2^k-1} f_{2^k-1} a_{2^k-1}$ where a_i are polynomials in $\mathcal{R}_{2^k}$ and $(deg(a_i) < deg(f_{i-1})) - deg(f_i))$. Since $f_{2^k-1}|...f_3|f_2|f_1|f_0|(x^n-1)$, there exists non-negative integers $m_i = \deg(f_0 a_0) - \deg(f_i a_i)$, and hence by Lemma 4.3, we have

$$\begin{aligned}
c(x)^* &= [f_0 a_0 + (1+u)f_1 a_1 + ... + (1+u)^{2^k-1} f_{2^k-1} a_{2^k-1}]^* \\
&= (f_0 a_0)^* + x^{m_1}((1+u)f_1 a_1)^* + ... + x^{m_{2^k-1}}((1+u)^{2^k-1} f_{2^k-1} a_{2^k-1})^* \\
&= f_0 (a_0)^* + x^{m_1}(1+u)f_1 (a_1)^* + ... + x^{m_{2^k-1}}(1+u)^{2^k-1} f_{2^k-1} (a_{2^k-1})^* \\
&\Rightarrow c(x)^* \in C.
\end{aligned}$$

Therefore, C is reversible. This means, by Theorem 3.5, that $\zeta^{-1}(C)$ is a reversible DNA code.

If moreover, f_0 divides $(x^n-1)/(x+1)$, then all f_i's divide $(x^n-1)/(x+1)$ since $f_{2^k-1}|...f_3|f_2|f_1|f_0|(x^n-1)$ over $\mathbb{F}_2$. This means $(1+u)^{2^k-1}(x^n-1)/(x+1) \in C$. Hence $\zeta^{-1}(C)$ is a reversible complement code.

The proofs for $C^\perp$, being similar to the above case, are omitted here. □

Example 4.5. *Let us consider the self reciprocal polynomials* $f_0 = x^5+x^4+x^3+x^2+x+1$, $f_1 = x^4+x^2+1$, $f_2 = x^2+x+1$, $f_3 = f_4 = ...f_7 = 0$ *with* $f_2|f_1|f_0|(x^6-1)$ *over* $\mathbb{F}_2$*. The code* $C = (f_0, (1+u)^4 f_1, (1+u)^6 f_2)$ *is generated by the following matrix* $\mathcal{R}_{2^3}$*:*

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ u^5+u^4+u+1 & 0 & u^5+u^4+u+1 & 0 & u^5+u^4+u+1 & 0 \\ u^6+u^4+u^2+1 & u^6+u^4+u^2+1 & u^6+u^4+u^2+1 & 0 & 0 & 0 \\ 0 & u^6+u^4+u^2+1 & u^6+u^4+u^2+1 & u^6+u^4+u^2+1 & 0 & 0 \end{pmatrix}.$$

By Theorem 4.4, C *is a reversible cyclic code over* $\mathcal{R}_{2^3}$ *and it corresponds to a reversible-complement DNA code of length 24.*
The dual is given by :
$C^\perp = (\bar{f}_2^*, (u+1)\bar{f}_2^*, (u+1)^2\bar{f}_1^*, (u+1)^3\bar{f}_1^*, (u+1)^4\bar{f}_0^*, (u+1)^5\bar{f}_0^*, (u+1)^6\bar{f}_0^*, (u+1)^7\bar{f}_0^*)$ *which can be expressed as* $C^\perp = (\bar{f}_2^*, (u+1)^2\bar{f}_1^*, (u+1)^4\bar{f}_0^*)$*, where*
$\bar{f}_2^* = 1+x+x^3+x^4$, $\bar{f}_1^* = 1+x^2$, $\bar{f}_0^* = 1+x$*. Then all these polynomials are self reciprocal but not all of them divide* $(x^n-1)/(x+1)$*. Thus* $C^\perp$ *corresponds to a reversible DNA code.*

Example 4.6. *Now consider self reciprocal polynomials* $f_0 = x^5+x^4+x^3+x^2+x+1$, $f_1 = x+1$, $f_2 = f_3 = f_4 = ...f_7 = 0$ *with* $f_1|f_0|(x^6-1)$ *over* $\mathbb{F}_2$*. The code* $C = (f_0, (1+u)^2 f_1,)$ *is generated by the following matrix over* $\mathcal{R}_{2^3}$*:*

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ u^2+1 & u^2+1 & 0 & 0 & 0 & 0 \\ 0 & u^2+1 & u^2+1 & 0 & 0 & 0 \\ 0 & 0 & u^2+1 & u^2+1 & 0 & 0 \\ 0 & 0 & 0 & u^2+1 & u^2+1 & 0 \end{pmatrix}.$$

By the same argument as in the previous example, C *is a reversible cyclic code over* $\mathcal{R}_{2^3}$ *and moreover, since the polynomials satisfy the complement properties,* C *corresponds to a reversible-complement DNA code of length 24.*
The dual of C *is given by* $C^\perp = (\bar{f}_1^*, (u+1)^6\bar{f}_0^*)$*, where* $\bar{f}_1^* = x^5+x^4+x^3+x^2+x+1$*, and* $\bar{f}_0^* = 1+x$*. Then we again have self reciprocal polynomials since they all divide* $(x^n-1)/(x+1)$*, we see that* $C^\perp$ *also corresponds to a reversible-complement DNA code.*

The next example illustrates the DNA reversibility in an explicit form.

Example 4.7. *Let* $f_0 = x^8+x^7+x^6+x^5+x^4+x^3+x^2+x+1$, $f_1 = x^6+x^3+1$ *and* $f_1|f_0|x^9-1$ *over* $\mathbb{F}_2$*, where we will consider the codes over* $\mathcal{R}_8$*. Suppose* $C = ((1+u)^3 f_0, (1+u)^4 f_1)$*, whose generator matrix can be written as follows: Let* $U_i = (1+u)^i$*,*

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} U_3 & U_3 & U_3 & U_3 & U_3 & U_3 & U_3 & U_3 & U_3 \\ U_4 & 0 & 0 & U_4 & 0 & 0 & U_4 & 0 & 0 \\ 0 & U_4 & 0 & 0 & U_4 & 0 & 0 & U_4 & 0 \end{pmatrix}.$$

We can easily observe that the reverse of v_2 *is given by* $v_2^r = (1+u)v_1 + v_2 + v_3$*, while* v_1 *and* v_3 *are self-reversible.*
Consider the codeword

$$\begin{aligned} c_1 = (1+u^2+u^3)v_2 =& (u^7+u^6+u^4+u^3+u^2+1, 0, 0, u^7+u^6+u^4+u^3+u^2+1, 0, 0, \\ & u^7+u^6+u^4+u^3+u^2+1, 0, 0) \end{aligned}$$

whose DNA-correspondence is given by

$$\begin{aligned} \zeta^{-1}(c_1) =& (T, G, T, G, A, A, A, A, A, A, A, A, T, G, T, G, A, A, A, A, A, A, A, A, T, G, T, G, \\ & A, A, A, A, A, A, A, A). \end{aligned}$$

Then if we take $c_2 = (1+u+u^2)((1+u)v_1 + v_2 + v_3)$, *we see that*

$$\begin{aligned} c_2 = (1+u+u^2)v_2^r = &(0,0,u^6+u^5+u^4+u^2+u+1,0,0,u^6+u^5+u^4+u^2+u+1,0,0,\\ &u^6+u^5+u^4+u^2+u+1). \end{aligned}$$

Note that

$$\begin{aligned} \zeta^{-1}(c_2) = &(A,A,A,A,A,A,A,A,G,T,G,T,A,A,A,A,A,A,A,A,G,T,G,T,A,A,A,A,\\ &A,A,A,A,G,T,G,T) \end{aligned}$$

and we have $(\zeta^{-1}(c_1))^r = \zeta^{-1}(c_2)$.

Next, consider $e_1 = uv_1 + (1+u^3)v_2 + (u^2+u^3)v_3$, *which is given as a vector by*

$$\begin{aligned} &(u^7+u^2+u+1,u^7+u^6+u^4+u,u^4+u^3+u^2+u,u^7+u^2+u+1,u^7+u^6+u^4+u,\\ &u^4+u^3+u^2+u,u^7+u^2+u+1,u^7+u^6+u^4+u,u^4+u^3+u^2+u), \end{aligned}$$

so that the DNA correspondence is given by

$$\begin{aligned} \zeta^{-1}(e_1) = &(C,A,G,T,T,G,A,C,A,G,T,C,C,A,G,T,T,G,A,C,A,G,T,C,C,A,G,T,\\ &T,G,A,C,A,G,T,C) \end{aligned}$$

Now, if we let $e_2 = ((u^2+u^3+u^4))v_1^r + (u+u^2)v_2^r + (1+u)v_3^r$, *which then would be written as a vector as*

$$\begin{aligned} &(u^7+u^5+u^4+u^2,u^7+u^2+u+1,u^7+u^6+u^4+u,u^7+u^5+u^4+u^2,u^7+u^2+u+1,\\ &u^7+u^6+u^4+u,u^7+u^5+u^4+u^2,u^7+u^2+u+1,u^7+u^6+u^4+u), \end{aligned}$$

we will have

$$\begin{aligned} \zeta^{-1}(e_2) = &(C,T,G,A,C,A,G,T,T,G,A,C,C,T,G,A,C,A,G,T,T,G,A,C,C,T,G,A,\\ &C,A,G,T,T,G,A,C). \end{aligned}$$

Note that, in this case we again have $(\zeta^{-1}(e_1))^r = \zeta^{-1}(e_2)$.

Example 4.8. *Let us consider self reciprocal polynomials* $f_0 = x^{14}+x^{13}+x^{12}+x^{11}+x^{10}+x^9+x^8+x^7+x^6+x^5+x^4+x^3+x^2+x+1$, $f_1 = 1+x^3+x^6+x^9+x^{12}$, $f_2 = x^4+x^3+x^2+x+1$, $f_3 = f_4 = ...f_7 = 0$ *and* $f_2|f_1|f_0|(x^{15}-1)$ *over* $\mathbb{F}_2$. *The code* $C = ((1+u)f_0,(1+u)^3f_1,(1+u)^4f_2)$ *is generated over* $\mathcal{R}_{2^4}$ *by the following matrix* ($U_i = (1+u)^i$)

$$\begin{pmatrix}
U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1 & U_1\\
U_3 & 0 & 0 & U_3 & 0 & 0 & U_3 & 0 & 0 & U_3 & 0 & 0 & U_3 & 0 & 0\\
0 & U_3 & 0 & 0 & U_3 & 0 & 0 & U_3 & 0 & 0 & U_3 & 0 & 0 & U_3 & 0\\
U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & U_4 & U_4 & U_4 & U_4 & U_4 & 0 & 0 & 0
\end{pmatrix}.$$

Then, C *is a reversible cyclic code over* $\mathcal{R}_{2^4}$ *and it corresponds to a reversible-complement DNA code of length 120 and minimum Hamming distance 2.*

Example 4.9. *Consider* $f_0 = x^3+x^2+x+1$ *and* $f_1 = 1+x^2$ *with* $f_1|f_0|(x^4-1)$ *over* $\mathbb{F}_2$. *Let* $C = ((1+u)f_0,(1+u)^7f_1)$ *be the code over* $\mathcal{R}_{2^3}$. *Then* C *is a cyclic code of length 4 and distance 2.* $\zeta^{-1}(C)$ *is a reversible-complement DNA code of length 16 and Hamming distance 4.*

Table 1. $f = 1 + x$, $g = 1 + x + x^2$. **Reversible cyclic codes over $\mathcal{R}_{2^3}$ of length 6 which correspond to reversible complement DNA codes of length 24. ($d(C)$: Minimum Hamming distance of the code C. $d(\zeta^{-1}(C))$: Minimum Hamming distance of the code $\zeta^{-1}(C)$)**

The code	Length	$d(C)$	$d(\zeta^{-1}(C))$
$(fg^2, (1+u)g^2)$	6	3	3
$(fg^2, (1+u)^4 g^2)$	6	3	6
$(fg^2, (1+u)^2 g^2)$	6	3	6

In Table 1 we give a table of some reversible-complement codes obtained from similar constructions.

We conclude this section by the following theoretical result whose proof is omitted, being similar to the above ones, with an application.

Theorem 4.10. *Let $C = (f_0, (1+u)f_1, (1+u)^2 f_2, ..., (1+u)^{2^k-1} f_{2^k-1})$ be a cyclic code over $\mathcal{R}_{2^k}$ of even length. If all f_i are self reciprocal polynomials and $f_{2^k-1-i} = f_i^*$ for $0 \leq i \leq 2^k - 1$, then C is a reversible cyclic self-dual code and $\zeta^{-1}(C)$ is a reversible DNA code. If $f_0|(x^n - 1)/(x + 1)$ then $\zeta^{-1}(C)$ is a reversible complement DNA code.*

Example 4.11. *Some self dual codes that correspond to reversible complement DNA codes are shown in Table 2:*

Table 2. $f = (x - 1)$ **and** $g = (x^2 + x + 1)$

Ring	Length	Generator of the code
$\mathcal{R}_{2^3}$	16	$(f^{13}, (u-1)^3 f^9, (u-1)^4 f^7, (u-1)^5 f^3)$
$\mathcal{R}_{2^3}$	16	$((u-1)^3 f^9, (u-1)^4 f^7, (u-1)^5)$
$\mathcal{R}_{2^3}$	8	$(f^7, (u-1) f^5, (u-1)^4 f^3, (u-1)^7 f)$
$\mathcal{R}_{2^3}$	6	$(f^2 g, (u-1)^4 g)$

5. Conclusion

In this work, we solve the reversibility problem on cyclic codes with ideal structures of the ring $\mathcal{R}_{2^k}$ by u^2-adic system. Both reversible cyclic codes on the ring $\mathcal{R}_{2^k}$ and reversible (or reversible complement) DNA codes are obtained from the same codes. The properties of DNA codes with the family of cyclic dual and self dual codes have been explored. Thus an extension of what was done in [20] has been established in the most general sense without puncturing. The idea that we have used can be used for different such extensions as well. However, the finite chain property of the ring, which is useful in both describing all the ideals of the ring and making it easier to study cyclic codes, has been essential in obtaining our main results. When the extension ring does not have the finite chain property, some partial results can still be obtained, however a complete characterization, similar to what we have done here, is very challenging. Hence, future studies on the general case i.e reversible and complement cyclic codes over different rings such as $\mathcal{R}_{2k}$ remains to be an open and interesting problem.

Acknowledgment: The work presented here was supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) (Grant No: 113F071). The authors wish to express their thanks to the reviewers for their valuable remarks and comments that improved the presentation of the paper.

References

[1] N. Aboluion, D. H. Smith, S. Perkins, Linear and nonlinear constructions of DNA codes with Hamming distance d, constant GC–content and a reverse–complement constraint, Discrete Math. 312(5) (2012) 1062–1075.

[2] T. Abulraub, A. Ghrayeb, X. N. Zeng, Construction of cyclic codes over $GF(4)$ for DNA computing, J. Frankl. Inst. 343(4–5) (2006) 448–457.

[3] L. Adleman, Molecular computation of solutions to combinatorial problems, Science 266(5187) (1994) 1021–1024.

[4] L. Adleman, P. W. K. Rothemund, S. Roweis, E. Winfree, On applying molecular computation to the Data Encryption Standard, J. Comput. Biol. 6(1) (1999) 53–63.

[5] R. Alfaro, S. Bennett, J. Harvey, C. Thornburg, On distances and self–dual codes over $F_q[u]/(u^t)$, Involv. J. Math. 2(2) (2009) 177–194.

[6] D. Boneh, C. Dunworth, R. Lipton, Breaking DES using molecular computer, Princeton CS Tech–Report, Number CS–TR-489–95, 1995.

[7] A. G. Frutos, Q. Liu, A. J. Thiel, A. M. W. Sanner, A. E. Condon, L. M. Smith, R. M. Corn, Demonstration of a word design strategy for DNA computing on surfaces, Nucleic Acids Res. 25(23) (1997) 4748–4757.

[8] P. Gaborit, O. D. King, Linear construction for DNA codes, Theoret. Comput. Sci. 334(1–3) (2005) 99–113.

[9] O. D. King, Bounds for DNA codes with constant GC–content, Electron. J. Comb. 10 (2003) 1–13.

[10] M. Li, H. J. Lee, A. E. Condon, R. M. Corn, DNA word design strategy for creating sets of non–interacting oligonucleotides for DNA microarrays, Langmuir 18(3) (2002) 805–812.

[11] M. Mansuripur, P. K. Khulbe, S. M. Kuebler, J. W. Perry, M. S. Giridhar, N. Peyghambarian, Information storage and retrieval using macromolecules as storage media, in Optical Data Storage, OSA Technical Digest Series (Optical Society of America) paper TuC2, 2003.

[12] A. Marathe, A. E. Condon, R. M. Corn, On combinatorial DNA word design, J. Comput. Biol. 8(3) (2001) 201–220.

[13] J. L. Massey, Reversible codes, Inform. and Control 7(3) (1964) 369–380.

[14] G. H. Norton, A. Salagean, On the structure of linear and cyclic codes over a finite chain ring, Appl. Algebr. Eng. Comm. 10 (2000) 489–506.

[15] E. S. Oztas, I. Siap, Lifted polynomials Over F_{16} and their applications to DNA codes, Filomat 27(3) (2013) 459–466.

[16] E. S. Oztas, I. Siap, B. Yildiz, Reversible codes and applications to DNA, Lect. Notes Comput. Sc. 8592 (2014) 124–128.

[17] E. S. Oztas, I. Siap, On a generalization of lifted polynomials over finite fields and their applications to DNA codes, Int. J. Comput. Math. 92(9) (2015) 1976–1988.

[18] I. Siap, T. Abulraub, A. Ghayreb, Similarity cyclic DNA codes over rings, International Conference on Bioinformatics and Biomedical Engineering, in Shanghai, iCBBE 2008, PRC, May 16–18th 2008.

[19] I. Siap, T. Abulraub, A. Ghrayeb, Cyclic DNA codes over the ring $\mathbb{F}_2[u]/(u^2-1)$ based on the deletion distance, J. Frankl. Inst. 346(8) (2009) 731–740.

[20] B. Yildiz, I. Siap, Cyclic codes over $\mathbb{F}_2[u]/(u^4-1)$ and applications to DNA codes, Comput. Math. Appl. 63(7) (2012) 1169–1176.