Indian J Pure Appl Math (2024) 55:1465–1476
https://doi.org/10.1007/s13226-023-00451-w

ORIGINAL RESEARCH

Construction of cyclic DNA codes over Z_4R

Tulay Yildirim

Received: 7 May 2022 / Accepted: 14 June 2023 / Published online: 23 June 2023
© The Indian National Science Academy 2023

Abstract This study is devoted to the construction of cyclic DNA codes of odd length over $\mathbb{Z}_4\mathbb{R}$. The generator polynomial of cyclic codes of this ring and the structure of its separable codes are studied in detail. A map ψ_β from $\mathbb{R}^\beta$ to $\{A, T, G, C\}^{2\beta}$ is defined and then constructed a one-to-one correspondence between DNA codons of alphabets and the elements of $\mathbb{R}$. Furthermore, necessary and sufficient circumstances of cyclic codes over $\mathbb{Z}_4$, $\mathbb{R}$ and $\mathbb{Z}_4\mathbb{R}$ to be reversible codes and reverse-complement constraints codes are determined. To support the theories, some important examples are illustrated.

Keywords DNA codes over rings · Gray map · Linear codes · Mixed alphabets · Reversible codes · Reversible-complement codes

Mathematics Subject Classification 94B15 · 94B05 · 94B60 · 94A45

1 Introduction

DNA (Deoxyribonucleic acid) contains the genetic program for biological development of life. It is a molecule consisting of two chains surrounding a double helix that carries the hereditary instructions used in the development of living things. DNA strands is a sequence of four possible nucleotides, adenine (A), thymine (T), guanine (G) and cytosine (C). The ends of a DNA strand are chemically polar with $5'$ and $3'$ ends, which implies that the strands are oriented. Hybridization occurs when a strand binds to another strand, forming a double strand of DNA. The DNA strand is held by a complementary base pairing that connects the Watson-Crick complementary bases with each other denoted by $\bar{A} = T, \bar{T} = A, \bar{G} = C$ and $\bar{C} = G$. For instance, the Watson-Crick complementary (WCC) strand of $3' - GATCGTTC - 5'$ is the strand $5' - CTAGCAAG - 3'$.

For the first time in 1994, Adleman introduced a successful experiment for DNA studies [2]. With this work, he brought a new perspective to the solution of mathematical problems by using DNA as an optimal computational resource due to its dense and self-replicating ability. His research was based upon the WCC feature of DNA strands. Henceforth, DNA computation studies have been expanded by many researchers with different approaches to solve mathematical problems [3,7,10,12,13,15].

In algebraic coding theory, researchers concerned with error correction codes over finite fields and rings of cardinality 4^n by mapping DNA nucleotides with elements of finite fields and rings. Especially, studying on DNA cyclic codes were quite interesting in terms of their applications. For instance, Abualrub et al. defined DNA codes over a finite field of order four and constructed linear codes and cyclic codes over $GF(4)$ for DNA

Communicated by Rahul Roy.

T. Yildirim (✉)
Eskipazar Vocational School, Karabuk University, Karabük, Turkey
E-mail: tulayturan@karabuk.edu.tr

computing [1]. Benenni et al discussed DNA cyclic codes over rings [6]. Bayram et al. studied DNA codes and also their applications over the ring $\mathbb{F}_4[v]/(v^2 - v)$ [5]. Oztas et al. introduced a new family of polynomials which generates reversible codes over a finite field with sixteen elements [16].

In [11], Gao et al. studied on linear codes over $\mathbb{Z}_4 + v\mathbb{Z}_4$ and their applications to construct well and new $\mathbb{Z}_4$-linear codes. In recent years, Liu investigated DNA codes of odd length over the ring $\mathbb{R} = \mathbb{Z}_4 + v\mathbb{Z}_4$ with $v^2 = v$ [14]. Also, Dinh et al. introduced a new vision by studying on DNA cyclic codes over mixed alphabets [8]. Thus, motivated by these ideas, in this paper, we first discussed cyclic DNA codes over single alphabets, $\mathbb{Z}_4$ and $\mathbb{R}$, and then extended this work to mixed alphabets $\mathbb{Z}_4\mathbb{R}$.

The aim of this paper is to introduce the theory of cyclic codes to construct codes that are suitable for DNA computing. Especially, we study linear cyclic codes over $\mathbb{Z}_4\mathbb{R}$ that is suitable for this application. We first determined the generator polynomials of $\mathbb{Z}_4\mathbb{R}$-cyclic codes and then consider the reversible and reversible-complement constraints of cyclic codes over the defined mixed alphabets.

This paper is constructed in the following way: Section 2 contains some basic definitions and structures of the ring $\mathbb{Z}_4\mathbb{R}$. This section also introduces the Gray map from $\mathbb{Z}_4\mathbb{R}$ to $\mathbb{Z}_4^3$ and discusses some important properties of this map. In Section 3, the algebraic construction of $\mathbb{Z}_4$, $\mathbb{R}$ and $\mathbb{Z}_4\mathbb{R}$-cyclic codes are investigated and their generator polynomials are defined, respectively. In Section 4, we concentrate the main results of this paper. In the beginning of the section, we introduce reversible and reversible-complement constraints over the ring $\mathbb{Z}_4$. Then we define a one-to-one relation from $\mathbb{R}^\beta$ to $\{A, T, G, C\}^{2\beta}$ and present necessary and sufficient circumstances for cyclic codes to be reversible and reversible-complement codes over the ring $\mathbb{R}$. To support our theory, we illustrate some examples and construct DNA codes. By the end of the section, we introduce our main results and consider necessary and sufficient circumstances for cyclic codes to be reversible and reversible-complement codes over the ring $\mathbb{Z}_4\mathbb{R}$. To illustrate obtained results, we also demonstrate some good examples. In Section 5, we conclude the paper.

2 Preliminaries and Definitions

In this study, we give our attention to the cyclic DNA codes over combined alphabets. Especially, we considered the cyclic DNA codes over the ring $\mathbb{Z}_4\mathbb{R}$, where $\mathbb{R} = \mathbb{Z}_4 + v\mathbb{Z}_4 = \{a + vb : a, b \in \mathbb{Z}_4\}$ where $v^2 = v$ is the commutative ring with 16 elements and characteristic 4. Its well known that the ring $\mathbb{R}$ is a non-chain ring with two maximal ideals $< 1 - v >$ and $< v >$. The Chinese Remainder Theorem then implies that $\mathbb{R} =< v > \oplus < 1 - v >$ [11]. More concretely, $\mathbb{R}$ can be uniquely expressed as $a + vb = v(a + b) + (1 - v)a$ for all $a, b \in \mathbb{Z}_4$.

A cyclic code of length $\beta \in \mathbb{N}$ over $\mathbb{R}$ is a linear code if $(c_0, c_1, \cdots, c_{\beta-1}) \in C_\beta$ then $(c_{\beta-1}, c_0, \cdots, c_{\beta-2}) \in C_\beta$. A cyclic code of length β over $\mathbb{R}$ can be identified with an ideal in $\mathbb{R}_\beta$ by $\mathbb{R}$-module isomorphism:

$$\varphi : \mathbb{R}^\beta \to \mathbb{R}_\beta$$
$$(c_0, c_1, \cdots, c_{\beta-1}) \mapsto c_0 + c_1 x + \cdots + c_{\beta-1} x^{\beta-1}$$

where $\mathbb{R}_\beta$ is the residue ring $\mathbb{R}[x]|(x^\beta - 1)$.

The Hamming weight $w_H(c)$ of a codeword $c \in C$ is the number of nonzero components in c. The Hamming distance between two codewords c_1 and c_2 in C is the Hamming weight of the codeword $c_1 - c_2$ and denoted by $d_H(c_1, c_2)$. The minimum Hamming distance d_H of C can be defined as

$$min\{d_H(c_1, c_2) : c_1, c_2 \in C, c_1 \neq c_2\}$$

Lee weight of $(a + vb)$ in $\mathbb{R}$ can be defined as

$$w_L(a + vb) = w_H(a) + w_H(a + b)$$

The Lee weight of a codeword is the rational sum of Lee weights of its components.

Let $x = (x_1, \cdots, x_\beta)$ and $y = (y_1, \cdots, y_\beta)$ be two vectors of $\mathbb{R}^\beta$. The Euclidean inner product of x and y is defined as follows:

$$x.y = \sum_{i=1}^{\beta} x_i y_i$$

The Euclidean dual code $C_\beta^\perp$ of C_β is defined as $C_\beta^\perp = \{x \in \mathbb{R}^\beta : x.c = 0,\ \forall c \in C_\beta\}$. C_β is Euclidean self-orthogonal if $C_\beta \subseteq C_\beta^\perp$ and Euclidean self-dual if $C_\beta = C_\beta^\perp$.

Now, we consider cyclic codes over $\mathbb{Z}_4\mathbb{R}$. Consider $\mathbb{R}_{\alpha,\beta} = \frac{\mathbb{Z}_4[x]}{x^\alpha - 1} \times \frac{\mathbb{R}[x]}{x^\beta - 1}$, and we define the set

$$\mathbb{Z}_4\mathbb{R} = \{(e, f) : e \in \mathbb{Z}_4, f \in \mathbb{R}\}$$

Let C be a cyclic code over $\mathbb{Z}_4\mathbb{R}$ and $\alpha(resp.\beta)$ is the set of $\mathbb{Z}_4(resp.\mathbb{R})$ coordinate positions and throughout the study, we assume that α and β are odd positive integers. Any codeword $c \in C$ has the form

$$c = (e_0, e_1, \cdots, e_{\alpha-1}, f_0, f_1, \cdots f_{\beta-1}) \in \mathbb{Z}_4^\alpha \mathbb{R}^\beta$$

where $f_i = a_i + vb_i \in \mathbb{R}^\beta$ for all $i = 0, 1, \cdots, \beta - 1$. The ring homomorphism map is defined as

$$\begin{aligned} \delta &: \mathbb{R} \to \mathbb{Z}_4 \\ a + vb &\mapsto a \end{aligned}$$

For any $r \in \mathbb{R}$, define a scalar multiplication $\star$ by $r \star (e, a + vb) = (\delta(r).e, r.(a + vb))$ where $e \in \mathbb{Z}_4$ and $a + vb \in \mathbb{R}$. It naturally extends to $\mathbb{Z}_4^\alpha \mathbb{R}^\beta$ as follows:

$$r \star c = (\delta(r)e_0, \delta(r)e_1, \cdots, \delta(r)e_{\alpha-1}, rf_0, rf_1, \cdots rf_{\beta-1}) \tag{1}$$

where $c = (e_0, e_1, \cdots, e_{\alpha-1}, f_0, f_1, \cdots f_{\beta-1}) \in \mathbb{Z}_4^\alpha \mathbb{R}^\beta$ for $\alpha, \beta \in \mathbb{N}$.

A nonempty subset C of $\mathbb{Z}_4^\alpha \mathbb{R}^\beta$ is called a $\mathbb{Z}_4\mathbb{R}$-linear code with length $\alpha + \beta$ if C is an $\mathbb{R}$-submodule of $\mathbb{Z}_4^\alpha \mathbb{R}^\beta$.

Lemma 2.1 *The set $\mathbb{Z}_4^\alpha \mathbb{R}^\beta$ is an $\mathbb{R}$-module with respect to the addition and scalar multiplication in Equation (1).*

An element $c = (e_0, e_1, \cdots, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-1}) \in \mathbb{Z}_4^\alpha \mathbb{R}^\beta$ can be identified with a module element consisting of two polynomials $c(x) = (e(x), f(x)) \in \mathbb{R}_{\alpha,\beta}$, where $e(x) = e_0 + e_1 x + \cdots + e_{\alpha-1}x^{\alpha-1}$ and $f(x) = f_0 + f_1 x + \cdots + f_{\beta-1}x^{\beta-1}$. This identification gives a one-to-one correspondence between elements in $\mathbb{Z}_4^\alpha \mathbb{R}^\beta$ and elements $\mathbb{R}_{\alpha,\beta}$.

Let $r(x) = r_0 + r_1 x + \cdots + r_s x^s \in \mathbb{R}[x]$ and $(e(x), f(x)) \in \mathbb{R}_{\alpha,\beta}$, multiplication operation on $\mathbb{Z}_4\mathbb{R}$ is defined as

$$r(x) \star (e(x), f(x)) = (\delta(r(x)).e(x), r(x).f(x))$$

where $\delta(r(x)) = \delta(r_0) + \delta(r_1)x + \cdots + \delta(r_s)x^s \in \mathbb{Z}_4[x]|(x^\alpha - 1)$. This multiplication is well-defined on $\mathbb{R}_{\alpha,\beta}$. Furthermore, $\delta(r(x)).e(x)$ and $r(x).f(x)$ are defined in $\mathbb{Z}_4[x]|(x^\alpha - 1)$ and $\mathbb{R}[x]|(x^\beta - 1)$, respectively.

A cyclic code of length $\alpha + \beta$ over $\mathbb{Z}_4\mathbb{R}$ is a linear code with a property that if $c = (e_0, e_1, \cdots, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-1}) \in C$, then $\varrho(c) := (e_{\alpha-1}, e_0, \cdots, e_{\alpha-2}, f_{\beta-1}, f_0, \cdots, f_{\beta-2}) \in C$.

The Euclidean inner product on the ring $\mathbb{Z}_4^\alpha \mathbb{R}^\beta$ is defined as

$$< x, y >= v \sum_{i=0}^{\alpha-1} e_i e_i' + \sum_{j=0}^{\beta-1} f_j f_j'$$

where $x = (e_0, e_1, \cdots, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-1},)$ and $y = (e_0', e_1' \cdots, e_{\alpha-1}', f_0', f_1', \cdots, f_{\beta-1}')$ are in $\mathbb{Z}_4^\alpha \mathbb{R}^\beta$.

The dual code of C, denoted by $C^\perp$, is also $\mathbb{Z}_4\mathbb{R}$ linear and defined by

$$C^\perp = \{c \in \mathbb{Z}_4^\alpha \mathbb{R}^\beta :< c, c' >= 0 \ for\ all\ c' \in C\}$$

Lemma 2.2 *Let C be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$. Then its dual code $C^\perp$ is also a $\mathbb{Z}_4\mathbb{R}$-cyclic code .*

Proof Let C be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha+\beta$ and $c=(e_0,e_1,\cdots,e_{\alpha-1},f_0,f_1,\cdots,f_{\beta-1})\in C$. Take $lcm(\alpha,\beta)=l\in\mathbb{N}$ and $c'=(e'_0,e'_1,\cdots,e'_{\alpha-1},f'_0,f'_1,\cdots,f'_{\beta-1})\in C^{\perp}$. Our aim is to show that $\varrho(c')\in C^{\perp}$. By the definition of Euclidean inner product,

$$\begin{aligned}c.\varrho(c') =& v(e_0e'_{\alpha-1}+e_1e'_0+\cdots+e_{\alpha-1}e'_{\alpha-2})\\ &+(f_0f'_{\beta-1}+f_1f'_0+\cdots+f_{\beta-1}f'_{\beta-2})\end{aligned}$$

As C is a $\mathbb{Z}_4\mathbb{R}$-cyclic code and $lcm(\alpha,\beta)=l$, $\varrho^{l-1}(c)\in C$ where $\varrho^{l-1}(c)=(e_1,e_2,\cdots,e_{\alpha-1},e_0,f_1,f_2,\cdots,f_{\beta-1},f_0)$. The inner product of $\varrho^{l-1}(c)$ and c' gives that $\varrho^{l-1}(c).c'=0$, so

$$\begin{aligned}e_1e'_0+e_2e'_1+\cdots+e_{\alpha-1}e'_{\alpha-2}+e_0e'_{\alpha-1}=0\\ f_1f'_0+f_2f'_1+\cdots+f_{\beta-1}f'_{\beta-2}+f_0f'_{\beta-1}=0\end{aligned}$$

Thus, $c.\varrho(c')=0$, this implies that $\varrho(c')\in C^{\perp}$. □

Theorem 2.3 *A linear code C is $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha+\beta$ if and only if C is an $\mathbb{R}[x]$-submodule of $\mathbb{R}_{\alpha,\beta}$.*

Proof Assume C is a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha+\beta$. Let

$$c=(e_0,e_1,\cdots,e_{\alpha-1},f_0,f_1,\cdots,f_{\beta-1})\in C$$

and its polynomial representation is $c(x)=(e(x),f(x))$. So we have $x\star c(x)=(e_{\alpha-1}+e_0x+\cdots+e_{\alpha-2}x^{\alpha-1},f_{\beta-1}+f_0x+\cdots+f_{\beta-2}x^{\beta-1})$ which corresponds to the cyclic shift $(e_{\alpha-1},e_0,\cdots,e_{\alpha-2},f_{\beta-1},f_0,\cdots,f_{\beta-2})$. So one gets $x\star c(x)\in C$. By the linearity property of C, $r(x)\star c(x)\in C$ for any $r(x)\in\mathbb{R}[x]$. Hence, C is an $\mathbb{R}[x]$-submodule of $\mathbb{R}_{\alpha,\beta}$. Converse follows by the definition. □

Now, we give our attention to the linear codes over $\mathbb{R}$ and $\mathbb{Z}_4\mathbb{R}$. By using the Chinese Remainder Theorem (CRT), we discuss the form of the linear codes. Moreover, we introduce a Gray map on $\mathbb{Z}_4\mathbb{R}$ and study some important properties of this Gray map.

Let A and B be codes over $\mathbb{R}$, we denote $A\oplus B=\{a+b:a\in A,b\in B\}$ and $A\otimes B=\{(a,b):a\in A,b\in B\}$. Define

$$C_1=\{a+b\in\mathbb{Z}_4^{\beta}:v(a+b)+(1-v)a\in C_{\beta},\ \ for\ some\ a,b\in\mathbb{Z}_4^{\beta}\}$$

and

$$C_2=\{a\in\mathbb{Z}_4^{\beta}:v(a+b)+(1-v)a\in C_{\beta},\ \ for\ some\ a,b\in\mathbb{Z}_4^{\beta}\}$$

It is clear that C_1 and C_2 are linear codes of length β over $\mathbb{Z}_4$. Any linear code C_{β} over $\mathbb{R}$ can be uniquely expressed as $C_{\beta}=vC_1\oplus(1-v)C_2$ [19].

The Gray map sends linear codes over $\mathbb{R}$ to linear codes over $\mathbb{Z}_4$: $\phi:\mathbb{R}\to\mathbb{Z}_4^2$ is given by $\phi(a+vb)=(a,a+b)$.

This map can be extended to $\mathbb{R}^{\beta}$ in natural way, for $c=(c_0,c_1,\cdots,c_{\beta-1})\in\mathbb{R}^{\beta}$ where $c_i=a_i+vb_i$, $i=0,1,\cdots,\beta-1$.

$$\begin{aligned}\phi:\mathbb{R}^{\beta}&\to\mathbb{Z}_4^{2\beta}\\ c&\mapsto(a(c),a(c)+b(c))\end{aligned}$$

where $a(c)=(a_0,a_1,\cdots,a_{\beta-1})$ and $b(c)=(b_0,b_1,\cdots,b_{\beta-1})$ and they are unique. For $a+vb\in\mathbb{R}$, $a,b\in\mathbb{Z}_4$. The following theories are studied in [4]:

Lemma 2.4 *The Gray map ϕ is a distance-preserving map from $\mathbb{R}^{\beta}$(Gray distance) to $\mathbb{Z}_4^{2\beta}$(Lee distance) and it is also $\mathbb{Z}_4$-linear.*

Lemma 2.5 *Let C_{β} be a $[\beta,M,d]$ linear code over $\mathbb{R}$, where β, M and d denote the length, the number of the codewords and the minimum Gray distance of C_{β}, respectively. Then $\phi(C_{\beta})$ is a $[2\beta,M,d]$ linear code over $\mathbb{Z}_4$.*

Theorem 2.6 *Let C_{β} be a linear code. Then $\phi(C_{\beta})^{\perp}=\phi(C_{\beta}^{\perp})$. Furthermore, if C_{β} is Euclidean self-dual code, so is $\phi(C_{\beta})$.*

Definition 2.7 The Gray map over $\mathbb{Z}_4\mathbb{R}$ defined as

$$\begin{aligned}\Phi\colon \mathbb{Z}_4\mathbb{R} &\to \mathbb{Z}_4^3\\ (e, a+vb) &\mapsto (e, \phi(a+vb)) = (e, a, a+b)\end{aligned}$$

The map Φ can be extended componentwise from $\mathbb{Z}_4^\alpha\mathbb{R}^\beta$ to $\mathbb{Z}_4^{\alpha+2\beta}$ as

$$\begin{aligned}&\Phi(e_0, e_1, \cdots, e_{\alpha-1}, a_0+vb_0, a_1+vb_1, \cdots, a_{\beta-1}+vb_{\beta-1})\\ &\quad= (e_0, e_1, \cdots, e_{\alpha-1}, \phi(a_0+vb_0), \phi(a_1+vb_1), \cdots, \phi(a_{\beta-1}+vb_{\beta-1}))\end{aligned}$$

where $(e_0, e_1, \cdots, e_{\alpha-1}) \in \mathbb{Z}_4^\alpha$ and $(a_0+vb_0, a_1+vb_1, \cdots, a_{\beta-1}+vb_{\beta-1}) \in \mathbb{R}^\beta$. If C is a $\mathbb{Z}_4\mathbb{R}$-linear code, then $\Phi(C)$ is also $\mathbb{Z}_4$-linear.

Let $(e, a+vb)$ be an element in $\mathbb{Z}_4\mathbb{R}$, where $e, a, b, \in \mathbb{Z}_4$. Then the Lee weight of $(e, a+vb)$ can be defined as

$$w_L((e, a+vb)) = w_H(e) + w_H(a) + w_H(a+b)$$

The Lee distance between two codewords c_1 and c_2 in $\mathbb{Z}_4^\alpha\mathbb{R}^\beta$ is defined as

$$d_L(c_1, c_2) = w_L(c_1 - c_2)$$

Proposition 2.8 *Φ is an $\mathbb{Z}_4$ linear distance preserving map from $\mathbb{Z}_4^\alpha\mathbb{R}^\beta$ (Lee distance) to $\mathbb{Z}_4^{\alpha+2\beta}$ (Hamming distance).*

Proof Let $c_1 = (e_0, e_1, \cdots, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-1})$ and $c_2 = (e_0', e_1', \cdots, e_{\alpha-1}', f_0,' f_1', \cdots, f_{\beta-1}')$ be two codewords in $\mathbb{Z}_4^\alpha\mathbb{R}^\beta$, where $f_i = a_i + vb_i$ and $f_i' = a_i' + vb_i'$ are in $\mathbb{R}^\beta$ for $i = 0, 1, \cdots, \beta-1$. So

$$\begin{aligned}\Phi(c_1+c_2) =&(e_0+e_0', e_1+e_1', \cdots, e_{\alpha-1}'+e_{\alpha-1}', a_0+a_0', a_1+a_1', \cdots, a_{\beta-1}+a_{\beta-1}',\\ &a_0+a_0'+b_0+b_0', a_1+a_1'+b_1+b_1', \cdots, a_{\beta-1}+a_{\beta-1}'+b_{\beta-1}+b_{\beta-1}')\\ =&(e_0, e_1, \cdots, e_{\alpha-1}, a_0, a_1, \cdots, a_{\beta-1}, a_0+b_0, a_1+b_1, \cdots, a_{\beta-1}+b_{\beta-1})+\\ &(e_0', e_1', \cdots, e_{\alpha-1}', a_0,' a_1,' \cdots, a_{\beta-1}', a_0'+b_0', b_1'+b_1', \cdots, a_{\beta-1}'+b_{\beta-1}')\\ =&\Phi(c_1)+\Phi(c_2)\end{aligned}$$

Moreover for $r \in \mathbb{Z}_4$, one gets

$$\begin{aligned}\Phi(rc_1) &= \Phi(re_0, re_1, \cdots, re_{\alpha-1}, rf_0, rf_1, \cdots, rf_{\beta-1})\\ &= (re_0, re_1, \cdots, re_{\alpha-1}, ra_0, ra_1, \cdots ra_{\beta-1}, r(a_0+b_0), r(a_1+b_1), \cdots, r(a_{\beta-1}+b_{\beta-1}))\\ &= r(e_0, e_1, \cdots, e_{\alpha-1}, a_0, a_1, \cdots a_{\beta-1}, (a_0+b_0), (a_1+b_1), \cdots, (a_{\beta-1}+b_{\beta-1}))\\ &= r\Phi(c_1)\end{aligned}$$

Thus the map Φ is an $\mathbb{Z}_4$-linear map.

Let $f = (a_0+vb_0, a_1+vb_1, \cdots, a_{\beta-1}+vb_{\beta-1})$, and $f' = (a_0'+vb_0', a_1'+vb_1', \cdots, a_{\beta-1}'+vb_{\beta-1}')$ are in $\mathbb{R}^\beta$. Then

$$\begin{aligned}\phi(f-f') = (a_0-a_0', a_1-a_1', \cdots, &a_{\beta-1}-a_{\beta-1}', a_0-a_0'+b_0-b_0', a_1-a_1'+b_1-b_1',\\ &\cdots, a_{\beta-1}-a_{\beta-1}'+b_{\beta-1}-b_{\beta-1}') = \phi(f)-\phi(f').\end{aligned}$$

Thus,

$$\begin{aligned}d_L(c_1, c_2) &= w_L(c_1-c_2)\\ &= w_H(e-e') + w_H(\phi(f-f'))\\ &= w_H(e-e') + w_H(\phi(f)-\phi(f'))\\ &= d_H(e, e') + d_H(\phi(f), \phi(f'))\\ &= d_H((e, \phi(f)), (e', \phi(f')))\\ &= d_H(\Phi(c_1), \Phi(c_2))\end{aligned}$$

This shows that Φ is a distance preserving map. □

Corollary 2.9 *If C is an $\mathbb{Z}_4\mathbb{R}$-linear code with parameters $(\alpha+\beta, M, d_L)$, then $\Phi(C)$ is a $\mathbb{Z}_4$-linear code with parameters $[\alpha+2\beta, log_4 M, d_L]$.*

3 Cyclic codes over $\mathbb{Z}_4\mathbb{R}$

This section is devoted to the structure of cyclic codes over $\mathbb{Z}_4\mathbb{R}$. We first give algebraic structure and the generator polynomials of cyclic codes over $\mathbb{R}$, then determine the generator polynomials of $\mathbb{Z}_4\mathbb{R}$-cyclic codes and separable codes of length $\alpha+\beta$.

Theorem 3.1 *[17] Let C_α be a cyclic code of length $\alpha \in \mathbb{N}$ over $\mathbb{Z}_4$, then $C_\alpha =< g(x) >=< f(x)[h(x)+2] >$, where $x^\alpha - 1 = f(x)k(x)h(x)$ in $\mathbb{Z}_4[x]$ and $f(x), k(x), h(x)$ are pairwise coprime polynomials. Moreover, $|C_\alpha| = 2^{2deg(k(x))+deg(h(x))}$.*

Theorem 3.2 *[14] Let $C_\beta = vC_1 \oplus (1-v)C_2$ be a cyclic code over $\mathbb{R}$ of length β. Then*

$$C_\beta =< vf_1(x)[h_1(x)+2] + (1-v)f_2(x)[h_2(x)+2] >$$

where $f_i(x)k_i(x)h_i(x) = x^\beta - 1$ and $C_i =< f_i(x)[h_i(x)+2] >$ over $\mathbb{Z}_4$ for $i = 1, 2$, respectively. Moreover, $|C_\beta| = 2^{\sum_{i=1}^{2} 2deg(k_i(x))+deg(h_i(x))}$.

Lemma 3.3 *Let $C_\alpha =< g(x) >$ be a submodule of $\mathbb{Z}_4[x]|(x^\alpha - 1)$. Then $g(x)$ is a divisor of $x^\alpha - 1$.*

Proof Let $g(x)$ be a polynomial of minimal degree in C_α. By the right divisor algorithm, there are two unique polynomials $p(x)$ and $r(x)$ in $\mathbb{Z}_4[x]$ such that

$$x^\alpha - 1 = p(x).g(x) + r(x)$$

where $deg(r(x)) < deg(g(x))$. Since $g(x)$ and $x^\alpha - 1 = 0$ are in C_α, we have $r(x) \in C$. But also $g(x)$ is of minimal degree in C_α, so $r(x) = 0$. Therefore, $g(x)$ is a divisor of $x^\alpha - 1$. □

Now, we determine the generator polynomias of $\mathbb{Z}_4\mathbb{R}$-cyclic codes by using the polynomials given in Theorem 3.2 and Lemma 3.3.

Theorem 3.4 *Let C be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha+\beta$. Then*

$$C =< (g(x), 0), (j(x), h(x)) >$$

where $g(x)|(x^\alpha - 1)$, $h(x) = vf_1(x)[h_1(x)+2] + (1-v)f_2(x)[h_2(x)+2]$ with $h(x)|(x^\beta - 1)$ and $f_i(x)k_i(x)h_i(x) = x^\beta - 1$, $j(x) \in \mathbb{Z}_4[x]$ for $i = 1, 2$. Moreover, $deg(j(x)) < deg(g(x))$.

Proof Since C and $\mathbb{R}_\beta$ are $\mathbb{R}[x]$-submodule of $\mathbb{R}_{\alpha,\beta}$, we can define a map

$$\theta : C \to \mathbb{R}_\beta$$
$$(z_1(x), z_2(x)) \mapsto z_2(x)$$

It is obvious that θ is an $\mathbb{R}[x]$-module homomorphism map. Moreover, one can check that $\theta(C)$ is an ideal of $\mathbb{R}_\beta$, i.e., $\theta(C)$ is a cyclic code of length β over $\mathbb{R}$. Thus, by Theorem 3.2, one gets $\theta(C) =< h(x) >=< vf_1(x)[h_1(x)+2] + (1-v)f_2(x)[h_2(x)+2] >$. We know that

$$Ker(\theta) = \{(z_1(x), 0) \in \mathbb{R}_{\alpha,\beta} \ : \ (z_1(x), 0) \in C\}$$

Define

$$I = \{z_1(x) \in \mathbb{Z}_4[x]|(x^\alpha - 1) \ : (z_1(x), 0) \in Ker(\theta)\}$$

Clearly, I is a $\mathbb{Z}_4[x]$-submodule of $\mathbb{Z}_4[x]|(x^\alpha - 1)$. By Lemma 3.3, $I =< g(x) >$ where $g(x)|(x^\alpha - 1)$. Thus, for any $(z_1(x), 0) \in Ker(\theta)$, $z_1(x) \in I$. So, for any polynomial $p(x) \in \mathbb{Z}_4[x]$ one gets $z_1(x) = p(x).g(x)$. This gives that

$$(z_1(x), 0) = p(x) \star (g(x), 0)$$

Therefore,

$$Ker(\theta) =< (g(x), 0) >$$

According to the first isomorphism theory of modules,

$$C/Ker(\theta) \cong \theta(C) =< h(x) >$$

Assume that $(j(x), h(x)) \in C$ with $\theta(j(x), h(x)) = h(x)$. So, any $\mathbb{Z}_4\mathbb{R}$-cyclic code can be generated by $(g(x), 0)$, $(j(x), h(x))$ as an $\mathbb{R}[x]$-submodule of $\mathbb{R}_{\alpha,\beta}$, where $g(x)|(x^\alpha - 1)$, $h(x) = vf_1(x)[h_1(x) + 2] + (1 - v)f_2(x)[h_2(x) + 2]$ with $h(x)|(x^\beta - 1)$ and $f_i(x)k_i(x)h_i(x) = x^\beta - 1$, $j(x) \in \mathbb{Z}_4[x]$ for $i = 1, 2$.

Now, we need to show that $deg(j(x)) < deg(g(x))$. Assume that $deg(j(x)) \geq deg(g(x))$ and $deg(j(x) - g(x)) = m \in \mathbb{N}$. Let

$$D =< (g(x), 0), (j(x) - x^m g(x), h(x)) >$$

Then, one gets $D \subseteq C$. Furthermore, we have

$$(j(x), h(x)) = (j(x) - x^m g(x), h(x)) + x^m \star (g(x), 0)$$

Thus, $C \subseteq D$ and so $C = D$. This yields a contradiction. Hence $deg(j(x)) < deg(g(x))$. □

Let C be a linear code over $\mathbb{Z}_4\mathbb{R}$ of length $\alpha+\beta$. Then C is called separable if $C = C'_\alpha \times C'_\beta$ while considering C'_α and C'_β as punctured codes of C by deleting the coordinates outside the α and β components, respectively.

Lemma 3.5 *Let $C =< (g(x), 0), (j(x), h(x)) >$ be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$. Then $C'_\alpha =< gcd(g(x), j(x)) >$ and $C'_\beta =< h(x) >$.*

Proof We only prove the case $C'_\alpha =< gcd(g(x), j(x)) >$. The proof process of $C'_\beta =< h(x) >$ follows similarly.

Let $p(x) \in C'_\alpha$, then there exists a polynomial $r(x) \in \mathbb{R}_\beta$ such that $(p(x), r(x)) \in C$. For some polynomials $a_1(x)$ and $a_2(x)$ in $\mathbb{R}[x]$, we have

$$(p(x), r(x)) = a_1(x) \star (g(x), 0) + a_2(x) \star (j(x), h(x))$$

So, one gets

$$p(x) = \delta(a_1(x))g(x) + \delta(a_2(x))j(x)$$

Thus, we have $gcd(g(x), j(x))|p(x)$ and so $p(x) \in< gcd(g(x), j(x)) >$. Hence $C'_\alpha \subseteq< gcd(g(x), j(x))$.

On the other hand, let $b_1(x)$ and $b_2(x)$ be two polynomials in $\mathbb{Z}_4[x]$. Then,

$$gcd(g(x), j(x)) = b_1(x)g(x) + b_2(x)j(x)$$

which gives that

$$(gcd(g(x), j(x)), b_2(x)h(x)) = b_1(x) \star (g(x), 0) + b_2(x) \star (j(x), h(x)) \subseteq C$$

Thus, $< gcd(g(x), j(x)) \subseteq C'_\alpha$. □

Lemma 3.6 *Let $C =< (g(x), 0), (j(x), h(x)) >$ be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$. Then $g(x)|j(x)$ if and only if $j(x) = 0$.*

Proof If $j(x) = 0$, then it is clear.
Assume that $g(x)|j(x)$, then $j(x) = a(x)g(x)$ for $a(x) \in \mathbb{Z}_4[x]$. Define

$$D =< g((x), 0), (0, h(x)) >$$

where $(0, h(x)) = (j(x), h(x)) - a(x) \star (g(x), 0) \in C$. Thus $D \subseteq C$. On the other hand, $(j(x), h(x)) = a(x) \star (g(x), 0) + (0, h(x)) \in D$ implies that $C \subseteq D$. Therefore, we can conclude that $j(x) = 0$. □

Theorem 3.7 *Let $C =< (g(x), 0), (j(x), h(x)) >$ be a $\mathbb{Z}_4\mathbb{R}$-cyclic code. Then the followings are equivalent:*

(1) C is a separable code
(2) $g(x)|j(x)$
(3) $C'_\alpha =< g(x) >$ and $C'_\beta =< h(x) >$
(4) $C =< (g(x), 0), (0, h(x)) >$

As a conclusion, we can deduce that $C_\alpha = C'_\alpha =< g(x) >$ and $C_\beta = C'_\beta =< h(x) >$. So we have the following theory:

Theorem 3.8 *Let $C = C_\alpha \times C_\beta$ be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$. Then C is a separable $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$ if and only if C_α, C_β are cyclic codes of length α and β over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively.*

Proof Let C be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$ and $(e_0, e_1, \cdots, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-1}) \in C$, where $(e_0, e_1, \cdots, e_{\alpha-1}) \in C_\alpha$ and $(f_0, f_1, \cdots, f_{\beta-1}) \in C_\beta$. By the hypothesis, we have $(e_{\alpha-1}, e_0, \cdots, e_{\alpha-2}, f_{\beta-1}, f_0, \cdots, f_{\beta-2}) \in C$. This gives that $(e_{\alpha-1}, e_0, \cdots, e_{\alpha-2}) \in C_\alpha$ and $(f_{\beta-1}, f_0, \cdots, f_{\beta-2}) \in C_\beta$. Thus C_α and C_β are cyclic codes of length α and β over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively.
On the other hand, assume that C_α, C_β are cyclic codes of length α and β over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively. Let $(e'_0, e'_1, \cdots, e'_{\alpha-1}) \in C_\alpha$ and $(f'_0, f'_1, \cdots, f'_{\beta-1}) \in C_\beta$, then $(e'_{\alpha-1}, e'_0, \cdots, e'_{\alpha-2}) \in C_\alpha$ and $(f'_{\beta-1}, f'_0, \cdots, f'_{\beta-2}) \in C_\beta$. Hence $(e'_{\alpha-1}, e'_0, \cdots, e'_{\alpha-2}, f'_{\beta-1}, f'_0, \cdots, f'_{\beta-2}) \in C_\alpha \times C_\beta = C$. This implies that C is a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$. □

Theorem 3.9 *Let $C = C_\alpha \times C_\beta$ be a $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha + \beta$. Then $C =< g(x) > \times < h(x) >$.*

Example 3.10 Let $\mathbb{R}_{\alpha,\beta} = \frac{\mathbb{Z}_4[x]}{x^\alpha - 1} \times \frac{\mathbb{R}[x]}{x^\beta - 1}$. Then

$$x^7 - 1 = (x-1)(x^3 + 3x^2 + 2x + 3)(x^3 + 2x^2 + x + 3)$$

Let $f(x) = x^3 + 2x^2 + x + 3$ and $h(x) = x^3 + 3x^2 + 2x + 3$. Then $C_\alpha =< (g(x) >=< f(x)[h(x) + 2] >$ is a cyclic code of length 7 over $\mathbb{Z}_4$ and $|C_\alpha| = 2^5$.

$$\begin{aligned} x^{21} - 1 =& (x-1)(x^2 + x + 1)(x^6 + 2x^5 + 3x^4 + 3x^2 + x + 1)(x^6 + x^5 + 3x^4 + 3x^2 + 2x + 1) \\ & (x^3 + 2x^2 + x + 3)(x^3 + 3x^2 + 2x + 3) \end{aligned}$$

Let $f_1(x) = f_2(x) = x^6 + 2x^5 + 3x^4 + 3x^2 + x + 1$ and $h_1(x) = h_2(x) = x^3 + 2x^2 + x + 3$. Then $C_i =< f_i(x)[h_i(x) + 2] >$ is a cyclic codes over $\mathbb{Z}_4$ for $i = 1, 2$. Thus C_β is a cyclic code of length 21 over $\mathbb{R}$ and $|C_\beta| = 2^{54}$.

Therefore, $C =< g(x) > \times < h(x) >$ is a separable $\mathbb{Z}_4\mathbb{R}$-cyclic code of length 28. Also, $|C| = 2^{59}$.

4 Cyclic DNA codes over $\mathbb{Z}_4\mathbb{R}$

In this chapter, we give our attention to the cyclic DNA codes over $\mathbb{Z}_4\mathbb{R}$. First of all, we examine the cyclic codes over $\mathbb{R}$, which satisfy the reversible and reversible-complement constraints. The theories over the ring $\mathbb{R}$ are inferred by [14]. However, in this study, some theories are considered from a different perspective.

Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ be a polynomial in $\mathbb{R}[x]$, with $a_n \neq 0$. The reciprocal polynomial of $f(x)$ is defined by

$$f^\star(x) = x^n f(x^{-1})$$

where $deg(f^\star(x)) \leq deg(f(x))$. If $a_0 \neq 0$, then $deg(f^\star(x)) = deg(f(x))$. Also, $f(x)$ is called self-reciprocal if $f^\star(x) = f(x)$.

Let $\mathbb{S}$ be a finite commutative ring and $\mathcal{C}$ be a linear code over $\mathbb{S}$ of length n. For any $s = (s_0, s_1, \cdots, s_{n-2}, s_{n-1}) \in \mathbb{S}$, the reverse of s is defined as $s^r = (s_{n-1}, s_{n-2}, \cdots, s_1, s_0)$, the complement of s is defined as $s^c = (s_0^c, s_1^c, \cdots, s_{n-2}^c, s_{n-1}^c)$, and the reverse-complement of s is defined as $s^{rc} = (s_{n-1}^c, s_{n-2}^c, \cdots, s_1^c, s_0^c)$. Note that the reverse-complement is also called the Watson-Crick complement (WCC).

Let $\mathcal{C}$ be a linear code over $\mathbb{S}$ of length n. For all $s \in \mathcal{C}$, if $s^r \in \mathcal{C}$, $s^c \in \mathcal{C}$ or $s^{rc} \in \mathcal{C}$, then $\mathcal{C}$ is called reversible code, complement code or reversible-complement code, respectively.

Definition 4.1 [18] Let $\mathcal{C}$ be a linear code over $\mathbb{S}$ of length n. Then $\mathcal{C}$ is called a cyclic DNA code if $\mathcal{C}$ is a cyclic and reversible-complement code with $s \neq s^{rc}$ for all $s \in \mathcal{C}$.

The folllowing two lemmas are from [9] and respective proofs are omitted for brevity.

Table 1 Identifying elements of the ring $\mathbb{R}$ with codons

Elements $r \in \mathbb{R}$	Gray images ϕ	DNA codons $\xi(r)$
0	(0, 0)	AA
1	(1, 1)	TT
2	(2, 2)	GG
3	(3, 3)	CC
v	(0, 1)	AT
2v	(0, 2)	AG
3v	(0, 3)	AC
1+v	(1, 2)	TG
1+2v	(1, 3)	TC
1+3v	(1, 0)	TA
2+v	(2, 3)	GC
2+2v	(2, 0)	GA
2+3v	(2, 1)	GT
3+v	(3, 0)	CA
3+2v	(3, 1)	CT
3+3v	(3, 2)	CG

Lemma 4.2 *Let $C_\alpha =< g(x) >$ be a cyclic code over $\mathbb{Z}_4$. Then C_α is reversible code if and only if $g(x)$ is self-reciprocal polynomial.*

Lemma 4.3 *Let $h_1(x), h_2(x)$ be any two polynomials in $\mathbb{R}$ with $deg(h_1(x)) \geq deg(h_2(x))$. Then*

(1) $[h_1(x)h_2(x)]^\star = h_1^\star(x)h_2^\star(x)$
(2) $[h_1(x) + h_2(x)]^\star = h_1^\star(x) + x^{deg(h_1(x))-deg(h_2(x))}h_2^\star(x)$

Lemma 4.4 *Let $C_\alpha =< g(x) >$ be a linear cyclic code of length α over $\mathbb{Z}_4$. Then C_α is complement code if and only if $g(x)$ is not divisible by $x - 1$.*

Proof Let $C_\alpha =< g(x) >$ be a cyclic code over $\mathbb{Z}_4$ and $e(x) = e_0 + e_1x + \cdots + e_{\alpha-1}x^{\alpha-1} \in C_\alpha$. So, we have $e^c(x) = e_0^c + e_1^c x + \cdots + e_{\alpha-1}^c x^{\alpha-1}$. Since $z + z^c = 1$ for any $z \in \mathbb{Z}_4$, one can obtain that $e(x) + e^c(x) = 1 + x + x^2 + \cdots + x^{\alpha-1}$. Since C_α is a linear and complement code, $e^c(x) \in C_\alpha$ if and only if $e(x) + e^c(x) = 1 + x + x^2 + \cdots + x^{\alpha-1} \in C_\alpha$ if and only if $(x^\alpha - 1)/(x - 1) \in C_\alpha$ if and only if $(x^\alpha - 1)/(x - 1) = j(x)g(x)$ if and only if $x^\alpha - 1 = (x - 1)j(x)g(x)$ for any polynomial $j(x) \in Z_4[x]$. Since α is odd, then $x^\alpha - 1$ factors uniquely into a product of distinct irreducible polynomials. Therefore, $e^c(x) \in C_\alpha$ if and only if $(x - 1)$ does not divide $g(x)$. □

By combining Lemma 4.2 and Lemma 4.4, we have the following useful lemma for reversible-complement codes over $\mathbb{Z}_4$.

Theorem 4.5 *Let $C_\alpha =< g(x) >$ be a cyclic code of length α over $\mathbb{Z}_4$. Then C_α is reversible-complement code if and only if $g(x)$ is self-reciprocal and $g(x)$ is not divisible by $x - 1$.*

Cyclic DNA codes over $\mathbb{Z}_4 + v\mathbb{Z}_4$ with $v^2 = v$ has 16 codons which are constructed by four nucleotides A, T, G, and C such that $P = \{AA, AT, AC, AG, TA, TT, TG, TC, CA, CT, CC, CG, GA, GT, GC, GG\}$. In the section 2, the Gray map, ϕ, of $\mathbb{R}$ is defined and by considering this map, we can construct one-to-one correspondence between $\mathbb{R}$ and DNA codons as follows:

Definition 4.6 Let C_β be a linear code of length β over $\mathbb{R}$ and $f = (f_0, f_1, \cdots, f_{\beta-1}) \in C_\beta$. By the Table 1, we define

$$\psi_\beta : C_\beta \to P^{2\beta}$$
$$(f_0, f_1, \cdots, f_{\beta-1}) \mapsto (\xi(f_0), \xi(f_1), \cdots \xi(f_{\beta-1}))$$

For instance, $(f_0, f_1) = (v, 1 + 3v)$ is mapped to $(\xi(v), \xi(1 + 3v)) = (ATTA)$.

Definition 4.7 Let C_β be a linear code of length β over $\mathbb{R}$. Then

(1) $\psi_\beta(C_\beta)$ is called a reversible DNA code if $\psi_\beta(f)^r \in \psi_\beta(C_\beta)$ for all $f \in C_\beta$.
(2) $\psi_\beta(C_\beta)$ is called a reversible-complement DNA code if $\psi_\beta(f)^{rc} \in \psi_\beta(C_\beta)$ for all $f \in C_\beta$.

The following theories introduce a relation between cyclic codes and reversible, reversible-complement codes over $\mathbb{R}$. Moreover, the proof of Theorem 4.8 was considered from a different perspective than Liu et al. [14].

Theorem 4.8 *Let $C_\beta = vC_1 \oplus (1-v)C_2$ be a cyclic code over $\mathbb{R}$ of lenght β. Then C_β is reversible code if and only if C_1 and C_2 are reversible codes over $\mathbb{Z}_4$*

Proof Let C_β be a reversible code over $\mathbb{R}$ and $f = (f_0, f_1, \cdots, f_{\beta-1}) \in C_\beta$ such that $f_i = v(a_i+b_i)+(1-v)a_i$ for $i = 0, 1, \cdots, \beta-1$ Let $a+b := (a_0+b_0, a_1+b_1, \cdots, a_{\beta-1}+b_{\beta-1}) \in C_1$ and $a := (a_0, a_1, \cdots a_{\beta-1}) \in C_2$. Then

$$\begin{aligned} f^r &= (f_{\beta-1}, f_{\beta-2}, \cdots, f_1, f_0) \\ &= (v(a_{\beta-1}+b_{\beta-1}) + (1-v)a_{\beta-1}, v(a_{\beta-2}+b_{\beta-2}) + (1-v)a_{\beta-2}, \\ &\cdots, v(a_1+b_1) + (1-v)a_1, v(a_0+b_0) + (1-v)a_0) \\ &= v(a_{\beta-1}+b_{\beta-1}, a_{\beta-2}+b_{\beta-2}, \cdots, a_1+b_1, a_0+b_0) + (1-v)(a_{\beta-1}, a_{\beta-2}, \cdots, a_1, a_0) \\ &= v(a+b)^r + (1-v)a^r \end{aligned}$$

By the hypothesis, we have $f^r = v(a+b)^r + (1-v)a^r \in C_\beta = C_1 \oplus C_2$. This implies that $(a+b)^r \in C_1$ and $a^r \in C_2$. Hence C_1 and C_2 are reversible over $\mathbb{Z}_4$.

On the other hand, assume that C_1 and C_2 be reversible codes over $\mathbb{Z}_4$ and let $a+b := (a_0+b_0, a_1+b_1, \cdots, a_{\beta-1}+b_{\beta-1}) \in C_1$ and $a := (a_0, a_1, \cdots a_{\beta-1}) \in C_2$. Then we have $(a+b)^r \in C_1$ and $a^r \in C_2$. Since $f^r = v(a+b)^r + (1-v)a^r \in C_\beta$, we can conclude that C_β is a reversible code over $\mathbb{R}$. □

Example 4.9 Let $\mathbb{R} = \mathbb{Z}_4 + v\mathbb{Z}_4$, where $v^2 = v$

$$x^7 - 1 = (x-1)(x^3+3x^2+2x+3)(3+x+2x^2+x^3)$$

Let $f_1(x) = f_2(x) = (x^3+3x^2+2x+3)(3+x+2x^2+x^3)$ and $h_1(x) = h_2(x) = x-1$. Then, by Theorem 3.2, $C_\beta =< vf_1(x)[h_1(x)+2]+(1-v)f_2(x)[h_2(x)+2] >$ is a cyclic code of length 7 over $\mathbb{R}$. Also by Lemma 4.2, we know that C_1 and C_2 are reversible codes over $\mathbb{Z}_4$ since both $f_i(x)[h_i(x)+2]$ are self-reciprocal polynomials for $i = 1, 2$. Therefore, by Theorem 4.8, C_β is a reversible code over $\mathbb{R}$.

The following two theories are studied in [14]:

Lemma 4.10 *Let $f = (f_0, f_1, \cdots f_{\beta-1})$ and $g = (g_0, g_1, \cdots g_{\beta-1})$ are in $\mathbb{R}^\beta$. Then we have*

(1) $f + f^c = (1, 1, \cdots, 1)$
(2) $(f+g)^c = f^c + g^c + 3(1, 1, \cdots, 1)$

Theorem 4.11 *[14] Let $C_\beta = vC_1 \oplus (1-v)C_2$ be a cyclic code over $\mathbb{R}$ of length β. Then C_β is reversible-complement code if and only if C_β is reversible and $(1, 1, \cdots, 1) \in C$.*

Example 4.12 Let $\mathbb{R} = \mathbb{Z}_4 + v\mathbb{Z}_4$, where $v^2 = v$

$$x^3 - 1 = (x-1)(x^2+x+1)$$

Let $f_1(x) = f_2(x) = x^2+x+1$ and $h_1(x) = h_2(x) = 1$. Then, by Theorem 3.2, $C_\beta =< vf_1(x)[h_1(x)+2]+(1-v)f_2(x)[h_2(x)+2] >$ is a cyclic code of length 3 over $\mathbb{R}$. Also by Lemma 4.2, we know that C_1 and C_2 are reversible codes over $\mathbb{Z}_4$ since both $f_i(x)[h_i(x)+2]$ are self-reciprocal polynomials for $i = 1, 2$. Therefore, by Theorem 4.8, C_β is a reversible code over $\mathbb{R}$. Furthermore, by Theorem 4.11, C_β is reversible-complement code over $\mathbb{R}$ since $(1, 1, 1) = (TTTTTT) \in C_\beta$. Hence, by Definition 4.1, we have C_β is a cyclic DNA code. The image of C_β under the map ϕ is a DNA code of length 6, size 16 and minimum Hamming distance 3. The corresponding DNA codewords obtained by using Table 1, are listed below.

Now, we are ready to study on reversible and reversible-complement constraints of $\mathbb{Z}_4\mathbb{R}$-cyclic codes.

Definition 4.13 For any $z = (e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in \mathbb{Z}_4\mathbb{R}$, the reverse of z is defined as $z^r = (e_{\alpha-1}, e_{\alpha-2}, \cdots, e_1, e_0, f_{\beta-1}, f_{\beta-2}, \cdots, f_1, f_0)$, the complement of z is defined as $z^c = (e_0^c, e_1^c, \cdots, e_{\alpha-2}^c, e_{\alpha-1}^c, f_0^c, f_1^c, \cdots, f_{\beta-2}^c, f_{\beta-1}^c)$ and the reverse-complement of z is defined as $z^{rc} = (e_{\alpha-1}^c, e_{\alpha-2}^c, \cdots, e_1^c, e_0^c, f_{\beta-1}^c, f_{\beta-2}^c, \cdots, f_1^c, f_0^c)$.

Table 2 DNA code of length 6 obtained from C_β

AAAAAA	TTTTTT	GGGGGG	CCCCCC
ATATAT	AGAGAG	ACACAC	TGTGTG
TCTCTC	TATATA	GCGCGC	GAGAGA
GTGTGT	CACACA	CTCTCT	CGCGCG

Definition 4.14 Let C be a $\mathbb{Z}_4\mathbb{R}$-linear code of length $\alpha+\beta$. Then C is called reversible code, complement code or reverse-complement code if $z^r \in C$, $z^c \in C$ or $z^{rc} \in C$ for any $z \in C$, respectively.

Theorem 4.15 *Let $C = C_\alpha \times C_\beta$ be a separable $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha+\beta$ where C_α and C_β are cyclic codes of length α and β over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively. Then C is reversible code if and only if C_α and C_β are reversible codes over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively.*

Proof Assume that C is reversible code and $z = (e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C$ where $(e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}) \in C_\alpha$ and $(f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C_\beta$. By the hypothesis, we have $z^r = (e_{\alpha-1}, e_{\alpha-2}, \cdots, e_1, e_0, f_{\beta-1}, f_{\beta-2}, \cdots, f_1, f_0) \in C$, which implies that $(e_{\alpha-1}, e_{\alpha-2}, \cdots, e_1, e_0) \in C_\alpha$ and $(f_{\beta-1}, f_{\beta-2}, \cdots, f_1, f_0) \in C_\beta$. So C_α and C_β are reversible codes over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively.

On the other hand, let $z = (e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C$, where $(e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}) \in C_\alpha$ and $(f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C_\beta$. By the converse hypothesis, we have $(e_{\alpha-1}, e_{\alpha-2}, \cdots, e_1, e_0) \in C_\alpha$ and $(f_{\beta-1}, f_{\beta-2}, \cdots, f_1, f_0) \in C_\beta$. This implies that $z^r = (e_{\alpha-1}, e_{\alpha-2}, \cdots, e_1, e_0, f_{\beta-1}, f_{\beta-2}, \cdots, f_1, f_0) \in C_\alpha \times C_\beta = C$ and C is a reversible code. □

Example 4.16 Let $C = C_\alpha \times C_\beta$ be a separable $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $(9, 7)$

$$x^9 - 1 = (x-1)(x^6 + x^3 + 1)(x^2 + x + 1) \in \mathbb{Z}_4[x]$$

Let $f(x) = x^2 + x + 1$ and $h(x) = x - 1$. Then by Theorem 3.1, $C_\alpha =< f(x)[h(x) + 2] >$ is a cyclic code of length 9 over $\mathbb{Z}_4$. As $f(x)[h(x) + 2]$ is a self-reciprocal, by Lemma 4.2, C_α is reversible over $\mathbb{Z}_4$. Moreover, the same processes follows for C_β as studied in Example 4.9, which is reversible over $\mathbb{R}$. Hence, by Theorem 4.15, we have $C = C_\alpha \times C_\beta$ is reversible code.

Theorem 4.17 *Let $C = C_\alpha \times C_\beta$ be a separable $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $\alpha+\beta$ where C_α and C_β are cyclic codes of length α and β over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively. Then C is reversible-complement code if and only if C_α and C_β are reversible-complement code over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively.*

Proof Assume that C is reversible-complement code and $z = (e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C$ where $(e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}) \in C_\alpha$ and $(f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C_\beta$. By the hypothesis, we have $z^{rc} = (e^c_{\alpha-1}, e^c_{\alpha-2}, \cdots, e^c_1, e^c_0, f^c_{\beta-1}, f^c_{\beta-2}, \cdots, f^c_1, f^c_0) \in C$, which implies that $(e^c_{\alpha-1}, e^c_{\alpha-2}, \cdots, e^c_1, e^c_0) \in C_\alpha$ and $(f^c_{\beta-1}, f^c_{\beta-2}, \cdots, f^c_1, f^c_0) \in C_\beta$. So C_α and C_β are reversible-complement codes over $\mathbb{Z}_4$ and $\mathbb{R}$, respectively.

On the other hand, let $z = (e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}, f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C$, where $(e_0, e_1, \cdots, e_{\alpha-2}, e_{\alpha-1}) \in C_\alpha$ and $(f_0, f_1, \cdots, f_{\beta-2}, f_{\beta-1}) \in C_\beta$. By the converse hypothesis, we have $(e^c_{\alpha-1}, e^c_{\alpha-2}, \cdots, e^c_1, e^c_0) \in C_\alpha$ and $(f^c_{\beta-1}, f^c_{\beta-2}, \cdots, f^c_1, f^c_0) \in C_\beta$. This implies that $z^{rc} = (e^c_{\alpha-1}, e^c_{\alpha-2}, \cdots, e^c_1, e^c_0, f^c_{\beta-1}, f^c_{\beta-2}, \cdots, f^c_1, f^c_0) \in C_\alpha \times C_\beta = C$ and C is a reversible-complement code. □

Example 4.18 Let $C = C_\alpha \times C_\beta$ be a separable $\mathbb{Z}_4\mathbb{R}$-cyclic code of length $(15, 3)$

$$x^{15} - 1 = (x-1)(x^4 + x^3 + x^2 + x + 1)(x^2 + x + 1)(x^4 + 2x^2 + 3x + 1)(x^4 + 3x^3 + 2x^2 + 1)$$

Let $f(x) = (x^4 + 2x^2 + 3x + 1)(x^4 + 3x^3 + 2x^2 + 1)$ and $h(x) = x - 1$. Then by Theorem 3.1, $C_\alpha =< f(x)[h(x) + 2] >$ is a cyclic code of length 5 over $\mathbb{Z}_4$. Also by Theorem 4.5, C_α is reversible-complement over $\mathbb{Z}_4$ since $f(x)[h(x) + 2]$ is a self-reciprocal and not divisible by $x - 1$. On the other hand, we have same procedures for the C_β code as given in Example 4.12, which is reversible complement over $\mathbb{R}$. As a conclusion, by Theorem 4.17, one gets $C = C_\alpha \times C_\beta$ is reversible-complement. Hence by the Definition 4.1, C is a cyclic DNA code.

5 Conclusion

In this paper, we studied the construction of cyclic DNA codes over the ring $\mathbb{Z}_4\mathbb{R}$ where $\mathbb{R} = \mathbb{Z}_4 + v\mathbb{Z}_4 = \{a + vb : a, b \in \mathbb{Z}_4\}$ with $v^2 = v$. In the begining of the study, we gave our attention to the algebraic structure of $\mathbb{R}$ and its linearity and generator polynomials. Then by using the different approach, we examined the reverse constraint and reverse-complement constraint of cyclic codes over $\mathbb{R}$. In the following of study, we especially considered cyclic DNA codes over the $\mathbb{Z}_4\mathbb{R}$ from the structure of separable codes. Furthermore, reverse constraint and reverse-complement constraint of $\mathbb{Z}_4\mathbb{R}$ -cyclic codes are discussed. To support our results, we demonstrated several examples and constructed cyclic DNA codes. In a future work, we will consider the algebraic structure of cyclic codes of even length and their relation to DNA codes over R. Also, structure of cyclic DNA codes over non-seperable codes will be open question for next studies.

References

1. Abualrub T, Ghrayeb A, Zeng XN Construction of cyclic codes over GF(4) for DNA computing. J. Frankl. Inst. 343: 448–457 (2006).
2. Adleman L Molecular computation of solutions to combinatorial problems. Science 266: 1021–1024 (1994).
3. Adleman LM, Rothmund PWK, Roweis S, Winfree E On applying molecular computation to the data encryption standard. J. Comput. Biol. 6, 53-63, (1999).
4. Bannd RK, Bhaintwal M Codes over $\mathbb{Z}_4 + v\mathbb{Z}_4$. "InternationalConferenceonAdvances in Computing, Communications and Informatics, pp. 422–427 (2014)"
5. Bayram A, Oztas E, Siap I Codes over $\mathbb{F}_4 + v\mathbb{F}_4$ and some DNA applications,. Des. Codes Cryptogr.80: 379–393 (2016).
6. Bennenni N, Guenda K, Mesnager S New DNA cyclic codes over rings. Adv. Math. Comp 11(1): (2017), 83-98.
7. Benenson Y, Gil B, Ben-Dor U., Adar R An autonomous molecular computer for logical control of gene expression. Nature 429: 423–429 (2004).
8. Dinh HQ, Pathak S, Upadhyay AK, Yamaka W New DNA codes from cyclic codes over mixed alphabers. Mathematics 8(11): 1-24 (2020).
9. Dinh HQ, Singh AK, Pattanayak S, Sriboonchitta S Cyclic DNA codes over the ring $\mathbb{F}_2 + u\mathbb{F}_2 + v\mathbb{F}_2 + uv\mathbb{F}_2 + v^2\mathbb{F}_2 + uv^2\mathbb{F}_2$. Des. Codes Cryptogr. 86: 1451-1467 (2018).
10. Gaborit P, King H Linear constructions for DNA codes. Theor. Comput. Sci 1(3): 99–113 (2005).
11. Gao J, Fu F, Gao Y Some classes of linear codes over $\mathbb{Z}_4 + v\mathbb{Z}_4$ and their applications to construct good and new $\mathbb{Z}_4$-linear codes. Appl. Algebra Engyg. Comm. Comput 28: 131-153 (2017).
12. Kari L, Gloor G, Yu S Using DNA to solve the bounded post correspondence problem. Theor. Comput. Sci 231: 193–203 (2000).
13. Lipton RJ DNA solution of hard computational problems. Science 2668: 542–545 (1995).
14. Liu J, Liu H Construction of cyclic DNA codes over the ring $\mathbb{Z}_4 + v\mathbb{Z}_4$. IEEE Access 8: 111200–111207 (2020).
15. Marathe A, Condon AE, Corn RM On combinatorial DNA word design. J. Comput. Biol 8: 201–219 (2001).
16. Oztas ES, Siap I Lifted polynomials over F16 and their applications to DNA codes, Filomat 27: 459–466 (2013).
17. Wan Z Quaternary codes. International Conference on Advances in Computing. Communications and Informatics 422-427 (2014).
18. Yildiz B, Siap I Cyclic DNA codes over the ring $\mathbb{F}_2[u]/<u^4 - 1>$ and applications to DNA codes. Comput. Math. Appl 63: 1169-1176 (2012).
19. Zhu S, Wang Y, Shi M Some results on cyclic codes over $\mathbb{F}_2 + v\mathbb{F}_2$. IEEE Trans. Inform. Theory 56: 1680-1684 (2010).

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.