DNA CYCLIC CODES OVER RINGS

NABIL BENNENNI AND KENZA GUENDA*
Faculty of Mathematics
University of Science and Technology
USTHB, Algeria

SIHEM MESNAGER
University of Paris VIII and XIII (Department of Mathematics)
and Telecom ParisTech, Paris, France

(Communicated by Steven Dougherty)

Abstract. In this paper we construct new DNA cyclic codes over rings. Firstly, we introduce a new family of DNA cyclic codes over the ring $R = \mathbb{F}_2[u]/(u^6)$. A direct link between the elements of such a ring and the 64 codons used in the amino acids of the living organisms is established. Using this correspondence we study the reverse-complement properties of our codes. We use the edit distance between the codewords which is an important combinatorial notion for the DNA strands. Next, we define the Lee weight, the Gray map over the ring $R$ as well as the binary image of the DNA cyclic codes allowing the transfer of studying DNA codes into studying binary codes. Secondly, we introduce another new family of DNA skew cyclic codes constructed over the ring $\tilde{R} = \mathbb{F}_2 + v\mathbb{F}_2 = \{0, 1, v, v + 1\}$, where $v^2 = v$. The codes obtained are cyclic reverse-complement over the ring $\tilde{R}$. Further we find their binary images and construct some explicit examples of such codes.

1. Introduction

DNA computing combines genetic data analysis with the computational science in order to tackle computationally difficult problems. This new field started by Leonard Adleman [3]. Adleman solved a hard (NP-complete) computational problem by DNA molecule in a test tube. Deoxyribonucleic acid (DNA) contains the genetic program for the biological development of life. DNA is formed by strands linked together and twisted in the shape of a double helix. Each strand is a sequence of four possible nucleotides, two purines; adenine ($A$), guanine ($G$) and two pyrimidines; thymine ($T$) and cytosine ($C$). Hybridization, known as base pairing, occurs when a strand binds to another strand, forming a double strand of DNA. The strands are linked following the Watson-Crick model, every ($A$) is linked with a ($T$), and every ($C$) with a ($G$), and vice versa. We denote by $\hat{X}$ the complement of $X$ defined as follows, $\hat{A} = T, \hat{T} = A, \hat{G} = C$ and $\hat{C} = G$ (for instance if $x = (AGCTAC)$ then its complement $\hat{x} = (TCGATG)$). The pairing is done in the opposite direction and the reverse order. Several authors have contributed to provide constructions of cyclic DNA codes over fixed rings. In [2][17], the authors gave DNA cyclic codes over finite field with four elements. Further, Siap et al. have
studied in [19] cyclic DNA codes over the ring $\mathbb{F}_2[u]/(u^2 - 1)$ using the deletion distance. More recently, Guenda et al. have studied in [10] cyclic DNA codes of arbitrary length over the ring $\mathbb{F}_2[u]/(u^4 - 1)$.

In this paper we consider the design of DNA codes of length $n$ over the ring $R = \mathbb{F}_2[u]/(u^6)$. The ring $R$ is a finite chain ring with 64 elements. With four possible bases, the three nucleotides can give $4^3 = 64$ different possibilities called codons. These combinations are used to specify the 20 different amino acids used in the living organisms [4]. To this end, we construct a one-to-one correspondence between the elements of $R$ and the 64 codons over the alphabet $\{A, G, C, T\}^3$ by the map $\phi$ such a correspondence is useful for the error protection (see [11]). We shall give the structure of the cyclic reversible-complement DNA codes over the ring $R$ with designed edit distance $D$. We also give some upper and lower bounds on $D$. The properties of our codes are the most required properties for DNA computing; namely, our codes are reversible-complement and also cyclic which is a very important property, since it can reduce the complexity of the dynamic programming when testing the strand from the unwanted secondary structures [14]. The edit distance is an important combinatorial notion for the DNA strands. It can be used for the correction of the insertion, deletion, substitution errors between the codewords. This is not the case for the Hamming, deletion, and the additive stem distances. We shall also define a Lee weight and a Gray map over $R$. The image of our DNA codes under the mapping are quasi-cyclic codes of index 6 and of length $6n$ over the alphabet $\{A, G, C, T\}$. There are several advantages in using codes over the ring $R$. We list some of them below:

1. There exists a one-to-one correspondence between the codons and the elements of the ring $R$. This is presented in section 2.
2. A code over $R$ contains more codewords than codes of the same length over fields.
3. The factorization of $x^n - 1$ is the same over the field $\mathbb{F}_2$ but in general is not the same over other rings. This fact simplifies the construction of cyclic codes over $R$.
4. The structure of the cyclic codes of any length over $R$ is well-known [9], whereas little is know concerning the structure of cyclic codes of any length over rings.
5. The cyclic character of the DNA strands is desired because the genetic code should represent an equilibrium status [18]. Another advantage of cyclic codes, as indicated by Milenkovic and Kashyap [14], is that the complexity of the dynamic programming algorithm for testing DNA codes for secondary structure will be less for cyclic codes.
6. The binary image of the cyclic codes over $R$ under our Gray map are linear quasi-cyclic codes.

In the second part of the paper, we study the DNA skew cyclic codes over the ring $\tilde{R} = \mathbb{F}_2 + v\mathbb{F}_2 = \{0, 1, v, v + 1\}$, where $v^2 = v$. The obtained codes are reverse-complement. Further we give the binary images of the DNA skew cyclic codes and provide some examples. The advantage of studying the reversible DNA code in skew polynomial rings is to exhibit several factorizations. Therefore, many reverse-complement DNA code could be obtained in a skew polynomial ring (which is not the case in a commutative ring).
This paper is organized as follows. In Section 2 we present some preliminaries results as well as the one-to-one correspondence between the elements of the ring \( R = \mathbb{F}_2[u]/(u^6) \) and their codons. Next, we give the algebraic structure of the cyclic codes over \( R = \mathbb{F}_2[u]/(u^6) \) and we study the DNA cyclic codes and reverse-complement of these codes. Moreover, we define the Lee weight related to such codes and give the binary image of the DNA cyclic code. Some explicit examples of such codes are presented. In Section 3, we describe a DNA skew cyclic codes over \( \bar{R} = \mathbb{F}_2 + v\mathbb{F}_2 = \{0, 1, v, v+1\} \) where \( v^2 = v \), study its property of being reverse-complement and provide explicit examples of such codes given with their minimum Hamming distances.

2. DNA cyclic codes over \( R = \mathbb{F}_2[u]/(u^6) \)

2.1. Preliminaries. The ring considered in this section is

\[
R = \mathbb{F}_2[u]/(u^6) = \{q_0 + a_1 u + a_2 u^2 + a_3 u^3 + a_4 u^4 + a_5 u^5; a_i \in \mathbb{F}_2, u^6 = 0\}.
\]

It is a commutative ring with 64 elements. It is a principal local ideal ring with maximal ideal \( (u) \). The ideals of \( R \) satisfy the following inclusions

\[
\langle 0 \rangle = \langle u^6 \rangle \subseteq \langle u^5 \rangle \subseteq \langle u^4 \rangle \subseteq \langle u^3 \rangle \subseteq \langle u^2 \rangle \subseteq \langle u \rangle \subseteq \langle R \rangle.
\]

Since the ring \( R \) is of the cardinality 64, then we can construct a one-to-one correspondence between the elements of \( R \) and the 64 codons over the alphabet \( \{A, G, C, T\}^3 \) by the map \( \phi \), this is given in Table 1. A simple verification shows that for all \( x \in R \), we have

\[
(1) \quad x + \hat{x} = u^5 + u^4 + u^3 + u^2 + u + 1.
\]

Now, since \( R^n \) is an \( R \)-module, a linear code over \( R \) of length \( n \) is a submodule \( C \) of \( R^n \). An \((n, k)\) linear block code of dimensions \( n = ml \), is called quasi-cyclic if every cyclic shift of a codeword by \( l \) symbol yields another codeword. For \( x \in R^n \), denote the number of the component of \( x \) equal to \( a_i \) by \( n_{a_i}(x) \). The Hamming weight of \( x \) is \( w_H(x) = \sum_{i=0}^{n-1} n_{a_i}(x) \), where \( a_i \in R^* \). The Hamming distance \( d_H(x, y) \) between the vector \( x \) and \( y \) equals \( w_H(x - y) \). Let \( x = x_0 x_1 \ldots x_{n-1} \) be a vector in \( R^n \). The reverse of \( x \) is defined as \( x^r = x_{n-1} x_{n-2} \ldots x_1 x_0 \), the complement of \( x \) is \( \hat{x} = \hat{x}_0 \hat{x}_1 \ldots \hat{x}_{n-1} \), and also called the Watson-Crick complement (WCC), the reverse-complement is defined as \( x^{rc} = \hat{x}_{n-1} \hat{x}_{n-2} \ldots \hat{x}_1 \hat{x}_0 \). A code \( C \) is said to be

| Table 1. Identifying codons with the elements of the ring \( R \) |
|---|
| CCC | u^5 + u^4 + u^3 + u^2 + u + 1 |
| GGA | u^5 + u^4 + u + u^2 + u |
| GCC | u^5 + u^4 + u^3 + u^2 + 1 |
| CGG | u^5 + u^4 + u^3 + u + 1 |
| GAG | u^5 + u^4 + u^3 + u |
| CCG | u^5 + u^4 + u + 1 |
| CAG | u^5 + u^4 |
| ATG | u^5 + u^4 + u^2 + 1 |
| TAC | u^5 + u^4 + u |
| ACT | u^5 + u^4 + u^3 + u |
| GCA | u^5 + u^4 + u^3 + u |
| GCC | u^5 + u^4 + u^3 + u |
| CGC | u^5 + u^4 + u^3 |
| CCG | u^5 + u^4 + u^2 |
| TGG | u^5 + u^4 + u |
| AGG | u^5 + u^4 |
| GGA | u^5 + u^4 + u |
| CCG | u^5 + u^4 + u |
| GCC | u^5 + u^4 + u |
| CAG | u^5 + u^4 + u |
| GCA | u^5 + u^4 + u|
| TAC | u^5 + u^4 + u |
| ACT | u^5 + u^4 + u |
| GCA | u^5 + u^4 + u |
| TGG | u^5 + u^4 + u |
| AGG | u^5 + u^4 + u |
| GGA | u^5 + u^4 + u |
| CCG | u^5 + u^4 + u |
| GCC | u^5 + u^4 + u |
| CAG | u^5 + u^4 + u |
| GCA | u^5 + u^4 + u |
| TAC | u^5 + u^4 + u |
| ACT | u^5 + u^4 + u |
reversible if for any \( x \in C \), we have \( x^r \in C \). Moreover, \( C \) is said to be reverse-complement if for any \( x \in C \), we have \( x^{rc} \in C \).

The edit distance is the minimum number of the operations (insertion, substitution and deletion) required to transform one string into another one. The edit distance can be defined as in [16].

Let \( A \) and \( B \) be finite sets of distinct symbols and let \( x^t \in A^t \) denotes an arbitrary string of length \( t \) over \( A \). Then \( x^t_i \) denotes the substring of \( x^t \) that begins at position \( i \) and ends at position \( j \). The edit distance is characterized by a triple \( (A, B, c) \) consisting of the finite sets \( A \) and \( B \), and the primitive function \( c : E \to \mathbb{R}_+ \), where \( \mathbb{R}_+ \) is the set of nonnegative reals, \( E = E_s \cup E_d \cup E_i \) is the set of primitive edit operations, \( E_s = A \ast B \) is the set of substitutions, \( E_d = A \ast E \) is the set of deletions and \( E_i = E \times B \) is the set of insertions. Each triple \( (A, B, c) \) induces a distance function \( d_c : A^* \times B^* \to \mathbb{R}_+ \) which maps a string \( x^t \) to a nonnegative value, defined as follows.

**Definition 2.1.** The edit distance \( d_c(x^t, y^v) \) between two strings \( x^t \in A^t \) and \( y^v \in B^v \) is defined recursively as

\[
d_c(x^t, y^v) = \min \left\{ \begin{array}{ll}
c(x_t, y_v) + d_c(x^t-1, y^{v-1}); \\
c(x_t, \epsilon) + d_c(x^t-1, y^v); \\
c(\epsilon, y_v) + d_c(x^t, y^{v-1});
\end{array} \right.
\]

where \( d_c(\epsilon, \epsilon) = 0 \) if \( \epsilon \) denotes the empty string of length \( n \).

It is easy to check the following bounds on the edit distance \( d_c \).

**Proposition 1.** Assume that \( X \) and \( Y \) are two strings in \( R^n \). Then the following holds:

(i) \( d_c(\phi(X), \phi(Y)) \leq n \);

(ii) \( d_c(\phi(X), \phi(Y)) \leq d_H(\phi(X), \phi(Y)) \);

(iii) \( d_c(\phi(X), \phi(Y)) = d_c(\phi(Y), \phi(X)) \).

2.2. Cyclic codes over \( R = \mathbb{F}_2[u]/(u^6) \). In this subsection we give the algebraic structure of the cyclic code of arbitrary length over \( R \). We start by giving the definition of cyclic code over this ring. Let \( C \) be a code over \( R \) of length \( n \). A codeword \( (c_0, c_1, \cdots, c_{n-1}) \) of \( C \) is viewed as a polynomial \( c_0 + c_1 x + \cdots + c_{n-1} x^{n-1} \) in \( R[x] \). Let \( \tau \) be the cyclic shift acting on the codewords of \( C \) in the following way:

\[
\tau(c_0, c_1, \cdots, c_{n-1}) = (c_{n-1}, c_0, c_1, \cdots, c_{n-2}).
\]

Recall that linear code \( C \) is cyclic if \( C \) is invariant under permutation \( \tau : c(x) \mapsto xc(x) \) (mod \( x^n - 1 \)).

The following theorem is a particular result of [9] which gives the structure of the cyclic codes of arbitrary lengths.

**Theorem 2.2.** Let \( C \) be a cyclic code of arbitrary length \( n \) over the ring \( R \).

(i) Assume \( n \) is odd. Then there exist six polynomials \( f_0, f_1, f_2, f_3, f_4, f_5 \) over \( R \), such that \( f_5 \mid f_4 \mid f_3 \mid f_2 \mid f_1 \mid f_0 \) and \( C = \langle f_0, uf_1, u^2f_2, u^3f_3, u^4f_4, u^5f_5 \rangle \).

(ii) Assume \( n = m2^n \) such that \( \gcd(m, 2) = 1 \). Then the cyclic codes of length \( n \) over \( R \) are the ideals generated by \( \langle f_0, uf_1, u^2f_2, u^3f_3, u^4f_4, u^5f_5 \rangle \), where \( f_i \mid f_0 \) and \( f_0 \) is divisor of \( x^n - 1 \) in \( \mathbb{F}_2 \).

Let denote by \( K \) the field \( R/(u) \). We have the following canonical ring morphism

\[
- : R[x] \to K[x]; \, f \mapsto \overline{f} = f \mod u.
\]
Thus we have the following tower of linear codes over \( R \).

Proof. (1) For \( d \) odd length. Then the minimum edit distance \( d \) satisfies the following inequalities.

(i) \( d_\ell(C) = \min\{d_\ell(Tor_i(C))\} \leq \min\{d_H(Tor_i(C))\} \), where \( i = 0, \cdots, 5 \);
(ii) \( d_\ell(C) \leq \min\{\deg(f_i)\} + 1 \), where \( i = 0, \cdots, 5 \);
(iii) \( d_\ell(C) \leq n - \text{rank}(C) + 1 \).

Proof. From [17] Lemma 5.1 and Proposition[11] we have \( d_\ell(C) = \min\{d_\ell(Tor_i(C))\} \leq \min\{d_H(Tor_i(C))\} \) for every \( i \in \{0, \cdots, 5\} \). Assertion (ii) comes from the fact that the code \( Tor_i(C) \) and \( Tor_0(C) \) are binary cyclic codes satisfying \( \langle f_i \rangle \subset C \). The dimension of \( Tor_i(C) \) is \( n - \deg(f_i) \). By the well-known Singleton bound, we have \( d_\ell(C) \leq \min\{\deg(f_i)\} + 1 \). Assertion (iii) follows from Proposition [17] using again the Singleton bound.

2.3. DNA cyclic codes. Now, we introduce a DNA cyclic code by constructing a DNA cyclic code over rings. A \([3n, d]\)-DNA cyclic code. Set \( \langle x \rangle := \langle x^n - 1 \rangle/(x - 1) \) and \( \alpha(u) := u^5 + u^4 + u^3 + u^2 + u + 1 \).

Definition 2.4. Let \( 1 \leq D \leq 3n - 1 \) be a positive real number. Then a cyclic code \( C \) of length \( n \) over \( R \) is called an \([n, D]\) DNA cyclic code if the following conditions hold:

(i) \( C \) is cyclic code, i.e., \( C \) is an ideal in \( R_n = R[x]/(x^n - 1) \);
(ii) for any codeword \( x \in C \), we have \( (x)^{rc} \neq (x) \) and \( (x)^{rc} \in C \);
(iii) \( d_\ell(x, y) \leq D \) for any \( x, y \in C \).

Condition (ii) given in Definition 2.4 shows that the defined DNA cyclic codes are reverse-complement cyclic codes.

Definition 2.5. Let \( f(x) \in R[x] \), denote \( f(x)^* = x^{\deg(f)} f(\frac{1}{x}) \) the reciprocal polynomial of \( f(x) \). The polynomial \( f \) is said to be self-reciprocal if \( f(x) = f^*(x) \).

The following statement can be obtained straightforwardly.

Lemma 2.6. Let \( f(x) \) and \( g(x) \) be a polynomials in \( R[x] \) with \( \deg(f(x)) \geq \deg(g(x)) \). Then the following conditions hold:

(i) \( (f(x)g(x))^* = f(x)^*g(x)^* \);
(ii) \( (f(x) + g(x))^* = f(x)^* + x^{\deg(f) - \deg(g)} g(x)^* \).
Theorem 2.7. Let \( C \) be a cyclic code of odd length \( n \) over \( R \) and assume that \( C \) is reverse-complement. Then we have:

(i) \( C \) contains all the codewords of the form \( \alpha(u)I(x) \):
(ii) \( C = \langle f_0, u^2f_1, u^3f_3, u^4f_4, u^5f_5 \rangle \) where \( f_i \) is self-reciprocal for every \( i \in \{0, \cdots , 5\} \).

Proof. (i) Since \( C \) is linear, \( (0, \cdots , 0) \in C \). Also \( C \) is revers-complement, so that \( (0, \cdots , 0)^{rc} \in C \). Then we have \( (0, \cdots , 0)^{rc} = (\alpha(u), \cdots , \alpha(u)) = \alpha(u)(x^n - 1)/(x - 1) \in C \).

(ii) Let us show that \( f_i^*(x) = f_i(x) \) for \( i \in \{0, \cdots , 5\} \).

Set \( f_0(x) = a_0 + a_1x + \cdots + a_{m-1}x^{m-1} + a_mx^m \) where \( f_0/(x^n - 1) \in F_2[x] \). One can assume that \( a_0 = a_m = 1 \). So that \( f_0(x) = 1 + a_1 + \cdots + a_{m-1}x^{m-1} + x^m \). Suppose that \( f_0(x) \) corresponds to the vector \( (1, a_1, \cdots , 1, 0, 0, \cdots , 0) \) and the reverse-complement of 0 in \( R \) is \( \alpha(u) \) then \( f_0^*(x) = \alpha(u)(1 + x + \cdots + x^{n-2} + (\alpha(u) + 1)x^{n-m-2} + a_{m-1}x^{n-m} + \cdots + a_1x^{n-2} + (\alpha(u) + 1)x^{n-1}) \in C \). Now, since \( C \) is a linear code, we get

\[
f_0(x)^{rc} + \alpha(u)(x^n - 1)/(x - 1) \in C.
\]

This implies that

\[
x^{n-m-1} + (\hat{a}_{m-1} + \alpha(u))x^{n-m} + \cdots + (\hat{a}_1 + \alpha(u))x^{n-2} + x^{n-1} = x^{n-m-1}[1 + (\hat{a}_{m-1} + \alpha(u))x + \cdots + (\hat{a}_1 + \alpha(u))x^{m-1} + x^m] \in C.
\]

Multiplying the last polynomial by \( x^{m+1} \) in \( R[x]/(x^n - 1) \), we obtain

\[
1 + (\hat{a}_{m-1} + \alpha(u)) + \cdots + (\hat{a}_1 + \alpha(u))x^{m-1} + x^m \in C.
\]

By Equation [5] we see that \( \hat{a} + \alpha(u) = a \). Therefore, we obtain

\[
f_0^*(x) = 1 + a_{m-1}x + \cdots + a_1x^{m-1} + x^m \in C.
\]

Consequently, we have

\[
f_i^*(x) = f_0k_0 + u_1f_1k_1 + \cdots + u_5f_5k_5,
\]

where \( f_i \) and \( k_i \) are all in \( F_2[x] \). Multiplying both sides of this equality by \( u^5 \) gives

\[
u^5f_0^*(x) = u^5k_0(x)f_0(x).
\]

Now, since \( f_0^*(x), f_0(x) \in F_2[x] \) have the same degree, leading coefficient and constant term, one necessary have \( k_0(x) = 1 \). Consequently, \( f_0(x) \) is self-reciprocal. The same argument can be used for \( f_1, f_2, f_3, f_4 \) and \( f_5 \) as well. \( \square \)

In the following, we are interested in providing sufficient conditions for a given code \( C \) to be reverse-complement.

Theorem 2.8. Assume that \( C = \langle f_0, u^1f_1, u^2f_2, u^3f_3, u^4f_4, u^5f_5 \rangle \) is a cyclic code of odd length \( n \) over \( R \) with \( f_1 | f_2 | f_3 | f_4 | f_5 \) \( f_0 | x^n - 1 \in F_2[x] \). If \( \alpha(u)I(x) \in C \) and \( f_i(x) \) are self-reciprocal, then \( C \) is a reverse-complement code.

Proof. Let \( c(x) \) be a codeword in \( C \), we have to prove that \( c(x)^{rc} \in C \). Since \( C = \langle f_0, u^1f_1, u^2f_2, u^3f_3, u^4f_4, u^5f_5 \rangle \) there exist \( \alpha_i(x) \in R[x] \) \(( \alpha_i(x) \in R[x] \) \(( \alpha_i(x) \in R[x] \) (i \in \{1, \cdots , 5\}) such that

\[
c(x) = f_0\alpha_0 + u_1f_1\alpha_1 + u^2f_2\alpha_2 + u^3f_3\alpha_3 + u^4f_4\alpha_4 + u^5f_5\alpha_5.
\]
Applying the reciprocal and using first Lemma 2.6 and next the fact that \(f_0(x), f_1(x), f_2(x), f_3(x), f_4(x)\) and \(f_5(x)\) are self-reciprocal, we obtain
\[
c^*(x) = (f_0^*)^* + (u f_1^* a_1)^* x^{m_1} + (u^2 f_2^* a_2)^* x^{m_2}
+ (u^3 f_3^* a_3)^* x^{m_3} + (u^4 f_4^* a_4)^* x^{m_4} + (u^5 f_5^* a_5)^* x^{m_5},
\]
proving that \(c(x)^*\) is in \(C\). Since \(C\) is cyclic, \(x^{n-t-1} c(x) = c_0 x^{n-t-1} + c_1 x n - t + \cdots + c_t x^{n-1} \in C\). It was also assumed that \(\alpha (u) + (\alpha (u) x + \cdots + \alpha (u) x^{n-1}) \in C\), which leads to
\[
\alpha (u) + \alpha (u) x + \cdots + \alpha (u) x^{n-1} + c_0 x^{n-t-1} + c_1 x n - t + \cdots + c_t x^{n-1} \in C.
\]
This is equal to
\[
\alpha (u) + \alpha (u) x + \cdots + \alpha (u) x^{n-t-2} + (\alpha (u) + c_0) x^{n-t-1} + \cdots + (\alpha (u) + c_t) x^{n-1}
= \alpha (u) + \alpha (u) x + \cdots + \alpha (u) x^{n-t-1} + \cdots + c_0 x^{n-t-1} + \cdots + c_t x^{n-1},
\]
which is precisely \((c^*(x))^* = c(x)^r \in C\).

Using similar arguments as in Theorem 2.8 one can prove the following statement.

**Theorem 2.9.** Assume that \(C = \langle f_0, u f_1, u^2 f_2, u^3 f_3, u^4 f_4, u^5 f_5 \rangle\) be a cyclic code of even length \(n = m 2^s\) over \(R\) such that \(f_i | f_0 \in \mathbb{F}_2[x]\). If \(\alpha (u) \mathbb{I} (x) \in C\) and \(f_i (x) \ (i \in \{1, \cdots , 5\})\) are self-reciprocal. Then \(C\) is a reverse-complement code.

**Corollary 1.** Let \(C\) be a cyclic code of length \(n = m 2^s\), \(s \geq 0\). If \(\alpha (u) \mathbb{I} (x) \in C\) and there exists an integer \(i\) such that
\[
2^i \equiv -1 \pmod{m}.
\]
Then \(C\) is a reverse-complement code.

**Proof.** The proof is similar to the proof of Corollary 4.13 in [10].

**Definition 2.10.** For a cyclic code \(C = \langle f_0, u f_1, u^2 f_2, u^3 f_3, u^4 f_4, u^5 f_5 \rangle\), we define the sub-code \(C_{u^2}\) consisting of all codewords in \(C\) that are a multiple of \(u^2\).

**Lemma 2.11.** Let \(C\) be a cyclic code \(C = \langle f_0, u f_1, u^2 f_2, u^3 f_3, u^4 f_4, u^5 f_5 \rangle\) of odd length, then we have: (i)
\[
\phi(u^2 R) = \{GGG, AGT, CGT, TGT, GAT, CAG, TAC, ATC, GTC, TCC, CTC, ACC, CTC, TCC, GGT, CAT\},
\]
\[
\phi(u^3 R) = \{GGG, TGT, CAG, TAC, CTC, GCC, AGT, TTAC\},
\]
\[
\phi(u^4 R) = \{GGG, TAC, CTC, TGT\};
\]
(ii) if \(C = \langle f_0, u f_1, u^2 f_2, u^3 f_3, u^4 f_4, u^5 f_5 \rangle\) is the cyclic code of odd length \(n\) over \(R\), then \(C_{u^2} = \langle u^2 f_5 \rangle\) and \(\phi(C_{u^2})\) is over the alphabet \(\phi(u^2 R)\).

**Proof.** The part (i) is obtained by a simple calculation. For the part (ii), assume that \(C = \langle f_0, u f_1, u^2 f_2, u^3 f_3, u^4 f_4, u^5 f_5 \rangle\). Since \(f_5 | f_4 | f_3 | f_2 | f_1 | f_0 | x^n - 1\), then we obtain \(\langle u^2 f_5 \rangle \subset C_{u^2}\). Conversely, assume that \(c(x) \in C\) such that \(c(x) = \alpha_0 (x) f_0 (x) + u \alpha_1 (x) f_1 (x) + u^2 \alpha_2 (x) f_2 (x) + u^3 \alpha_3 (x) f_3 (x) + u^4 \alpha_4 (x) f_4 (x) + u^5 \alpha_5 (x) f_5 (x)\) for all \(\alpha \in \mathbb{F}_2[x]\). If \(c(x)\) is a multiple of \(u^2\) then \(x^n - 1\) divides \(\alpha_0 (x) f_0 (x)\) and \(x^n - 1\) divides \(\alpha_1 (x) f_1 (x)\). Hence, \(c(x) = u^2 \alpha_2 (x) f_2 (x) + u^3 \alpha_3 (x) f_3 (x) + u^4 \alpha_4 (x) f_4 (x) + u^5 \alpha_5 (x) f_5 (x)\). Therefore, \(C_{u^2} \subset \langle u^2 f_5 \rangle\). Consequently \(C_{u^2} = \langle u^2 f_5 \rangle\), which completes the proof.
Remark 1. The DNA cyclic codes which are obtained in the Lemma 2.11 are stable across the error in the DNA strands by the usage of the codons, see [5]. Any codeword of sub-code $\phi(C_{u^2})$ over $\phi(u^2R)$ contains the nucleotide $C$ and $G$. This is an interesting thermodynamic property of the DNA strand. For its importance, we send the reader to [10].

Example 2.12. Let us consider the following polynomial in $F_2[x]$, (4) \[ x^7 - 1 = (x - 1)(x^3 + x + 1)(x^3 + x^2 + 1) = f_0f_1f_2. \]

In Table 2, we present the associate DNA cyclic codes of length 7 given with their corresponding size and their minimal Hamming distance. Table 3 and Table 4 present all codewords of DNA cyclic code associate to $C = \langle u^4f_0f_1 \rangle$ and to $C = \langle f_1f_2 \rangle$, respectively.

Table 2. DNA cyclic codes of length 7

| Code C   | Size of C | $d_H$ |
|----------|-----------|-------|
| $\langle u^2f_0 \rangle$ | 4096     | 2     |
| $\langle u^2f_1 \rangle$ | 256      | 3     |
| $\langle u^2f_2 \rangle$ | 256      | 3     |
| $\langle u^2f_1f_2 \rangle$ | 4        | 7     |
| $\langle u^2f_0f_1 \rangle$ | 64       | 4     |
| $\langle u^2f_0f_2 \rangle$ | 64       | 4     |
| $\langle u^4f_0f_1 \rangle$ | 64       | 4     |

Table 3. A DNA Cyclic Code associate to $C = \langle u^4f_0f_1 \rangle$ given in (4)

Example 2.13. We have that $x^{17} - 1 = (x - 1)(x^8 + x^5 + x^4 + x^3 + 1)(x^8 + x^7 + x^6 + x^4 + x^2 + x + 1) = f_0f_1f_2f_3$ in $F_2[x]$. In Table 5 we present the DNA cyclic codes associate to $C = \langle f_0, u_1, u_2, u_3, u_4, u_5 \rangle$. 

Advances in Mathematics of Communications Volume 11, No. 1 (2017), 83–98
2.4. Binary Image of DNA Codes. In this Section we will define a Gray map which allows us to translate the properties of the suitable DNA codes for DNA computing to the binary cases. Table 7 gives a binary image of the DNA cyclic code of length 7 given by Table 2. Any element \( c \) computing to the binary cases. Table 7 gives a binary image of the DNA cyclic which allows us to translate the properties of the suitable DNA codes for DNA from \( R \) to \( \mathbb{F}_2 \) where the Lee weight over the ring \( R \) is defined as follows

\[
\varphi : R^n \rightarrow \mathbb{F}_2^n
\]

where \( a_i \in \mathbb{F}_2, 0 \leq i \leq 5 \). We have for example \( \varphi(1 + u) = (1, 1, 0, 0, 0, 0) \). We define the Lee weight over the ring \( R \) by

\[
w_{Lee}(a_0 + a_1 u + a_2 u^2 + a_3 u^3 + a_4 u^4 + a_5 u^5) = \sum_{i=0}^{5} a_i.
\]

The Lee distance \( d_L(x, y) \) between the vector \( x \) and \( y \) is \( w_{Lee}(x - y) \). According to the definition of the Gray map, it is easy to check that the image of a linear code over \( R \) by \( \varphi \) is a binary linear code. We can obtain the binary image of the DNA code by the map \( \varphi \) and the map \( \psi \). In Table 6 we give the binary image of the codons. The binary image of DNA code resolved the problem of construction

Table 4. DNA Cyclic associate to \( C = \langle f_1 f_2 \rangle \) given in [4]

| The Code \( C \) | Size of the code \( C \) |
|-----------------|-------------------------|
| \( \langle u^6 f_1, u^2 f_2, u^3 f_3, u^4 f_4, u^5 f_5 \rangle \) | 11,258,999,068,426,240 |
| \( \langle u^3 f_3, u^4 f_4 \rangle \) | 512 |
| \( \langle u^3 f_4, u^5 f_5 \rangle \) | 46,116,860,184,273,879,040 |
| \( \langle u^3 f_5, u^5 f_5 \rangle \) | 85,899,345,920 |

Table 5. DNA cyclic codes associate to \( C = \langle f_0, u f_1, u^2 f_2, u^3 f_3, u^4 f_4, u^5 f_5 \rangle \)
of DNA codes with some properties, see [14]. The following property of the binary

Table 6. Binary image of the codons given by Table 1

| Codon   | Binary Image | Codon   | Binary Image |
|---------|--------------|---------|--------------|
| GGG     | 000000       | CCC     | 111111       |
|         |              |         |              |
| GGA     | 011111       | CCA     | 110110       |
|         |              |         |              |
| GGG     | 101111       | CCT     | 010000       |
|         |              |         |              |
| GGC     | 100111       | CCA     | 110110       |
|         |              |         |              |
| AGG     | 010011       | TCA     | 011000       |
|         |              |         |              |
| CGG     | 110011       | CGA     | 001100       |
|         |              |         |              |
| CGA     | 011011       | TGT     | 101000       |
|         |              |         |              |
| CGC     | 100111       | TCG     | 011000       |
|         |              |         |              |
| CGT     | 000111       | TCA     | 111000       |
|         |              |         |              |
| TGG     | 110011       | ACC     | 001100       |
|         |              |         |              |
| TGA     | 010011       | GCC     | 110100       |
|         |              |         |              |
| TGC     | 100011       | ACC     | 001100       |
|         |              |         |              |
| TGG     | 000011       | ACA     | 111100       |

image of the DNA codes comes from the definition.

Lemma 2.14. The Gray map $\varphi$ is a linear weight preserving $(R^n, \text{Lee distance}) \rightarrow (F_2^{6n}, \text{Hamming distance})$.

Further, if $C$ is a DNA cyclic code of length $n$ over $R$, then $\varphi(C)$ is a binary DNA quasi-cyclic code of length $6n$ over $F_2$ and of index 6.

Proof. Let $C$ be a DNA cyclic code of length $n$ over $R$. Hence $\varphi(C)$ is a set of length $6n$ over the alphabet $F_2$ which is a quasi-cyclic code of index 6. It is easy to verify that the Gray map is a linear weight preserving.

Remark 2. The usual Gray map from the ring $R = \{0, 1, u, u+1\}$ to $F_2$, have the same isometric properties.

Table 7. A binary image of DNA cyclic codes of length 7 given Table 2

| Code   | Length of $\varphi(C)$ | $d_H(\varphi(C))$ | Size of the Code $\varphi(C)$ |
|--------|------------------------|------------------|-------------------------------|
| $\langle u^2 f_0 \rangle$ | 42 | 12 | 4096 |
| $\langle u^2 f_1 \rangle$ | 42 | 18 | 256 |
| $\langle u^2 f_2 \rangle$ | 42 | 18 | 256 |

Remark 3. The codes of rows 2 and 3 given by Table 7 are optimal according to [20].

3. DNA skew cyclic codes over $\tilde{R} = F_2 + vF_2$

3.1. Notation and preliminaries. The ring considered in this section is the non-commutative ring $\tilde{R} = F_2 + vF_2$ where $\theta$ is an automorphism of $\tilde{R}$. The structure of the latter ring depends on the element of the commutative ring $\tilde{R} = F_2 + vF_2 = \{0, 1, v, v+1\}$, where $v^2 = v$ and the automorphism $\theta$ on $\tilde{R}$, defined by $\theta(0) = 0$, $\theta(1) = 1$, $\theta(v) = v + 1$, $\theta(v + 1) = v$. Note that $\theta^2(a) = \theta(\theta(a)) = a$ for all $a \in \tilde{R}$. This implies that $\theta$ is a ring automorphism of order 2. The skew polynomial ring $\tilde{R}[x; \theta]$ is the set of polynomials $\tilde{R}[x; \theta] = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mid a_i \in \tilde{R}\}$ endowed with the usual addition of polynomials and the multiplication $*$ (which
is not commutative) is defined by the basic rule \((ax) \ast (bx) = a\theta^i(b)x^{i+j}\) (the distributive and the associative laws occur).

There is a one-to-one map \(\psi\) between the elements of \(\tilde{R}\) and the DNA nucleotide base \(\{A, T, C, G\}\) given by \(0 \mapsto G, v \mapsto C, v + 1 \mapsto T\) and \(1 \mapsto A\). A simple verification shows that for all \(x \in \tilde{R}\), we have

\[
\theta(x) + \theta(\dot{x}) = v + 1.
\]

In the following, we only consider codes with even lengths.

**Definition 3.1.** Let \(\tilde{R} = \mathbb{F}_2 + v\mathbb{F}_2 = \{0, 1, v, v + 1\}\) be a ring where \(v^2 = v\) and the automorphism \(\theta\) defined previously. A subset \(\tilde{C}\) of \(\tilde{R}^n\) is called a skew cyclic code \((\theta\text{-cyclic code})\) of length \(n\) if the two following conditions hold

1. \(\tilde{C}\) is a \(\tilde{R}\)-submodule of \(\tilde{R}^n\);
2. if \(c = (c_0, c_1, \cdots, c_{n-1}) \in \tilde{C}\) then \((\theta(c_{n-1}), \theta(c_0), \cdots, \theta(c_{n-2})) \in \tilde{C}\).

The ring \(\tilde{R}_n = \tilde{R}[x; \theta]/(x^n - 1)\) denotes the quotient ring of \(\tilde{R}[x; \theta]\) by the (left) ideal \((x^n - 1)\). Let \(f(x) \in \tilde{R}_n\) and \(r(x) \in \tilde{R}[x; \theta]\), we define the multiplication from the left as follows.

\[
r(x) \ast (f(x) + (x^n - 1)) = r(x) \ast f(x) + (x^n - 1)
\]

for any \(r(x) \in \tilde{R}[x; \theta]\). Define a map as follows

\[
\xi : \tilde{R}^n \to \tilde{R}[x; \theta]/(x^n - 1)
\]

\[
(c_0, c_1, \cdots, c_{n-1}) \to c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1}.
\]

Clearly, \(\xi\) is an \(\tilde{R}\)-module isomorphism map which implies that each element \((c_0 + c_1 \cdots + c_{n-1})\) of \(\tilde{R}^n\) can be identified with the polynomial \(c(x) = c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1}\) of \(\tilde{R}_n\).

**Lemma 3.2 (\([1]\) Lemma 1)).** If \(n\) is even, and \(x^n - 1 = g(x) \ast f(x)\) in \(\tilde{R}[x; \theta]\), then we have:

\[
x^n - 1 = g(x) \ast f(x) = f(x) \ast g(x).
\]

The following proposition gives the structure of the skew cyclic codes over \(\tilde{R}_n\).

**Proposition 2 (\([1]\) Corollary 5)].** Let \(\tilde{C}\) be a skew cyclic code in \(\tilde{R}_n\). Then

1. If a polynomial \(g(x)\) of least degree in \(\tilde{C}\) is a monic then \(\tilde{C} = (g(x))\), where \(g(x)\) is (skew) right divisor of \(x^n - 1\).
2. If \(\tilde{C}\) contains some monic polynomials but no polynomial \(f(x)\) of least degree in \(\tilde{C}\) is monic, then \(\tilde{C} = (f(x), g(x))\), where \(g(x)\) is a monic polynomial of least degree in \(\tilde{C}\) and \(f(x) = vf_1(x)\) or \(f(x) = (v + 1)f_1(x)\) for some binary polynomial \(f_1(x)\).
3. If \(\tilde{C}\) does not contain any monic polynomials. Then \(\tilde{C} = (f(x))\) where \(f(x) = vf_1(x)\) or \(f(x) = (v + 1)f_1(x)\) and \(f_1(x)\) is a binary polynomial that divides \(x^n - 1\).

Now, we are interesting in constructions of \([n, d]\)-DNA skew cyclic codes. To this end, we start by defining such codes.

**Definition 3.3.** Let \(1 \leq d \leq n - 1\) be a positive real number. A skew cyclic code \(\hat{C}\) over \(\tilde{R}\) is said to be a \([n, d]\)-DNA skew cyclic code if the following conditions hold.

1. \(\hat{C}\) is a skew cyclic code, that is, \(\hat{C}\) is a \(\tilde{R}\)-submodule of \(\tilde{R}_n\);
2. for any codeword \(X \in \hat{C}\): \((X)^{rc} \neq (X)\) and \((X)^{rc} \in (C)\);
3. \(d_H(X, Y) \leq d\) for any \(X, Y \in C\).
3.2. The reverse-complement DNA skew cyclic codes over $\tilde{R}$. In this subsection, we give conditions on the existence of the reverse-complement cyclic codes of an even length $n$ over the ring $\tilde{R}$. In Table 8 we present all codewords of the DNA skew cyclic code of length 10 and minimal Hamming distance 2.

Let $v = (a_0, a_1, \ldots, a_{n-2}, a_{n-1})$ be a vector in $\tilde{R}_n$, the reverse of the vector $v$ is $v^r = (a_{n-1}, a_{n-2}, \ldots, a_1, a_0)$. Let $f(x)$ be the polynomial corresponding of the vector $v$ such that $f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$. To get the polynomial corresponding of the vector $v^r$ in $R[x; \theta]$, we multiply the right of the polynomial $f(x^{-1})$ by $x^{n-1}$ leading to $f(x^{-1})x^{n-1} = a_0x^{n-1} + a_1\theta(1)x^{n-2} + \cdots + a_{n-2}\theta^{n-2}(1)x + a_{n-1}\theta^{n-1}(1) = a_{n-1} + a_{n-2}x + \cdots + a_1x^{n-2} + a_0x^{n-1}$. The polynomial corresponding of the vector $v^r$ denoted by $f^*(x)$.

**Definition 3.4.** Let $f(x)^* = f(x^{-1}) * x^{\deg(f)}$ be the reciprocal polynomial of a given $f(x)$ in $\tilde{R}[x; \theta]$. Then the polynomial $f$ is called self-reciprocal if $f$ coincides with $f^*$.

**Example 3.5.** Let $f$ be a polynomial in $\tilde{R}[x; \theta]$ given by $f(x) = x^3 + vx^2 + (v + 1)x + v$. The polynomial $f$ represents the DNA sequence $X(ACTC)$. We get the reverse of the sequence $X$ via $f^*(x)$

\[
\begin{align*}
  f^*(x) &= f(x^{-1})x^3 \\
  &= \theta^3(1) + v\theta^2(1)x + (v + 1)\theta(1)x^2 + v\theta^0(1)x^3 \\
  &= vx^3 + (v + 1)x^2 + vx + 1.
\end{align*}
\]

The reverse of the DNA sequence of $X$ is given by $(CTCA)$.

Notice that the definition of reciprocal polynomial over $\tilde{R}[x, \theta]$ is being different from the one defined over a commutative ring. Indeed, in the non-commutative ring $\tilde{R}[x, \theta]$, we use the right multiplication over the automorphism $\theta$ and the multiplication over $\tilde{R}[x, \theta]$.

**Lemma 3.6.** Let $f(x)$ and $g(x)$ be polynomials in $\tilde{R}[x, \theta]$ with $\deg(f(x)) \geq \deg(g(x))$. Then the following assertions hold.

(i) $(f(x)g(x))^* = f(x)^*g(x)^*$;

(ii) $(f(x) + g(x))^* = f(x)^* + g(x)x^{\deg(f) - \deg(g)}$.

**Proof.** Let us prove Assertion (i). One have $f(x) = \sum_{i=0}^{n}a_ix^i$ and $g(x) = \sum_{j=0}^{p}b_jx^j$, with $\deg(f) \geq \deg(g)$. Therefore, $f(x)g(x) = \sum_{k=0}^{n+p} \sum_{i=0}^{k}a_i\theta^i(b_{k-i})x^k$. From Definition 3.2 $(f(x)g(x))^* = (\sum_{k=0}^{n+p} \sum_{i=0}^{k}a_i\theta^i(b_{k-i})x^{-k})x^{n+p}$. Thus $(f(x)g(x))^* = \sum_{k=0}^{n+p} \sum_{i=0}^{k}a_i\theta^i(b_{k-i})x^{n+p-k}$. Again from Definition 3.2 we have

\[
\begin{align*}
  f(x)^* &= \sum_{i=0}^{n}a_i\theta^i(1)x^{n-i} = \sum_{i=0}^{n}a_ix^{n-i} \\
  g(x)^* &= \sum_{j=0}^{p}b_j\theta^j(1)x^{p-j} = \sum_{j=0}^{p}b_jx^{p-j}. \quad \text{Consequently we have}
\end{align*}
\]

\[
\begin{align*}
  f(x)^*g(x)^* &= \sum_{k=0}^{n+p} \sum_{i=0}^{k}a_i\theta^i(b_{k-i})x^{n+p-k}.
\end{align*}
\]

The result follows.
Now let us prove Assertion (ii). From the Definition \[3.2\] we have
\[
(f(x) + g(x))^* = (f + g)^*(x) = ((f + g)(x^{-1}))x^{\deg(f)}
\]
\[
= (f(x^{-1}) + g(x^{-1}))x^{\deg(f)} = (f(x^{-1})x^{\deg(f)} + g(x^{-1})x^{\deg(f)})
\]
\[
= (f^*(x) + g(x^{-1})x^{\deg(f)}) = f^*(x) + g(x^{-1})x^{\deg(g)x^{\deg(f)} - \deg(g)}
\]
\[
= f^*(x) + g^*(x)x^{\deg(f) - \deg(g)},
\]
which completes the proof. \[\square\]

In the following we are interested in providing necessary conditions for \(\tilde{C}\) to be a reverse-complement code.

**Theorem 3.7.** Let \(\tilde{C} = \langle f(x) \rangle\) be a skew cyclic code in \(\tilde{R}_n\), where \(f(x)\) is monic polynomial of minimal degree. If \(\tilde{C}\) is reverse-complement then the polynomial \(f(x)\) is self-reciprocal and \(v(x^n - 1)/(x - 1) \in \tilde{C}\).

**Proof.** Let \(\tilde{C} = \langle f(x) \rangle\) be a skew cyclic code over \(\tilde{R}\), where \(f(x)\) is monic polynomial of minimal degree in \(\tilde{C}\). We know that \((0, 0, \cdots, 0) \in \tilde{C}\), since \(\tilde{C}\) is reverse-complement then \((0, 0, \cdots, 0)^{rc} \in \tilde{C}\), i.e.; \((0, v, \cdots, v) \in \tilde{C}\), this vector correspond of the polynomial \(v + vx + \cdots + vx^{n-1} = v(x^n - 1)/(x - 1) \in \tilde{C}\). We have that \(f(x)\) is monic polynomial of minimal degree in \(\tilde{C}\), where \(f(x) = 1 + a_1x + \cdots + x^t\), the vector correspond to the polynomial \(f(x)\) is \((1, a_1, \cdots, 0, 0, \cdots, 0)\), since \(\tilde{C}\) is reverse-complement and linear, then \((1, a_1, \cdots, 0, 0, \cdots, 0)^{rc} \in \tilde{C}\), i.e.,
\[
f^{rc}(x) = v + vx + \cdots + vx^{n-t-2} + (v + 1)x^{n-t-1}
\]
\[
+ a_{t-1}x^{n-t} + \cdots + a_1x^{n-2} + vx^{n-1}
\]
\[
= f^{rc}(x) + v(x^n - 1)/(x - 1) \in \tilde{C}.
\]
This implies that
\[
x^{n-t-1} + (a_{t-1} + v)x^{n-t} + \cdots + (a_1 + v)x^{n-2} + x^{n-1} \in \tilde{C}.
\]
Multiplying on the right by \(x^{t+1-n}\), we obtain
\[
(1 + (a_{t-1} + v)\theta(1)x + \cdots + (a_1 + v)\theta^{t-1}(1)x^{t-1} + \theta^{t}(1)x^t)x^{t-n-1} \in \tilde{C}.
\]
Hence, \((1 + (a_{t-1} + v)x + \cdots + (a_1 + v)x^{t-1} + x^t) \in \tilde{C}\), which implies (thanks to Equation \[6\]) that \(f^*(x) = 1 + a_{t-1}x + \cdots + x^t \in \tilde{C}\). Since \(\tilde{C} = \langle f(x) \rangle\), there exists \(q(x) \in R[x, \theta]\) such that \(f^*(x) = q(x)f(x)\), one necessary have \(q(x) = 1\), that is \(f^*(x) = f(x)\). \[\square\]

**Theorem 3.8.** Let \(\tilde{C} = \langle vf_1(x) \rangle\) be a skew cyclic code in \(\tilde{R}_n\), where \(f_1(x)\) is a monic binary polynomial of lowest degree with \(f_1(x)/(x^n - 1)\). If \(\tilde{C}\) is a reverse-complement code then \(f_1(x)\) is self-reciprocal.

**Proof.** Let \(f_1(x) = 1 + a_1x + a_2x + \cdots + x^r\) be a binary polynomial. The vector corresponds to \(f_1(x)\) is \(v = (1, a_1, \cdots, a_r, 1, 0, 0, 0, \cdots, 0)\). Hence \(v^{rc} = (\hat{0}, \hat{0}, \hat{0}, \cdots, 0, 1, \hat{a}_r, \cdots, \hat{a}_1)\). These vectors correspond of the polynomial
\[
f_1^{rc}(x) = v + vx + \cdots + vx^{n-r-2} + (v + 1)x^{n-r-1}
\]
\[
+ \hat{a}_r x^{n-r} + \cdots + \hat{a}_1 x^{n-2} + (v + 1)x^{n-1}
\]
\[
= f_1^{rc} + v(x^n - 1)(x - 1).
\]
Since \( \hat{C} \) is a linear code, then \( f_1^c + v(x^n - 1)(x - 1) \in \hat{C} \). Therefore
\[
x^{n-r-1} + (\hat{a}_{r-1} + v)x^{n-1} + \cdots + (\hat{a}_1 + v)x^{n-2} + x^{n-1} \in \hat{C},
\]
multiplying by \( x^{-n+r+1} \), we obtain
\[
1 + (\hat{a}_{r-1} + v)\theta(1)x^1 + \cdots + (\hat{a}_1 + v)\theta^{r-1}(1)x^{r-1} + 1\theta^r(1)x^r \in \hat{C}.
\]
Then
\[
1 + (\hat{a}_{r-1} + v)x^1 + \cdots + (\hat{a}_1 + v)x^{r-1} + x^r \in \hat{C}.
\]
By Equation (5), we obtain \( f_1^*(x) = 1 + (a_{r-1}x^1 + \cdots + a_1x^{r-1} + x^r) \in \hat{C} \) by Corollary 2, we have \( v f_1^*(x) = v f_1(x)q(x) \), one necessary have \( q(x) = 1 \). Then \( f_1^*(x) = f_1(x) \).

\[\textbf{Theorem 3.9.}\]
Let \( \hat{C} = (f(x), g(x)) \) be a skew cyclic code in \( \hat{R}_n \), where \( f(x) \) is a polynomial of minimal degree in \( \hat{C} \) and is not a monic polynomial, \( g(x) \) is a polynomial of least degree among the monic polynomials in \( \hat{C} \). If \( \hat{C} \) is a reverse-complement code then \( f(x) \) and \( g(x) \) are self-reciprocal.

\[\textbf{Proof.}\]
The proof is similar to the proof of Theorem 3.7 and of Theorem 3.8.

In the following, we provide sufficient conditions for \( \hat{C} \) being reverse-complement.

\[\textbf{Theorem 3.10.}\]
Let \( \hat{C} = (f(x)) \) be a skew cyclic codes in \( \hat{R}_n \), where \( f(x) \) is monic polynomial of the degree minimal in \( \hat{C} \). If \( v(x^n - 1)/(x - 1) \in \hat{C} \) and \( f(x) \) is self-reciprocal then \( \hat{C} \) is reverse-complement.

\[\textbf{Proof.}\]
Let \( f(x) = 1 + a_1x + a_2x^2 + \cdots + a_{r-1}x^{r-1} + x^r \) be a monic polynomial of the degree minimal in \( \hat{C} \) and \( c(x) \in \hat{C} \), we have \( c(x) = q(x)f(x) \) where \( q(x) \in \hat{R}[x, \theta] \). \( c(x)^* = (q(x)f(x))^* \), by the Lemma 3.6 we have \( c(x)^* = q(x)^*f(x)^* \), since \( f(x) \) is self-reciprocal then \( c(x)^* = q(x)^*f(x) \in \hat{C} \) for all \( c(x) \in \hat{C} \). Recall that we have
\[
v + vx + \cdots + vx^{n-1} \in \hat{C}.
\]
Now, let \( c(x) = c_0 + c_1x + c_2x + \cdots + c_i x^i \), we multiply the right polynomial \( c(x) \) by \( x^{n-t-1} \) we obtain \( c(x) \cdot x^{n-t-1} = c_0 + c_1\theta(1)x + c_2\theta^2(1)x^2 + \cdots + c_i\theta^i(1)x^i \), then
\[
c(x) \cdot x^{n-t-1} = c_0x^{n-t-1} + c_1x^{n-t} + \cdots + c_i x^{n-1} \in \hat{C}.
\]
Combining (7) and (8) we obtain
\[
(v + vx + \cdots + vx^{n-t-2} + (c_0 + v)x^{n-t-1} + \cdots + (c_i + v)x^{n-1}) \in \hat{C},
\]
leading to the following equality (using Equation 5) we have that \( c_i + v = \hat{c}_i \), Then we obtain
\[
v + vx + \cdots + vx^{n-t-2} + \hat{c}_0x^{n-t-1} + \hat{c}_1x^{n-t} + \cdots + \hat{c}_t x^{n-2} + \hat{c}_t x^{n-1} = (c(x))^{rc}.
\]
Therefore, \( (c^{*}(x))^{rc} = c(x)^{rc} \in \hat{C} \).

Using similar arguments as those used in the proof of Theorem 3.10, we can prove the two following statements.

\[\textbf{Theorem 3.11.}\]
Let \( C = (v f_1(x)) \) be a skew cyclic code in \( \hat{R}_n \), where \( f_1(x) \) is a monic binary polynomial of lowest degree with \( f_1(x)|(x^n - 1) \). If \( v(x^n - 1)/(x - 1) \in \hat{C} \) and \( f_1(x) \) is self-reciprocal, then \( \hat{C} \) is reverse-complement.
Theorem 3.12. Let $C = (f(x), g(x))$ be a skew cyclic codes in $\tilde{R}_n$, where $f(x)$ is a polynomial of degree minimal in $\tilde{C}$ and is not monic polynomial. Let $g(x)$ be a polynomial of least degree among monic polynomial in $\tilde{C}$. If $v(x^n - 1)/(x - 1) \in \tilde{C}$ and $f(x)$ and $g(x)$ are self-reciprocal, then $\tilde{C}$ is reverse-complement.

3.3. Binary image of DNA skew cyclic codes. Recall that the Gray map $\varphi$ from $F_2 + vF_2$ to $F_2$, is defined as follows: for each element of $F_2 + vF_2$ expressed as $a + vb$, where $a, b \in F_2$ maps $\varphi(a + vb) = (a + b, a)$, that is, $0 \mapsto (0, 0)$, $1 \mapsto (1, 1)$, $v + 1 \mapsto (0, 1)$, $v \mapsto (1, 0)$. The linearity of $\varphi$ comes straightforwardly from the definition of the Gray map.

We can obtain the binary image of the DNA code from the maps $\varphi$ and $\psi$ as well as the DNA alphabet onto the set of length 2 binary word given by $G \mapsto (0, 0)$, $A \mapsto (1, 1)$, $T \mapsto (0, 1)$, $C \mapsto (1, 0)$.

We have the following property of binary image of the DNA skew cyclic code.

Corollary 2. The map $R \mapsto F_2^n$ is distance preserving linear isometry, hence if $\tilde{C}$ is DNA skew cyclic code over $R$, then $\varphi(\tilde{C})$ is a DNA skew quasi-cyclic code of length $2n$ and of index 2.

Proof. The proof is similar to the one of Lemma 2.14 \hfill $\square$

Table 8. DNA skew cyclic code of length 10 and minimal distance 2

| GGGGGGGGGG | CCCCCCCCCC | GGGGGGGGGC | GGGGGGGGCG |
| GGTTTTTTTT | CCAAAACCCC | GGGGGGGCAG | GGGGGGGGCAG |
| GGTTAAACCC | CCAAAATTTG | GGGGGCCCAG | GGGGGGCAG |
| GCCCAAACTT | CCAAAACGGG | GGGGCCCAGG | GGGGGCAG |
| GTAAAAACCGG | CAGTATTCAG | GGGCATGCCC | GGGCATGCCC |
| GTAACCATGTC | TATGATGATC | GGGCATTGAC | GGGCATTGAC |
| GATCCATCGC | CATGCCATGC | GGGCCATCAG | GGGCCATCAG |
| GATCCCATGG | TATGGATTGG | GGGCCATTGG | GGGCCATTGG |
| GATTTTGCTT | TAAAACCCCA | GGGCTACGGC | GGGCTACGGC |
| GATTTTACCC | TAAATACCGG | GGGCTACGGC | GGGCTACGGC |
| GTTCCAAGCC | CACCACTTGG | GGGCTAGTTC | GGGCTAGTTC |
| GTTGGGTTTG | CAACTTACCG | GGGCTAGTTC | GGGCTAGTTC |

4. Conclusion

The structure of DNA is used as a model for constructing good error correcting codes and conversely error correcting codes that enjoy similar properties with DNA structure are also used to understand DNA itself. Some attention has been made on DNA cyclic codes but currently, our knowledge on those codes is not enough according to their importance. In this paper, we have contributed to the knowledge of DNA codes by designing new families of DNA cyclic codes constructed over rings and by studying their corresponding properties and algebraic structures.

References

[1] T. Abualrub, N. Aydin and P. Seneviratne, On $\Theta$-cyclic codes over $F_2 + vF_2$, Austral. J. Combin., 54 (2012), 115–126.
[2] T. Abualrub, A. Ghrayeb and X. N. Zeng, Construction of cyclic codes over $F_4$ for DNA computing J. Franklin Ins., 343 (2006), 488–457.
[3] L. Adleman, Molecular computation of the solutions to combinatorial problems Science, 266 (1994), 1021–1024.
[4] C. Alf-Steinberger, The genetic code and error transmission, Proc. Natl. Acad. Sci. USA, 64 (1969), 584–591.
[5] M. B. Bechet, Bias de codons et Régulation de la Traduction chez les Bactéries et le Phages, Ph.D thesis, Univ. Paris 7, 2007.
[6] H. Q. Dinh and S. R. Lopez-Permouth, Cyclic and negacyclic codes over finite chain rings, IEEE Trans. Inf. Theory, 50 (2004), 1728–1744.
[7] S. T. Dougherty, J. Lark Kim and H. Kulosman, MDS code over finit principal ideal rings, Des. Codes Cryptogr., 50 (2009), 77–92.
[8] K. Guenda and T. A. Gulliver, Construction of cyclic codes over $F_2 + uF_2$ for DNA computing, Appl. Algebra Eng. Commun. Comput., 24 (2013), 445–459.
[9] K. Guenda and T. A. Gulliver, Repeated root constacyclic codes of length $mp^r$ over $F_p + uF_p + \cdots + u^{e-1}F_p$, J. Alg. App., to appear.
[10] K. Guenda, T. A. Gulliver and P. Solé, On cyclic DNA codes in Proc. IEEE Int. Symp. Inform. Theory, Istanbul, 2013, 121–125.
[11] A. K. Konopka, Theory of the degenerate coding and information parameters of the protein coding genes, Biochimie, 67 (1985), 455–468.
[12] M. Mansuripur, P. K. Khulbe, S. M. Kuebler, J. W. Perry, M. S. Giridhar and N. Peyghambarian, Information Storage and Retrieval Using Macromolecules as Storage Media, Univ. Arizona Technical Report, 2003.
[13] J. L. Massey, Reversible codes, Inf. Control, 7 (1964), 369–380.
[14] O. Milenkovic and N. Kashyap, On the design of codes for DNA computing in IEEE Int. Symp. Inf. Theory (ISIT), 2006.
[15] G. H. Norton and A. Salagean, On the structure of linear and cyclic codes over finite chain ring, AAECC, 10 (2000), 489–506.
[16] E. S. Ristad and P. N. Yianilos, Learning string-edit distance IEEE Trans. Anal. Mach. Intell., 20 (1998), 522–532.
[17] V. Rykov, A. J. Macula, D. Torney and P. White, DNA sequences and quaternary cyclic codes in IEEE Int. Symp. Inf. Theory (ISIT), 2001.
[18] R. Sanchez, E. Morgado and R. Grau, Gene algebra from a genetic code algebraic structure J. Math. Biol., 51 (2005), 431–475.
[19] I. Siap, T. Abualrub and A. Ghrayeb, Cyclic DNA codes over ring $F_2[u]/(u^2 − 1)$ based on the deletion distance, Franklin Institute, 36 (2009), 731–740.
[20] http://www.codetables.de

Received June 2015; revised December 2015.

E-mail address: nabil.bennenni@gmail.com
E-mail address: ken.guenda@gmail.com
E-mail address: sihem.mesnager@univ-paris8.fr