

ANKARA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

CEBİRSEL SAYILAR TEORİSİ KULLANARAK DNA KODU İNŞA  
ETME

Emre GÜDAY

MATEMATİK ANABİLİM DALI

ANKARA  
2020

Her hakkı saklıdır

# ÖZET

Yüksek Lisans Tezi

## CEBİRSEL SAYILAR TEORİSİ KULLANARAK DNA KODU İNŞA ETME

Emre GÜDAY

Ankara Üniversitesi  
Fen Bilimleri Enstitüsü  
Matematik Anabilim Dalı

Danışman: Prof. Dr. Ali Bülent EKİN

Bir DNA kodu,  $S = \{A, C, G, T\}$  sembol kümesi üzerinde, sabit  $n$  uzunluğundaki kelimelerin bir kümesi olarak tanımlanır. DNA kodları; DNA hesaplaması, DNA mikrodizilim teknolojileri ve kimyasal kütüphaneler gibi alanlara uygulanmaktadır. Bu uygulamalarda belli kombinatorik kısıtlamaları sağlayan DNA kodları istenmektedir. DNA kodu inşa etmedeki esas problem, olabildiğince fazla kısıtlamayı sağlayan ve olabildiğince fazla kodkelimesine sahip olan DNA kodları oluşturmaktır.

Bu tezde, indirgenmez devirli kodların ağırlıkları ile Gauss periyodu arasındaki ilişkiden yararlanılarak kısıtlamaların tamamını sağlayan DNA kodları inşa edilmiştir. Tezin giriş bölümünde literatür taramasına yer verilmiş ve DNA kodlarının sağlaması istenen kısıtlamalar açıklanmıştır. İkinci bölümde sonlu cisimlerden ve kodlama teorisinden bazı temel kavramlar verilmiştir. Üçüncü bölümde Gauss toplamı tanımlanmış ve bazı önemli özellikleri çalışılmıştır. Dördüncü bölümde tezin en önemli aracı olan Gauss periyodu çalışılmıştır. Son bölümde belli kombinatorik koşulları sağlayan DNA kodları inşa edilip parametreleri hesaplanmıştır.

Bu tezdeki hesaplamalarda MAGMA Computer Algebra programı kullanılmıştır.

**Ağustos 2020, 56 sayfa**

**Anahtar Kelimeler:** Gauss Toplamı, Gauss Periyodu, İndirgenmez Devirli Kodlar, İndirgenmez Devirli Kodların Ağırlıkları, DNA Kodu İnşası

# ABSTRACT

Master Thesis

## CONSTRUCTION OF DNA CODES BY USING ALGEBRAIC NUMBER THEORY

Emre GÜDAY

Ankara University  
Graduate School of Natural and Applied Sciences  
Department of Mathematics

Supervisor: Prof. Dr. Ali Bülent EKİN

A DNA code is defined as a set of words of fixed length  $n$  over the symbol set  $S = \{A, C, G, T\}$ . DNA codes are implemented in DNA computations, DNA microarray technologies and chemical libraries. In these applications, DNA codes satisfying certain combinatorial constraints are desired. The main problem in construction of a DNA code is to find a DNA code that satisfies as many constraints as possible and contains as many codewords as possible.

In this thesis, DNA codes which satisfy certain constraints are constructed by using a relation between weights of irreducible cyclic codes and the Gaussian periods. In introduction of the thesis, a literature survey is presented and constraints that are desired to be satisfied by DNA codes are introduced. In the second chapter, some basic concepts from finite fields and coding theory are presented. In the third chapter, Gauss sums are defined and some important properties of them are studied. In the fourth chapter, Gaussian periods which is one of the most important tools in the thesis are studied. In the last chapter, DNA codes satisfying the combinatorial constraints given in Introduction are constructed. Furthermore, their parameters are computed with the help of MAGMA Computer Algebra System.

**August 2020, 56 pages**

**Key Words:** Gauss Sums, Gaussian Periods, Irreducible Cyclic Codes, Weights of Irreducible Cyclic Codes, Construction of DNA Codes

## TEŞEKKÜR

Lisansüstü eğitimim boyunca benden bilgisini, deneyimlerini ve emeğini asla esirgemeyen, karşılaştığım zorluklar karşısında bana her zaman destek olan danışman hocam Prof. Dr. Ali Bülent EKİN'e, tezimi yazarken bana katlanan ve her zaman yanımda olan ailem ve dostlarıma sonsuz teşekkürlerimi sunarım.

Emre GÜDAY  
Ankara, Ağustos 2020



## İÇİNDEKİLER

|                                                                                                                               |     |
|-------------------------------------------------------------------------------------------------------------------------------|-----|
| TEZ ONAY SAYFASI                                                                                                              |     |
| ETİK.....                                                                                                                     | i   |
| ÖZET .....                                                                                                                    | ii  |
| ABSTRACT .....                                                                                                                | iii |
| TEŞEKKÜR.....                                                                                                                 | iv  |
| ŞEKİLLER DİZİNİ.....                                                                                                          | vi  |
| ÇİZELGELER DİZİNİ.....                                                                                                        | vii |
| 1. GİRİŞ .....                                                                                                                | 1   |
| 2. TEMEL KAVRAMLAR .....                                                                                                      | 4   |
| 2.1 Sonlu Cisimler.....                                                                                                       | 4   |
| 2.2 Kodlama Teorisi.....                                                                                                      | 7   |
| 2.3 Devirli Kodlar .....                                                                                                      | 9   |
| 2.4 İndirgenmez Devirli Kodlar.....                                                                                           | 13  |
| 2.5 Ters Sırada Yazılabilen Kodlar .....                                                                                      | 16  |
| 3. GAUSS TOPLAMLARI .....                                                                                                     | 20  |
| 3.1 Toplamsal ve Çarpımsal Karakterler .....                                                                                  | 20  |
| 3.2 Gauss Toplamları .....                                                                                                    | 23  |
| 4. GAUSS PERİYODU .....                                                                                                       | 27  |
| 5. DNA KODLARININ İNŞASI .....                                                                                                | 32  |
| 5.1 Hamming Uzaklığı ve GC-İçeriği Kısıtlamalarını Sağlayan DNA<br>Kodlarının İnşası.....                                     | 35  |
| 5.2 Hamming Uzaklığı, Ters-Eşlenik ve GC-İçeriği Kısıtlamalarını<br>Sağlayan DNA Kodlarının İnşası.....                       | 39  |
| 5.3 Hamming Uzaklığı, Ters Sırada Yazılış, Ters-Eşlenik ve GC-<br>İçeriği Kısıtlamalarını Sağlayan DNA Kodlarının İnşası..... | 46  |
| 6. TARTIŞMA VE SONUÇ .....                                                                                                    | 53  |
| KAYNAKLAR .....                                                                                                               | 54  |
| ÖZGEÇMİŞ.....                                                                                                                 | 56  |

## ŞEKİLLER DİZİNİ

|           |                                                        |   |
|-----------|--------------------------------------------------------|---|
| Şekil 2.1 | $\mathbb{F}_{3^{30}}$ cisminin alt cisimleri . . . . . | 5 |
|-----------|--------------------------------------------------------|---|



## ÇİZELGELER DİZİNİ

|             |                                                                     |    |
|-------------|---------------------------------------------------------------------|----|
| Çizelge 4.1 | Gauss Periyotları . . . . .                                         | 31 |
| Çizelge 5.1 | (1) ve (4) kısıtlamalarını sağlayan DNA kodları . . . . .           | 39 |
| Çizelge 5.2 | (1), (3) ve (4) kısıtlamalarını sağlayan DNA kodları . . . . .      | 46 |
| Çizelge 5.3 | (1), (2), (3) ve (4) kısıtlamalarını sağlayan DNA kodları . . . . . | 52 |



## 1. GİRİŞ

DNA kodu inşasının ortaya çıkışı, Amerikan bilgisayar bilimci Leonard Adleman'ın 1994 yılındaki çalışması olarak kabul görmektedir (Adleman 1994). DNA kodları; DNA hesaplamada (Marathe vd. 2001, Frutos vd. 1997), DNA mikrodizilim teknolojilerinde (Shena vd. 1995) ve kimyasal kütüphanelerde (Branner ve Larner 1992) uygulanmaktadır. Bu uygulamalarda başarılı çalışmalar gerçekleştirmek için belli kombinatorik kısıtlamaları sağlayan DNA kodlarına ihtiyaç duyulmaktadır.

DNA'nın doğal yapısı dört ana unsurdan oluşmaktadır. Bunlar, adenin (A), sitozin (C), guanin (G) ve timin (T) 'dir. Bu unsurlar arasında Watson-Crick eşlenik prensibi olarak bilinen bir ilişki vardır. Adenin'in eşleniği timin, guaninin eşleniği de sitozindir ( $\overline{A} = T$ ,  $\overline{T} = A$ ,  $\overline{C} = G$  ve  $\overline{G} = C$ ) (D'yachkov vd. 2005).

Bir DNA kodu,  $(n, M, d, w)$  ile gösterilirken;  $n$  ile kodkelimesi uzunluğu,  $M$  ile kodkelimesi sayısı,  $d$  ile minimum Hamming uzaklığı ve  $w$  ile de Guanin ve Sitozin içeriği sayısı gösterilmektedir.

DNA kodu inşa etmekte esas problem, kodkelimesi uzunluğu sabit  $n$  sayısı olan ve olabildiğince fazla kısıtlamayı sağlayan en geniş DNA kodunu bulmaktır (Hong vd. 2016).

### DNA Kodları Üzerindeki Kısıtlamalar

DNA kodları üzerinde, uygulamaya yönelik, çeşitli kısıtlamalar vardır. Genel olarak bu kısıtlamalar aşağıdaki gibi iki sınıfa ayrılabilir:

- A) Kombinatorik kısıtlamalar
- B) Termodinamik kısıtlamalar

Bu tezde kombinatorik kısıtlamaları sağlayan DNA kodları inşa edilecektir. DNA kodlarının sağlaması istenen kombinatorik kısıtlamalar aşağıda açıklanmıştır.

1) Hamming Uzaklığı Kısıtlaması:

Herhangi bir Hamming uzaklığı  $d$  ve  $x \neq y$  olacak biçimde herhangi iki  $x$  ve  $y$  kodkelimesi için

$$H(x, y) \geq d$$

sağlanmalıdır (Gaborit ve King 2005). İki kodkelimesi arasındaki Hamming uzaklığı, kodkelimelerinin birbirlerinden farklı olan yerlerinin toplam sayısı olarak hesaplanır.

Örnek:  $x = ACGTACGT$  ve  $y = TCGGATCC$  olsun.  $H(x, y) = 5$  olur.

2) Ters Sırada Yazılış (Reverse) Kısıtlaması:

Herhangi bir Hamming uzaklığı  $d$  ve  $y := (y_0, y_1, \dots, y_{n-1})$  sıralı  $n$ -lisi için  $y^R := (y_{n-1}, y_{n-2}, \dots, y_0)$  olmak üzere herhangi iki  $x$  ve  $y$  kodkelimesi için

$$H(x, y^R) \geq d$$

sağlanmalıdır (Gaborit ve King 2005). Burada  $x = y$  olabilir.

Örnek:  $x = TCGGATCC$  ve  $y = TCGGATCC$  olsun.

Bu durumda  $y^R = CCTAGGCT$  olup,  $H(x, y^R) = 6$ 'dır.

3) Ters-Eşlenik(Reverse-Complement) Kısıtlaması:

Herhangi bir Hamming uzaklığı  $d$  ve  $y := (y_0, y_1, \dots, y_{n-1})$  sıralı  $n$ -lisi için  $\overline{y^R} := (\overline{y_{n-1}}, \overline{y_{n-2}}, \dots, \overline{y_0})$  olmak üzere herhangi iki  $x$  ve  $y$  kodkelimesi için

$$H(x, \overline{y^R}) \geq d$$

sağlanmalıdır (Gaborit ve King 2005).

Örnek:  $x = GTACGCGT$  ve  $y = GTACGCGT$  olsun.

Bu durumda  $\overline{y^R} = ACGCGTAC$  olup,  $H(x, \overline{y^R}) = 6$ 'dır.

4) GC-İçeriği Kısıtlaması:

Her bir kodkelimesindeki Guanin veya Sitozin bulunan yerlerin toplam sayısı aynı ve  $w$ -tane olmalıdır (Gaborit ve King 2005).

Örnek:  $\mathbf{C} = \{AGGA, GTCA, CTTC, ACTG, GGTT, TTGG\}$  olsun.

$\mathbf{C}$ 'nin her bir kodkelimesi için  $w = 2$  olur.

Büyük  $n$  sayıları için bu dört koşulu aynı anda sağlayan bir DNA kodu inşa etmek oldukça zordur. Ayrıca uygulamalarda olabildiğince fazla bilgi depolamak için, olabildiğince fazla kodkelimesi istenir. Birçok durumda en fazla kodkelimesi sayısına  $w \approx \lfloor \frac{n}{2} \rfloor$  olduğunda ulaşılır (King 2003).

2005'te  $n \leq 20$  için (3) ve (4) koşullarını sağlayan bir DNA kodu inşa etmek için hata düzelten kodlar kullanılmıştır (Gaborit ve King 2005). 2008'de  $n \leq 20$  için (1), (3) ve (4) koşullarını sağlayan (Montemanni ve Smith 2008), 2011'de  $n \leq 30$  için (1) ve (4) koşullarını sağlayan (Smith vd. 2011), 2012'de  $n \leq 30$  için (1), (3) ve (4) koşullarını sağlayan DNA kodları inşa edilmiştir (Aboluion vd. 2012).

DNA kodu inşa etme problemi üzerine birçok yaklaşım mevcuttur. Bu tezde daha büyük  $n$  sayıları için iyi özelliklere sahip yeni DNA kodları, cebirsel sayılar teorisinden metotlar kullanılarak inşa edilecektir. İlk olarak (1) ve (4) koşullarını sağlayan bir DNA kodu inşa edilecektir. İnşa edilen DNA kodlarının parametrelerini hesaplamak için Gauss periyodu ile indirgenmez devirli kodların ağırlıkları arasındaki bir bağıntıdan yararlanılacaktır. Daha sonra (1), (3) ve (4) koşullarını sağlayan ve son olarak kısıtlamaların tamamını sağlayan yeni DNA kodu inşa edilip bu kodların parametrelerini hesaplamak için MAGMA Computer Algebra program dilindeki kodlar verilecektir.

## 2. TEMEL KAVRAMLAR

Kodlama teorisinin temel taşı sonlu cisimlerdir. İlk kısımda sonlu cisimler tanıtılacaktır. Ardından kodlama teorisinden, önemli kavramlar, tanımlar ve teoremler verilecektir. Daha sonra, devirli kodlar ve indirgenmez devirli kodlar tanımlanacak ve indirgenmez devirli kodlara bir örnek gösterilecektir. Son kısımda ise ters sırada yazılabilen ve yazılamayan kodlar tanıtılıp, örnekler verilecektir.

### 2.1 Sonlu Cisimler

Sonlu sayıda elemanı olan cisme sonlu cisim denir. Bir  $p$  asal için  $\mathbb{Z}_p = \{\overline{0}, \overline{1}, \dots, \overline{p-1}\}$  kümesi sonlu cisimlere bir örnektir.

**Tanım 2.1** Bir  $p$  asal sayısı için  $\mathbb{F}_p := \{0, 1, \dots, p-1\}$  kümesi tanımlansın.

$$\begin{aligned}\phi : \mathbb{Z}_p &\longrightarrow \mathbb{F}_p \\ \bar{a} &\longmapsto \phi(\bar{a}) := a\end{aligned}$$

şeklinde tanımlanan izomorfizm ile  $\mathbb{Z}_p$ 'nin cisim yapısı  $\mathbb{F}_p$ 'ye taşınır. Bu  $\mathbb{F}_p$  cismine,  $p$  elemanlı Galois cismi denir (Lidl ve Niederreiter 1997).

Burada, her  $a, b \in \mathbb{F}_p$  için  $a + b := \phi(\bar{a} + \bar{b})$  ve  $a.b := \phi(\bar{a}.\bar{b})$  şeklinde tanımlanır.

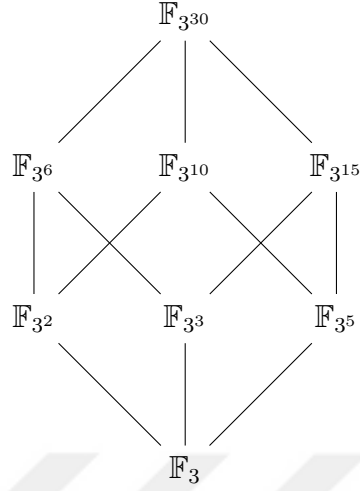
**Uyarı 2.1** Sonlu bir cismin karakteristiği asaldır.

**Teorem 2.1**  $\mathbb{F}_q$  bir sonlu cisim olsun. Bu takdirde her  $a \in \mathbb{F}_q$  için  $a^q = a$  olur (Lidl ve Niederreiter 1997).

**Teorem 2.2 (Varlık ve Teklik)** Her  $p$  asal ve her pozitif  $r$  tamsayısı için  $p^r$  elemanlı bir sonlu cisim vardır ve bu cisim  $x^{p^r} - x \in \mathbb{F}_p[x]$  polinomunun parçalanış cismine izomorftur (Lidl ve Niederreiter 1997).

**Teorem 2.3 (Alt Cisim Kriteri)**  $\mathbb{F}_{p^r}$  bir sonlu cisim olsun.  $\mathbb{F}_{p^r}$ 'nin  $p^t$  elemanlı her alt cismi için  $t|r$ 'dir. Tersine  $r$ 'nin her pozitif  $t$  böleni için,  $\mathbb{F}_{p^r}$ 'nin bir tek  $\mathbb{F}_{p^t}$  alt cismi vardır (Lidl ve Niederreiter 1997).

**Örnek 2.1**  $\mathbb{F}_{3^{30}}$  cisminin bütün alt cisimleri 30'un pozitif bölenleri listelenerek belirlenebilir.



Şekil 2.1.  $\mathbb{F}_{3^{30}}$  cisminin alt cisimleri

Bundan sonra  $q = p^t$  elemanlı herhangi bir sonlu cisim  $\mathbb{F}_q$  ile, onun  $r$ . dereceden genişlemesi de  $\mathbb{F}_{q^r}$  ile gösterilecektir.

**Teorem 2.4**  $\mathbb{F}_q$  bir sonlu cisim olmak üzere,  $\mathbb{F}_q^* := \mathbb{F}_q - \{0\}$  çarpımsal grubu devirlidir (Lidl ve Niederreiter 1997).

**Tanım 2.2**  $\mathbb{F}_q^*$  grubunun bir üreticine  $\mathbb{F}_q$  sonlu cisminin primitif elemanı denir (Lidl ve Niederreiter 1997).

**Tanım 2.3**  $f(x) \in \mathbb{F}_q[x]$  polinomunun derecesi  $r \geq 1$  olsun. Eğer  $f(x)$ ,  $\mathbb{F}_{q^r}$  cisminin bir primitif elemanının minimal polinomu ise  $f(x)$  polinomuna  $\mathbb{F}_q$  üzerinde bir primitif polinom denir (Lidl ve Niederreiter 1997).

**Uyarı 2.2**  $\Phi$ , Euler phi fonksiyonu olmak üzere,  $\mathbb{F}_q$  sonlu cisminin tam  $\Phi(q-1)$  tane primitif elemanı vardır.

**Teorem 2.5** Her  $\mathbb{F}_q$  sonlu cismi ve her  $r \in \mathbb{Z}^+$  için  $\mathbb{F}_q[x]$  polinom halkasında derecesi  $r$  olan bir indirgenmez polinom vardır (Lidl ve Niederreiter 1997).

**Teorem 2.6**  $f(x) \in \mathbb{F}_q[x]$  indirgenmez polinomunun derecesi  $r \geq 1$  olsun. Bu durumda  $f(x)$  polinomunun  $\theta \in \mathbb{F}_{q^r}$  olacak biçimde bir kökü vardır. Dahası,  $f(x)$  polinomunun

tüm kökleri  $\theta, \theta^q, \theta^{q^2}, \dots, \theta^{q^{r-1}}$  şeklindedir. Burada  $\theta, \theta^q, \theta^{q^2}, \dots, \theta^{q^{r-1}} \in \mathbb{F}_{q^r}$  elemanlarına  $\theta$ 'nın  $\mathbb{F}_q$ -eşlenikleri denir (Lidl ve Niederreiter 1997).

Sonlu cisimlerin yapısı karakterize edildikten sonra Gauss toplamının tanımında ve indirgenmez devirli kodların inşasında önemli bir yere sahip olan,  $\mathbb{F}_q$ 'daki bir elemanın izi kavramı tanımlanacak ve bazı özellikleri verilecektir.

**Tanım 2.4**  $\mathbb{F}_q < \mathbb{F}_{q^r}$  olsun.  $\alpha \in \mathbb{F}_{q^r}$  için iz dönüşümü

$$Tr_q(\alpha) := \alpha + \alpha^q + \alpha^{q^2} + \dots + \alpha^{q^{r-1}}$$

şeklinde tanımlanır (Lidl ve Niederreiter 1997).  $q = p$  durumunda kısaca  $Tr$  ile gösterilir.

**Teorem 2.7**  $\forall \alpha, \beta \in \mathbb{F}_{q^r}$  ve  $\forall c \in \mathbb{F}_q$  olsun. Bu durumda,

- 1)  $Tr_q(\alpha) \in \mathbb{F}_q$ ,
- 2)  $Tr_q : \mathbb{F}_{q^r} \longrightarrow \mathbb{F}_q$  örten bir lineer dönüşümdür,
- 3)  $Tr_q(\alpha^q) = Tr_q(\alpha)$ ,
- 4)  $Tr_q(c) = r.c$ ,

özellikleri sağlanır (Lidl ve Niederreiter 1997).

**Örnek 2.2**  $\mathbb{F}_4 := \{0, 1, \alpha, \alpha^2\}$ ,  $f(x) := x^2 + x + \alpha \in \mathbb{F}_4[x]$  primitif polinom ve  $f(\theta) = 0$  olsun. Bu durumda  $\theta$ ,  $\mathbb{F}_{4^2}$  cisminin bir primitif elemanıdır. Her  $\beta \in \mathbb{F}_{4^2}$  için iz dönüşümü,

$$Tr_4(\beta) := \beta + \beta^4$$

şeklinde tanımlanır. Buradan  $\mathbb{F}_{4^2}$  cisminin elemanlarının izi,

$$\begin{aligned} Tr_4(0) &= Tr_4(1) = Tr_4(\theta^5) = Tr_4(\theta^{10}) = 0, \\ Tr_4(\theta) &= Tr_4(\theta^2) = Tr_4(\theta^4) = Tr_4(\theta^8) = 1, \\ Tr_4(\theta^6) &= Tr_4(\theta^7) = Tr_4(\theta^9) = Tr_4(\theta^{13}) = \alpha, \\ Tr_4(\theta^3) &= Tr_4(\theta^{11}) = Tr_4(\theta^{12}) = Tr_4(\theta^{14}) = \alpha^2, \end{aligned}$$

olarak hesaplanır.

## 2.2 Kodlama Teorisi

Kodlama teorisi, Claude Elwood Shannon'ın The Bell System Technical Journal tarafından 1948'de yayınlanan "A Mathematical Theory of Communication" adlı çığır açan makalesi ile ortaya çıkmıştır. Shannon bu makalesinde, bilginin sıkıştırılarak iletilmesini çalışmış ve bunu gerçekleştirecek kodların var olduğunu kanıtlamıştır (Shannon 1948).

$\mathbb{F}_q$  bir sonlu cisim olsun. Bu durumda  $\mathbb{F}_q^n$ ,  $n$ -boyutlu bir  $\mathbb{F}_q$ -vektör uzayıdır.  $\mathbb{F}_q^n$  vektör uzayının boştan farklı herhangi bir altkümesine  $\mathbb{F}_q$  üzerinde bir  $q$ -lu kod denir. Kodun elemanlarına ise kodkelimesi adı verilir. Örneğin 1 ve 0'lardan oluşan koda ikili kod denir.

**Tanım 2.5** Her  $x := (x_1, x_2, \dots, x_n), y := (y_1, y_2, \dots, y_n) \in \mathbb{F}_q^n$  için

$$\delta(x_i, y_i) = \begin{cases} 1 & x_i \neq y_i \\ 0 & x_i = y_i \end{cases}$$

olmak üzere,

$$H(x, y) := \sum_{i=1}^n \delta(x_i, y_i)$$

fonksiyonuna Hamming uzaklığı denir (Hill 1986).

**Örnek 2.3**  $x = (1, 0, 1, 0, 1), y = (1, 1, 0, 1, 1) \in \mathbb{F}_2^5$  için  $H(x, y) = 3$ 'tür.

Bir kod için önemli bir parametre, minimum Hamming uzaklığıdır ve şöyle tanımlanır:

**Tanım 2.6**  $\mathbf{C}$  herhangi bir kod olsun.

$$H(\mathbf{C}) := \min\{H(x, y) \mid x, y \in \mathbf{C} \text{ ve } x \neq y\}$$

sayısına  $\mathbf{C}$  kodunun minimum Hamming uzaklığı denir (Hill 1986).

**Örnek 2.4**  $\mathbf{C} = \{(0, \alpha^2, \alpha), (\alpha^2, 0, \alpha^2), (\alpha, \alpha^2, 0), (\alpha, \alpha, \alpha^2), (\alpha^2, \alpha, \alpha)\} \subset \mathbb{F}_4^3$  kodu göz önüne alındığında  $H(\mathbf{C}) = 2$ 'dir. Bunu bulmak için  $\binom{5}{2} = 10$  tane kıyaslama yapılmıştır.

Kodkelimesi uzunluğu  $n$ , kodkelimesi sayısı  $M$  ve minimum Hamming uzaklığı  $d$  olan bir kod,  $(n, M, d)$  ile gösterilir. İyi bir kod; mesajların hızlı iletilebilmesi için küçük bir  $n$ , olabildiğince çeşitli mesajlar iletilmesi için büyük bir  $M$  ve olabildiğince çok sayıda hatanın düzeltilmesi için büyük bir  $d$  değerine sahip olmalıdır.

Kodlama teorisinin ana problemi, bu üç parametreden ikisi verildiğinde, diğer parametre için en iyi değeri bulmaktır. Yaygın problem ise,  $n$  ve  $d$  değerleri verildiğinde en büyük  $M$  sayısını bulmaktır.

Bu tezde inşa edilecek olan DNA kodları lineer kodların önemli bir alt ailesi olan devirli kodlardan, özellikle indirgenmez devirli kodlardan elde edilecektir. Dolayısıyla lineer kodlardan bazı önemli kavramlara değinerek devirli ve indirgenmez devirli kodların teorisini geliştirmekte fayda vardır. İnşa edilecek olan ve bahsi geçen kısıtlamaları sağlayan DNA kodlarının lineer olmadığına dikkat edilmelidir.

**Tanım 2.7**  $\mathbb{F}_q^n$  vektör uzayının  $k$ -boyutlu herhangi bir altuzayına,  $\mathbb{F}_q$  üzerinde bir  $[n, k]$  lineer kod denir (Hill 1986). Bazen bir lineer kod  $[n, k, d]$  ile de gösterilebilir.

Yukarıdaki örnekte verilen **C** kodu,  $\mathbb{F}_4$  üzerinde bir  $(3, 5, 2)$  kodudur.

**Uyarı 2.3** Bir  $[n, k, d]$  lineer kodu, aynı zamanda bir  $(n, q^k, d)$  kodudur.

$M$  tane kodkelimesine sahip bir kodun minimum Hamming uzaklığını bulmak için  $\binom{M}{2}$  tane kıyaslama yapılmalıdır. Ancak lineer kodlarla çalışılıyorsa, daha az zaman harcayarak kodun minimum Hamming uzaklığını bulmak mümkündür. Bu avantaj aşağıdaki ağırlık kavramı ile elde edilir.

**Tanım 2.8**  $x \in \mathbb{F}_q^n$  elemanının sıfırdan farklı koordinatlarının sayısına  $x$  vektörünün ağırlığı denir ve  $\mathbf{w}(x)$  ile gösterilir.

Örneğin,  $(1, 0, 1, 0, 1)$  vektörünün ağırlığı 3'tür.

**Teorem 2.8** **C** bir lineer kod olsun ve **C**'deki sıfırdan farklı kodkelimelerinin ağırlıklarının en küçüğü  $\mathbf{w}(\mathbf{C})$  ile gösterilsin. Bu durumda,

$$H(\mathbf{C}) = \mathbf{w}(\mathbf{C})$$

eşitliği sağlanır.

Teorem 2.8 gereği bir  $\mathbf{C}$  lineer kodunun minimum Hamming uzaklığını bulmak için  $M - 1$  tane kodkelimesine bakmak yeterlidir. Ancak yine de  $M$ 'nin büyük değerleri için bu yöntem çok zaman alabilir. Bu yüzden ağırlık formülleri geliştirilmiştir. Bu tezde inşa edilecek kodlar  $\mathbb{F}_4$  üzerinde indirgenmez devirli kodlardır. Dolayısıyla bu kodların minimum Hamming uzaklıklarını bulmak için  $\mathbb{F}_4$  üzerinde indirgenmez devirli kodların ağırlık formülünden yararlanılacaktır.

**Tanım 2.9**  $\mathbf{C}$  bir  $[n, k]$  lineer kodu olsun. Satırları  $\mathbf{C}$ 'nin baz elemanları olan  $k \times n$  tipindeki matrise  $\mathbf{C}$  için bir üreteç matrisi denir (Hill 1986).

Herhangi iki  $u = (u_1, u_2, \dots, u_n)$  ve  $v = (v_1, v_2, \dots, v_n) \in \mathbb{F}_q^n$  vektörlerinin iç çarpımı,

$$\langle u, v \rangle := u_1v_1 + u_2v_2 + \dots + u_nv_n$$

şeklinde tanımlanır. Eğer,  $\langle u, v \rangle = 0$  ise  $u$  ile  $v$  vektörlerine ortogonal vektörler denir.

**Tanım 2.10**  $\mathbf{C}$  bir  $[n, k]$  lineer kodu olsun.  $\mathbf{C}$ 'nin her bir kodkelimesi ile ortogonal olan vektörlerin kümesine  $\mathbf{C}$ 'nin dual kodu denir ve  $\mathbf{C}^\perp$  ile gösterilir. Yani  $\mathbf{C}$ 'nin dual kodu,

$$\mathbf{C}^\perp := \{v \in \mathbb{F}_q^n \mid \langle u, v \rangle = 0, \forall u \in \mathbf{C}\}$$

kümesidir (Hill 1986).

**Teorem 2.9**  $\mathbf{C}$ , lineer  $[n, k]$ -kodu olsun. Bu takdirde  $\mathbf{C}^\perp$  dual kodu lineer  $[n, n - k]$ -kodu olur (Hill 1986).

**Sonuç 2.1**  $\mathbf{C}$  lineer kod olsun.  $(\mathbf{C}^\perp)^\perp = \mathbf{C}$ 'dir (Hill 1986).

**Tanım 2.11**  $\mathbf{C}^\perp$  kodunun üreteç matrisine  $\mathbf{C}$  için bir benzerlik kontrol matrisi denir (Hill 1986).

## 2.3 Devirli Kodlar

Devirli kodlar, zengin cebirsel yapısı ile kodlama ve kod çözümleme tekniklerinin kolay uygulanabilmesi açısından lineer kodların önemli bir alt ailesidir. Hamming

kodları, Golay kodları ve BCH kodları devirli kodlara denk olan önemli kodlardan birkaçıdır. Bu tezde DNA kodlarının inşası devirli kodların özel bir türü olan indirgenmez devirli kodlara dayanmaktadır. Dolayısıyla indirgenmez devirli kodlar çalışılmadan önce devirli kodları incelemek faydalı olacaktır.

**Tanım 2.12** Bir  $\mathbf{C}$  kodu lineer ve

$$\forall (c_0, c_1, \dots, c_{n-1}) \in \mathbf{C} \text{ için } (c_{n-1}, c_0, \dots, c_{n-2}) \in \mathbf{C}$$

oluyorsa  $\mathbf{C}$  koduna devirli kod denir (MacWilliams ve Sloane 1977).

**Örnek 2.5**  $\mathbf{C} = \{(0, 0, 0), (1, 1, 0), (0, 1, 1), (1, 0, 1)\} \subset \mathbb{F}_2^3$  bir devirli koddur.

$\mathbb{F}_q^n$  ve  $\mathbb{F}_q[x]/(x^n - 1)$  birbirine izomorf olduğundan  $a := (a_0, a_1, \dots, a_{n-1}) \in \mathbb{F}_q^n$  vektörüne  $a(x) := a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in \mathbb{F}_q[x]/(x^n - 1)$  polinomu gözüyle bakmak, devirli kodlarda avantaj sağlamaktadır. Esasında bu avantaj kodlama ve kod çözümleme eyleminin "feed-back shift register" diye adlandırılan donanımlar vasıtasıyla kolayca uygulanabilmesini sağlar; ancak burada devirli kodların cebirsel özelliklerini geliştirmek amacına hizmet edecektir.

Şimdi  $a(x)$  polinomu  $x$  ile çarpılırsa:

$$\begin{aligned} xa(x) &= a_0x + a_1x^2 + \dots + a_{n-1}x^n \\ &= a_{n-1} + a_0x + a_1x^2 + \dots + a_{n-2}x^{n-1} \end{aligned}$$

$a$  vektörünün dairesel bir kayması olan  $(a_{n-1}, a_0, a_1, \dots, a_{n-2})$  vektörü elde edilir.

**Teorem 2.10**  $\mathbb{F}_q[x]/(x^n - 1)$ 'in bir  $\mathbf{C}$  altkümesinin devirli kod olması için gerek ve yeter şart  $\mathbf{C}$ 'nin bir ideal olmasıdır (Lint 1999).

**Teorem 2.11**  $f(x) \in \mathbb{F}_q[x]/(x^n - 1)$  olmak üzere  $\langle f(x) \rangle$  bir devirli koddur (Hill 1986).

**Örnek 2.6**  $f(x) = x^6 + x^5 + x^3 + 1 \in \mathbb{F}_2[x]/(x^7 - 1)$  olsun.

$$\mathbf{C} = \langle x^6 + x^5 + x^3 + 1 \rangle = \{0, x^6 + x^5 + x^3 + 1, x^5 + x^2 + x + 1, x^6 + x^5 + x^4 + x^2, x^6 + x^3 + x^2 + x, x^4 + x^3 + x^2 + 1, x^6 + x^4 + x + 1, x^5 + x^4 + x^3 + x\}$$

kümesi  $\mathbb{F}_2$  üzerinde bir devirli koddur.  $\mathbf{C}$ 'nin elemanlarına  $\mathbb{F}_2^7$ 'nin elemanları gözüyle bakılarak,

$C = \{0000000, 1101001, 0100111, 1110100, 1001110, 0011101, 1010011, 0111010\}$  kümesinin devirli bir kod olduğu daha kolay gözlemlenebilir.

Aşağıdaki teorem devirli kodların nasıl inşa edileceğini göstermesi açısından önem teşkil etmektedir.

**Teorem 2.12**  $\{0\} \neq C \subset \mathbb{F}_q[x]/(x^n - 1)$  bir devirli kod olsun. Aşağıdakiler sağlanır:

- (i)  $C$  'de en küçük dereceli bir tek monik  $g(x)$  polinomu vardır,
- (ii)  $C = \langle g(x) \rangle$ ,
- (iii)  $g(x) | x^n - 1$ .

(MacWilliams ve Sloane 1977).

**Örnek 2.7**  $C = \langle x^4 + x^3 + x^2 + 1 \rangle$  devirli kodu göz önüne alınsın.

$$\begin{aligned} x^7 - 1 &= (x + 1)(x^3 + x + 1)(x^3 + x^2 + 1) \\ &= (x^4 + x^3 + x^2 + 1)(x^3 + x^2 + 1) \end{aligned}$$

olduğundan  $x^4 + x^3 + x^2 + 1 | x^7 - 1$ 'dir.

Dahası,  $\langle x^4 + x^3 + x^2 + 1 \rangle = \langle x^6 + x^5 + x^3 + 1 \rangle$  'dir.

O halde  $x^n - 1 \in \mathbb{F}_q[x]$  polinomunun  $\mathbb{F}_q$  üzerindeki çarpanlara ayrılışı biliniyorsa, kodkimesi uzunluğu  $n$  olan tüm devirli kodlar kolaylıkla elde edilebilir.

Her bir devirli koda ilişkin iki önemli polinom vardır. Bu polinomlar, üreteç polinomu ve kontrol polinomudur.

**Tanım 2.13**  $C$  bir devirli kod olsun.  $C$  'yi üreten en küçük dereceli ve monik  $g(x)$  polinomuna  $C$ 'nin üreteç polinomu denir (Hill 1986).

Dikkat edilmelidir ki, devirli bir kodu üreten, üreteç polinomundan başka polinomlar da mevcut olabilir; ancak üreteç polinomu tektir.

**Tanım 2.14**  $h(x) := (x^n - 1)/g(x)$  polinomuna  $C$ 'nin kontrol polinomu denir (MacWilliams ve Sloane 1977).

Aşağıdaki teorem göstermektedir ki bir polinomun aslında bir kodkelimesi olup olmadığı  $h(x)$  polinomu aracılığı ile öğrenilmektedir. Dolayısıyla bu polinoma kontrol polinomu denilmiştir.

**Teorem 2.13**  $\{0\} \neq \mathbf{C} \subset \mathbb{F}_q[x]/(x^n - 1)$  devirli kodu için  $g(x)$  üreteç polinomu,  $h(x)$  ise kontrol polinomu olsun. Bir  $c(x) \in \mathbb{F}_q[x]/(x^n - 1)$  polinomu için

$$c(x) \in \mathbf{C} \Leftrightarrow c(x)h(x) = 0$$

dır (Hill 1986).

**Örnek 2.8** Bir önceki örnek ele alınırsa (Örnek 2.7),

$x^3 + x^2 + 1$  polinomu  $\mathbf{C}$ 'nin kontrol polinomudur.

Gerçekten,  $x^6 + x^5 + x^3 + 1 \in \mathbf{C}$  için  $(x^6 + x^5 + x^3 + 1)(x^3 + x^2 + 1) = 0$  olup,  $x^3 + x^2 + x + 1 \notin \mathbf{C}$  için  $(x^3 + x^2 + x + 1)(x^3 + x^2 + 1) \neq 0$  olduğu gözlemlenebilir.

Devirli kodlar tanımları gereği lineerdir. Dolayısıyla akla gelen ilk soru bu kodun boyutu ve üreteç matrisinin ne olduğudur. Aşağıdaki teorem, devirli bir kodun üreteç polinomuna bakarak hem boyutunun hem de üreteç matrisinin elde edilebileceğini göstermesi bakımından önemlidir.

**Teorem 2.14**  $\mathbf{C}$  bir devirli kod ve

$$g(x) := g_0 + g_1x + \cdots + g_kx^k$$

polinomu,  $\mathbf{C}$ 'nin üreteç polinomu olsun. Bu durumda,

$$1) \dim(\mathbf{C}) = n - k,$$

$$2)$$

$$G = \begin{bmatrix} g_0 & g_1 & \cdots & g_k & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_k & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & \cdots & g_k \end{bmatrix}$$

$\mathbf{C}$  için bir üreteç matrisidir (MacWilliams ve Sloane 1977).

**Teorem 2.15**  $C$  bir devirli  $[n, r]$  kodu ve

$$h(x) := h_0 + h_1x + \cdots + h_rx^r$$

$C$  kodunun kontrol polinomu olsun. Bu durumda,

1)

$$H = \begin{bmatrix} h_r & h_{r-1} & \cdots & h_0 & 0 & 0 & \cdots & 0 \\ 0 & h_r & h_{r-1} & \cdots & h_0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & h_r & h_{r-1} & \cdots & h_0 \end{bmatrix}$$

$C$  için bir benzerlik kontrol matrisidir.

2)  $C^\perp$  dual kodu

$$h_R(x) := h_r + h_{r-1}x + \cdots + h_0x^r$$

polinomu tarafından üretilir (Hill 1986).

## 2.4 İndirgenmez Devirli Kodlar

**Tanım 2.15**  $D$  bir devirli kod ve kontrol polinomu  $h(x)$  olsun. Eğer  $h(x)$  indirgenmez ise  $D$ 'ye indirgenmez devirli kod denir (Lint 1999).

**Örnek 2.9**  $C = \langle x^4 + x^3 + x^2 + 1 \rangle$  kodu indirgenmez devirli kodlara örnek gösterilebilir.

Aşağıdaki teorem indirgenmez devirli kodları inşa etmede büyük kolaylık sağlaması bakımından önemlidir.

**Teorem 2.16**  $\mathbb{F}_q$  ve  $\mathbb{F}_{q^r}$  iki sonlu cisim,  $\gamma \in \mathbb{F}_{q^r}$  birin  $n$ .primitif kökü ve  $q^r \equiv 1 \pmod{n}$  olacak biçimdeki en küçük pozitif tamsayı  $r$  olsun. Bu durumda

$$\mathbf{D} := \{c_\beta := (Tr_q(\beta), Tr_q(\beta\gamma), \dots, Tr_q(\beta\gamma^{n-1})) \mid \beta \in \mathbb{F}_{q^r}\}$$

kümesi  $\mathbb{F}_q$  üzerinde bir  $[n, r]$  indirgenmez devirli kodudur (Lint 1999).

**İspat.** Önce  $\mathbf{D}$  kümesinin devirli kod olduğu, sonra da onun kontrol polinomunun indirgenmez bir polinom olduğu gösterilecektir. Her  $c_\beta, c_\alpha \in \mathbf{D}$  ve  $a \in \mathbb{F}_q$  için,

$$c_\beta + c_\alpha = (Tr_q(\beta + \alpha), Tr_q((\beta + \alpha)\gamma), \dots, Tr_q((\beta + \alpha)\gamma^{n-1})) \in \mathbf{D}$$

$$ac_\beta = (Tr_q(a\beta), Tr_q((a\beta)\gamma), \dots, Tr_q((a\beta)\gamma^{n-1})) \in \mathbf{D}$$

olduğundan  $\mathbf{D}$  lineerdir.

Her  $c_\beta \in \mathbf{D}$  için,

$$\begin{aligned} c_{\beta\gamma} &= (Tr_q(\beta\gamma), Tr_q((\beta\gamma)\gamma), \dots, Tr_q((\beta\gamma)\gamma^{n-1})) \in \mathbf{D} \\ &= (Tr_q(\beta\gamma), Tr_q(\beta\gamma^2), \dots, Tr_q(\beta\gamma^{n-1}), Tr_q(\beta)) \in \mathbf{D} \end{aligned}$$

olur.

Yani  $\mathbf{D}$ 'nin her kodkelimesinin devirsel her kayması yine  $\mathbf{D}$ 'nin bir kodkelimesidir. Dolayısıyla  $\mathbf{D}$  bir devirli koddur.

Şimdi,  $\gamma \in \mathbb{F}_{q^r}$  birin  $n$ . primitif kökü ve  $\gamma$ 'nın  $\mathbb{F}_q$  üzerindeki minimal polinomu  $h(x) := h_0 + h_1x + \dots, h_{r-1}x^{r-1} + h_rx^r$  olsun.  $q^r \equiv 1 \pmod{n}$  olacak biçimdeki en küçük pozitif tamsayı  $r$  olduğundan  $h(x)$  polinomunun tüm kökleri,  $\gamma, \gamma^q, \dots, \gamma^{q^{r-1}}$ 'dir. Diğer yandan  $\gamma$  birin  $n$ . primitif kökü olduğundan  $x^n - 1$  polinomunun da bir köküdür. Sonuç olarak,  $h(x)$  indirgenmez,  $h(\gamma) = 0$  ve  $\gamma^n - 1 = 0$  olduğundan  $h(x)|x^n - 1$  elde edilir.

Şimdi,  $h(x)$  ile üretilen devirli kodun üreteç matrisi göz önüne alınsın.

$$H = \begin{bmatrix} h_0 & h_1 & \dots & h_r & 0 & 0 & \dots & 0 \\ 0 & h_0 & h_1 & \dots & h_r & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & h_0 & h_1 & \dots & h_r \end{bmatrix}$$

Buradan  $\langle h(x) \rangle$  bir  $[n, n - r]$  devirli kodudur.

$\mathbf{D}$ 'nin her bir kodkelimesi  $c_\beta := (c_0, c_1, \dots, c_{n-1})$  şeklinde gösterilirse,

$$\begin{aligned}
c_0 h_0 + c_1 h_1 + \dots + c_r h_r &= h_0 \text{Tr}_q(\beta) + h_1 \text{Tr}_q(\beta\gamma) + \dots + h_r \text{Tr}_q(\beta\gamma^r) \\
&= \text{Tr}_q(\beta(h_0 + h_\gamma + \dots + h_r \gamma^r)) \\
&= \text{Tr}_q(\beta h(\gamma)) \\
&= 0 \\
c_1 h_0 + c_2 h_1 + \dots + c_{r+1} h_r &= h_0 \text{Tr}_q(\beta\gamma) + h_1 \text{Tr}_q(\beta\gamma^2) + \dots + h_r \text{Tr}_q(\beta\gamma^{r+1}) \\
&= \text{Tr}_q(\beta\gamma h(\gamma)) \\
&= 0 \\
&\vdots \\
c_{n-r-1} h_0 + \dots + c_{n-1} h_r &= h_0 \text{Tr}_q(\beta\gamma^{n-r-1}) + \dots + h_r \text{Tr}_q(\beta\gamma^{n-1}) \\
&= \text{Tr}_q(\beta\gamma^{n-r-1} h(\gamma)) \\
&= 0
\end{aligned}$$

elde edilir ki bu da  $\mathbf{D}$  'nin her bir kodkelimesi ile,  $H$  matrisinin satırlarının ortogonal olduğunu gösterir. Buradan  $\langle h(x) \rangle$  devirli kodu ile  $\mathbf{D}$  devirli kodu birbirlerinin dualidir. Sonuç olarak  $H$  matrisi  $\mathbf{D}$  için bir benzerlik kontrol matrisidir. O halde  $x^r h(x^{-1})$  polinomu  $\mathbf{D}$  için kontrol polinomudur. Ayrıca  $h(x)$  indirgenmez olduğundan  $x^r h(x^{-1})$  polinomu da indirgenmezdir. Sonuç olarak  $\mathbf{D}$ ,  $[n, r]$  indirgenmez devirli kodudur. ■

**Örnek 2.10**  $\theta \in \mathbb{F}_{16}$  primitif eleman,  $\gamma = \theta^3$  birin 5.primitif kökü olsun. Teorem 2.16 gereğince  $\mathbb{F}_4$  üzerinde  $[5, 2]$  indirgenmez devirli kod,

$$\mathbf{D} := \{c_\beta = (\text{Tr}_4(\beta), \text{Tr}_4(\beta\gamma), \text{Tr}_4(\beta\gamma^2), \text{Tr}_4(\beta\gamma^3), \text{Tr}_4(\beta\gamma^4)) \mid \beta \in \mathbb{F}_{q'}\}$$

$$\begin{aligned}\beta = 0 \Rightarrow c_0 &= (Tr_4(0), Tr_4(0), Tr_4(0), Tr_4(0), Tr_4(0)) \\ &= (0, 0, 0, 0, 0)\end{aligned}$$

$$\begin{aligned}\beta = 1 \Rightarrow c_1 &= (Tr_4(1), Tr_4(\theta^3), Tr_4(\theta^6), Tr_4(\theta^9), Tr_4(\theta^{12})) \\ &= (0, \alpha^2, \alpha, \alpha, \alpha^2)\end{aligned}$$

$$\begin{aligned}\beta = \theta \Rightarrow c_\theta &= (Tr_4(\theta), Tr_4(\theta^4), Tr_4(\theta^7), Tr_4(\theta^{10}), Tr_4(\theta^{13})) \\ &= (1, 1, \alpha, 0, \alpha)\end{aligned}$$

$$\begin{aligned}\beta = \theta^2 \Rightarrow c_{\theta^2} &= (Tr_4(\theta^2), Tr_4(\theta^5), Tr_4(\theta^8), Tr_4(\theta^{11}), Tr_4(\theta^{14})) \\ &= (1, 0, 1, \alpha^2, \alpha^2)\end{aligned}$$

olmak üzere

$$\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha^2 & \alpha & \alpha & \alpha^2 \\ \alpha^2 & 0 & \alpha^2 & \alpha & \alpha \\ \alpha & \alpha^2 & 0 & \alpha^2 & \alpha \\ \alpha & \alpha & \alpha^2 & 0 & \alpha^2 \\ \alpha^2 & \alpha & \alpha & \alpha^2 & 0 \\ 1 & 1 & \alpha & 0 & \alpha \\ \alpha & 1 & 1 & \alpha & 0 \\ 0 & \alpha & 1 & 1 & \alpha \\ \alpha & 0 & \alpha & 1 & 1 \\ 1 & \alpha & 0 & \alpha & 1 \\ 1 & 0 & 1 & \alpha^2 & \alpha^2 \\ \alpha^2 & 1 & 0 & 1 & \alpha^2 \\ \alpha^2 & \alpha^2 & 1 & 0 & 1 \\ 1 & \alpha^2 & \alpha^2 & 1 & 0 \\ 0 & 1 & \alpha^2 & \alpha^2 & 1 \end{bmatrix}$$

şeklindedir.

## 2.5 Ters Sırada Yazılabilen Kodlar

Bu kısımda ters sırada yazılabilen kodlar tanımlanacak ve bir devirli kodun ne zaman ters sırada yazılabilen kod olduğuna dair bir teorem verilecektir. DNA kodlarının

sağlaması istenen kısıtlamalardan birisi ters sırada yazılış kısıtlamasıdır. Bu kodların tezdeki önemi, bahsi geçen kısıtlamanın hangi durumda sağlanabileceğini, hangi durumda sağlanamayacağını göstermektir.

**Tanım 2.16**  $\mathbf{C}$ ,  $\mathbb{F}_q$  üzerinde bir  $(n, M, d)$  kodu olsun.

$$\forall c := (c_0, c_1, \dots, c_{n-1}) \in \mathbf{C} \text{ için } c^R := (c_{n-1}, c_{n-2}, \dots, c_0) \in \mathbf{C}$$

oluyorsa  $\mathbf{C}$  koduna ters sırada yazılabilen kod denir (Massey 1964).

**Örnek 2.11** Aşağıda  $\mathbb{F}_2$  üzerinde ters sırada yazılabilen ve yazılamayan kodlara bazı örnekler verilmiştir.

- 1)  $\mathbf{A} = \{(1, 1, 0), (0, 0, 1)\}$  ters sırada yazılabilen kod değildir,
- 2)  $\mathbf{B} = \{(1, 0, 1), (0, 1, 0)\}$  ters sırada yazılabilen koddur,
- 3)  $\mathbf{C} = \{(1, 1, 1), (1, 0, 0), (0, 1, 1), (0, 0, 0)\}$  ters sırada yazılabilen kod değildir,
- 4)  $\mathbf{D} = \{(1, 1, 0), (0, 1, 1), (1, 0, 1), (0, 0, 0)\}$  ters sırada yazılabilen koddur.

Dikkat edilirse  $\mathbf{A}$  ve  $\mathbf{B}$  kodu lineer değil,  $\mathbf{C}$  ve  $\mathbf{D}$  kodu lineerdir. Dikkat edilirse ters sırada yazılabilen kodlar lineer olmak zorunda değildir.

Şimdi, ters sırada yazılabilen devirli kodlar incelenecektir.

**Tanım 2.17**  $f(x) \in \mathbb{F}_q[x]$ , derecesi  $n$  olan bir polinom olsun.

$$f_R(x) := x^n f(x^{-1})$$

polinomuna,  $f(x)$  polinomunun resiprokal polinomu denir (Meyn 1990).

$c := (c_0, c_1, \dots, c_{n-1})$  kodkelimesinin ters sırada yazılışı olan  $c^R := (c_{n-1}, c_{n-2}, \dots, c_0)$  sıralı  $n$ -lisine karşılık gelen polinom  $c_R(x)$  ile gösterilirse bu polinom,

$$\begin{aligned} c_R(x) &= x^{n-1} c(x^{-1}) = x^{n-1} (c_0 + c_1 x^{-1} + \dots + c_{n-1} x^{-n+1}) \\ &= x^{n-1} (c_0 + c_1 \frac{1}{x} + \dots + c_{n-1} \frac{1}{x^{n-1}}) \\ &= x^{n-1} \left( \frac{c_0 x^{n-1} + c_1 x^{n-2} + \dots + c_{n-1}}{x^{n-1}} \right) \\ &= c_0 x^{n-1} + c_1 x^{n-2} + \dots + c_{n-1} \end{aligned}$$

şeklinde elde edilir.

Şimdi,  $\mathbf{C}$  bir  $[n, n - k]$  devirli kod ve  $g(x) := g_0 + g_1x + \cdots + g_{k-1}x^{k-1} + x^k \in \mathbb{F}_q[x]$  polinomu da  $\mathbf{C}$ 'nin üreteç polinomu olsun. Buradan,  $c(x) = g(x)j(x)$  ve  $\deg j \leq n - k - 1$

$$\begin{aligned} c_R(x) &= x^{n-1}g(x^{-1})j(x^{-1}) \\ &= x^k g(x^{-1})x^{n-k-1}j(x^{-1}) \end{aligned}$$

yazılır. Burada  $g_R(x) = x^k g(x^{-1})$  olduğuna dikkat edilmelidir. O halde,  $q^{n-k}$  farklı  $j(x)$  polinomu yazılabildiğinden,  $g_R(x)$  bir  $[n, n - k]$  devirli kodunu üretir.

**Tanım 2.18**  $g(x) \in \mathbb{F}_q[x]$  polinomu için,

$$g(x) = g_R(x)$$

ise  $g(x)$  polinomuna self-resiprokal polinom denir (Massey 1964).

O halde bir devirli kodun ne zaman ters sırada yazılabilen bir kod olduğu aşağıdaki teorem ile verilebilir.

**Teorem 2.17** Üreteç polinomu  $g(x)$  olan devirli kodun ters sırada yazılabilen bir kod olması için gerek ve yeter şart  $g(x) = g_R(x)$  olmasıdır (Massey 1964).

**Örnek 2.12** Örnek 2.10'da  $\mathbf{D}$ 'nin bir  $[5, 2]$  indirgenmez devirli kodu gösterilmiştir.  $\gamma$ , birin 5. primitif kökü ve  $4^r \equiv 1 \pmod{5}$  olacak biçimdeki en küçük pozitif tamsayı  $r = 2$  olduğundan,

$$\begin{aligned} h(x) &= (x - \gamma)(x - \gamma^4) \\ &= x^2 + \alpha^2 x + 1 \in \mathbb{F}_4[x] \end{aligned}$$

elde edilir. Teorem 2.16 'dan  $h_R(x)$  polinomu  $\mathbf{D}$  indirgenmez devirli kodu için bir kontrol polinomudur. Buradan  $\mathbf{D}$ 'nin üreteç polinomu,

$$\begin{aligned} g(x)(x^2 + \alpha^2 x + 1) &= x^5 - 1 \\ g(x) &= (x^5 - 1)/(x^2 + \alpha^2 x + 1) \\ g(x) &= x^3 + \alpha^2 x^2 + \alpha^2 x + 1 \end{aligned}$$

şeklinde bulunur.  $g_R(x) = g(x)$  olduğundan  $\mathbf{D}$ , ters sırada yazılabilen bir koddur.

**Örnek 2.13**  $\mathbf{D} = \{c_\beta := (Tr_4(\beta), Tr_4(\beta\gamma), Tr_4(\beta\gamma^2), \dots, Tr_4(\beta\gamma^{20})) | \beta \in \mathbb{F}_{64}\}$  bir  $[21, 3]$  indirgenmez devirli kodu için  $\gamma$ , birin 21. primitif kökü ve  $4^r \equiv 1 \pmod{21}$  olacak biçimdeki en küçük pozitif tamsayı  $r = 3$ 'tür. Buradan,

$$\begin{aligned} h(x) &= (x - \gamma)(x - \gamma^4)(x - \gamma^{4^2}) \\ &= x^3 + \alpha x^2 + 1 \in \mathbb{F}_4[x] \end{aligned}$$

olup  $\mathbf{D}$ 'nin üreteç polinomu

$$\begin{aligned} g(x) &= (x^{21} - 1)/(x^3 + \alpha x + 1) \\ &= x^{18} + \alpha x^{16} + x^{15} + \alpha^2 x^{14} + \alpha^2 x^{11} + x^9 + \alpha^2 x^8 + \alpha x^7 + \alpha x^4 + \alpha^2 x^2 + \alpha x + 1 \end{aligned}$$

elde edilir. Buradan  $g(x)$  polnomu self-resiprokal polinom olmadığından  $\mathbf{D}$ , ters sırada yazılabilen kod değildir.

Son bölümde görülecektir ki, ters sırada yazılabilen indirgenmez devirli kodlar (1), (3) ve (4) kısıtlamalarını sağlayan DNA kodları inşa etmek için, ters sırada yazılamayan indirgenmez devirli kodlar ise (1), (2), (3) ve (4) kısıtlamalarının tamamını sağlayan DNA kodları inşa etmek için kullanışlı olacaktır.

### 3. GAUSS TOPLAMLARI

Gauss toplamları, sayılar teorisindeki önemli enstrümanlardan biridir. Gauss toplamı esasında birin bazı köklerinin, çarpımlarının bir toplamıdır. Dolayısıyla Gauss toplamını tanımlamak ve Gauss toplamının özelliklerini geliştirmek için karakter adı verilen özel bir homomorfizm tipinin özelliklerini incelemek gereklidir. Bu bölüm iki kısma ayrılmıştır. Birinci kısımda toplamsal ve çarpımsal karakterler, ikinci kısımda ise Gauss toplamları tanımlanıp özellikleri incelenmiştir.

#### 3.1 Toplamsal ve Çarpımsal Karakterler

$\mathbb{F}_q$  üzerinde Gauss toplamları, karakter adı verilen özel bir grup homomorfizmi ile tanımlanır. Bir cisim göz önüne alındığında o cismin toplamsal ve çarpımsal grubundan söz edilebilir. Dolayısıyla  $\mathbb{F}_q$  cismi üzerinde iki tip karakter tanımlanmaktadır. Birincisi,  $\mathbb{F}_q$  cisminin toplamsal grubuna karşılık gelen toplamsal karakter, ikincisi ise  $\mathbb{F}_q$  cisminin çarpımsal grubuna karşılık gelen çarpımsal karakterdir.

**Tanım 3.1**  $\mathbb{F}_q$  bir sonlu cisim olsun.

$$\psi : \mathbb{F}_q \longrightarrow \mathbb{C}^*$$

homomorfizmine  $\mathbb{F}_q$  cisminin toplamsal karakteri denir (Lidl ve Niederreiter 1997).

Her  $\alpha \in \mathbb{F}_q$  için  $\psi_0(\alpha) := 1$  şeklinde tanımlanan  $\psi_0$  karakterine, aşık toplamsal karakter denir.

**Teorem 3.1** Toplamsal karakterler için aşağıdaki özellikler sağlanır (Lidl ve Niederreiter 1997).

- 1)  $\psi(0) = 1$ ,
- 2)  $\forall \alpha \in \mathbb{F}_q$  için  $1 = \psi^p(\alpha)$ , yani  $\psi(\alpha)$  birin  $p$ . köküdür.

Teorem 3.1 'deki özellikler homomorfizm tanımından gelmektedir.

**Örnek 3.1** Her  $\alpha \in \mathbb{F}_q$  için  $\psi_1(\alpha) := e^{2\pi i Tr(\alpha)/p}$  şeklinde tanımlanan fonksiyon,  $\mathbb{F}_q$ 'nin bir toplamsal karakteridir ve  $\mathbb{F}_q$ 'nin kanonik toplamsal karakteri diye adlandırılır.

**Uyarı 3.1** Toplamsal karakter tanımında kullanılan  $Tr$  dönüşümü  $\mathbb{F}_q$  sonlu cisiminden, onun asal cismi olan  $\mathbb{F}_p$  sonlu cismine tanımlanır.

Aşağıdaki teorem  $\mathbb{F}_q$  cisminin toplamsal karakterlerini verir.

**Teorem 3.2**  $\beta \in \mathbb{F}_q$  için

$$\begin{aligned}\psi_\beta : \mathbb{F}_q &\longrightarrow \mathbb{C}^* \\ \alpha &\longmapsto \psi_\beta(\alpha) := \psi_1(\beta\alpha)\end{aligned}$$

şekline tanımlanan fonksiyon,  $\mathbb{F}_q$  cisminin bir toplamsal karakteridir ve  $\mathbb{F}_q$ 'nin bütün toplamsal karakterleri bu şekilde elde edilir (Lidl ve Niederreiter 1997).

**Tanım 3.2**  $\mathbb{F}_q$  bir sonlu cisim olsun.

$$\chi : \mathbb{F}_q^* \longrightarrow \mathbb{C}^*$$

homomorfizmine  $\mathbb{F}_q$  cisminin çarpımsal karakteri denir (Lidl ve Niederreiter 1997).

Örneğin, her  $\alpha \in \mathbb{F}_q^*$  için  $\varepsilon(\alpha) := 1$  şeklinde tanımlanan  $\varepsilon$  homomorfizmi,  $\mathbb{F}_q$  cisminin bir çarpımsal karakteridir ve bu karaktere aşık çarpımsal karakter denir.

**Teorem 3.3** Çarpımsal karakterler için aşağıdaki özellikler sağlanır (Lidl ve Niederreiter 1997).

- 1)  $\chi(1) = 1$ ,
- 2)  $\forall \alpha \in \mathbb{F}_q^*$  için  $\chi(\alpha)^{q-1} = 1$ , yani  $\chi(\alpha)$  birin  $(q-1)$ . köküdür,
- 3)  $\chi(\alpha^{-1}) = \chi(\alpha)^{-1} = \overline{\chi(\alpha)}$ .

Teorem 3.4 'teki özellikler homomorfizm tanımından gelmektedir.

**Teorem 3.4**  $\mathbb{F}_q$  sonlu cisminin çarpımsal karakterlerinin kümesi bir devirli gruptur ve  $\hat{\mathbb{F}}_q$  ile gösterilir. Dahası,  $\mathbb{F}_q^*$  çarpımsal grubu ile  $\hat{\mathbb{F}}_q$  karakter grubu izomorftur (Lidl ve Niederreiter 1997).

Aşağıdaki teorem  $\mathbb{F}_q$  sonlu cisminin çarpımsal karakterlerini verir.

**Teorem 3.5**  $\theta, \mathbb{F}_q$  sonlu cisminin bir primitif elemanı olsun.

$j = 0, 1, \dots, q-1$  ve  $k = 0, 1, \dots, q-1$  için,

$$\begin{aligned}\chi_j : \mathbb{F}_q^* &\longrightarrow \mathbb{C}^* \\ \theta^k &\longmapsto \chi_j(\theta^k) := e^{2\pi i j k / (q-1)}\end{aligned}$$

şeklinde tanımlanan fonksiyon  $\mathbb{F}_q$  sonlu cisminin bir çarpımsal karakteridir ve  $\mathbb{F}_q$ 'nin bütün çarpımsal karakterleri bu şekilde elde edilir (Lidl ve Niederreiter 1997).

Şimdi, karakterler için ortogonalite bağıntısı olarak adlandırılan bir bağıntı aşağıda Teorem 3.6 ve Teorem 3.7 olarak verilecektir. Bu bağıntı, bir toplamsal karakterden çarpımsal karaktere ya da bir çarpımsal karakterden toplamsal karaktere geçiş bağıntısını elde etmede kullanılacaktır.

**Teorem 3.6**  $\alpha, \beta, \theta, \gamma \in \mathbb{F}_q$  ve  $\psi_\alpha, \psi_\gamma$  karakterleri de  $\mathbb{F}_q$  sonlu cisminin herhangi iki toplamsal karakteri olsun. Bu durumda aşağıdaki özellikler sağlanır (Lidl ve Niederreiter 1997).

- 1)  $\alpha \neq 0$  için  $\sum_{\beta \in \mathbb{F}_q} \psi_\alpha(\beta) = 0$ ,
- 2)  $\sum_{\beta \in \mathbb{F}_q} \psi_\alpha(\beta) \overline{\psi_\gamma(\beta)} = \begin{cases} 0 & , \alpha \neq \gamma \\ q & , \alpha = \gamma \end{cases}$ ,
- 3)  $\sum_{\alpha \in \mathbb{F}_q} \psi_\alpha(\beta) \overline{\psi_\alpha(\theta)} = \begin{cases} 0 & , \beta \neq \theta \\ q & , \beta = \theta \end{cases}$ .

**Teorem 3.7**  $\alpha, \beta \in \mathbb{F}_q^*$  ve  $\chi, \lambda \mathbb{F}_q$  sonlu cismi üzerinde herhangi iki çarpımsal karakter olsun. Bu durumda aşağıdaki özellikler sağlanır (Lidl ve Niederreiter 1997).

- 1)  $\varepsilon \neq \chi$  için  $\sum_{\beta \in \mathbb{F}_q^*} \chi(\beta) = 0$
- 2)  $\sum_{\beta \in \mathbb{F}_q^*} \chi(\beta) \overline{\lambda(\beta)} = \begin{cases} 0 & , \chi \neq \lambda \\ q-1 & , \chi = \lambda \end{cases}$

$$3) \sum_{\chi \in \hat{\mathbb{F}}_q} \chi(\beta) \overline{\chi(\alpha)} = \begin{cases} 0 & , \beta \neq \alpha \\ q-1 & , \beta = \alpha \end{cases}$$

### 3.2 Gauss Toplamları

Sonlu cisimler üzerindeki denklemlerin çözüm sayılarını bulmak için kullanılan en önemli araçlardan birisi Gauss toplamlarıdır. Kuadratik, kübik ve bikuadratik resiprositi kurallarının ispatı bu toplamlara dayanmaktadır. Bu kısımda Gauss toplamı, indirgenmez devirli kodların ağırlıklarının hesaplanmasında önem teşkil eden Gauss periyodu kavramını vermek için incelenecektir.

**Tanım 3.3**  $\mathbb{F}_q$  bir sonlu cisim,  $\chi$  ve  $\psi$  sırasıyla  $\mathbb{F}_q$ 'nin çarpımsal ve toplamsal karakterleri olsun. Bu durumda  $\alpha \in \mathbb{F}_q$  elemanı ve  $\chi \in \hat{\mathbb{F}}_q$  karakterinin  $\mathbb{F}_q$  üzerindeki Gauss toplamı,

$$g_\alpha(\chi) := \sum_{\beta \in \mathbb{F}_q^*} \chi(\beta) \psi(\alpha\beta)$$

şeklinde tanımlanır (Ireland ve Rosen 1990).

Buradan sonra  $\alpha = 1$  durumunda kısaca  $g(\chi)$  ile gösterilecektir.

**Teorem 3.8**  $\mathbb{F}_q$  cisminin bir çarpımsal karakteri  $\chi$  ve aşikar çarpımsal karakteri de  $\varepsilon$  ile gösterilsin. Bu durumda Gauss toplamı için aşağıdaki özellikler sağlanır.

1)  $\chi = \varepsilon$  ise

$$g_\alpha(\chi) = \begin{cases} -1 & , \alpha \neq 0 \\ q-1 & , \alpha = 0 \end{cases}$$

2)  $\chi \neq \varepsilon$  ise

$$g_\alpha(\chi) = \begin{cases} \overline{\chi(\alpha)} g(\chi) & , \alpha \neq 0 \\ 0 & , \alpha = 0 \end{cases}$$

Teorem 3.8, Gauss toplamının tanımından kolayca elde edilir (Lidl ve Niederreiter 1997).

**Teorem 3.9** Gauss toplamı için aşağıdaki özellikler sağlanır (Berndt vd. 1998).

- 1)  $g(\bar{\chi}) = \chi(-1)\overline{g(\chi)}$ ,
- 2)  $|g(\chi)| = \sqrt{q}$ ,
- 3)  $g(\chi)g(\bar{\chi}) = \chi(-1)q$ .

**Tanım 3.4**  $p$  bir asal sayı olsun.  $p^t \equiv -1 \pmod{e}$  olacak biçimde pozitif bir  $t$  tamsayısı varsa  $e$ 'ye self-eşlenik sayı denir (Hong vd. 2016).

**Teorem 3.10**  $\mathbb{F}_q$  karakteristiği  $p$  olan sonlu bir cisim,  $\chi$  karakteri  $\mathbb{F}_q$  cisminin mertebesi  $e$  olan bir çarpımsal karakteri ve  $e$  self-eşlenik sayı olsun. Bu durumda

$$g(\chi^p) = g(\chi) = g(\bar{\chi}) = g(\chi^{p^t})$$

eşitlikleri sağlanır (Hong vd. 2016).

**Teorem 3.11 (Hasse-Davenport Teoremi)**  $\chi$  ve  $\psi$  sırasıyla  $\mathbb{F}_q$ 'nin çarpımsal ve toplamsal karakterleri olsun ve  $\mathbb{F}_{q^r}$ 'nin  $\chi'$  ve  $\psi'$  karakterlerine çekilsin. Bu durumda,

$$g(\chi', \psi') = (-1)^{r-1} g(\chi, \psi)^r$$

eşitliği sağlanır (Lidl ve Niederreiter 1997).

**Örnek 3.2**  $\mathbb{F}_{16}$  cisminin mertebesi 3 olan çarpımsal karakterlerinin Gauss toplamlarını hesaplamak için  $\mathbb{F}_4$  cisminin mertebesi 3 olan çarpımsal karakterlerinin Gauss toplamlarını bulmak yeterlidir.  $\mathbb{F}_4 = \{0, 1, \alpha, \alpha^2\}$  olsun.

$\mathbb{F}_4$ 'ün çarpımsal karakterleri

$$\chi_k : \mathbb{F}_4^* \longrightarrow \mathbb{C}^*$$

$$\alpha \longmapsto \chi_k(\alpha) := \zeta_3^k, \quad \zeta_3 := e^{2\pi i/3}, \quad k = 0, 1, 2$$

ve toplamsal karakterleri

$$\psi_\beta : \mathbb{F}_4 \longrightarrow \mathbb{C}^*$$

$$\gamma \longmapsto \psi_\beta(\gamma) := \xi_2^{Tr(\beta\gamma)}, \quad \xi_2 := e^{2\pi i/2} = -1$$

şeklindedir.

$\beta \in \mathbb{F}_4$  için  $Tr(\beta) = \beta + \beta^2$  olduğundan

$$Tr(0) = 0,$$

$$Tr(1) = 0,$$

$$Tr(\alpha) = 1,$$

$$Tr(\alpha^2) = 1,$$

elde edilir.

Buradan  $\mathbb{F}_4$  üzerinde Gauss toplamı,

$$g(\varepsilon) = \sum_{\beta} \varepsilon(\beta) \psi(\beta) = \psi(1) + \psi(\alpha) + \psi(\alpha^2) = 1 - 1 - 1 = -1$$

$$g(\chi_1) = \chi_1(1)\psi(1) + \chi_1(\alpha)\psi(\alpha) + \chi_1(\alpha^2)\psi(\alpha^2) = 1 - \zeta_3 - \zeta_3^2 = 2$$

$$g(\chi_2) = \chi_2(1)\psi(1) + \chi_2(\alpha)\psi(\alpha) + \chi_2(\alpha^2)\psi(\alpha^2) = 1 - \zeta_3^2 - \zeta_3 = 2$$

olarak bulunur.

Burada  $\chi_1$  ve  $\chi_2$  karakterlerinin mertebesi 3'tür. Şimdi,  $\mathbb{F}_{16}$ 'nın  $\chi_1$  ve  $\chi_2$  karakterlerine karşılık gelen, mertebesi 3 olan karakterleri  $\chi_1^{(1)}$  ve  $\chi_2^{(2)}$  ile gösterilirse,

$$g(\chi_1^{(1)}) = (-1)^{2-1} g(\chi_1)^2 = (-1)2^2 = -4,$$

$$g(\chi_1^{(2)}) = (-1)^{2-1} g(\chi_2)^2 = (-1)2^2 = -4,$$

şeklinde hesaplanır.

**Uyarı 3.2** Yukarıdaki örnekte görüldüğü gibi  $\mathbb{F}_{16}$ 'nın mertebesi 5 olan bir çarpımsal karakterinin Gauss toplamını hesaplamak için  $\mathbb{F}_4$  üzerindeki Gauss toplamı kullanılamaz. Çünkü  $\mathbb{F}_4$ 'ün mertebesi 5 olan çarpımsal karakteri yoktur.

**Teorem 3.12**  $e$  self-eşlenik bir asal sayı,  $q = 4^t$  ve  $q' = 4^{rt}$  olmak üzere  $\chi$  karakteri  $\mathbb{F}_q$ 'nın,  $\chi'$  karakteri ise  $\mathbb{F}_{q'}$ 'nin mertebesi  $e$  olan birer çarpımsal karakteri olsun. Bu takdirde

$$1) \ g(\chi) = 2^t,$$

$$2) \ g(\chi') = 2^{rt}(-1)^{r-1},$$

özellikleri sağlanır (Hong vd. 2016).

**İspat.**

$$1) \ 4^t = |g(\chi)|^2 = g(\chi)\overline{g(\chi)} = g(\chi)^2 \Rightarrow g(\chi) = 2^t.$$

$$2) \ g(\chi') = (-1)^{r-1}g(\chi)^r = (-1)^{r-1}2^{rt}.$$

■



#### 4. GAUSS PERİYODU

Bu bölümde tezin en önemli aracı olan Gauss periyodu geliştirilecektir. Daha sonra, ilerde inşa edilecek DNA kodlarının parametrelerinin hesaplanabilmesi adına gerekli olan Gauss periyotlarını bulmak için MAGMA kodları verilecektir. Son olarak, elde edilen sonuçlar bir tabloda gösterilecektir.

**Tanım 4.1**  $q - 1 = e.n$ ,  $\mathbb{F}_q^* = \langle \theta \rangle$ ,  $\gamma = \theta^e$ ,  $\langle \gamma \rangle < \mathbb{F}_q^*$  olsun.  $\lambda = 0, 1, \dots, e - 1$  olmak üzere  $C_\lambda = \theta^\lambda \langle \gamma \rangle$  kosetlerine  $e$ -daireseel sınıf denir (Myerson 1981).

**Tanım 4.2**  $\psi$  karakteri  $\mathbb{F}_q$  cisminin kanonik toplamsal karakteri ve  $C_\lambda$  kümesi  $e$ -daireseel sınıf olmak üzere

$$\eta_\lambda := \sum_{\beta \in C_\lambda} \psi(\beta), \quad \lambda = 0, 1, \dots, e - 1$$

toplamına Gauss periyodu denir (Myerson 1981).

**Örnek 4.1**  $\theta \in \mathbb{F}_{16}$  primitif eleman ve  $\gamma = \theta^3$  ( $e = 3$ ) olsun. Bu durumda

$$\begin{aligned} C_0 &= \{\theta^3, \theta^6, \theta^9, \theta^{12}, 1\}, \\ C_1 &= \{\theta^4, \theta^7, \theta^{10}, \theta^{13}, \theta\}, \\ C_2 &= \{\theta^5, \theta^8, \theta^{11}, \theta^{14}, \theta^2\}, \end{aligned}$$

olur. Buradan Gauss periyotları,

$$\begin{aligned} \eta_0 &= \sum_{\beta \in C_0} \psi(\beta) = \psi(\theta^3) + \psi(\theta^6) + \psi(\theta^9) + \psi(\theta^{12}) + \psi(1) = -3, \\ \eta_1 &= \sum_{\beta \in C_1} \psi(\beta) = \psi(\theta^4) + \psi(\theta^7) + \psi(\theta^{10}) + \psi(\theta^{13}) + \psi(\theta) = 1, \\ \eta_2 &= \sum_{\beta \in C_2} \psi(\beta) = \psi(\theta^5) + \psi(\theta^8) + \psi(\theta^{11}) + \psi(\theta^{14}) + \psi(\theta^2) = 1. \end{aligned}$$

olarak hesaplanır.

**Teorem 4.1**  $q - 1 = e.n$  ve  $C_\lambda$  kümesi  $e$ -daireseel sınıf olmak üzere,

$$\sum_{\lambda=0}^{e-1} \eta_\lambda = -1$$

dir (Ding and Yang 2013).

**İspat.**  $\sum_{\lambda=0}^{e-1} \eta_\lambda = \sum_{\beta \in C_0} \psi(\beta) + \sum_{\beta \in C_1} \psi(\beta) + \cdots + \sum_{\beta \in C_{e-1}} \psi(\beta) = \sum_{\beta \in \mathbb{F}_q^*} \psi(\beta) = -1 \quad \blacksquare$

Gauss periyodunun tanımı oldukça anlaşılır ve kolay gibi görünse de esasında Gauss periyodunu hesaplamak oldukça zor bir problem olarak görülmektedir. Hasse-Davenport teoremi ve Gauss toplamı yardımı ile bu hesaplama kısmen kolaylaşmıştır. Bunun için Gauss toplamı ile Gauss periyodu arasında bir bağıntı verilecektir.

Öncelikle  $\mathbb{F}_q$  cisminin toplamsal ve çarpımsal karakterleri arasında;

$\alpha \in \mathbb{F}_q^*$  için

$$\begin{aligned} \psi(\alpha) &= \frac{1}{q-1} \sum_{\beta} \psi(\beta) \sum_{\chi} \chi(\beta) \overline{\chi(\alpha)} \\ &= \frac{1}{q-1} \sum_{\chi} \overline{\chi(\alpha)} \sum_{\beta} \psi(\beta) \chi(\beta) \\ &= \frac{1}{q-1} \sum_{\chi} \overline{\chi(\alpha)} g(\chi) \end{aligned}$$

şeklinde, çarpımsal karakterden toplamsal karaktere bir geçiş bağıntısı vardır.

Buradan hareketle Gauss toplamı ile Gauss periyodu arasındaki bağıntı verilebilir.

**Teorem 4.2**  $q - 1 = e.n$  ve  $\lambda = 0, 1, \dots, e-1$  için  $\eta_\lambda$ , Gauss periyodu olmak üzere,

$$\eta_\lambda = \frac{1}{e} \sum_{l=0}^{e-1} g(\chi^l) \overline{\chi(\theta^{\lambda l})}$$

eşitliği sağlanır (Ding and Yang 2013).

**İspat.**

$$\begin{aligned}
\eta_\lambda &= \psi(\theta^\lambda \theta^e) + \psi(\theta^\lambda \theta^{2e}) + \dots + \psi(\theta^\lambda \theta^{e(n-1)}) + \psi(\theta^\lambda) \\
&= \frac{1}{q-1} \sum_{\chi} \overline{\chi(\theta^\lambda \theta^e)} g(\chi) + \frac{1}{q-1} \sum_{\chi} \overline{\chi(\theta^\lambda \theta^{2e})} g(\chi) + \dots + \frac{1}{q-1} \sum_{\chi} \overline{\chi(\theta^\lambda)} g(\chi) \\
&= \frac{1}{q-1} \sum_{\chi} g(\chi) \overline{\chi(\theta^\lambda)} (1 + \overline{\chi(\theta^e)} + \overline{\chi(\theta^{2e})} + \dots + \overline{\chi(\theta^{(n-1)e})}) \\
&= \frac{1}{e} \sum_{\chi^e = \varepsilon} g(\chi) \overline{\chi(\theta^\lambda)} \\
&= \frac{1}{e} \sum_{l=0}^{e-1} g(\chi^l) \overline{\chi(\theta^{\lambda l})}
\end{aligned}$$

olarak elde edilir. Burada,

$$1 + \overline{\chi(\theta^e)} + \overline{\chi(\theta^{2e})} + \dots + \overline{\chi(\theta^{(n-1)e})} = \begin{cases} n & , \chi^e = \varepsilon \\ 0 & , \chi^e \neq \varepsilon \end{cases}$$

şeklindedir. ■

**Lemma 4.1**  $e \neq 2$  bir self-eşlenik asal sayı ise  $\eta_1 = \eta_2 = \dots = \eta_{e-1}$ 'dir (Hong vd. 2016).

**İspat.**

$$\eta_1 = \frac{1}{e} (1 + g(\chi)(\bar{\chi}(\theta) + \bar{\chi}(\theta^2) + \dots + \bar{\chi}(\theta^{e-1})))$$

ve  $\lambda = 2, 3, \dots, e-1$  için

$$\eta_\lambda = \frac{1}{e} (1 + g(\chi)(\bar{\chi}(\theta^\lambda) + \bar{\chi}(\theta^{2\lambda}) + \dots + \bar{\chi}(\theta^{(e-1)\lambda})))$$

şeklinde ifade edilebilir.

Ayrıca  $\{0, 1, \dots, e-1\}$  kümesi bir tam kalan sistemi ve  $(\lambda, e) = 1$  olduğundan  $\{0, \lambda, \dots, \lambda(e-1)\}$  kümesi de bir tam kalan sistemidir.

Dolayısıyla  $\bar{\chi}(\theta^\lambda) + \bar{\chi}(\theta^{2\lambda}) + \dots + \bar{\chi}(\theta^{(e-1)\lambda})$  toplamı,  $\bar{\chi}(\theta) + \bar{\chi}(\theta^2) + \dots + \bar{\chi}(\theta^{e-1})$  toplamının sadece terimlerinin yerleri değişmiş halidir.

Sonuç olarak  $\eta_1 = \eta_2 = \dots = \eta_{e-1}$ 'dir. ■

**Uyarı 4.1** O halde  $e \neq 2$  bir self eşlenik asal sayı ise sadece  $\eta_0$  ve  $\eta_1$  periyotlarının hesaplanması yeterlidir.

**Teorem 4.3**  $e$  bir self-eşlenik asal sayı,  $q = 4^t$ ,  $q' = 4^r$  ve  $q' - 1 = e.n$  olmak üzere  $\chi$  karakteri  $\mathbb{F}_{q'}$  cisminin mertebesi  $e$  olan bir çarpımsal karakteri olsun. Bu durumda,

$$1) \quad \eta_0 = \frac{1}{e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1}),$$

$$2) \quad \eta_1 = \frac{1}{e}(-1 + 2^r(-1)^{\frac{r}{t}}),$$

eşitlikleri sağlanır (Hong vd. 2016).

**İspat.**

1)

$$\begin{aligned} \eta_0 &= \frac{1}{e} \sum_{l=0}^{e-1} g(\chi^l) = \frac{1}{e}(-1 + g(\chi) + g(\chi^2) + \cdots + g(\chi^{e-1})) \\ &= \frac{1}{e}(-1 + (e-1)(2^t)^{\frac{r}{t}}(-1)^{\frac{r}{t}-1}) \\ &= \frac{1}{e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1}) \end{aligned}$$

2)

$$\begin{aligned} \eta_1 &= \frac{1}{e} \sum_{l=0}^{e-1} g(\chi^l) \overline{\chi^l(\theta)} = \frac{1}{e}(-1 + g(\chi) \overline{\chi(\theta)} + g(\chi^2) \overline{\chi^2(\theta)} + \cdots + g(\chi^{e-1}) \overline{\chi^{e-1}(\theta)}) \\ &= \frac{1}{e}(-1 + g(\chi)(\overline{\chi(\theta)} + \overline{\chi^2(\theta)} + \cdots + \overline{\chi^{e-1}(\theta)})) \\ &= \frac{1}{e}(-1 + (2^t)^{\frac{r}{t}}(-1)^{\frac{r}{t}-1}(-1)) \\ &= \frac{1}{e}(-1 + 2^r(-1)^{\frac{r}{t}}) \end{aligned}$$

■

Aşağıdaki magma kodları ile bazı  $\mathbb{F}_{4^r}$  sonlu cisimleri için Gauss periyotları hesaplanıp, Çizelge 4.1’de verilmiştir.

Şimdi,  $4^r - 1 = e.n$  ve  $2^t \equiv -1 \pmod{e}$  olacak biçimdeki en küçük pozitif  $t$  tamsayısı için  $e$  self-eşlenik asal sayı olsun.

```

eta0:=func<e,r,t|1/e*(-1+(e-1)*2^r*(-1)^(r/t-1))>;
eta0(e,r,t);
eta1:=func<e,r,t|1/e*(-1+2^r*(-1)^(r/t))>;
eta1(e,r,t);

```

Burada bazı  $t$  deęerleri kabaca, ařaęıdaki magma kodları ile elde edilmiřtir.

```

k:={1..100 by 1};
for t in k do
if 2^t mod e eq e-1 then
t;
end if;
end for;

```

Daha sonra inřa edilecek indirgenmez devirli kodların aęırlıklarının hesaplanmasında gerekli olan Gauss periyotları ařaęıda verilmiřtir.

Çizelge 4.1. Gauss Periyotları

| e  | r | t | $\eta_0$ | $\eta_1$ |
|----|---|---|----------|----------|
| 3  | 2 | 1 | -3       | 1        |
| 3  | 3 | 1 | -3       | 1        |
| 3  | 4 | 1 | -11      | 5        |
| 5  | 4 | 2 | -13      | 3        |
| 3  | 5 | 1 | 21       | -11      |
| 11 | 5 | 5 | 29       | -3       |
| 5  | 6 | 2 | 51       | -13      |
| 13 | 6 | 6 | 59       | -5       |

## 5. DNA KODLARININ İNŞASI

Birinci bölümde, inşa edilecek olan DNA kodlarının sağlaması istenen kısıtlamalar sebepleri ile açıklanmıştır. Bu bölümde ilk olarak hangi kısıtlamanın göz önüne alınacağı, akabinde diğer kısıtlamaların nasıl dahil edileceği ve bu kısıtlamaları sağlayan DNA kodlarının nasıl inşa edileceği gösterilecektir. Daha sonra, arzu edilen DNA kodları inşası için MAGMA programı dilindeki kodlar ile bazı örnekler verilecektir. Burada uygulanacak atağın kilit noktası önceden de belirtildiği gibi, indirgenmez devirli kodların ağırlıklarının hesaplanmasıdır. Bu metot ile hem DNA kodlarını inşa etmek hem de inşa edilen DNA kodlarının parametrelerini hesaplamak çok kolaydır.

DNA kodu inşa etmeden önce,  $\mathbb{F}_4$  cisminden  $S = \{A, C, G, T\}$  kümesine bir eşleme yapılmalıdır. Bu, DNA'nın yapıtaşlarına sonlu cisim elemanı gözüyle bakarak, arzu edilen DNA kodlarını inşa etmek için gereklidir. 24 olası eşleme mevcuttur; ancak burada tamamen iyi bir GC-içeriğine erişebilme amacına yönelik bir eşleme yapılacaktır. Dolayısıyla,

$$A \longrightarrow 0,$$

$$C \longrightarrow \alpha,$$

$$G \longrightarrow \alpha^2,$$

$$T \longrightarrow 1,$$

şeklinde bir eşleme göz önüne alınacaktır (Hong vd. 2016).

Teorem 2.16'dan

$$\mathbf{D} = \{c_\beta = (Tr_4(\beta), Tr_4(\beta\gamma), \dots, Tr_4(\beta\gamma^{n-1})) \mid \beta \in \mathbb{F}_{4^r}\}$$

kümesinin  $\mathbb{F}_4$  üzerinde bir  $[n, r]$  indirgenmez devirli kodu olduğu biliniyor. Açıkça  $\mathbf{D}$ 'nin kodkelimleri,  $c_0, c_1, c_\theta, \dots, c_{\theta^{e-1}}$  ile bunların devirsel kaymalarıdır. Diğer yandan,  $\mathbb{F}_4$  üzerinde bir indirgenmez devirli kodun herhangi bir kodkelimindeki  $a \in \mathbb{F}_4$  elemanının bulunduğu yerlerin sayısı aşağıdaki teorem ile kolaylıkla elde edilebilmektedir.

**Teorem 5.1**  $D$  kodu  $\mathbb{F}_4$  sonlu cismi üzerinde bir  $[n, r]$  indirgenmez devirli kod olsun.  $\beta := \theta^\lambda$ ,  $\lambda = 0, 1, \dots, e-1$ ,  $\frac{4^r-1}{3} \equiv l \pmod{e}$  olmak üzere  $N_\beta(a)$  ile  $c_\beta$ 'daki  $a \in \mathbb{F}_4$  elemanının sayısı gösterilsin. Bu durumda,

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_\lambda + \eta_{l+\lambda} + \eta_{2l+\lambda}), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_\lambda - \eta_{l+\lambda} - \eta_{2l+\lambda}), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_\lambda - \eta_{l+\lambda} + \eta_{2l+\lambda}), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_\lambda + \eta_{l+\lambda} - \eta_{2l+\lambda}). \end{aligned}$$

şeklinde hesaplanır (Hong vd. 2016, McEliece 1975).

**Sonuç 5.1**  $4^r - 1 = e.n$  olsun. Burada  $3|4^r - 1$  'dir. Buradan ya  $e = 3$  ya da  $e \neq 3$  olur.  $e \neq 3$  ise  $l = 0$ ,  $e = 3$  ise  $l = 0, 1, 2$  olabilir.

O halde aşağıdaki sonuçlar elde edilir (Hong vd. 2016).

1)  $e \neq 3$  ve  $\lambda = 0$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_0 + \eta_0 + \eta_0) = \frac{1}{4}(n + 3\eta_0), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_0 - \eta_0 - \eta_0) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_0 - \eta_0 + \eta_0) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_0 + \eta_0 - \eta_0) = \frac{1}{4}(n - \eta_0). \end{aligned}$$

2)  $e \neq 3$  ve  $\lambda = 1, 2, \dots, e-1$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_1 + \eta_1 + \eta_1) = \frac{1}{4}(n + 3\eta_1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_1 - \eta_1 - \eta_1) = \frac{1}{4}(n - \eta_1), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_1 - \eta_1 + \eta_1) = \frac{1}{4}(n - \eta_1), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_1 + \eta_1 - \eta_1) = \frac{1}{4}(n - \eta_1). \end{aligned}$$

3)  $e = 3, l = 0$  ve  $\lambda = 0$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_0 + \eta_0 + \eta_0) = \frac{1}{4}(n + 3\eta_0), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_0 - \eta_0 - \eta_0) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_0 - \eta_0 + \eta_0) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_0 + \eta_0 - \eta_0) = \frac{1}{4}(n - \eta_0). \end{aligned}$$

4)  $e = 3, l = 0$  ve  $\lambda = 1, 2$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_1 + \eta_1 + \eta_1) = \frac{1}{4}(n + 3\eta_1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_1 - \eta_1 - \eta_1) = \frac{1}{4}(n - \eta_1), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_1 - \eta_1 + \eta_1) = \frac{1}{4}(n - \eta_1), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_1 + \eta_1 - \eta_1) = \frac{1}{4}(n - \eta_1). \end{aligned}$$

5)  $e = 3, l = 1$  ve  $\lambda = 0$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_0 + \eta_1 + \eta_2) = \frac{1}{4}(n - 1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_0 - \eta_1 - \eta_2) = \frac{1}{4}(n + \eta_0 - 2\eta_1), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_0 - \eta_1 + \eta_2) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_0 + \eta_1 - \eta_2) = \frac{1}{4}(n - \eta_0). \end{aligned}$$

6)  $e = 3, l = 1$  ve  $\lambda = 1$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_1 + \eta_2 + \eta_0) = \frac{1}{4}(n - 1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_1 - \eta_2 - \eta_0) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_1 - \eta_2 + \eta_0) = \frac{1}{4}(n + \eta_0 - 2\eta_1), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_1 + \eta_2 - \eta_0) = \frac{1}{4}(n - \eta_0). \end{aligned}$$

7)  $e = 3, l = 1$  ve  $\lambda = 2$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_2 + \eta_0 + \eta_1) = \frac{1}{4}(n - 1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_2 - \eta_0 - \eta_1) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_2 - \eta_0 + \eta_1) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_2 + \eta_0 - \eta_1) = \frac{1}{4}(n + \eta_0 - 2\eta_1). \end{aligned}$$

8)  $e = 3, l = 2$  ve  $\lambda = 0$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_0 + \eta_2 + \eta_1) = \frac{1}{4}(n - 1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_0 - \eta_2 - \eta_1) = \frac{1}{4}(n + \eta_0 - 2\eta_1), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_0 - \eta_2 + \eta_1) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_0 + \eta_2 - \eta_1) = \frac{1}{4}(n - \eta_0). \end{aligned}$$

9)  $e = 3, l = 2$  ve  $\lambda = 1$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_1 + \eta_0 + \eta_2) = \frac{1}{4}(n - 1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_1 - \eta_0 - \eta_2) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_1 - \eta_0 + \eta_2) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_1 + \eta_0 - \eta_2) = \frac{1}{4}(n + \eta_0 - 2\eta_1). \end{aligned}$$

10)  $e = 3, l = 2$  ve  $\lambda = 2$  için;

$$\begin{aligned} N_\beta(0) &= \frac{1}{4}(n + \eta_2 + \eta_1 + \eta_0) = \frac{1}{4}(n - 1), \\ N_\beta(1) &= \frac{1}{4}(n + \eta_2 - \eta_1 - \eta_0) = \frac{1}{4}(n - \eta_0), \\ N_\beta(\alpha) &= \frac{1}{4}(n - \eta_2 - \eta_1 + \eta_0) = \frac{1}{4}(n + \eta_0 - 2\eta_1), \\ N_\beta(\alpha^2) &= \frac{1}{4}(n - \eta_2 + \eta_1 - \eta_0) = \frac{1}{4}(n - \eta_0). \end{aligned}$$

## 5.1 Hamming Uzaklığı ve GC-İçeriği Kısıtlamalarını Sağlayan DNA Kodlarının İnşası

Herhangi bir kodkelimesindeki sembollerin görünme sayıları belirlenebildiğinden, ilk olarak (1) ve (4) kısıtlamalarını sağlayan DNA kodları elde etmeye çalışmak akla

yatkın olacaktır.

Sonuç 5.1 'den,  $c_1$  kodkelimesinin  $GC$ -içeriği,

$$w_1 := N_{\theta^0}(\alpha) + N_{\theta^0}(\alpha^2) = \frac{1}{2}(n - \eta_0)$$

ve  $\lambda \neq 0$  olmak üzere  $c_{\theta^\lambda}$  kodkelimelerinin  $GC$ -içeriği,

$$w_2 := N_{\theta^\lambda}(\alpha) + N_{\theta^\lambda}(\alpha^2) = \frac{1}{2}(n - \eta_1)$$

olarak bulunur. Dolayısıyla indirgenmez devirli  $\mathbf{D}$  kodunun iki alt kümesi göz önüne alınır. Bunlar,

$$\mathbf{C}_1 := \{c_{\gamma^i} \in \mathbf{D} \mid i = 0, 1, \dots, n-1\}$$

ve

$$\mathbf{C}_2 := \{c_{\theta^j \gamma^i} \in \mathbf{D} \mid j = 1, 2, \dots, e-1, i = 0, 1, \dots, n-1\}$$

kodlarıdır. O halde (1) ve (4) kısıtlamalarını sağlayan DNA kodları için aşağıdaki teorem verilebilir.

**Teorem 5.2**  $\mathbf{D}$  bir  $[n, r]$  indirgenmez devirli kod ve  $2^t \equiv -1 \pmod{e}$  olacak biçimdeki en küçük pozitif tamsayı  $t$  olsun. Bu durumda (1) ve (4) şartlarını sağlayan DNA kodları,  $l \equiv \frac{q-1}{3} \pmod{e}$  olmak üzere  $\mathbf{C}_1$ , bir  $(n, n, d = \min\{d_1, d_2\}, w_1)_4$  kodu ve  $\mathbf{C}_2$ , bir  $(n, (e-1)n, d = \min\{d_1, d_2\}, w_2)$  kodu olup bu kodların parametreleri

1)  $l \equiv 0 \pmod{e}$  ise,

$$\begin{aligned} w_1 &= \frac{n}{2} - \frac{1}{2e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1}), \\ w_2 &= \frac{n}{2} - \frac{1}{2e}(-1 + 2^r(-1)^{\frac{r}{t}}), \\ d_1 &= \frac{3n}{4} - \frac{3}{4e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1}), \\ d_2 &= \frac{3n}{4} - \frac{3}{4e}(-1 + 2^r(-1)^{\frac{r}{t}}). \end{aligned}$$

2)  $l \equiv 1, 2 \pmod{e}$  ise,

$$\begin{aligned} w_1 &= \frac{n}{2} - \frac{1}{2e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1}), \\ w_2 &= \frac{n}{2} - \frac{1}{2e}(-1 + 2^r(-1)^{\frac{r}{t}}), \\ d_1 &= d_2 = \frac{3n}{4} + \frac{1}{4}, \end{aligned}$$

şeklinde (Hong vd. 2016).

**İspat.**  $c_{\theta^0}$  kodkelimesi ve onun devirsel kaymalarının oluşturduğu kod  $\mathbf{C}_1$  ile,  $\lambda = 1, 2, \dots, e-1$  için  $c_{\theta^\lambda}$  kodkelimeleri ve onların devirsel kaymalarının oluşturduğu kod  $\mathbf{C}_2$  ile gösterilsin.  $\mathbf{C}_1$  kodu  $n$  tane kodkelimesine,  $\mathbf{C}_2$  kodu ise  $(e-1)n$  tane kodkelimesine sahiptir. Sonuç 5.1'den,  $\lambda = 0$  için

$$N_{\theta^0}(\alpha) = \frac{1}{4}(n - \eta_0 - \eta_l + \eta_{2l}) \text{ ve } N_{\theta^0}(\alpha^2) = \frac{1}{4}(n - \eta_0 + \eta_l - \eta_{2l})$$

olup buradan

$$w_1 = N_{\theta^0}(\alpha) + N_{\theta^0}(\alpha^2) = \frac{1}{2}(n - \eta_0) = \frac{n}{2} - \frac{1}{2e}(-1 + (e-1)2^r(-1)^{\frac{r}{e}-1})$$

elde edilir.

$\lambda \neq 0$  için,

$$N_{\theta^\lambda}(\alpha) = \frac{1}{4}(n - \eta_\lambda - \eta_{\lambda+l} + \eta_{\lambda+2l}) \text{ ve } N_{\theta^\lambda}(\alpha^2) = \frac{1}{4}(n - \eta_\lambda + \eta_{\lambda+l} - \eta_{\lambda+2l})$$

olup buradan da

$$w_2 = N_{\theta^\lambda}(\alpha) + N_{\theta^\lambda}(\alpha^2) = \frac{1}{2}(n - \eta_1) = \frac{n}{2} - \frac{1}{2e}(-1 + 2^r(-1)^{\frac{r}{e}})$$

olur. Sonuç olarak  $w_1$  ve  $w_2$  sabittir.  $c_{\theta^0}$  kodkelimesindeki sıfırdan farklı yerlerin sayısı  $d_1$  ile,  $c_{\theta^\lambda}$  kodkelimesindeki sıfırdan farklı yerlerin sayısı da  $d_2$  ile gösterilsin.

Yani  $d_1 := n - N_{\theta^0}(0)$  ve  $d_2 := n - N_{\theta^\lambda}(0)$  olsun.  $l \equiv 0 \pmod{e}$  ise,

$$\begin{aligned} d_1 &= n - N_{\theta^0}(0) = n - \frac{1}{4}(n + 3\eta_0) = \frac{3n}{4} - \frac{3}{4e}(-1 + (e-1)2^r(-1)^{\frac{r}{e}-1}) \\ d_2 &= n - N_{\theta^\lambda}(0) = n - \frac{1}{4}(n + 3\eta_1) = \frac{3n}{4} - \frac{3}{4e}(-1 + 2^r(-1)^{\frac{r}{e}}) \end{aligned}$$

$l \equiv 1, 2 \pmod{e}$  ise

$$N_{\theta^0}(0) = N_{\theta^\lambda}(0) = \frac{1}{4}(n - 1)$$

olduğundan

$$d_1 = d_2 = \frac{3n}{4} + \frac{1}{4}$$

olur.  $\mathbf{C}_1$  ve  $\mathbf{C}_2$ , (1) ve (4) koşullarını sağlayan DNA kodlarıdır. ■

Aşağıdaki Magma kodları kullanılarak  $\mathbf{C}_1$  ve  $\mathbf{C}_2$  DNA kodlarının parametreleri kolayca elde edilir.  $l \equiv 0 \pmod{e}$  ise,

```

eta0:=func<e,r,t|1/e*(-1+(e-1)*2^r*(-1)^(r/t-1))>;
eta0(e,r,t);
eta1:=func<e,r,t|1/e*(-1+2^r*(-1)^(r/t))>;
eta1(r,r,t);
w1:=func<n,e|n/2-1/2*eta0(e,r,t)>;
w1(n,e);
w2:=func<n,e|n/2-1/2*eta1(e,r,t)>;
w2(n,e);
d1:=func<n,e|3*n/4-3/4*eta0(e,r,t)>;
d1(n,e);
d2:=func<n,e|3*n/4-3/4*eta1(e,r,t)>;
d2(n,e);

l ≡ 1,2 (mod e) ise,

eta0:=func<e,r,t|1/e*(-1+(e-1)*2^r*(-1)^(r/t-1))>;
eta0(e,r,t);
eta1:=func<e,r,t|1/e*(-1+2^r*(-1)^(r/t))>;
eta1(e,r,t);
w1:=func<n,e|n/2-1/2*eta0(e,r,t)>;
w1(n,e);
w2:=func<n,e|n/2-1/2*eta1(e,r,t)>;
w2(n,e);
d:=func<n|3*n/4+1/4>;
d(n);

```

(1) ve (4) kısıtlamalarını sağlayan DNA kodlarının parametreleri aşağıdaki çizelgede verilmiştir:

Çizelge 5.1. (1) ve (4) kısıtlamalarını sağlayan DNA kodları

| e  | n   | $M(\mathbf{C}_1/\mathbf{C}_2)$ | $d(\mathbf{C}_1/\mathbf{C}_2)$ | $w(\mathbf{C}_1/\mathbf{C}_2)$ |
|----|-----|--------------------------------|--------------------------------|--------------------------------|
| 3  | 5   | 5/10                           | 4/4                            | 4/2                            |
| 3  | 21  | 21/42                          | 12/12                          | 8/12                           |
| 5  | 51  | 51/204                         | 36/36                          | 32/24                          |
| 3  | 85  | 85/170                         | 64/64                          | 48/40                          |
| 3  | 341 | 341/682                        | 256/256                        | 160/176                        |
| 11 | 93  | 93/930                         | 48/48                          | 32/48                          |
| 13 | 315 | 315/3780                       | 192/192                        | 128/160                        |
| 5  | 819 | 819/3276                       | 576/576                        | 384/416                        |

## 5.2 Hamming Uzaklığı, Ters-Eşlenik ve GC-İçeriği Kısıtlamalarını Sağlayan DNA Kodlarının İnşası

Eğer  $\mathbf{D}$  indirgenmez devirli kodunun üreteç polinomu self-resiprokal ise yani  $\mathbf{D}$  ters sırada yazılabilen bir kod ise her  $c \in \mathbf{D}$  kodkelimesi için  $c^R \in \mathbf{D}$  olacağından ters sırada yazılış kısıtlaması sağlanmaz. Bu durumda (1), (3) ve (4) kısıtlamalarını sağlayan DNA kodları,

$$\overline{\mathbf{C}}_1^R := \{\overline{c_{\gamma^i}}^R \mid i = 0, 1, \dots, n-1\}$$

ve

$$\overline{\mathbf{C}}_2^R := \{\overline{c_{\theta^j \gamma^i}}^R \mid j = 1, 2, \dots, e-1, i = 0, 1, \dots, n-1\}$$

olmak üzere,

$$\mathbf{C}_3 := \mathbf{C}_1 \cup \overline{\mathbf{C}}_1^R \text{ ve } \mathbf{C}_4 := \mathbf{C}_2 \cup \overline{\mathbf{C}}_2^R$$

kodlarıdır.

**Örnek 5.1**  $f(x) = x^2 + x + \alpha \in \mathbb{F}_4[x]$  primitif polinom,  $\mathbb{F}_{16}^* = \langle \theta \rangle$  ve  $\gamma := \theta^3$  olsun.  $r = 2$  ve  $e = 3$  olmak üzere

$$\mathbf{D} = \{c_\beta = (Tr_4(\beta), Tr_4(\beta\gamma), Tr_4(\beta\gamma^2), Tr_4(\beta\gamma^3), Tr_4(\beta\gamma^4)) \mid \beta \in \mathbb{F}_{16}\}$$

$[5, 2]$  indirgenmez devirli kodu,

$$c_0 = (0, 0, 0, 0, 0),$$

$$c_{\theta^0} = (0, \alpha^2, \alpha, \alpha, \alpha^2),$$

$$c_{\theta^1} = (1, 1, \alpha, 0, \alpha),$$

$$c_{\theta^2} = (1, 0, 1, \alpha^2, \alpha^2),$$

kodkelimleri ile bunların devirsel kaymalarının kümesidir.  $\mathbf{D}$ 'nin iki alt kümesi,

$$\mathbf{C}_1 := \{c_{\gamma^i} \in \mathbf{D} \mid i = 0, 1, 2, 3, 4\}$$

ve

$$\mathbf{C}_2 := \{c_{\theta^j \gamma^i} \in \mathbf{D} \mid j = 1, 2, i = 0, 1, 2, 3, 4\}$$

olsun. Teorem 5.2 gereği,  $\frac{q-1}{3} = 5 \equiv 2 \pmod{3}$  olup parametreler

$$w_1 = \frac{n}{2} - \frac{1}{2e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1})$$

$$w_2 = \frac{n}{2} - \frac{1}{2e}(-1 + 2^r(-1)^{\frac{r}{t}})$$

$$d_1 = d_2 = \frac{3n}{4} + \frac{1}{4}$$

şeklindedir.  $n = 5$ ,  $e = 3$ ,  $r = 2$  ve  $t = 1$  parametrelerdeki yerlerine yazılırsa,  $w_1 = 4$ ,  $w_2 = 2$ ,  $d_1 = d_2 = 4$  elde edilir. Buradan (1) ve (4) şartlarını sağlayan DNA kodları,

$$\mathbf{C}_1 = \begin{bmatrix} 0 & \alpha^2 & \alpha & \alpha & \alpha^2 \\ \alpha^2 & 0 & \alpha^2 & \alpha & \alpha \\ \alpha & \alpha^2 & 0 & \alpha^2 & \alpha \\ \alpha & \alpha & \alpha^2 & 0 & \alpha^2 \\ \alpha^2 & \alpha & \alpha & \alpha^2 & 0 \end{bmatrix}$$

ve

$$\mathbf{C}_2 = \begin{bmatrix} 1 & 1 & \alpha & 0 & \alpha \\ \alpha & 1 & 1 & \alpha & 0 \\ 0 & \alpha & 1 & 1 & \alpha \\ \alpha & 0 & \alpha & 1 & 1 \\ 1 & \alpha & 0 & \alpha & 1 \\ 1 & 0 & 1 & \alpha^2 & \alpha^2 \\ \alpha^2 & 1 & 0 & 1 & \alpha^2 \\ \alpha^2 & \alpha^2 & 1 & 0 & 1 \\ 1 & \alpha^2 & \alpha^2 & 1 & 0 \\ 0 & 1 & \alpha^2 & \alpha^2 & 1 \end{bmatrix}$$

kodları olup,  $\mathbf{C}_1$  bir (5, 5, 4, 4) kodu,  $\mathbf{C}_2$  bir (5, 10, 4, 2) kodudur. Diğer taraftan Örnek 2.12'de  $\mathbf{D}$  'nin ters sırada yazılabilen bir kod olduğu gösterilmişti. Dolayısıyla ters sırada yazılış kısıtlaması sağlanamayacaktır. Şimdi,

$$\overline{\mathbf{C}}_1^R := \{\overline{c_{\gamma^i}}^R \mid i = 0, 1, 2, 3, 4\}$$

ve

$$\overline{\mathbf{C}}_2^R := \{\overline{c_{\theta^j \gamma^i}}^R \mid j = 1, 2, i = 0, 1, 2, 3, 4\}$$

olsun. Buradan (1), (3) ve (4) şartlarını sağlayan yeni DNA kodları  $\mathbf{C}_3 := \mathbf{C}_1 \cup \overline{\mathbf{C}}_1^R$  ve  $\mathbf{C}_4 := \mathbf{C}_2 \cup \overline{\mathbf{C}}_2^R$  şeklinde elde edilir.

$$\mathbf{C}_3 = \begin{bmatrix} 0 & \alpha^2 & \alpha & \alpha & \alpha^2 \\ \alpha^2 & 0 & \alpha^2 & \alpha & \alpha \\ \alpha & \alpha^2 & 0 & \alpha^2 & \alpha \\ \alpha & \alpha & \alpha^2 & 0 & \alpha^2 \\ \alpha^2 & \alpha & \alpha & \alpha^2 & 0 \\ \alpha & \alpha^2 & \alpha^2 & \alpha & 1 \\ 1 & \alpha & \alpha^2 & \alpha^2 & \alpha \\ \alpha & 1 & \alpha & \alpha^2 & \alpha^2 \\ \alpha^2 & \alpha & 1 & \alpha & \alpha^2 \\ \alpha^2 & \alpha^2 & \alpha & 1 & \alpha \end{bmatrix}$$

ve

$$C_4 = \begin{bmatrix} 1 & 1 & \alpha & 0 & \alpha \\ \alpha & 1 & 1 & \alpha & 0 \\ 0 & \alpha & 1 & 1 & \alpha \\ \alpha & 0 & \alpha & 1 & 1 \\ 1 & \alpha & 0 & \alpha & 1 \\ 1 & 0 & 1 & \alpha^2 & \alpha^2 \\ \alpha^2 & 1 & 0 & 1 & \alpha^2 \\ \alpha^2 & \alpha^2 & 1 & 0 & 1 \\ 1 & \alpha^2 & \alpha^2 & 1 & 0 \\ 0 & 1 & \alpha^2 & \alpha^2 & 1 \\ \alpha^2 & 1 & \alpha^2 & 0 & 0 \\ 0 & \alpha^2 & 1 & \alpha^2 & 0 \\ 0 & 0 & \alpha^2 & 1 & \alpha^2 \\ \alpha^2 & 0 & 0 & \alpha^2 & 1 \\ 1 & \alpha^2 & 0 & 0 & \alpha^2 \\ \alpha & \alpha & 0 & 1 & 0 \\ 0 & \alpha & \alpha & 0 & 1 \\ 1 & 0 & \alpha & \alpha & 0 \\ 0 & 1 & 0 & \alpha & \alpha \\ \alpha & 0 & 1 & 0 & \alpha \end{bmatrix}$$

Bu kodların minimum uzaklıkları MAGMA Computer Algebra kullanılarak aşağıdaki gibi elde edilebilir:

$C_3$  kodunun minimum uzaklığı:

1)  $\mathbb{F}_{16}$  cismi ve  $Tr_4$  dönüşümü tanımlanır. Burada  $g$  primitif elemanı göstermektedir.

```
F<g>:=GF(16);
```

```
Fset:={x: x in F};
```

```
Tr4:=func<t|t+t^4>;
```

2)  $\gamma$ ,  $\gamma^2$ ,  $\gamma^3$ ,  $\gamma^4$  ve 1 dizi terimleri olarak atanır. Buradan, indirgenmez devirli  $\mathbf{D}$  kodundaki  $c_{\theta^0}$  ile  $\overline{c_{\theta^0}}^R$  kodkelimleri ve bunların devirsel kaymaları elde edilir.

```
Fseq=[x^3:x in Fset];
S:={1..5 by 1};
L:={0..4 by 1};
a:=func<l|[Tr4(g^(3*1)*Fseq[n]):n in S]>;
b:=func<l|[Tr4(g^(3*1)*Fseq[6-n])+1:n in S]>;
c0set:={a(l): l in L};
c0:=[c:c in c0set];
c0RCset:={b(l): l in L};
c0RC:=[c:c in c0RCset];
```

Burada "c0set" ile  $C_1$  kodu, "c0RCset" ile de  $\overline{C_1}^R$  kodu gösterilmektedir.

3)  $C_3 := C_1 \cup \overline{C_1}^R$  kodu oluşturulur.

```
C3:= c0set join c0RCset;
C3seq:=[x: x in C3];
```

4) İki kodkelimesi arasındaki Hamming uzaklık fonksiyonu tanımlanır.

```
S2:={1..10 by 1};
H:=function(x,y)
d:=0;
  for n in S do
    if C3seq[x][n] ne C3seq[y][n] then
      d:=d+1;
    end if;
  end for;
return d;
end function;
```

5)  $C_1$  kodunun minimum uzaklığı olan  $d_1$ ,  $d_3$  'e atanır.

```
d3:=4;
```

6)  $C_3$  kodundaki kodkelimleri arasındaki uzaklık ikişer ikişer hesaplanıp  $d_3$  ile kıyaslanır, ta ki  $d_3$  'ten daha küçük bir değer bulana kadar. Eğer böyle bir değer varsa,  $C_3$  kodunun minimum Hamming uzaklığı  $d_3$  bulunmuş olur. Kıyaslama bu şekilde tüm kodkelimleri ikişerli kıyaslanana kadar devam eder. Eğer 4 'ten daha küçük bir değer yoksa,  $C_3$  kodunun minimum Hamming uzaklığı  $d_3 = 4$  olarak bulunmuş olur.

```
for x in S2 do
  for y in S2 do
    if x ne y then
      if H(x,y) lt d3 then
        d3:=H(x,y);
      end if;
    end if;
  end for;
end for;
print d3;
```

$C_3$  kodu için minimum Hamming uzaklığı  $d_3 = 3$  bulunur.

$C_4$  kodunun minimum uzaklığı:

1)  $\mathbb{F}_{16}$  cismi ve  $Tr_4$  dönüşümü tanımlanır.

```
F<g>:=GF(16);
Fset:={x: x in F};
Tr4:=func<t|t+t^4>;
```

2) İndirgenmez devirli  $D$  kodundaki  $c_{\theta^1}$ ,  $\overline{c_{\theta^1}}^R$ ,  $c_{\theta^2}$ ,  $\overline{c_{\theta^2}}^R$  kodkelimleri ve bunların devirsel kaymaları elde edilir.

```
Fseq:=[x^3:x in Fset];
S:={1..5 by 1};
L:={0..4 by 1};
a1:=func<l|[Tr4(g^(3*1+1)*Fseq[n]):n in S]>;
```

```

b1:=func<l|[Tr4(g^(3*l+1)*Fseq[6-n])+1:n in S]>;
a2:=func<l|[Tr4(g^(3*l+2)*Fseq[n]):n in S]>;
b2:=func<l|[Tr4(g^(3*l+2)*Fseq[6-n])+1:n in S]>;
c1set:={a1(1): 1 in L};
c1:=[c:c in c1set];
c1RCset:={b1(1): 1 in L};
c1RC:=[c:c in c1RCset];
c2set:={a2(1): 1 in L};
c2:=[c:c in c2set];
c2RCset:={b2(1): 1 in L};
c2RC:=[c:c in c2RCset];

```

3)  $C_2$  ve  $\overline{C_2}^R$  kodları oluşturulur.

```

C2set:= c1set join c2set;
C2:=[c: c in C2set];
C2RCset:= c1RCset join c2RCset;
C2RC:=[c:c in C2RCset];

```

Burada "C2set" ile  $C_2$  kodu, "C2RCset" ile de  $\overline{C_2}^R$  kodu gösterilmektedir.

4) İki kodkelimesi arasındaki Hamming uzaklık fonksiyonu tanımlanır.

```

S2:={1..10 by 1};
H:=function(x,y)
d:=0;
  for n in S do
    if C2[x][n] ne C2RC[y][n] then
      d:=d+1;
    end if;
  end for;
return d;
end function;

```

5)  $C_2$  kodunun minimum uzaklığı olan  $d_2$ ,  $d_4$  'e atanır.

d4:=4;

6) Minimum Hamming uzaklığı hesaplanır.

```
for x in S2 do
  for y in S2 do
    if H(x,y) lt d4 then
      d4:=H(x,y);
    end if;
  end for;
end for;
print d4;
```

$C_4$  kodu için minimum Hamming uzaklığı  $d_4 = 3$  bulunur.

Sonuç olarak  $C_3$  bir  $(5, 10, 3, 4)_4$  kodu,  $C_4$  bir  $(5, 20, 3, 2)_4$  kodudur.

(1), (3) ve (4) kısıtlamalarını sağlayan DNA kodlarının parametreleri aşağıdaki çizelgede verilmiştir:

Çizelge 5.2. (1), (3) ve (4) kısıtlamalarını sağlayan DNA kodları

| e  | n   | $M(C_3/C_4)$ | $d(C_3/C_4)$ | $w(C_3/C_4)$ |
|----|-----|--------------|--------------|--------------|
| 3  | 5   | 10/20        | 3/3          | 4/2          |
| 3  | 21  | 42/84        | 12/11        | 8/12         |
| 5  | 51  | 102/408      | 33/31        | 32/24        |
| 3  | 85  | 170/340      | 55/53        | 48/40        |
| 3  | 341 | 682/1364     | 221/231      | 160/176      |
| 11 | 93  | 186/1860     | 48/48        | 32/48        |
| 13 | 315 | 630/7560     | 188/189      | 128/160      |
| 5  | 819 | 1638/6552    | 556/554      | 384/416      |

### 5.3 Hamming Uzaklığı, Ters Sırada Yazılış, Ters-Eşlenik ve GC-İçeriği Kısıtlamalarını Sağlayan DNA Kodlarının İnşası

Eğer  $D$  indirgenmez devirli kodunun üreteç polinomu self-resiprokal polinom değil ise yani  $D$  ters sırada yazılabilen bir kod değil ise her  $c \in D$  kodkelimesi için  $c^R \notin D$

olacaktır. Dolayısıyla  $H(c, c^R) > 0$  olur. O halde (1), (2), (3) ve (4) kısıtlamalarını sağlayan DNA kodları,

$$\mathbf{C}_1^R := \{c_{\gamma^i}^R \mid i = 0, 1, \dots, n-1\}$$

ve

$$\mathbf{C}_2^R := \{c_{\theta^j \gamma^i}^R \mid j = 1, 2, \dots, e-1 \ i = 0, 1, \dots, n-1\}$$

olmak üzere

$$\mathbf{C}_5 := \mathbf{C}_1 \cup \overline{\mathbf{C}_1}^R \cup \mathbf{C}_1^R \text{ ve } \mathbf{C}_6 := \mathbf{C}_2 \cup \overline{\mathbf{C}_2}^R \cup \mathbf{C}_2^R$$

kodlarıdır.

**Örnek 5.2**  $f(x) = x^3 + x^2 + x + \alpha^2 \in \mathbb{F}_4[x]$  primitif polinom,  $\mathbb{F}_{64}^* = \langle \theta \rangle$  ve  $\gamma := \theta^3$  olsun.  $r = 3$  ve  $e = 3$  olmak üzere

$$\mathbf{D} = \{c_\beta := (Tr_4(\beta), Tr_4(\beta\gamma), Tr_4(\beta\gamma^2), \dots, Tr_4(\beta\gamma^{20})) \mid \beta \in \mathbb{F}_{64}\}$$

$[21, 3]$  indirgenmez devirli kodu,

$$c_0 = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$$

$$c_{\theta^0} = (1, \alpha, \alpha^2, 0, \alpha, 0, 0, \alpha, \alpha^2, 1, 0, \alpha^2, 0, 0, \alpha^2, 1, \alpha, 0, 1, 0, 0)$$

$$c_{\theta^1} = (1, 1, 0, 1, \alpha^2, 1, \alpha^2, \alpha, \alpha, 0, \alpha, 1, \alpha, 1, \alpha^2, \alpha^2, 0, \alpha^2, \alpha, \alpha^2, \alpha)$$

$$c_{\theta^2} = (1, 1, 1, \alpha^2, 0, 1, 1, \alpha, \alpha, \alpha, 1, 0, \alpha, \alpha, \alpha^2, \alpha^2, \alpha^2, \alpha, 0, \alpha^2, \alpha^2)$$

kodkelimeleri ile bunların devirsel kaymalarının kümesidir.

Buradan (1) ve (4) şartlarını sağlayan DNA kodları,

$$\mathbf{C}_1 := \{c_{\gamma^i} \in \mathbf{D} \mid i = 0, 1, \dots, 20\}$$

ve

$$\mathbf{C}_2 := \{c_{\theta^j \gamma^i} \in \mathbf{D} \mid j = 1, 2, \ i = 0, 1, \dots, 20\}$$

kodlarıdır.

Teorem 5.1 gereği,  $\frac{q-1}{3} = 21 \equiv 0 \pmod{3}$  olup parametreler

$$\begin{aligned} w_1 &= \frac{n}{2} - \frac{1}{2e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1}) \\ w_2 &= \frac{n}{2} - \frac{1}{2e}(-1 + 2^r(-1)^{\frac{r}{t}}) \\ d_1 &= \frac{3n}{4} - \frac{3}{4e}(-1 + (e-1)2^r(-1)^{\frac{r}{t}-1}) \\ d_2 &= \frac{3n}{4} - \frac{3}{4e}(-1 + 2^r(-1)^{\frac{r}{t}}) \end{aligned}$$

şeklindedir.

$n = 21$ ,  $e = 3$ ,  $r = 3$  ve  $t = 1$  parametrelerdeki yerlerine yazılırsa,  $w_1 = 8$ ,  $w_2 = 12$ ,  $d_1 = 12$ ,  $d_2 = 18$  elde edilir. Buradan  $\mathbf{C}_1$  bir  $(21, 21, 12, 8)_4$  kodu,  $\mathbf{C}_2$  bir  $(21, 42, 12, 12)_4$  kodudur.

Şimdi,

$$\overline{\mathbf{C}}_1^R := \{\overline{c_{\gamma^i}}^R \mid i = 0, 1, \dots, 20\}$$

ve

$$\overline{\mathbf{C}}_2^R := \{\overline{c_{\theta^j \gamma^i}}^R \mid j = 1, 2, i = 0, 1, \dots, 20\}$$

olsun.

Buradan (1), (3) ve (4) şartlarını sağlayan yeni DNA kodları

$$\mathbf{C}_3 := \mathbf{C}_1 \cup \overline{\mathbf{C}}_1^R \text{ ve } \mathbf{C}_4 := \mathbf{C}_2 \cup \overline{\mathbf{C}}_2^R$$

şeklinde elde edilir. Örnek 5.1 'deki MAGMA kodları kullanılarak yapılan hesaplama ile  $d_3 = 12$  ve  $d_4 = 11$  bulunur. Sonuç olarak  $\mathbf{C}_3$  bir  $(21, 42, 12, 8)_4$  kodu,  $\mathbf{C}_4$  bir  $(21, 84, 11, 12)$  kodudur.

Öte yandan, Örnek 2.13'te  $\mathbf{D}$  'nin ters sırada yazılamayan bir kod olduğu gösterilmişti. Dolayısıyla buradan dört kısıtlamayı da sağlayan DNA kodları inşa edilebilir. O halde,

$$\mathbf{C}_1^R := \{c_{\gamma^i}^R \mid i = 0, 1, \dots, n-1\}$$

ve

$$\mathbf{C}_2^R := \{c_{\theta^j \gamma^i}^R \mid j = 1, 2, i = 0, 1, \dots, 20\}$$

olmak üzere (1), (2), (3) ve (4) kısıtlamalarını sağlayan DNA kodları,

$$\mathbf{C}_5 := \mathbf{C}_1 \cup \overline{\mathbf{C}}_1^R \cup \mathbf{C}_1^R \text{ ve } \mathbf{C}_6 := \mathbf{C}_2 \cup \overline{\mathbf{C}}_2^R \cup \mathbf{C}_2^R$$

kodlarıdır.

Aşağıdaki MAGMA kodları ile  $C_5$  kodunun minimum Hamming uzaklıkları bulunabilir.

```
F<g>:=GF(64);
Fset:={x: x in F};
Fseq:=[x^3:x in Fset];
Tr4:=func<t|t+t^4+t^16>;
S:={1..21 by 1};
L:={0..20 by 1};
a:=func<l|[Tr4(g^(3*l)*Fseq[n]):n in S]>;
aRC:=func<l|[Tr4(g^(3*l)*Fseq[22-n])+1:n in S]>;
aR:=func<l|[Tr4(g^(3*l)*Fseq[22-n]):n in S]>;

c0set:={a(l): l in L};
c0:=[c:c in c0set];
c0RCset:={aRC(l): l in L};
c0RC:=[c:c in c0RCset];
c0Rset:={aR(l): l in L};
c0R:=[c:c in c0Rset];

C5:= c0set join c0RCset join c0Rset;
C5seq:=[x: x in C1];

S2:={1..63 by 1};
S3:={1..63 by 1};

H:=function(x,y)
    d:=0;
    for n in S do
        if C5seq[x][n] ne C5seq[y][n] then
            d:=d+1;
        end if;
    end for;
end function;
```

```

        end if;
    end for;
return d;
end function;
d5:=12;
for x in S2 do
    for y in S3 do
        if x ne y then
            if H(x,y) lt d5 then
                d5:=H(x,y);
            end if;
        end if;
    end for;
end for;
print d5;

```

$C_5$  kodu için minimum Hamming uzaklığı  $d_5 = 12$  bulunur.

Aşağıdaki MAGMA kodları ile  $C_6$  kodunun minimum Hamming uzaklıkları bulunabilir.

```

F<g>:=GF(64);
Fset:={x: x in F};
Fseq:=[x^3:x in Fset];
Tr4:=func<t|t+t^4+t^16>;
S:={1..21 by 1};
L:={0..20 by 1};
a1:=func<l|[Tr4(g^(3*l+1)*Fseq[n]):n in S]>;
aRC1:=func<l|[Tr4(g^(3*l+1)*Fseq[22-n])+1:n in S]>;
aR1:=func<l|[Tr4(g^(3*l+1)*Fseq[22-n]):n in S]>;
a2:=func<l|[Tr4(g^(3*l+2)*Fseq[n]):n in S]>;
aRC2:=func<l|[Tr4(g^(3*l+2)*Fseq[22-n])+1:n in S]>;
aR2:=func<l|[Tr4(g^(3*l+2)*Fseq[22-n]):n in S]>;

```

```

c1set:={a1(l): l in L};
c1:=[c:c in c1set];
c1RCset:={aRC1(l): l in L};
c1RC:=[c:c in c1RCset];
c1Rset:={aR1(l): l in L};
c1R:=[c:c in c1Rset];

c2set:={a2(l): l in L};
c2:=[c:c in c2set];
c2RCset:={aRC2(l): l in L};
c2RC:=[c:c in c2RCset];
c2Rset:={aR2(l): l in L};
c2C:=[c:c in c2RCset];

C6set:=c1set join c1RCset join c1Rset
      join c2set join c2RCset join c2Rset;

C6:=[c: c in C5set];

S2:={1..126 by 1};

H:=function(x,y)
  d:=0;
  for n in S do
    if C6[x][n] ne C6[y][n] then
      d:=d+1;
    end if;
  end for;
  return d;
end function;

d6:=12;

```

```

for x in S2 do
  for y in S2 do
    if x ne y then
      if H(x,y) lt d6 then
        d6:=H(x,y);
      end if;
    end if;
  end for;
end for;
print d6;

```

$C_6$  kodu için minimum Hamming uzaklığı  $d_6 = 11$  bulunur.

Sonuç olarak  $C_5$  bir  $(21, 63, 12, 8)_4$  kodu,  $C_6$  bir  $(21, 126, 11, 12)_4$  kodudur.

(1), (3) ve (4) kısıtlamalarını sağlayan DNA kodlarının parametreleri aşağıdaki çizelgede verilmiştir:

Çizelge 5.3. (1), (2), (3) ve (4) kısıtlamalarını sağlayan DNA kodları

| e  | n   | $M(C_3/C_4)$ | $d(C_3/C_4)$ | $w(C_3/C_4)$ |
|----|-----|--------------|--------------|--------------|
| 3  | 21  | 63/126       | 12/11        | 8/12         |
| 5  | 51  | 153/612      | 32/24        | 32/24        |
| 3  | 85  | 255/510      | 48/48        | 48/40        |
| 3  | 341 | 1023/2046    | 221/228      | 160/176      |
| 11 | 93  | 279/2790     | 48/48        | 32/48        |
| 13 | 315 | 945/11340    | 188/189      | 128/160      |
| 5  | 819 | 2457/9828    | 555/554      | 384/416      |

## 6. TARTIŞMA VE SONUÇ

Bu tezde, önceki DNA kodu inşa etme metotlarından farklı olarak Gauss periyodu ve  $\mathbb{F}_4$  üzerinde indirgenmez devirli kodlar arasındaki ilişkiden yararlanılmıştır. Bu metot ile ilk olarak Hamming uzaklığı ve sabit  $GC$ -içeriği kısıtlamalarını sağlayan DNA kodları inşa edilip, parametreleri  $\mathbb{F}_4$  üzerinde indirgenmez devirli kodların ağırlık formülleri kullanılarak kolayca hesaplanabilmektedir. Daha sonra, Hamming uzaklığı, ters-eşlenik ve sabit  $GC$ -içeriği kısıtlamalarını sağlayan DNA kodları ve son olarak dört kısıtlamayı birden sağlayan DNA kodları inşa edilmiş ve parametreleri hesaplanmıştır.

Literatürde dört kısıtlamayı birden sağlayan DNA kodu inşa etmenin zor olduğu ileri sürülmektedir. Ancak 2015 yılına kadar tezde uygulanan metot kullanılmamıştır. Burada kullanılan metotlar ve ters sırada yazılamayan devirli kodlar birlikte göz önüne alındığında, kısıtlamaların tamamını sağlayan DNA kodları inşa etmek mümkün olmuştur. Elde edilen DNA kodlarının minimum uzaklıkları MAGMA Computer Algebra programı ile hesaplanmıştır.

## KAYNAKLAR

- Aboluion, N., Smith, D. H. and Perkins, S. 2012. *Linear and Nonlinear Construction of DNA Codes with Hamming Distance  $d$ , Constant GC-Content and a Reverse-Complement Constraint*. Discrete Mathematics, 312; 1062-1075.
- Adleman, L. M. 1994. *Molecular Computation of Solution to Combinatorial Problems*. American Association for the Advancement of Science, 266; 1021-1024.
- Berndt, B. C., Evans, R. J. and Williams, K. S. 1998. *Gauss and Jacobi Sums*. John Wiley and Sons.
- Branner, S. and Larner, R. A. 1992. *Encoded Combinatorial Chemistry*. Proc. Natl. Acad. Sci. USA, 89; 5381-5383
- D'yachkov, A. G., Vilenkin, P. A., Ismagilov, I. K., Sarbaev, R. S., Macula, A., Torney, D. and White, S. 2005. *On DNA Codes*. Problems of Information Transmission, 41(4); 349-367.
- Ding, C. and Yang, J. 2013. *Hamming Weights in Irreducible Cyclic Codes*. Discrete Mathematics, 313; 434-446.
- Frutos, A. G., Liu, Q., Thiel, A. J., Sanner, A. M. W., Condon, A. E., Smith, L. M. and Corn, R. M. 1997. *Demonstration of a Word Design Strategy for DNA Computing on Surfaces*. Nucleic Acids Res., 25(23); 4748-4757.
- Gaborit, P. and King, O. D. 2005. *Linear Constructions for DNA Codes*. Theoretical Computer Science, 334; 99-113.
- Hill, R. 1986. *A First Course in Coding Theory*. Oxford University Press, New York.
- Hong, H., Wang, L., Ahmad, H., Li, J., Yang, Y. and Wu, C. 2016. *Construction of DNA Codes by Using Algebraic Number Theory*. Finite Fields and Their Applications, 37; 328-343.
- Ireland, K. and Rosen, M. 1990. *A Classical Introduction to Modern Number Theory*. Springer-Verlag, New York.
- King, O. D. 2003. *Bounds for DNA Codes with Constant GC-Content*. J. Comb., 10; R33.
- Lidl, R. and Niederreiter, H. 1997. *Finite Fields*. Cambridge University Press, New York.
- Lint, J. H. V. 1999. *Introduction to Coding Theory*. Springer-Verlag, Berlin, Heidelberg.
- MacWilliams, F. J. and Sloane, N. J. A. 1977. *The Theory of Error-Correcting Codes*. Nort-Holland Publishing Company, Amsterdam.
- Marathe, A., Condon, A. E. and Corn, R. M. 2001. *On Combinatorial DNA Word Deesign*. Journal of Computational Biology, 8; 201-219.

- Massey, J. L. 1964. *Reversible Codes*. Information and Control, 7; 369-380.
- McEliece, R. J. 1975. *Irreducible Cyclic Codes and Gauss Sums*. In: Combinatorics. Hall, M. J. and Lint, J. H. V. (eds), D. Reidel Publishing Company, 185-202, Boston.
- Meyn, H. 1990. *On the Construction of Irreducible Self-Reciprocal Polynomials Over Finite Fields*. Applicable Algebra in Engineering, Communication and Computing, 1; 43-53.
- Montemanni, R. and Smith, D. H. 2008. *Construction of GC-Content DNA Codes via a Variable Neighborhood Search Algorithm*. J. Math. Model. Algorithms, 7; 311-326.
- Myerson, G. 1981. *Period Polynomials and Gauss Sums for Finite Fields*. ACTA Arithmetica, 39; 251-264.
- Shannon, C. E. 1948. *A Mathematical Theory of Communication*. The Bell System Technical Journal, 27; 379-423.
- Shena, M., Shalon, D., Davis, R. W. and Brown, P. O. 1995. *Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray*. Science, 270; 467-470.
- Smith, D. H., Abolun, N., Montemanni, R. and Perkins, S. 2011. *Linear and Non-linear Constructions of DNA Codes with Hamming Distance  $d$  and Constant GC-Content*. Discrete Mathematics, 311; 1207-1219.

## ÖZGEÇMİŞ

**Adı Soyadı** : Emre GÜDAY

**Doğum Yeri** : Uğurludağ

**Doğum Tarihi** : 15/01/1993

**Medeni Hali** : Bekar

**Yabancı Dili** : İngilizce

**Eğitim Durumu (Kurum ve Yıl):**

**Lise** : Ayrancı Aysel Yücetürk Lisesi, 2011

**Lisans** : Ankara Üniversitesi, Matematik Bölümü, 2016

**Yüksek Lisans** : Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Matematik Anabilim Dalı, 2020

**Çalıştığı Kurum/Kurumlar ve Yıl:** Bilecik Şeyh Edebali Üniversitesi, Matematik Anabilim Dalı, Arş. Gör. (2018-2020)

Ankara Üniversitesi, Matematik Anabilim Dalı, Arş. Gör. (2020-Günümüz)

**Yayınları:** .