Universidade do Estado do Rio de Janeiro
Centro de Educação e Humanidades
Faculdade de Formação de Professores

Alexandre Peixoto Marcet

Códigos Corretores de Erros BCH: uma aplicação de polinômios em Corpos Finitos

São Gonçalo
2019

Alexandre Peixoto Marcet

Códigos Corretores de Erros BCH: uma aplicação de polinômios em Corpos Finitos

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT) da Universidade do Estado do Rio de Janeiro.

Orientador: Profª. Dra. Clícia Valladares Peixoto Friedmann

São Gonçalo

2019

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CEH/D

M314 Marcet, Alexandre Peixoto
Códigos Corretores de Erros BCH: uma aplicação de polinômios em Corpos Finitos / Alexandre Peixoto Marcet - 2019.
77 f.

Orientador: Profª. Dra. Clícia Valladares Peixoto Friedmann
Dissertação (Mestrado) – Mestrado Profissional em Matemática em Rede Nacional (PROFMAT) - Universidade do Estado do Rio de Janeiro, Faculdade de Formação de Professores.

1. Polinômios. 2. Códigos Corretores de Erros (Teoria da informação).
I. Friedmann, ClíciaValladares Peixoto . II. Universidade do Estado do Rio de Janeiro. III. Faculdade de Formação de Professores. IV. Título

CRB-7 - 4994 CDU 519.725

Assinatura Data

Alexandre Peixoto Marcet

Códigos Corretores de Erros BCH: uma aplicação de polinômios em Corpos Finitos

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT) da Universidade do Estado do Rio de Janeiro.

Aprovada em 31 de outubro de 2019.
Banca Examinadora:

Profª. Dra. Clícia Valladares Peixoto Friedmann (Orientador)
Faculdade de Formação de Professores – UERJ

Prof. Dr. Abel Rodolfo García Lozano
Faculdade de Formação de Professores – UERJ

Profª. Dra. Andréa Gomes Guimarães
Instituto de Matemática – Universidade Federal Fluminense

São Gonçalo

2019

DEDICATÓRIA

Dedico esse trabalho à minha mãe por sempre acreditar em mim, ao meu pai (*in memorian*), à minha esposa Tatiana por sua infinita compreensão e incentivo e ao meu mestre Dr. Daisaku Ikeda por ter propagado o Budismo Nichiren para o Brasil.

AGRADECIMENTOS

Agradeço à minha mãe por sempre estar me apoiando e incentivando carinhosamente.

Agradeço ao meu pai (*in memorian*) por ter me apoiado em diversas situações da minha vida.

Agradeço à minha esposa pelo seu amor, apoio, companheirismo e dedicação.

Agradeço à minha orientadora Clícia pela paciência, compreensão e apoio que foram importantes para a realização deste trabalho.

Agradeço aos professores e colegas do Profmat da Uerj - São Gonçalo por terem participado positivamente da minha vida.

Agradeço à todas as demais pessoas que fizeram ou fazem parte da minha vida e que contribuíram para o meu crescimento individual em todos os aspectos.

Agradeço aos professores da banca pelo apoio.

Agradeço à CAPES pelo apoio financeiro, que foi fundamental para a realização deste mestrado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original.

Albert Einstein

RESUMO

MARCET, A.P.M. *Códigos Corretores de Erros BCH: uma aplicação de polinômios em Corpos Finitos*. 2019. 77 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional - PROFMAT) – Faculdade de Formação de Professores, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

Códigos corretores de erros são usados para recuperar informações que, no processo de armazenamento ou trnasmissão, sofreram algum tipo de alteração. Pretendemos mostrar neste trabalho a relevância dos polinômios, conteúdo matemático abordado desde o Ensino Fundamental. E para isso, apresentaremos sua aplicação nos códigos corretores de erros BCH (Bose-Chaudhuri-Hocquenghem), que trabalham sobre Corpos Finitos. Introduziremos alguns conceitos e resultados de Álgebra e exploraremos exemplos que mostram como é feita a utilização dos polinômios nos códigos BCH.

Palavras-chave: Polinômios. Códigos Corretores de Erros. BCH. Corpos Finitos.

ABSTRACT

MARCET, A.P.M. *BCH Error-Correcting Codes: an application of polynomials in Finite Fields*. 2019. 77 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional - PROFMAT) – Faculdade de Formação de Professores, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

Error-correcting codes are used to retrieve information that has changed in the process of storing or transmitting data. This TCC intends to show the relevance of polynomials, mathematical content approached since elementary school. And for that, we will present its application in the error correcting codes BCH (Bose-Chaudhuri-Hocquenghem) er-ror correcting codes, which work on Finite Fields. We will present some concepts and results of Algebra and explore examples showing how polynomials are used in BCH codes.

Keywords: Polynomials. Error Correcting Codes. BCH. Finite Fields.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ALGORITMOS

SUMÁRIO

INTRODUÇÃO

No Brasil, o estudo de polinômios na escola se inicia nos anos finais do Ensino Fundamental, onde são vistos, entre outros, os seguintes conteúdos: expressões algébricas, produtos notáveis, fatoração e polinômios sobre o conjunto dos números reais. Depois, no Ensino Médio, o tema polinômios é retomado e aprofundado, sendo abordado nos conteúdos: equações polinomiais e polinômios sobre o conjunto dos números complexos. Nos dois níveis de ensino – Fundamental e Médio – os polinômios são estudados sobre os seguintes conjuntos: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ e $\mathbb{C}$. Contudo, várias aplicações que surgiram com o desenvolvimento de Ciência da Computação, o uso maciço de computadores e as comunicações via satétile trabalham com polinômios sobre Corpos Finitos e suas extensões, como é o caso de algumas aplicações relacionadas à Criptografia e Teoria dos Códigos Corretores de Erros. Dessa forma, a abordagem de polinômios sobre Corpos Finitos passou a ser valorizada nas graduações da área tecnológica, principalmente em algumas engenharias como: elétrica, de computação, eletrônica, de telecomunicações e etc.., além da Matemática. Com isso, é natural pensarmos que, em algum momento, necessitaremos trabalhar, ainda no Ensino Médio, com polinômios sobre os Corpos Finitos, o que nos remete ao Corpo dos inteiros módulo m ($\mathbb{Z}_m$), sendo m um número primo.

A constatação de que precisaremos abordar polinômios sobre $\mathbb{Z}_m$ no Ensino Médio é uma provável consequência da valorização dos conteúdos de congruência modular e noções de Criptografia na Educação Básica, o que já pode ser visto em trabalhos acadêmicos que se relacionam ao ensino de Matemática, inclusive dentro do PROFMAT, como por exemplo (CARVALHO, 2014), (NICOLETTI, 2015) e (ARAGÃO, 2017). Essa constatação e o desejo de trabalhar com códigos corretores de erros em nossa pesquisa serviram de motivação para a elaboração deste trabalho.

Começamos então a pesquisar algum tipo de código corretor de erros que valorizasse o uso de polinômios de uma forma mais direta e que trabalhasse sobre Corpos Finitos, o que confluiu para a escolha dos códigos BCH binários, que foram descobertos por R.C. Bose e D.K. Chaudhuri (1960) e independentemente por A. Hocquenghem (1959). Tais códigos pertencem à classe dos cíclicos, os quais admitem uma representação em termos de polinômios sobre extensões de $\mathbb{Z}_2$.

Este trabalho está dividido em três capítulos. No primeiro, fizemos uma revisão de definições e propriedades dos corpos em geral e o Anel dos Inteiros módulo m, com ênfase para a condição de ser um corpo. Também mostramos os principais resultados que envolvem polinômios sobre corpos. No segundo capítulo, apresentamos algumas das principais propriedades e resultados de Corpos Finitos, com destaque para os que envolvem extensões do corpo $\mathbb{Z}_2$. Já no terceiro capítulo, tratamos de Códigos Corretores de Erros. Além de introduzirmos o assunto, a ênfase foi dada aos códigos BCH binários, mas antes

apresentamos os códigos lineares e os cíclicos binários, que formam, respectivamente, a classe e subclasse as quais os BCH binários pertencem.

O leitor poderá observar que os polinômios, embora tratados no capítulo 1, transitam nos demais. Advertimos que muitos resultados apresentados no último capítulo foram focados apenas sobre corpos finitos com 2^m elementos e algumas demonstrações foram omitidas, por fugirem ao escopo deste trabalho, mas podem ser encontradas nas seguintes referências bibliográficas (LIN; DANIEL, 1983) e (HEFEZ; VILLELA, 2017). Optamos por essa forma de proceder para tratarmos dos códigos BCH binários de uma maneira mais prática e por meio de exemplos que valorizassem a utilização de polinômios.

1 FUNDAMENTOS MATEMÁTICOS

Neste capítulo apresentaremos, de forma breve, alguns conceitos e resultados que dão embasamento teórico a este trabalho, tais como a definição de corpo e suas propriedades, o Anel dos inteiros módulo m e polinômios sobre corpos finitos. A ênfase será dada ao estudo dos polinômios, porque o código corretor de erros BCH trabalha com polinômios sobre corpos finitos.

1.1 Conceitos Básicos

Definição 1.1.1. Seja A um conjunto, com $A \neq \emptyset$. Um anel é a terna $(A, +, \cdot)$ que possui duas operações em A; chamadas de **adição** e **multiplicação**, denotadas respectivamente da seguinte forma:

$$\begin{array}{rccc} + : & A \times A & \longrightarrow & A \\ & (a,b) & \longmapsto & a+b \end{array} \quad \text{e} \quad \begin{array}{rccc} \cdot : & A \times A & \longrightarrow & A \\ & (a,b) & \longmapsto & a \cdot b \end{array}$$

e que possuem as seguintes propriedades:

A1) Associatividade da adição:

$$\forall a, b, c \in A, \quad (a+b)+c = a+(b+c) = a+b+c.$$

A2) Existência do elemento neutro para a adição:
Existe um elemento neutro, denotado por 0, denominado **elemento neutro do anel**, tal que

$$\forall a \in A, \quad a+0 = 0+a = a.$$

A3) Todos os elementos de A são simetrizáveis em relação à adição:
$\forall a \in A$, existe um elemento chamado **simétrico** de a, denotado por $-a$, tal que

$$a+(-a) = (-a)+a = 0$$

A4) Comutatividade da adição:

$$\forall a, b \in A, \quad a+b = b+a.$$

M1) Associatividade da multiplicação:

$$\forall a, b, c \in A, \quad (a \cdot b) \cdot c = a \cdot (b \cdot c) = a \cdot b \cdot c.$$

AM) Distributividade da multiplicação com relação à adição:

$$\forall a, b, c \in A, \quad a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{e} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

Os elementos $a+b$ e $a \cdot b$ (ou ab) são chamados respectivamente de **soma** e **produto** de a e b. Nesse texto, muitas vezes iremos nos referir ao anel $(A, +, \cdot)$ apenas por A.

Observação 1.1.2. Sabemos que todo elemento a de um anel tem inverso para a adição $(-a)$ e ele é único. Podemos assim, definir uma operação chamada de **subtração** como:

$$a - b = a + (-b), \forall a, b \in A$$

Os conjuntos dos números inteiros $\mathbb{Z}$, dos números racionais $\mathbb{Q}$, dos números reais $\mathbb{R}$ e dos números complexos $\mathbb{C}$, juntamente com as suas operações de adição e de multiplicação, são exemplos de anéis.

Um anel A é um **Anel com Unidade** se for verificada a seguinte propriedade.
M2) Existência do elemento neutro para a multiplicação:
Existe um elemento chamado **unidade** e denotado por 1 tal que

$$\forall a \in A, \quad a \cdot 1 = 1 \cdot a = a.$$

Um anel A é um **Anel Comutativo** se for verificada a seguinte propriedade.
M3) Comutatividade da multiplicação:

$$\forall a, b \in A, \quad a \cdot b = b \cdot a.$$

Um anel A é um **Anel sem Divisores de Zero** se for verificada a seguinte propriedade.
M4) $\forall a, b \in A, a \neq 0 \quad \text{e} \quad b \neq 0 \Longrightarrow a \cdot b \neq 0$.
A propriedade acima é equivalente à seguinte:

$$\forall a, b \in A, a \cdot b = 0 \Longrightarrow a = 0 \quad \text{ou} \quad b = 0.$$

Um anel A é um **Domínio de Integridade** se forem verificadas as propriedades M2, M3 e M4.

Definição 1.1.3. Seja A um anel comutativo com unidade. Um elemento $a \neq 0$ de A será dito **invertível**, se tiver inverso multiplicativo a^{-1}, isto é, $a \cdot a^{-1} = 1$. Além disso $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1} = a^{-1} \cdot b^{-1}$.

Se a é invertível, então a^{-1} é invertível e $(a^{-1})^{-1} = a$.

Definição 1.1.4. Um domínio de integridade em que todo elemento diferente do elemento neutro aditivo do domínio é invertível, é chamado de **corpo**.

Os anéis $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ são corpos porque, nesses conjuntos, todo elemento diferente de 0 é invertível. Por outro lado, o anel $\mathbb{Z}$ não é um corpo, pois seus únicos elementos invertíveis são -1 e 1.

Um corpo é finito se o conjunto de seus elementos é finito. No capítulo 2 voltaremos a falar sobre esses corpos.

A seguir apenas enunciaremos algumas propriedades básicas que envolvem anéis, domínios de integridade e corpos.

Sejam A um anel e $a, b, c \in A$, então:

1) $0 \cdot a = a \cdot 0 = 0$
2) $-(a \cdot b) = (-a) \cdot b = a \cdot (-b)$
3) $(-a) \cdot (-b) = a \cdot b$
4) $a \cdot (b - c) = a \cdot b - a \cdot c$
5) $(b - c) \cdot a = b \cdot a - c \cdot a$
6) $(-1) \cdot a = -a$
7) $(-1) \cdot (-1) = 1$
8) $(-1) \cdot (-a) = a$

Além das propriedades acima, se o anel A for um domínio de integridade e se $a \neq 0$, $a \cdot b = a \cdot c$ então $b = c$. (Lei do Cancelamento da Multiplicação).

1.2 Anel dos Inteiros Módulo m

Definição 1.2.1. Sejam o anel $\mathbb{Z}$ e $m \in \mathbb{Z}$, com $m > 1$. Dois elementos $a, b \in \mathbb{Z}$ são **congruentes módulo m** se m divide $b - a$. Neste caso, denotaremos por $a \equiv b \pmod{m}$.

É imediato da definição que para quaisquer $a, b, c, a', b' \in \mathbb{Z}$:

i) $a \equiv a \pmod{m}$;

ii) Se $a \equiv b \pmod{m}$ então $b \equiv a \pmod{m}$;

iii) Se $a \equiv b \pmod{m}$ e $b \equiv c \pmod{m}$ então $a \equiv c \pmod{m}$;

iv) Se $a \equiv a' \pmod{m}$ e $b \equiv b' \pmod{m}$ então $a + b \equiv a' + b' \pmod{m}$ e $a \cdot b \equiv a' \cdot b' \pmod{m}$.

As três primeiras propriedades acima indicam que a congruência módulo m é uma relação de equivalência. A propriedade (iv) é compatível com as operações aditiva e multiplicativa de $\mathbb{Z}$.

Definição 1.2.2. Sejam o anel $\mathbb{Z}$ e $m \in \mathbb{Z}$, $m > 1$. A **classe residual** de um elemento $a \in \mathbb{Z}$ módulo m é o conjunto

$$\overline{a} = \{a + km, k \in \mathbb{Z}\}.$$

O elemento a é denominado um **representante** da classe residual $\overline{a}$.

O conjunto formado por todas as classes residuais em $\mathbb{Z}$ módulo m é denotado por $\mathbb{Z}_m$. Observamos que

$$\mathbb{Z}_m = \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$$

e que se $i, j = 0, \ldots, m-1$ com $i \neq j$, então $\overline{i} \neq \overline{j}$.

De fato, dado $a \in \mathbb{Z}$, pelo algoritmo da divisão euclidiana, temos que existem inteiros q e r univocamente determinados tais que: $a = mq + r$ e $0 \leq r \leq m-1$. Portanto, há um único inteiro r com $0 \leq r \leq m-1$, tal que $\overline{a} = \overline{r}$.

Logo, $\mathbb{Z}_m$ é um anel finito com exatamente m elementos.

Como a congruência módulo m em $\mathbb{Z}$ é uma relação de equivalência, então $\mathbb{Z}_m$ é uma partição de $\mathbb{Z}$. Além disso, podemos definir as operações de adição e multiplicação em $\mathbb{Z}_m$.

Definição 1.2.3. Sejam $\overline{a}, \overline{b}$ quaisquer classes residuais módulo m. A adição e a multiplicação em $\mathbb{Z}_m$ são definidas por:

$$\overline{a} + \overline{b} = \overline{a+b} \qquad \text{e} \qquad \overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

Podemos observar que as definições acima não dependem dos representantes das classes residuais. De fato, sabemos que se $a \equiv a' \pmod{m}$ e $b \equiv b' \pmod{m}$ então $a+b \equiv a'+b' \pmod{m}$ e $a \cdot b \equiv a' \cdot b' \pmod{m}$ e, portanto, $\overline{a+b} = \overline{a'+b'}$, o que acarreta que $\overline{a} + \overline{b} = \overline{a'} + \overline{b'}$.

É imediato que $\mathbb{Z}_m$ munido das operações de adição e multiplicação definidas acima formam um anel comutativo com unidade com $\overline{0}$ e $\overline{1}$. Sendo estes os elementos neutros da adição e da multiplicação respectivamente.

Exemplo 1.2.4. Seja $m = 2$. Logo, $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ com as operações da tabela 1 é um anel.

Tabela 1 - Operações em $\mathbb{Z}_2$

Adição em $\mathbb{Z}_2$

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

Multiplicação em $\mathbb{Z}_2$

$\cdot$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

Fonte: (HEFEZ; VILLELA, 2017)

Veremos que nem todo $\mathbb{Z}_m$ é um corpo. Então, com o objetivo de determinar quais são corpos, precisaremos inicialmente caracterizar os elementos invertíveis desses anéis.

Proposição 1.2.5. $\overline{a} \in \mathbb{Z}$ é invertível se, e somente se, $\text{MDC}(a, m) = 1$.

Demonstração: Suponhamos que $\overline{a}$ seja invertível. Consequentemente, existe $b \in \mathbb{Z}$ tal que $\overline{a} \cdot \overline{b} = \overline{1}$. Logo, $\overline{a \cdot b} = \overline{1}$ e, portanto, $a \cdot b \equiv 1 \pmod{m}$. Isso implica que existe um inteiro s tal que $a \cdot b + s \cdot m = 1$, logo $\text{MDC}(a, m) = 1$.

Reciprocamente, suponhamos que $\text{MDC}(a, m) = 1$. Então, pelo teorema de Bézout, existem inteiros b e c tais que $b \cdot a + c \cdot m = 1$. Logo, $b \cdot a \equiv 1 \pmod{m}$. Consequentemente, $\overline{a} \cdot \overline{b} = \overline{a \cdot b} = \overline{1}$ e, portanto, $\overline{a}$ é invertível. □

Teorema 1.2.6. O anel $\mathbb{Z}_m$ é um corpo se, e somente se, m é um número primo.

Demonstração: $\mathbb{Z}_m$ é um corpo se, e somente se, todos os elementos $\overline{1}, \overline{2}, \ldots, \overline{m-1}$ são invertíveis, o que, pela Proposição 1.2.5, equivale ao fato de que $\text{MDC}(1, m) = \text{MDC}(2, m) = \ldots = \text{MDC}(m-1, m) = 1$; e isso é, portanto, equivalente a m ser primo. □

Temos assim garantida a existência de um corpo com m elementos para todo número primo positivo m.

O anel $\mathbb{Z}_2$ do Exemplo 1.2.4 é um corpo.

1.3 Polinômios sobre um Corpo

O código corretor de erros BCH trabalha com polinômios sobre corpos finitos. Por esse motivo, faremos uma breve apresentação de polinômios sobre corpos.

Definição 1.3.1. Sejam K um corpo e $\boldsymbol{x}$ uma **indeterminada ou variável**. Um **polinômio $\boldsymbol{p(x)}$** com coeficientes em K, na indeterminada x, é uma expressão do tipo

$$p(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n,$$

onde $n \in \mathbb{Z}^+$ e $a_i \in K, \forall i = 0, 1, \ldots, n.$. Omitiremos, na expressão do polinômio, o termo $a_i x^i$, toda vez que $a_i = 0$.

Na definição acima usamos a notação formal de polinômios sobre um corpo, mas um polinômio $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ está associado à uma n-upla $(a_0, a_1, \ldots, a_n)$, em que $a_i \neq 0$ para um número de índices. Ao longo deste trabalho, poderemos usar indistintamente as duas notações, dependendo do contexto.

Dados dois polinômios $p(x) = a_0 + a_1 x^1 + \cdots + a_n x^n$ e $q(x) = b_0 + b_1 x^1 + \cdots + b_m x^m$, dizemos que $p(x) = q(x)$ se $a_i = b_i$, para todo i.

O conjunto de todos os polinômios na indeterminada x com coeficientes em K será denotado por $K[x]$. Observamos que $K \subset K[x]$.

Se $p(x) = 0 + 0x + 0x^2 + \ldots + 0x^n$ então dizemos que $p(x)$ é o **polinômio identicamente nulo** e será denotado por $p(x) = 0$. Se $p(x) = a_0$, com $a_0 \in K$, então $p(x)$ é o **polinômio constante** a_0.

Definição 1.3.2. Seja $p(x) = a_0 + a_1x^1 + \cdots + a_nx^n \in K[x]$ um polinômio não nulo, $a_n \neq 0$. O **grau** de $p(x)$, que será denotado por $gr(p(x))$, é o inteiro não negativo n. E, nesse caso, a_n é chamado **coeficiente líder** de $p(x)$.

Se $p(x)$ é o polinômio nulo, $gr(p(x))$ não está definido. Além disso, notemos que $gr(p(x)) = 0$ se, e somente se, $p(x)$ é um polinômio constante não nulo.

Definição 1.3.3. Se $gr(p(x)) = n$, dizemos que $p(x) \in K[x]$ é um **polinômio mônico** se $a_n = 1$.

Sejam os polinômios $p(x) = a_0 + a_1x^1 + \cdots + a_nx^n$ e $q(x) = b_0 + b_1x^1 + \cdots + b_mx^m$ pertencentes a $K[x]$, definimos as operações de **adição** e de **multiplicação** em $K[x]$, como a seguir:

$$p(x) + q(x) = \sum_{i=0}^{máx\{n,m\}} (a_i + b_i)x^i,$$

$$p(x) \cdot q(x) = \sum_{i=0}^{n+m} c_ix^i, \text{ onde } c_i = \sum_{j=0}^{i} a_j \cdot b_{i-j}$$

As operações de adição e multiplicação definidas acima, em $K[x]$, possuem propriedades semelhantes às das respectivas operações em K. Isso será mostrado no teorema a seguir.

Teorema 1.3.4. Com as operações de adição e de multiplicação definidas acima, $K[x]$ é um anel.

A demonstração deste teorema pode ser encontrada em (BIAZZI, 2014, pp. 19–23).

De acordo com as definições das operações em $K[x]$, dados os polinômios $p(x), q(x) \in K[x]$, com $gr(p(x)) = n$ e $gr(q(x) = m$, temos que $gr(p(x)+q(x)) = máx(gr(p(x)), gr(q(x)))$ e que $gr(p(x) \cdot q(x)) = n + m$.

Definição 1.3.5. Sejam K um corpo e os polinômios $f(x) \neq 0, g(x) \in K[x]$. Dizemos que $f(x)$ **divide** $g(x)$ ou $g(x)$ é **divisível** por $f(x)$, se existe $h(x) \in K[x]$ tal que

$$g(x) = f(x) \cdot h(x).$$

Se $f(x)$ divide $g(x)$ denotaremos por $f(x) \mid g(x)$.

Teorema 1.3.6 (Algoritmo da Divisão)**.** Sejam $f(x), g(x) \in K[x]$, com $g(x) \neq 0$. Existem únicos $q(x), r(x) \in K[x]$ tais que:

$$f(x) = q(x) \cdot g(x) + r(x),$$

onde $r(x) = 0$ ou $gr(r(x)) < gr(g(x))$.

Demonstração. Sejam os polinômios $f(x) = a_0 + a_1x + \ldots + a_nx^n$ e $g(x) = b_0 + b_1x + \ldots + b_mx^m$, com $gr(g(x)) = m$.
Existência: Se $f(x) = 0$ então $q(x) = r(x) = 0$. Suponhamos $f(x) \neq 0$ e $gr(f(x)) = n$. Se $n < m$ então $q(x) = 0$ e $r(x) = f(x)$. Agora, assumiremos que $n \geq m$.

Demonstraremos o teorema por indução sobre $gr(f(x)) = n$. Se $n = 0$, então teremos $m = 0$ pois supomos $n \geq m$. Nesse caso, $f(x) = a_0 \neq 0, g(x) = b_0 \neq 0$. Logo, obtemos $q(x) = a_0b_0^{-1}$ e $r(x) = 0$.

Assumiremos a validade do resultado para polinômios $g(x)$ de grau menor do que $n = gr(f(x))$.

Seja o polinômio $f_1(x) = f(x) - a_nb_m^{-1}x^{n-m} \cdot g(x)$, onde $gr(f_1(x)) < gr(f(x))$. Logo, pela hipótese de indução, existem únicos $q_1(x), r_1(x) \in K[x]$, tais que:

$$f_1(x) = q_1(x) \cdot g(x) + r_1(x),$$

onde $r_1(x) = 0$ ou $gr(r_1(x)) < gr(g(x))$. Então, temos que

$$f(x) = q_1(x)g(x) + r_1(x) + a_nb_n^{-1}x^{n-m}g(x)$$

e, consequentemente

$$f(x) = (q_1(x) + a_nb_m^{-1}x^{n-m})g(x) + r_1(x).$$

Portanto, tomando $q(x) = q_1(x) + a_nb_m^{-1}x^{n-m}$ e $r_1(x) = r(x)$, provamos a existência dos polinômios $q(x), r(x) \in K[x]$ tais que $f(x) = q(x) \cdot g(x) + r(x)$, onde $r(x) = 0$ ou $gr(r(x)) < gr(g(x))$.

Unicidade: Sejam os polinômios $q_1(x), q_2(x), r_1(x), r_2(x) \in K[x]$, tais que:

$$f(x) = q_1(x) \cdot g(x) + r_1(x) = q_2(x) \cdot g(x) + r_2(x),$$

onde $r_1(x) = 0$ ou $gr(r_1(x)) < gr(g(x))$ e $r_2(x) = 0$ ou $gr(r_2(x)) < gr(g(x))$. Então, temos que

$$(q_1(x) - q_2(x)) \cdot g(x) = r_2(x) - r_1(x).$$

Se $q_1(x) = q_2(x)$, então $r_1(x) = r_2(x)$ e a unicidade está provada. Por outro lado,

se $q_1(x) \neq q_2(x)$, temos que $gr(r_2(x) - r_1(x)) \geq gr(g(x))$, o que é uma contradição, pois devemos ter $gr(r_2(x) - r_1(x)) < gr(g(x))$ por hipótese. Logo, $q_1(x) = q_2(x)$ e, consequentemente, temos $r_1(x) = f(x) - q_1(x)g(x) = f(x) - q_2(x)g(x) = r_2(x)$. □

Exemplo 1.3.7. Vamos efetuar a divisão de $f(x) = x^4 + x^3 + x + 1$ por $g(x) = x + 1$, onde $f(x), g(x) \in \mathbb{Z}_2[x]$.

$$\begin{array}{rrrrrrrr|rrrr}
x^4 & + & x^3 & & + & x & + & 1 & x & + & 1 \\
\cline{9-11}
x^4 & + & x^3 & & & & & & x^3 & + & 1 \\
\cline{1-8}
 & & & & & x & + & 1 & & & \\
 & & & & & x & + & 1 & & & \\
\cline{5-8}
 & & & & & & & 0 & & &
\end{array}$$

Temos que: $x^4 + x^3 + x + 1 = (x + 1) \cdot (x^3 + 1)$, onde $q(x) = x^3 + 1$ e $r(x) = 0$. Portanto, $g(x) = x + 1$ divide $f(x) = x^4 + x^3 + x + 1$.

Definição 1.3.8. Sejam K um corpo e $p(x) \in K[x]$ um polinômio. Dizemos que $x_0 \in K$ é uma **raiz** do polinômio $p(x)$ se $p(x_0) = 0$.

Proposição 1.3.9. Sejam K um corpo, $p(x) \in K[x]$ um polinômio e $x_0 \in K$. Temos que x_0 é uma raiz de $p(x)$ se, e somente se, $x - x_0$ divide $p(x)$.

Demonstração: Sejam K um corpo, os polinômios $f(x), g(x) \in K[x]$, x_0 uma raiz de $f(x)$ e $g(x) = x - x_0$. Pelo algoritmo da divisão, existem dois únicos polinômios $q(x), r(x) \in K[x]$, tais que

$$f(x) = g(x) \cdot q(x) + r(x), \text{ com } r(x) = 0 \text{ ou } gr(r(x)) = 0.$$

Se $r(x) = 0$ temos que $f(x) = g(x) \cdot q(x)$ e, portanto, $g(x) = x - x_0$ divide $f(x)$. Por outro lado, se $gr(r(x)) = 0$ então $r(x) = r_0$ constante não nula. Logo, x_0 é raiz de $f(x)$ se, e somente se,

$$0 = f(x_0) = (x_0 - x_0) \cdot q(x) + r_0 = r_0,$$

ou seja, se, e somente se, $r_0 = 0$. Portanto, temos que

$$f(x) = g(x) \cdot q(x),$$

que significa $g(x) = x - x_0$ dividir $f(x)$. □

O Exemplo 1.3.7 mostra esse fato, onde $f(1) = 1^4 + 1^3 + 1 + 1 = 0$ em $\mathbb{Z}_2$, e $x - 1 = x + 1 = g(x)$ divide $f(x)$.

Definição 1.3.10. Sejam K um corpo e $p(x) \in K[x]$ tal que $gr(p(x)) \geq 1$. Dizemos que $p(x)$ é um **polinômio irredutível** sobre K se $p(x) = s(x) \cdot t(x)$, tal que $s(x), t(x) \in K[x]$, então $s(x) = c_1 \in K \backslash \{0\}$ ou $t(x) = c_2 \in K \backslash \{0\}$.

Como consequência da definição acima, dado $f(x) \in K[x]$, se $f(x)$ divide o polinômio irredutível $p(x) \in K[x]$, então ou $f(x) \in K\backslash\{0\}$ ou existe $c \in K\backslash\{0\}$ tal que $f(x) = c \cdot p(x)$.

Definição 1.3.11. Seja K um corpo. Dados $f(x), g(x) \in K[x]$, um polinômio $d(x) \in K[x]$ é o **máximo divisor comum** de $f(x)$ e $g(x)$ se:

i) $d(x) \mid f(x)$ e $d(x) \mid g(x)$

ii) $\forall d_1(x) \in K[x]$, se $d_1(x) \mid f(x)$ e $d_1(x) \mid g(x)$ então $d_1(x) \mid d(x)$.

O máximo divisor comum de $f(x)$ e $g(x)$ é denotado por $\text{MDC}(f(x), g(x))$.

Definição 1.3.12. Seja K um corpo. Dados os polinômios não nulos $f(x), g(x) \in K[x]$, um polinômio $m(x) \in K[x]$ é o **mínimo múltiplo comum** de $f(x)$ e $g(x)$ se:

i) $m(x)$ é mônico

ii) $f(x) \mid m(x)$ e $g(x) \mid m(x)$

iii) $\forall m_1(x) \in K[x]$, se $f(x) \mid m_1(x)$ e $g(x) \mid m_1(x)$ então $m(x) \mid m_1(x)$.

O mínimo múltiplo comum de $f(x)$ e $g(x)$ é denotado por $\text{MMC}(f(x), g(x))$.

Teorema 1.3.13. Um polinômio de grau n com coeficientes num corpo possui, no máximo, n raízes distintas.

Demonstração: Sejam o corpo K, o polinômio $p(x) \in K[x]$ e $\alpha_1, \ldots, \alpha_m \in K$ raízes distintas de $p(x)$. Seja $g_1(x) \in K[x]$. Pela Proposição 1.3.9,

$$f(x) = (x - \alpha_1) \cdot g_1(x).$$

Como α_2 também é raiz de $f(x)$. Então $f(\alpha_2) = 0$, logo $f(\alpha_2) = (\alpha - \alpha_1) \cdot g_1(\alpha_2)$. Como $\alpha_1 \neq \alpha_2$, $g_1(\alpha_2) = 0$, portanto α_2 é raiz de $g_1(x)$. Então, existe $g_2(x) \in K[x]$ tal que

$$g_1(x) = (x - \alpha_2) \cdot g_2(x).$$

Logo,

$$f(x) = (x - \alpha_1) \cdot (x - \alpha_2) \cdot g_2(x).$$

Podemos repetir esse raciocínio até que tenhamos:

$$f(x) = (x - \alpha_1) \cdot (x - \alpha_2) \cdot \ldots \cdot (x - \alpha_m) \cdot g_m(x)$$

então $gr(f(x)) = n = m + gr(g_m(x))$. Portanto, $m \leq n$. □

Voltaremos a definições e resultados envolvendo polinômios nos demais capítulos deste trabalho sempre que for necessário.

2 CORPOS FINITOS

2.1 Conceitos Básicos

Definição 2.1.1. Seja K um corpo finito com elemento unidade 1. Definimos a **característica** de K, como o menor inteiro positivo λ, tal que

$$\sum_{i=1}^{\lambda} 1 = 0.$$

Exemplo 2.1.2. A característica do corpo $\mathbb{Z}_2$ é 2, pois $\sum_{i=1}^{2} 1 = 1 + 1 = 0$.

Teorema 2.1.3. Sejam K um corpo finito e λ a característica de K, então λ é um número primo.

Demonstração: Suponhamos que λ não seja primo. Então existem $\lambda_1, \lambda_2 \in \mathbb{Z}$ tais que $1 < \lambda_1 < \lambda$ e $1 < \lambda_2 < \lambda$, onde $\lambda = \lambda_1 \cdot \lambda_2$. Logo, temos que:

$$0 = \lambda \cdot 1 = (\lambda_1 \cdot \lambda_2) \cdot 1 = \lambda_1 \cdot (\lambda_2 \cdot 1) = (\lambda_1 \cdot 1) \cdot (\lambda_2 \cdot 1).$$

Como K é também um domínio de integridade, temos então que $\lambda_1 \cdot 1 = 0$ ou $\lambda_2 \cdot 1 = 0$, o que é uma contradição, pois λ é o menor inteiro positivo tal que $\lambda \cdot 1 = 0$. Portanto, λ é primo. □

Teorema 2.1.4. Sejam K um corpo finito com q elementos e $\alpha \in K$, onde $\alpha \neq 0$. Então, $\alpha^{q-1} = 1$.

Demonstração: Sejam K um corpo finito com q elementos, $\beta_1, \ldots, \beta_{q-1}$ os $q-1$ elementos não nulos de K e α um elemento não nulo de K. Como a multiplicação é fechada em K, os $q-1$ elementos $\alpha \cdot \beta_1, \ldots, \alpha \cdot \beta_{q-1}$ são não nulos e distintos entre si, dois a dois. Então, temos que

$$(\alpha \cdot \beta_1) \cdot \ldots \cdot (\alpha \cdot \beta_{q-1}) = \beta_1 \cdot \ldots \cdot \beta_{q-1}$$

$$\alpha^{q-1} \cdot (\beta_1 \cdot \ldots \cdot \beta_{q-1}) = \beta_1 \cdot \ldots \cdot \beta_{q-1}$$

Mas como, por hipótese, temos $\alpha \neq 0$ e $\beta_1 \cdot \ldots \cdot \beta_{q-1} \neq 0$, devemos ter $\alpha^{q-1} = 1$. □

Como consequência do Teorema 2.1.4, temos que $\alpha^q = \alpha$, $\alpha \neq 0 \in K$.

Definição 2.1.5. Sejam $K^* = K \backslash \{0\}$ um corpo finito e $\alpha \in K^*$. A **ordem** de α é o menor inteiro positivo n tal que $\alpha^n = 1$.

Teorema 2.1.6. Sejam K um corpo finito com q elementos e $\alpha \in K$, onde $\alpha \neq 0$. Seja n a ordem de α. Então, n divide $q-1$.

Demonstração: Suponhamos que n não divida $q-1$. Ao dividirmos $q-1$ por n, obtemos $q-1 = ns + r$, onde $n > r > 0$. Então, temos que

$$\alpha^{q-1} = \alpha^{ns+r} = \alpha^{ns} \cdot \alpha^r = (\alpha^n)^s \cdot \alpha^r.$$

Mas como $\alpha^{q-1} = 1$ e $\alpha^n = 1$, pois n é a ordem de α, então devemos ter $\alpha^r = 1$. Mas isto é um absurdo porque $n > r > 0$ e n é o menor inteiro tal que $\alpha^n = 1$. portanto, n divide $q-1$. □

Definição 2.1.7. Seja K um corpo finito com q elementos. Chamamos de **elemento primitivo** a um elemento não nulo $\alpha \in K$ cuja ordem é $q-1$

Uma consequência direta da definição acima é que as potências de um elemento primitivo de um corpo finito geram todos os elementos não nulos desse corpo finito.

Teorema 2.1.8. Todo corpo finito possui elementos primitivos.

A demonstração desse teorema pode ser consultada em (HEFEZ; VILLELA, 2017, p. 78).

O exemplo a seguir ilustra o que foi dito nos teoremas e definições acima sobre elemento primitivo e ordem de um elemento de um corpo finito.

Exemplo 2.1.9. Considere o corpo finito $\mathbb{Z}_5$, cujas operações de adição e multiplicação estão ilustradas na tabela 2 a seguir.

Tabela 2 - Operações em $\mathbb{Z}_5$

Adição em $\mathbb{Z}_5$

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Multiplicação em $\mathbb{Z}_5$

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	1	2	3	4
4	0	4	3	2	1

Fonte: O autor, 2019

O corpo finito $\mathbb{Z}_5$ tem característica $\lambda = 5$ e possui $q = 5$ elementos. Se usarmos a tabela da multiplicação para calcularmos as potências do elemento 2, obteremos: $2^1 = 2$, $2^2 = 4$, $2^3 = 3$ e $2^4 = 1$. Observamos então que o elemento 2 é primitivo, pois $2^{5-1} = 2^4 = 1$. Isto é, a ordem do elemento 2 é $q - 1 = 5 - 1 = 4$. Portanto, as potências do elemento 2 geram todos os elementos não nulos do corpo finito $\mathbb{Z}_5$.

Por outro lado, o elemento 4 não é um elemento primitivo em $\mathbb{Z}_5$. Isto pode ser verificado analisando as potências desse elemento, que são as seguintes: $4^1 = 4$, $4^2 = 1$, $4^3 = 4$ e $4^4 = 1$. A ordem do elemento 4 é $2 \neq 4 = 5-1 = q-1$ e 2 divide $4 = 5-1 = q-1$. Logo, as potências do elemento 4 não geram todos os elementos do corpo finito $\mathbb{Z}_5$.

Nas duas próximas seções apresentaremos um método para construir extensões do corpo $\mathbb{Z}_2$ e, em seguida, daremos algumas propriedades de corpos finitos que sejam relevantes para o embasamento do código corretor de erros BCH binários. Alertamos que muitos dos resultados serão apenas mencionados no texto, sem que haja um aprofundamento por exceder à finalidade deste trabalho.

2.2 Extensão de um Corpo Finito

Definição 2.2.1. Sejam F e K dois corpos finitos. F é uma **extensão** de K se $K \subset F$, $F \neq K$.

Observação 2.2.2. Não é difícil verificar que se F é uma extensão de K, então F é um espaço vetorial sobre K. Para isso, basta termos as seguintes operações:

$$\begin{array}{rccc} + : & F \times F & \longrightarrow & F \\ & (u, v) & \longmapsto & u + v \end{array} \qquad \text{e} \qquad \begin{array}{ccc} K \times F & \longrightarrow & F \\ (\lambda, u) & \longmapsto & \lambda \cdot u \end{array}$$

Definição 2.2.3. O **grau** de extensão do corpo F é sua dimensão como espaço vetorial sobre K, que será denotada por $[F : K]$.

Como F e K são finitos, então $[F : K]$ é um número inteiro positivo.

Proposição 2.2.4. Sejam F e K corpos finitos, K com q elementos. Se F é extensão de K então F possui q^m elementos, onde $m = [F : K]$.

Demonstração: Como F pode ser visto como um espaço vetorial sobre K, logo a dimensão deste espaço vetorial é finita e igual a m, sendo $m = [F : K]$.
Temos que F possui uma base $b = (b_1, b_2, \ldots, b_m)$ e qualquer elemento $y \in F$; y pode ser escrito de forma única como $y = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_m b_m$, sendo $\lambda_1, \lambda_2, \ldots, \lambda_m$ pertencentes a K. Como K tem q elementos então F possui exatamente q^m elementos. □

Exemplo 2.2.5. Se $K = \mathbb{Z}_p$ então qualquer extensão finita F de K terá p^m elementos, sendo $m = [F : K]$ a dimensão do espaço vetorial F.

Exemplo 2.2.6. Seja $K = \mathbb{Z}_2$ e $m = 3$ então a dimensão de F, como espaço vetorial, é 3. E F tem 2^3 elementos que são: (0,0,0),(0,0,1),(0,1,0),(1,0,0),(0,1,1),(1,1,0),(1,0,1),(1,1,1).

2.3 Construção de Extensões de Corpos Finitos com 2^m elementos

O código corretor de erros que mostraremos neste texto trabalha com polinômios sobre extensões F de $\mathbb{Z}_2$, ou seja, sobre corpos com 2^m elementos, sendo m a dimensão de F. Portanto, nesta seção apresentaremos um método para a construção de corpos finitos de 2^m elementos, com $m > 1$. E usaremos como base a referência bibliográfica (LIN; DANIEL, 1983).

Teorema 2.3.1. Seja $p(x)$ um polinômio irredutível em $\mathbb{Z}_2$ com $gr(p(x)) = m$, então $p(x) \mid (x^{2^m-1} + 1)$.

Demonstração: A demonstração deste teorema encontra-se em (BERLEKAMP, 2015, p. 103). □

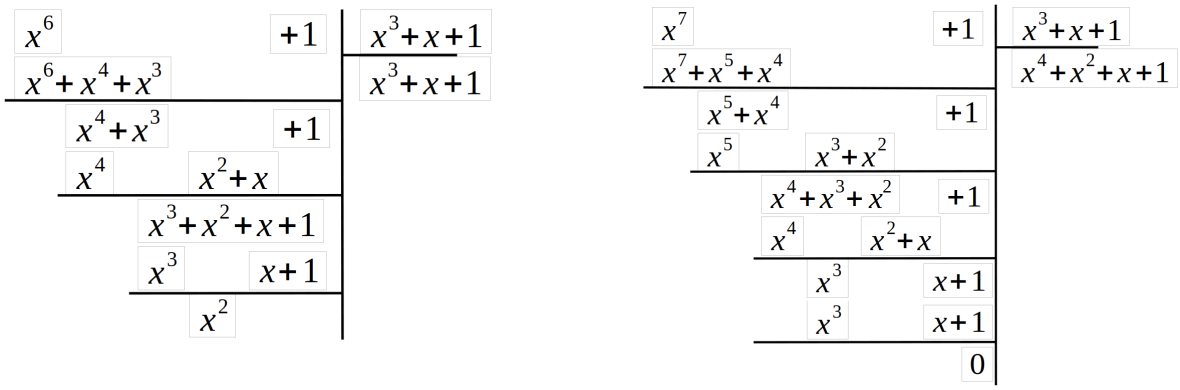
Definição 2.3.2. Seja $p(x)$ um polinômio irredutível em $\mathbb{Z}_2$, com $gr(p(x)) = m$. Dizemos que $p(x)$ é um **polinômio primitivo** se $p(x)$ não divide nenhum polinômio da forma $x^t + 1$ para $1 \leq t < 2^m - 1$.

Exemplo 2.3.3. Seja o corpo $K = \mathbb{Z}_2$. Sejam os polinômios $f(x), g(x) \in \mathbb{Z}_2$. O polinômio $g(x) = x^3 + x + 1$ é um polinômio primitivo sobre $\mathbb{Z}_2$, pois ele divide $f(x) = x^7 + 1$ com $n = 2^3 - 1 = 7$, mas não divide nenhum dos seguintes polinômios $h(x) = x^6 + 1$, $u(x) = x^5 + 1$, $v(x) = x^4 + 1$ e $w(x) = x^3 + 1$, onde $h(x), u(x), v(x), w(x) \in \mathbb{Z}_2$.
Obviamente que $g(x)$ não divide $w(x)$. Porém, os resultados das outras divisões podem ser verificadas nas figuras 1 e 2 a seguir.

Figura 1 - Divisões de $v(x) = x^4 + 1$ e $u(x) = x^5 + 1$ por $g(x) = x^3 + x + 1$

Fonte: O autor, 2019

Figura 2 - Divisões de $h(x) = x^6 + 1$ e $f(x) = x^7 + 1$ por $g(x) = x^3 + x + 1$



Fonte: O autor, 2019

Em geral, não é fácil identificar um polinômio primitivo. Além disso, dado um valor de m, podem existir mais de um polinômio primitivo de grau m. Na tabela 3 a seguir estão listados alguns polinômios primitivos sobre $\mathbb{Z}_2$, sendo um para cada valor de m, onde $3 \leq m \leq 10$, com o menor número possível de termos.

Tabela 3 - Polinômios primitivos de grau m $(3 \leq m \leq 10)$

m	polinômio primitivo	m	polinômio primitivo
3	$x^3 + x + 1$	7	$x^7 + x^3 + 1$
4	$x^4 + x + 1$	8	$x^8 + x^4 + x^3 + x^2 + 1$
5	$x^5 + x^2 + 1$	9	$x^9 + x^4 + 1$
6	$x^6 + x + 1$	10	$x^{10} + x^3 + 1$

Fonte: (LIN; DANIEL, 1983)

Consideremos os elementos $0, 1 \in \mathbb{Z}_2$ e um novo símbolo α. Construiremos um conjunto formado por 0,1 e potências de α e, para isso, definimos a multiplicação " $\cdot$ " da seguinte forma:

$$\begin{aligned} &0 \cdot 0 = 0, \\ &0 \cdot 1 = 1 \cdot 0 = 0, \\ &1 \cdot 1 = 1, \\ &0 \cdot \alpha = \alpha \cdot 0 = 0, \\ &1 \cdot \alpha = \alpha \cdot 1 = \alpha, \\ &\alpha^i = \overbrace{\alpha \cdot \ldots \cdot \alpha}^{i \text{ vezes}}, \text{ com } i \geq 2. \end{aligned} \tag{1}$$

Como consequência da definição de multiplicação dada acima, temos que

$$\begin{aligned} &0 \cdot \alpha^i = \alpha^i \cdot 0 = 0, \\ &1 \cdot \alpha^i = \alpha^i \cdot 1 = \alpha^i, \\ &\alpha^i \cdot \alpha^j = \alpha^j \cdot \alpha^i = \alpha^{i+j}. \end{aligned} \tag{2}$$

Dessa forma, temos o conjunto $E = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^i, \ldots\}$ no qual a operação de multiplicação está definida. Notemos que o elemento 1 será denotado por α^0.

Necessitamos impor uma condição ao elemento α para que o conjunto E tenha apenas 2^m elementos e seja fechado em relação à operação de multiplicação definida em (1).

Seja $p(x)$ um polinômio primitivo de grau m sobre $\mathbb{Z}_2$. Logo, $p(x)$ é irredutível em $\mathbb{Z}_2$ e divide $x^{2^m-1} + 1$. Então , temos que

$$x^{2^m-1} + 1 = p(x) \cdot q(x), \text{ com } q(x) \in \mathbb{Z}_2. \tag{3}$$

Assumimos que α é raiz de $p(x)$ em E, ou seja, $p(\alpha) = 0$. Substituindo x por α em (3) obtemos

$$\alpha^{2^m-1} + 1 = p(\alpha) \cdot q(\alpha)$$

Mas como $p(\alpha) = 0$, temos

$$\alpha^{2^m-1} + 1 = 0 \cdot q(\alpha)$$

O que resulta na seguinte igualdade, pelas operações definidas em (1)

$$\alpha^{2^m-1} + 1 = 0.$$

Adicionando 1 a ambos os lados dessa igualdade, usando a adição em $\mathbb{Z}_2$, obtemos

$$\alpha^{2^m-1} = 1. \tag{4}$$

Portanto, sob a condição de que α é raiz de $p(x)$, obtemos a igualdade (4) que juntamente com as operações em (1), torna E finito contendo os seguintes elementos

$$E^* = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{2^m-2}\}.$$

Observemos que o conjunto E^* tem 2^m elementos, é fechado em relação a operação de multiplicação definida em (1) e ainda pode-se provar que $E^* \backslash \{0\}$ forma um grupo comutativo com a operação " $\cdot$ ".

A seguir, definiremos a operação de adição " + " em E^*.

Para $0 \leq i < 2^m - 1$, dividimos o polinômio x^i por $p(x)$ e obtemos

$$x^i = q_i(x)p(x) + a_i(x) \tag{5}$$

sendo $q_i(x)$ e $a_i(x)$ polinômios sobre $\mathbb{Z}_2$ e $gr(a_i(x)) \leq m-1$. Assim,

$$a_i(x) = a_{i,0} + a_{i,1}x + a_{i,2}x^2 + \ldots + a_{i,m-1}x^{m-1}. \tag{6}$$

Como x e $p(x)$ não têm nenhum fator em comum diferente de 1, então x não é divisível por $p(x)$, logo

$$a_i(x) \neq 0 \tag{7}$$

para todo $i \geq 0$.

Além disso, mostraremos que $a_i(x) \neq a_j(x)$ para todo $0 \leq i, j < 2^m - 1$.

Suponhamos que $a_i(x) = a_j(x)$ para algum i e j, $i \neq j$. Vamos considerar $i < j$. Temos que

$$x^i + x^j = [q_i(x) + q_j(x)]p(x) + a_i(x) + a_j(x) = [q_i(x) + q_j(x)]p(x),$$

logo $p(x) \mid (x^i + x^j)$ e, portanto, $p(x) \mid x^i(1 + x^{j-i})$. Como x^i e $p(x)$ não têm fator em comum diferente de 1, então $p(x) \mid (1 + x^{j-i})$. O que é um absurdo, pois $p(x)$ é um polinômio primitivo de grau m, logo $p(x)$ não divide $(1 + x^{j-i})$, então

$$a_i(x) \neq a_j(x) \tag{8}$$

Podemos observar que, desta forma, existirão $2^m - 1$ polinômios não nulos $a_0(x), a_1(x), \ldots, a_{2^m-2}(x)$, todos distintos tais que $gr(a_i(x)) \leq m-1$.

Se substituirmos x por α em (5)

$$\alpha^i = q_i(\alpha)p(\alpha) + a_i(\alpha).$$

Como α é raiz de $p(x)$ então

$$\alpha^i = a_i(\alpha) \tag{9}$$

De (6) e (9), temos que

$$\alpha^i = a_{i,0} + a_{i,1}\alpha + a_{i,2}\alpha^2 + \ldots + a_{i,m-1}\alpha^{m-1}. \tag{10}$$

Além disso, de (7), (8) e (10), vemos que os elementos não nulos $\alpha^0, \alpha^1, \alpha^2, \ldots, \alpha^{2^m-2}$ de E^* podem ser representados por $2^m - 1$ polinômios não nulos de α sobre $\mathbb{Z}_2$, com graus menores ou iguais do que $m-1$.

O elemento $0 \in E^*$ será representado pelo polinômio identicamente nulo. Dessa forma, os 2^m elementos de E^*, que são distintos entre si, serão representados por 2^m

polinômios distintos de α sobre $\mathbb{Z}_2$, com grau menor ou igual a $m-1$.

Definimos a operação de adição " + " em E^* da seguinte forma:

$$\begin{aligned} 0+0 &= 0, \\ 0+\alpha^i &= \alpha^i+0=\alpha^i, \\ \alpha^i+\alpha^j &= (a_{i,0}+a_{i,1}\alpha+\cdots+a_{i,m-1}\alpha^{m-1})+(a_{j,0}+a_{j,1}\alpha+\cdots+a_{j,m-1}\alpha^{m-1}) \\ &= (a_{i,0}+a_{j,0})+(a_{i,1}+a_{j,1})\alpha+\cdots+(a_{i,m-1}+a_{j,m-1})\alpha^{m-1}. \end{aligned} \tag{11}$$

sendo 0(zero) o polinômio nulo em $\mathbb{Z}_2$ e para $0 \leq i, j < 2^m-1$, cada coeficiente dado por $a_{i,k}+a_{j,k}$ pertence a $\mathbb{Z}_2$ $(0 \leq k \leq m-1)$.
Como consequência da definição dada acima, observamos que:

$$\alpha^i+\alpha^j = \begin{cases} 0 \quad \text{se} \quad i=j, \\ (a_{i,0}+a_{j,0})+(a_{i,1}+a_{j,1})\alpha+\cdots+(a_{i,m-1}+a_{j,m-1})\alpha^{m-1} \quad \text{se} \quad i \neq j. \end{cases}$$

A expressão polinomial $(a_{i,0}+a_{j,0})+(a_{i,1}+a_{j,1})\alpha+\cdots+(a_{i,m-1}+a_{j,m-1})\alpha^{m-1}$ é não nula e é a representação de algum $\alpha^k \in E^*$. Portanto, conforme (LIN; DANIEL, 1983) o conjunto E^* é fechado em relação à operação de adição " + " definida em (11) e forma um grupo comutativo.

Não é difícil mostrar que a multiplicação em E^* é distributiva. Basta usarmos a notação polinomial para os elementos de E^*. Como a multiplicação em $E^*\backslash\{0\}$ é um grupo comutativo, então o conjunto $E^* = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{2^m-2}\}$ é um corpo finito de 2^m elementos, que será denotado por $\mathbb{F}_{2^m}$. Notemos que as operações de adição e multiplicação definidas em $E^* = \mathbb{F}_{2^m}$ são as operações de adição e multiplicação em $\mathbb{Z}_2$. Portanto, $\mathbb{Z}_2$ é um subcorpo de $\mathbb{F}_{2^m}$, cuja característica é 2.

Durante o processo de construção deste corpo a partir de $\mathbb{Z}_2$, foram desenvolvidas duas representações para os elementos não nulos de $\mathbb{F}_{2^m}$, que são as seguintes: a representação por meio de potências e a representação polinomial. Nas operações de multiplicação é conveniente utilizarmos a primeira representação. Enquanto que, nas operações de adição, é oportuno trabalharmos com a segunda representação. Vejamos o exemplo abaixo.

Exemplo 2.3.4. Vamos construir o corpo finito $\mathbb{F}_{2^3}$, por exemplo. Sejam $m=3$ e $p(x) = 1+x+x^3 \in \mathbb{Z}_2[x]$ um polinômio primitivo. Seja α uma raiz de $p(x)$. Então, $p(\alpha) = 1+\alpha+\alpha^3 = 0$. Logo, $\alpha^3 = \alpha+1$. Usando esta igualdade, podemos construir os elementos de $\mathbb{F}_{2^3}$ cujos expoentes são maiores que 3. Vejamos:

$$\begin{aligned} &\alpha^4 = \alpha\cdot\alpha^3 = \alpha\cdot(1+\alpha) = \alpha+\alpha^2, \\ &\alpha^5 = \alpha^2\cdot\alpha^3 = \alpha^2\cdot(1+\alpha) = \alpha^2+\alpha^3 = 1+\alpha+\alpha^2, \\ &\alpha^6 = \alpha^3\cdot\alpha^3 = (1+\alpha)\cdot(1+\alpha) = 1+\alpha+\alpha+\alpha^2 = 1+\alpha^2. \end{aligned}$$

Todos os elementos de $\mathbb{F}_{2^3}$, gerados por $p(x) = 1 + x + x^3$, estão representados de três formas distintas na tabela 4.

Tabela 4 - Representações dos elementos de $\mathbb{F}_{2^3}$

Potência	Polinômio	Vetor
0	0	(0 0 0)
1	1	(1 0 0)
α	α	(0 1 0)
α^2	α^2	(0 0 1)
α^3	$1+\alpha$	(1 1 0)
α^4	$\alpha+\alpha^2$	(0 1 1)
α^5	$1+\alpha+\alpha^2$	(1 1 1)
α^6	$1+\alpha^2$	(1 0 1)

Fonte: O autor, 2019

Observação 2.3.5. No corpo finito $\mathbb{F}_{2^3}$, para efetuarmos a multiplicação de dois elementos, α^i e α^j, $0 \leq i, j \leq 6$, somamos seus respectivos expoentes e usamos $\alpha^7 = 1$ caso essa soma seja maior do que 7. Por exemplo, $\alpha^5 \cdot \alpha^6 = \alpha^{11} = \alpha^4 \cdot \alpha^7 = \alpha^4$. Para efetuarmos a divisão de α^i por α^j, multiplicamos α^i pelo inverso multiplicativo de α^j, que é α^{7-j}. Assim, $\alpha^2/\alpha^3 = \alpha^2 \cdot \alpha^{7-3} = \alpha^2 \cdot \alpha^4 = \alpha^6$. Para efetuarmos a adição de dois elementos α^i e α^j, usamos a representação polinomial desses elementos dadas na tabela 4. Então, para efetuarmos a adição $\alpha^3 + \alpha^5$, procedemos da seguinte forma:

$$\alpha^3 + \alpha^5 = (1 + \alpha) + (1 + \alpha + \alpha^2) = \alpha^2.$$

2.4 Algumas Propriedades de Corpos Finitos com 2^m elementos

Ao nos referirmos ao corpo $\mathbb{R}$, sabemos que existem polinômios com coeficientes reais cujas raízes não pertencem a esse corpo, e sim ao conjunto $\mathbb{C}$, que é uma extensão de $\mathbb{R}$. Para exemplificar, consideremos o polinômio $p(x) = x^2 + 2x + 2 \in \mathbb{R}[x]$, que não possui raízes em $\mathbb{R}$. As raízes de $p(x)$ são os números complexos $-1-i$ e $-1+i$, que estão em $\mathbb{C}$. Isto também ocorre com os polinômios com coeficientes em $\mathbb{Z}_2$. Há polinômios em $\mathbb{Z}_2[x]$ cujas raízes não pertencem ao corpo $\mathbb{Z}_2$, mas estão em um corpo que é uma extensão de $\mathbb{Z}_2$. Por exemplo, consideremos o polinômio $q(x) = x^3 + x^2 + 1 \in \mathbb{Z}_2[x]$, que é irredutível em $\mathbb{Z}_2$. Entretanto, $q(x)$ possui três raízes no corpo $\mathbb{F}_{2^3}$, que é uma extensão de $\mathbb{Z}_2$. É imediato que 0, 1 e α de $\mathbb{F}_{2^3}$ não são raízes de $q(x)$. Mas, ao substituirmos os demais elementos desse corpo em $q(x)$, dados pela tabela 4, verificamos que α^3, α^5 e α^6 são as

raízes de $q(x) = x^3 + x^2 + 1$, como podemos ver a seguir:

$$q(\alpha^2) = (\alpha^2)^3 + (\alpha^2)^2 + 1 = \alpha^6 + \alpha^4 + 1 = \alpha^2 + 1 + \alpha^2 + \alpha + 1 = \alpha,$$
$$q(\alpha^3) = (\alpha^3)^3 + (\alpha^3)^2 + 1 = \alpha^9 + \alpha^6 + 1 = \alpha^2 + \alpha^2 + 1 + 1 = 0,$$
$$q(\alpha^4) = (\alpha^4)^3 + (\alpha^4)^2 + 1 = \alpha^{12} + \alpha^8 + 1 = \alpha^5 + \alpha + 1 = \alpha^2 + \alpha + 1 + \alpha + 1 = \alpha^2,$$
$$q(\alpha^5) = (\alpha^5)^3 + (\alpha^5)^2 + 1 = \alpha^{15} + \alpha^{10} + 1 = \alpha + \alpha^3 + 1 = \alpha + \alpha + 1 + 1 = 0,$$
$$q(\alpha^6) = (\alpha^6)^3 + (\alpha^6)^2 + 1 = \alpha^{18} + \alpha^{12} + 1 = \alpha^4 + \alpha^5 + 1 = \alpha^2 + \alpha + \alpha^2 + \alpha + 1 + 1 = 0.$$

Como $q(x) = x^3 + x^2 + 1$ é um polinômio de grau 3, ele só pode ter, no máximo, três raízes que são α^3, α^5 e α^6. Então, temos que $q(x) = x^3 + x^2 + 1$ é igual ao produto $(x + \alpha^3)(x + \alpha^5)(x + \alpha^6)$.

Definição 2.4.1. Seja o corpo $\mathbb{F}_{2^m}$ e $\beta \in \mathbb{F}_{2^m}$. O elemento β^{2^l}, para qualquer $l \geq 0$, é chamado de **conjugado de β**.

Teorema 2.4.2. Sejam $p(x) \in \mathbb{Z}_2[x]$, o corpo $\mathbb{F}_{2^m}$ e $\beta \in \mathbb{F}_{2^m}$. Se β é uma raiz de $p(x)$, então todos os conjugados de β também são raízes de $p(x)$.

A demonstração deste teorema pode ser vista em (LIN; DANIEL, 1983, pp. 34–35).

Exemplo 2.4.3. O elemento α^3 do corpo $\mathbb{F}_{2^3}$ é uma raiz de $q(x) = x^3 + x^2 + 1$. De acordo com o Teorema 2.4.2, os conjugados de α^3 também são raízes de $q(x) = x^3 + x^2 + 1$. Isto pode ser verificado, usando a tabela 4 e $\alpha^7 = 1$. Portanto, os conjugados de α^3 são:

$$(\alpha^3)^2 = \alpha^6, \qquad (\alpha^3)^{2^2} = \alpha^{12} = \alpha^5, \qquad (\alpha^3)^{2^3} = \alpha^{24} = \alpha^3.$$

Notemos que $(\alpha^3)^{2^4} = \alpha^{48} = \alpha^6$. Portanto, não calcularemos os valores dos próximos conjugados de α^3, pois obteremos os já encontrados α^6, α^5 e o próprio α^3.

Teorema 2.4.4. Seja o corpo finito $\mathbb{F}_{2^m}$. Então, os $2^m - 1$ elementos não nulos de $\mathbb{F}_{2^m}$ geram todas as raízes de $x^{2^m-1} + 1$.

Demonstração: Sejam um corpo finito $\mathbb{F}_{2^m}$ e $\beta \in \mathbb{F}_{2^m}$, com $\beta \neq 0$. De acordo com o Teorema 2.1.4, temos que

$$\beta^{2^m-1} = 1.$$

Somando 1 a ambos os lados dessa igualdade, obtemos

$$\beta^{2^m-1} + 1 = 0.$$

Isto significa que β é uma raiz do polinômio $x^{2^m-1} + 1$. Portanto, todo elemento não nulo do corpo $\mathbb{F}_{2^m}$ é uma raiz de $x^{2^m-1} + 1$. Como o grau de $x^{2^m-1} + 1$ é $2^m - 1$, então os $2^m - 1$ elementos não nulos de $\mathbb{F}_{2^m}$ geram todas as raízes do polinômio $x^{2^m-1} + 1$. □

Corolário 2.4.5. Seja o corpo finito $\mathbb{F}_{2^m}$. Os elementos de $\mathbb{F}_{2^m}$ geram todas as raízes de $x^{2^m} + x$.

Demonstração: Seja o corpo finito $\mathbb{F}_{2^m}$. Considerando que o elemento 0 (zero) de $\mathbb{F}_{2^m}$ é a raiz do polinômio $p(x) = x$, e de acordo com o Teorema 2.4.4, temos que os elementos não nulos de $\mathbb{F}_{2^m}$ geram todas as raízes de $x^{2^m-1} + 1$. Portanto, os elementos de $\mathbb{F}_{2^m}$ geram todas as raízes de $x^{2^m} + x = x(x^{2^m-1} + 1)$. □

Definição 2.4.6. Sejam os corpos finitos $\mathbb{Z}_2$ e $\mathbb{F}_{2^m}$. Sejam $\beta \in \mathbb{F}_{2^m}$ e $min_\beta(x) \in \mathbb{Z}_2[x]$ o polinômio de menor grau, tal que $min_\beta(\beta) = 0$. O polinômio $min_\beta(x)$ é chamado de **polinômio mínimo de β**.

Exemplo 2.4.7. O polinômio mínimo do elemento 0 (zero) de $\mathbb{F}_{2^m}$ é o polinômio $min_0(x) = x \in \mathbb{Z}_2[x]$, pois $min_0(0) = 0$. E o polinômio mínimo do elemento 1 (unitário) de $\mathbb{F}_{2^m}$ é o polinômio $min_1(x) = x + 1 \in \mathbb{Z}_2[x]$, pois $min_1(1) = 1 + 1 = 0$.

Teorema 2.4.8. Sejam o corpo finito $\mathbb{F}_{2^m}$ e $\beta \in \mathbb{F}_{2^m}$. O polinômio mínimo de β, $min_\beta(x)$, é um polinômio irredutível.

Demonstração: Sejam o corpo finito $\mathbb{F}_{2^m}$ e $\beta \in \mathbb{F}_{2^m}$. Suponhamos que $min_\beta(x)$ não seja irredutível e, portanto, pode ser escrito como produto dos dois polinômios $p_1(x)$ e $p_2(x)$. Então, $min_\beta(x) = p_1(x) \cdot p_2(x)$, onde os graus de ambos $p_1(x)$ e $p_2(x)$ são maiores do que 0 e menores do que o grau de $min_\beta(x)$. Como $min_\beta(\beta) = p_1(\beta) \cdot p_2(\beta) = 0$, então temos que $p_1(\beta) = 0$ ou $p_2(\beta) = 0$. E, isto contradiz a hipótese de que $min_\beta(x)$ é o polinômio de menor grau, tal que $min_\beta(\beta) = 0$. Logo, $min_\beta(x)$ é irredutível. □

Teorema 2.4.9. Sejam os corpos finitos $\mathbb{Z}_2$ e $\mathbb{F}_{2^m}$. Sejam um polinômio $p(x) \in \mathbb{Z}_2[x]$, $\beta \in \mathbb{F}_{2^m}$ e $min_\beta(x)$. Se $p(\beta) = 0$, então $min_\beta(x)$ divide $p(x)$.

Demonstração: Na divisão de $p(x)$ por $min_\beta(x)$, temos que $p(x) = min_\beta(x) \cdot q(x) + r(x)$, onde $q(x)$ é o quociente e $r(x)$ é o resto da divisão. Logo, devemos ter $gr(r(x)) < gr(min_\beta(x))$ ou $r(x) = 0$. Como β é raiz de $p(x)$ e $min_\beta(x)$ é o polinômio mínimo de β, então ao substituirmos x por β, obtemos $p(\beta) = min_\beta(\beta) \cdot q(\beta) + r(\beta) = 0$. E, consequentemente, temos que $r(\beta) = 0$. Se $r(x)$ é o polinômio identicamente nulo, o Teorema 2.4.9 está provado. Por outro lado, se $r(x) \neq 0$, então $r(x)$ é o polinômio mínimo de β. O que contradiz a hipótese de que $min_\beta(x)$ é o polinômio mínimo de β. Portanto, $r(x)$ é o polinômio identicamente nulo e $min_\beta(x)$ divide $p(x)$. □

Teorema 2.4.10. Sejam o corpo finito $\mathbb{F}_{2^m}$, $\beta \in \mathbb{F}_{2^m}$ e $min_\beta(x)$ o polinômio mínimo de β. O polinômio $x^{2^m} + x$ é divisível por $min_\beta(x)$.

Demonstração: Segue do Corolário 2.4.5 e do Teorema 2.4.9. □

Considerando β um elemento do corpo finito $\mathbb{F}_{2^m}$, o teorema acima nos diz que todas as raízes de um polinômio mínimo de β pertencem a $\mathbb{F}_{2^m}$. Os dois seguintes teoremas nos dirão como encontrar essas raízes.

Teorema 2.4.11. Sejam o corpo finito $\mathbb{F}_{2^m}$, $\beta \in \mathbb{F}_{2^m}$ e $p(x)$ irredutível em $\mathbb{Z}_2[x]$. Seja $min_\beta(x)$ o polinômio mínimo de β. Se $p(\beta) = 0$, então $min_\beta(x) = p(x)$.

Demonstração: De acordo com o Teorema 2.4.9, temos que $min_\beta(x)$ divide $p(x)$. Como $min_\beta(x) \neq 1$ e $p(x)$ é irredutível, devemos ter $min_\beta(x) = p(x)$. □

Teorema 2.4.12. Sejam o corpo finito $\mathbb{F}_{2^m}$, $\beta \in \mathbb{F}_{2^m}$ e $p(x) \in \mathbb{Z}_2[x]$. Seja e o menor inteiro não negativo tal que $\beta^{2^e} = \beta$. Então

$$p(x) = \prod_{i=0}^{e-1}(x + \beta^{2^i})$$

é um polinômio irredutível em $\mathbb{Z}_2[x]$.

A demonstração desse teorema pode ser encontrada em (LIN; DANIEL, 1983).

Teorema 2.4.13. Sejam o corpo finito $\mathbb{F}_{2^m}$, $\beta \in \mathbb{F}_{2^m}$ e $min_\beta(x)$ o polinômio mínimo de β. Seja e o menor inteiro não negativo tal que $\beta^{2^e} = \beta$. Então

$$min_\beta(x) = \prod_{i=0}^{e-1}(x + \beta^{2^i}). \tag{12}$$

sendo β^{2^i} os conjugados de β.

Demonstração. Esse teorema é uma consequência direta dos teoremas 2.4.11 e 2.4.12. □

Exemplo 2.4.14. Consideremos o corpo finito $\mathbb{F}_{2^3}$, cujos elementos estão ilustrados na tabela 4 (página 31). Seja $\beta = \alpha$. Então, os conjugados de β são

$$\beta^{2^1} = \alpha^2 \quad \text{e} \quad \beta^{2^2} = \alpha^4.$$

O polinômio mínimo de $\beta = \alpha$ é dado por

$$min_\beta(x) = (x + \alpha) \cdot (x + \alpha^2) \cdot (x + \alpha^4).$$

Efetuando as multiplicações no lado direito dessa igualdade de acordo com as informações da tabela 4 (página 31), e considerando que $\alpha^7 = 1$, obtemos

$$min_\beta(x) = (x^2 + (\alpha + \alpha^2)x + \alpha^3) \cdot (x + \alpha^4)$$

$$= (x^3 + (\alpha + \alpha^2 + \alpha^4)x^2 + (\alpha^3 + \alpha^5 + \alpha^6)x + \alpha^7$$

$$= x^3 + x + 1.$$

Todos os polinômios mínimos dos elementos de $\mathbb{F}_{2^3}$ são dados na tabela 5.

Tabela 5 - Polinômios mínimos dos elementos de $\mathbb{F}_{2^3}$ gerado por $p(x) = x^3 + x + 1$

Raízes conjugadas	Polinômios mínimos
0	x
1	$x + 1$
$\alpha, \alpha^2, \alpha^4$	$x^3 + x + 1$
$\alpha^3, \alpha^5, \alpha^6$	$x^3 + x^2 + 1$

Fonte: O autor, 2019

Teorema 2.4.15. Sejam o corpo finito $\mathbb{F}_{2^m}$, $\beta \in \mathbb{F}_{2^m}$ e $min_\beta(x)$ o polinômio mínimo de β. Seja e o grau de $min_\beta(x)$. Então e é o menor inteiro tal que $\beta^{2^e} = \beta$. Além disso, $e \leq m$.

Demonstração: O Teorema 2.4.15 é consequência direta do Teorema 2.4.13. □

Teorema 2.4.16. Sejam o corpo finito $\mathbb{F}_{2^m}$ e $\beta \in \mathbb{F}_{2^m}$. Se β é um elemento primitivo de $\mathbb{F}_{2^m}$, então todos os conjugados de β também são elementos primitivos de $\mathbb{F}_{2^m}$.

Demonstração: Sejam o corpo finito $\mathbb{F}_{2^m}$ e β um elemento primitivo de $\mathbb{F}_{2^m}$. Seja n a ordem de β^{2^l}, para $l > 0$. Então $(\beta^{2^l})^n = 1$, portanto

$$\beta^{n \cdot 2^l} = 1. \tag{13}$$

Além disso, de acordo com o Teorema 2.1.6, n divide $2^m - 1$. Então, podemos dizer que

$$2^m - 1 = k \cdot n. \tag{14}$$

Como β é um elemento primitivo de $\mathbb{F}_{2^m}$, a ordem de β é $2^m - 1$. Logo, temos que

$$\beta^{2^m - 1} = 1. \tag{15}$$

De (13) e (15) temos que

$$\beta^{n \cdot 2^l} = \beta^{2^m - 1},$$

e portanto, $n \cdot 2^l$ é um múltiplo de $2^m - 1$. Como 2^l e $2^m - 1$ são primos entre si, n é divisível por $2^m - 1$. Então

$$n = q \cdot (2^m - 1). \tag{16}$$

De (14) e (16) concluimos que $n = 2^m - 1$. Portanto, β^{2^l} é também um elemento primitivo de $\mathbb{F}_{2^m}$. □

Exemplo 2.4.17. Considere o corpo finito $\mathbb{F}_{2^3}$, dado pela tabela 4 (página 31), e o elemento $\beta = \alpha^3 \in \mathbb{F}_{2^3}$. As potências de β são as seguintes: $\beta^0 = 1$, $\beta^1 = \alpha^3$, $\beta^2 = (\alpha^3)^2 = \alpha^6$, $\beta^3 = (\alpha^3)^3 = \alpha^9 = \alpha^2$, $\beta^4 = (\alpha^3)^4 = \alpha^{12} = \alpha^5$, $\beta^5 = (\alpha^3)^5 = \alpha^{15} = \alpha$, $\beta^6 = (\alpha^3)^6 = \alpha^{18} = \alpha^4$, $\beta^7 = (\alpha^3)^7 = \alpha^{21} = 1$.

Notemos que as potências de $\beta = \alpha^3$ geram todos os elementos não nulos de $\mathbb{F}_{2^3}$, então $\beta = \alpha^3$ é um elemento primitivo de $\mathbb{F}_{2^3}$.

Como já vimos anteriormente, os conjugados de α^3 são os elementos α^5 e α^6. As potências desses dois elementos também geram todos os elementos não nulos de $\mathbb{F}_{2^3}$.

O teorema a seguir é uma forma generalizada do Teorema 2.4.16.

Teorema 2.4.18. Sejam o corpo finito $\mathbb{F}_{2^m}$ e $\beta \in \mathbb{F}_{2^m}$. Se a ordem de β é igual a n, então todos os conjugados de β têm ordem igual a n.

3 CÓDIGOS CORRETORES DE ERROS

Conforme nos referimos anteriormente, neste texto pretendemos mostrar a utilização de polinômios em códigos corretores de erros, mais especificamente, nos códigos BCH binários, que pertencem à classe dos códigos lineares e à subclasse dos códigos cíclicos. Para tanto, faremos uma introdução sobre os códigos corretores de erros. Em seguida, veremos os códigos lineares e algumas propriedades dos códigos cíclicos.

Neste capítulo, tomaremos principalmente como base as referências bibliográficas (HEFEZ; VILLELA, 2017) e (MILIES, 2009).

3.1 **Introdução**

Os códigos corretores de erros fazem parte do nosso dia a dia nas mais variadas formas, tanto na transmissão como no armazenamento de informações digitalizadas. São exemplos disso: o uso do celular, navegar pela internet, assistir a um filme em DVD e usar uma rede Wi-Fi.

Uma determinada informação pode sofrer algum tipo de alteração durante seu armazenamento ou transmissão. Essa alteração é chamada de **ruído** e pode ser gerada por um erro humano de digitação ou por uma interferência eletromagnética durante o envio de uma mensagem, por exemplo. A finalidade dos códigos corretores de erros é justamente detectar e corrigir tais erros.

Em 1948, o matemático C. E. Shannon criou a Teoria dos Códigos Corretores de Erros, que foi bastante desenvolvida nas duas décadas seguintes por outros matemáticos. Por causa das pesquisas espaciais e da grande popularização dos computadores, houve grande interesse por essa teoria a partir da década de 70. Atualmente, a utilização dos códigos corretores de erros visa garantir a confiabilidade das informações ao se transmitir ou armazenar dados.

O exemplo abaixo ilustra a utilização de códigos corretores de erros.

Exemplo 3.1.1. Vamos supor que um semáforo (sinal de trânsito) com as três cores (vermelho, amarelo e verde) receba um comando da central de tráfego. Cada comando acenderá uma dessas cores e irá apagar as duas restantes. Os três comandos acima podem ser codificados da seguinte forma:

$$\begin{aligned} \text{Vermelho} &\longmapsto 00 \\ \text{Amarelo} &\longmapsto 01 \\ \text{Verde} &\longmapsto 10 \end{aligned}$$

O conjunto {00,01,10} é denominado **código da fonte**. A central de tráfego vai transmitir esses códigos através de um canal que pode sofrer com interferências, as quais podem causar apenas um erro no código transmitido. Vamos imaginar que o comando 01 foi enviado ao semáforo, que recebeu o comando 10. Nesse caso, nenhum erro seria percebido pelo código, pois o comando 10 pertence ao código da fonte, e o semáforo iria acender a luz verde ao invés de acender a amarela e apagar as outras duas.

Por outro lado, se o comando 01 fosse enviado pela central de tráfego e o sinal de trânsito recebesse o comando 11. O código verificaria que houve um erro, pois o comando 11 não pertence ao código da fonte. Mesmo assim, ele não conseguiria corrigir esse erro, já que não teria como saber qual foi o comando realmente enviado pela central de tráfego.

Nesse caso, o que pode ser feito, então, é acrescentar redundâncias que permitem a detecção e correção de erros, gerando um novo código. Portanto, o código seria modificado como mostramos a seguir:

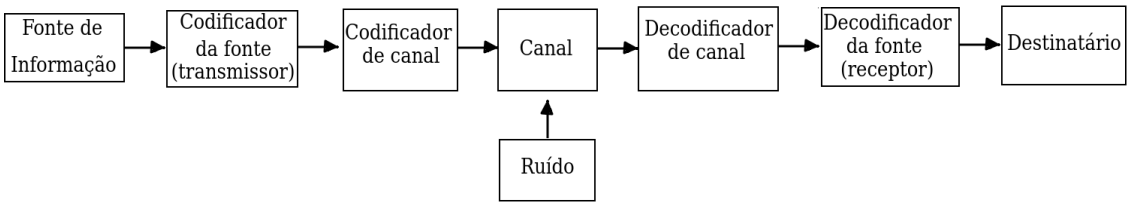
$$\begin{aligned} 00 &\longmapsto 00000 \\ 01 &\longmapsto 01101 \\ 10 &\longmapsto 10011 \end{aligned}$$

O conjunto {00000,01101,10011} é denominado **código do canal**. Com esse novo código, se algum erro for causado pelo canal de transmissão, o erro poderá ser detectado e corrigido.

Suponhamos agora, que o comando 01101 foi enviado ao semáforo, mas este recebeu o comando 01111, o qual não pertence ao código da fonte. Portanto, o código detectou o erro. Sabendo que esse canal pode causar apenas um erro a cada comando enviado, o código corrige o que foi recebido constatando que o comando efetivamente enviado foi 01101. Isto foi verificado porque o comando 01101 é o que tem o menor número de dígitos distintos em relação ao comando 01111.

A figura 3 descreve o procedimento adotado acima.

Figura 3 - Diagrama simplificado de transmissão ou armazenamento de dados



Fonte: Adaptada de (LIN; DANIEL, 1983)

Ao trabalharmos com códigos corretores de erros, temos como objetivo acrescentar redundâncias ao código da fonte transformando-o em código de canal. Com isso,

poderemos detectar e corrigir os possíveis erros que surgiram durante a trasnsmissão da informação. E, finalmente, decodificar o código de canal em código da fonte.

A seguir mostraremos os conceitos básicos dos códigos corretores de erros.

A construção de um código corretor de erros é feita a partir de um conjunto finito A, que será denominado **alfabeto**. Denotaremos por $q = |A|$ o número de elementos de A. Dessa forma, dizemos que A é um código **q-ário**. Além disso, um código corretor de erros é formado por **palavras**, que são sequências finitas de símbolos do alfabeto. O **comprimento** de cada palavra do código corresponde ao número de símbolos dessa palavra. Um código corretor de erros é um subconjunto próprio qualquer de A^n, para algum número natural n.

No Exemplo 3.1.1, utilizamos a quantidade de dígitos distintos entre dois comandos para decodificar e corrigir um erro. Nesse caso, o comando é uma palavra do código. Essa quantidade de dígitos diferentes entre palavras pode ser expressa como veremos a seguir.

Definição 3.1.2. Dados dois elementos $x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ de um conjunto finito A^n, chama-se **distância de Hamming** de x a y ao número de coordenadas em que estes elementos diferem entre si; isto é:

$$d(x, y) = |\{i : \ x_i \neq y_i, \ 1 \leq i \leq n\}|.$$

Exemplo 3.1.3. Usando a definição acima, podemos calcular as distâncias de Hamming entre as palavras do código usado no Exemplo 3.1.1. Sendo assim, temos:
$d(00000, 01101) = 3$, $d(01101, 10011) = 4$ e $d(00000, 10011) = 3$.

Observamos que a distância de Hamming satisfaz as seguintes propriedades:
i) Positividade: $d(x, y) \geq 0$, $\forall x, y \in A^n$, valendo a igualdade se, e somente se, $x = y$.
ii) Simetria: $d(x, y) = d(y, x)$, $\forall x, y \in A^n$.
iii) Desigualdade Triangular: $d(x, y) \leq d(x, z) + d(z, y)$.

***Demonstração**:* A terceira propriedade (desigualdade triangular) não decorre diretamente da definição como as duas primeiras. Portanto, é a única que será demonstrada.

De fato, dados $x, y \in A^n$, a contribuição das i-ésimas coordenadas de x e y para $d(x, y)$ é igual a zero se $x_i = y_i$, e igual a um se $x_i \neq y_i$. Assim, dado $z \in A^n$, se a contribuição das i-ésimas coordenadas para $d(x, y)$ for zero, teremos que ela é sempre menor ou igual às contribuições das i-ésimas coordenadas para $d(x, z)+d(z, y)$, que podem ser iguais a 0,1 ou 2. Por outro lado, temos que $x_i \neq y_i$, de modo que não é possível termos $x_i = z_i$ e $y_i = z_i$, o que nos dá que a contribuição das i-ésimas coordenadas para $d(x, z) + d(z, y)$ é sempre maior ou igual a um, que é a contribuição das i-ésimas coordenadas para $d(x, y)$. □

As propriedades citadas acima caracterizam uma **métrica**. Portanto, a distância de Hamming entre elementos de A^n é também chamada de **métrica de Hamming**.

Usaremos a distância de Hamming para definir disco e esfera em A^n.

Definição 3.1.4. Dados um elemento $x \in A^n$ e um inteiro positivo r, chama-se **disco de centro em x e raio r**, ao conjunto

$$D(x, r) = \{y \in A^n;\ d(x, y) \leq r\},$$

e **esfera de centro em x e raio r**, ao conjunto

$$S(x, r) = \{y \in A^n;\ d(x, y) = r\}.$$

Definição 3.1.5. Dado um código $C \in A^n$ chama-se **distância mínima** de C ao número

$$d = \min\{d(x, y);\ x, y \in C \text{ e } x \neq y\}.$$

Dado um código C com distância mínima d, define-se

$$k = \left[\frac{d-1}{2}\right],$$

onde $[t]$ representa a parte inteira de um número real t.

Lema 3.1.6. Seja C um código com distância mínima d. Se c e c' são palavras distintas de C, então

$$D(c, k) \cap D(c', k) = \emptyset.$$

Demonstração: Vamos supor que existe $x \in D(c, k) \cap D(c', k)$. Então, teríamos $d(x, c) \leq k$ e $d(x, c') \leq k$, e, portanto, pela simetria e pela desigualdade triangular,

$$d(c, c') \leq d(c, x) + d(x, c') \leq 2k \leq d - 1,$$

o que é um absurdo, pois $d(c, c') \geq d$. □

Teorema 3.1.7. Seja C um código com distância mínima d. Então C pode corrigir até $k = \left[\frac{d-1}{2}\right]$ erros e detectar até $d - 1$ erros.

Demonstração: Se durante a transmissão de uma palavra c do código C forem cometidos t erros, com $t \leq k$, e por causa disso, for recebida a palavra r, então $d(r, c) = t \leq k$. Pelo Lema 3.1.6, a distância de r a qualquer outra palavra de C é maior do que k, o que determina c univocamente a partir de r.

Por outro lado, dada uma palavra c do código, podemos nela introduzir até $d - 1$ erros sem encontrar outra palavra do código, e assim, será possível detectar o erro. □

Exemplo 3.1.8. O código do Exemplo 3.1.1 possui distância mínima $d = 3$, pois esta é a menor distância de Hamming entre duas palavras distintas do código. Ele pode corrigir até $k = \left[\frac{3-1}{2}\right] = 1$ erro e detectar até $3 - 1 = 2$ erros.

O teorema acima evidencia a importância de se calcular a distância mínima, visto que quanto maior for a distância mínima de um código, maior será a sua capacidade de detecção e correção de erros .

Definição 3.1.9. Seja $C \subset A^n$ um código com distância mínima d e seja $k = [\frac{d-1}{2}]$. O código C será dito **perfeito** se

$$\bigcup_{c \in C} D(c, k) = A^n.$$

Assim, podemos dizer que um código C sobre um alfabeto A possui 3 parâmtetros fundamentais $[n, M, d]$, que são, respectivamente, o comprimento n, o número de elementos M e a distância mínima d.

Muitas vezes queremos saber se dois códigos são equivalentes, para isso daremos a seguir a definição de isometria em A^n.

Definição 3.1.10. Sejam A um alfabeto e n um número natural. Diremos que uma função $F : A^n \to A^n$ é uma **isometria** de A^n se ela preserva distâncias de Hamming. Isto é,

$$d(F(x), F(y)) = d(x, y); \quad \forall x, y \in A^n.$$

Teorema 3.1.11. Toda isometria de A^n é uma bijeção em A^n.

Demonstração: Seja F uma isometria em A^n. Suponhamos que $F(x) = F(y)$, então $d(F(x), F(y)) = 0 = d(x, y)$. Logo $x = y$. Assim, provamos que F é injetora, e como toda aplicação injetora de um conjunto finito nele próprio é sobrejetora, temos que F é uma bijeção. □

Definição 3.1.12. Sejam C e C' dois códigos em A^n, diremos que C' é equivalente a C se existir uma isometria F de A^n tal que $F(C) = C'$.

A seguir, enunciaremos um teorema que indica quando dois códigos são equivalentes.

Teorema 3.1.13. Sejam C e C' dois códigos em A^n. Temos que C e C' são equivalentes se, e somente se, existem uma permutação π de $\{1, \ldots, n\}$ e bijeções $f_1, \ldots, f_n$ de A tais que

$$C' = \{(f_{\pi(1)}(x_{\pi(1)}), \ldots, f_{\pi(n)}(x_{\pi(n)})); \ (x_1, \ldots, x_n) \in C\}.$$

Demonstração: A demonstração deste teorema pode ser vista no apêndice 2 de (HEFEZ; VILLELA, 2017). □

Como consequência do teorema acima, temos que dois códigos C e C', de comprimento n sobre um alfabeto A, são equivalentes se, e somente se, um deles pode ser obtido a partir do outro através de uma sequência das seguintes operações:

(i) Substituição dos símbolos numa dada posição fixa em todas as palavras do código

por meio de uma bijeção de A.

(ii) Permutação das posições dos símbolos em todas as palavras do código, mediante uma permutação fixa de $\{1, 2, \ldots, n\}$.

Exemplo 3.1.14. Consideremos o código $C = \{00000, 01101, 10011\}$ usado no Exemplo 3.1.1. Utilizaremos as operações citadas acima para gerarmos códigos equivalentes a C. Aplicando a operação (i), substituindo os símbolos da segunda coordenada em cada palavra de C da seguinte forma: $0 \to 1$ e $1 \to 0$, obtemos o código $C' = \{01000, 00101, 11011\}$, que é equivalente a C. E, aplicando a operação (ii), permutando os símbolos da primeira coordenada com os da terceira coordenada em cada palavra de C, obtemos o código $C'' = \{00000, 11001, 00111\}$, que é equivalente a C.

Os códigos que apresentaremos na próxima seção são subespaços vetoriais. Portanto, partiremos do princípio que o leitor esteja familiarizado com as principais definições e resultados a respeito de espaços vetoriais.

3.2 Códigos Lineares

Incialmente, faremos uma introdução aos códigos lineares apresentando algumas de suas propriedades básicas.

Consideremos um corpo finito K com q elementos, que denominaremos alfabeto. Seja $n \in \mathbb{N}$, temos que K^n é um espaço vetorial de dimensão n sobre K.

Definição 3.2.1. Um código $C \subset K^n$ será chamado de **código linear** se for um subespaço vetorial de K^n.

Observação 3.2.2. As palavras de um código linear são representadas por vetores (n-uplas) de elementos de K. Tais vetores são também chamados de **palavra código**.

Exemplo 3.2.3. Consideremos o alfabeto $K = \{0, 1\}$, que é o próprio $\mathbb{Z}_2$, e o código linear $C = \{(0000), (1011), (0110), (1101)\}$, que é um subespaço vetorial de $\mathbb{Z}_2^4$. Portanto, temos que: $C \subset \mathbb{Z}_2^4$.

Observação 3.2.4. O código do Exemplo 3.2.3 é **binário**, pois seu alfabeto é o conjunto $\mathbb{Z}_2$, que possui dois elementos. Tal código é um subconjunto de $\mathbb{Z}_2^4$ porque tem comprimento $n = 4$. Isto é, cada palavra código de C possui 4 símbolos.

De acordo com a definição acima, todo código linear é um espaço vetorial de dimensão finita. Seja $C \in K$ um código linear de dimensão k com q elementos. E seja $\{v_1, v_2, \ldots, v_k\}$ uma base de C. Então, C possui $M = |C| = q^k$ elementos e todo elemento de C é escrito de maneira única como

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k,$$

onde os $a_i,\ i = 1, \ldots, k$, são elementos de K. Portanto,

$$\dim C = k = \log_q q^k = \log_q M.$$

Além disso, se denotarmos por 0 o elemento neutro da soma no espaço vetorial, temos que $0 \in C$. Abaixo, definimos o peso de um elemento pertencente a um espaço vetorial.

Definição 3.2.5. Dado $x \in K^n$, chama-se **peso de x** ao número inteiro

$$w(x) = d(x, 0),$$

onde d é a distância de Hamming.

Definição 3.2.6. O peso de um código linear C é o inteiro

$$w(C) = \min\{w(x);\ x \in C \setminus \{0\}\}.$$

Proposição 3.2.7. Seja $C \subset K^n$ um código linear com distância mínima d. Temos que
i) $\forall x, y \in K^n, \quad d(x, y) = w(x - y)$.
ii) $d = w(C)$.

Demonstração: A partir das definições de distância de Hamming e de peso, a prova do item (i) é torna trivial. Para provar o item (ii), vamos considerar um código linear C de distância mínima d. Então, para quaisquer elementos x e y de C com $x \neq y$, temos que $x - y$ também pertence a $C\backslash\{0\}$. Além disso, $d(x, y) = w(x - y)$, pois o número de coordenadas diferentes entre x e y é igual ao número de coordenadas não nulas da diferença entre x e y. Portanto, minimizar o peso do código é equivalente a minimizar a distância entre quaisquer duas palavras dele. □

Cabe a observação de que para conhecermos a distância mínima de um código com M elementos, necessitaríamos (teoricamente) avaliar $\binom{M}{2}$ distâncias mínimas. No entanto, a proposição acima garante que tal número é reduzido para $(M - 1)$ cálculos, pelo fato de o peso do código ser igual a sua distância mínima.

Exemplo 3.2.8. Seja o código linear $C = \{(00000), (01011), (10110), (11101)\} \subset \mathbb{Z}_2^5$ de distância mínima $d = 3$. Os pesos das palavras não nulas de C são: $w(01011) = 3, w(10110) = 3$ e $w(11101) = 4$. De acordo com a Definição 3.2.6, o peso do código linear C é $w(C) = 3$, pois este é o menor valor entre os pesos das palavras não nulas de C. Conhecendo sua distância mínima, poderemos calcular o peso de C de outra forma. Isto é, de acordo com a Proposição 3.2.7, temos que: $w(C) = d = 3$.

Como códigos lineares são subespaços vetoriais de K^n, as isometrias a serem consideradas são lineares.

Definição 3.2.9. Um código linear $C \in K^n$ é **linearmente equivalente** ao código $C' \in K^n$ se existir uma isometria linear $T : K^n \to K^n$ tal que

$$T(C) = C'$$

.

Desta definição, podemos concluir que dois códigos lineares são linearmente equivalentes se, e somente se, cada um deles puder ser obtido a partir do outro mediante uma sequência de operações do tipo:

(i) Multiplicação dos elementos numa dada posição fixa por um escalar não nulo em todas as palavras.

(ii) Permutação das posições de todas as palavras do código, mediante uma permutação fixa de $\{1, 2, \ldots, n\}$.

Definição 3.2.10. Seja $C \in K^n$ um código linear. Chamaremos de **parâmetros** do código linear C à terna (n, k, d), onde n é o comprimento de C, k é a dimensão de C sobre K e d é a distância mínima de C.

Observação 3.2.11. Em algumas situações utilizaremos a notação código linear (n, k), para nos referirmos a um código linear de comprimento n, distância mínima d e dimensão k de C sobre K.

Exemplo 3.2.12. O código linear dado no Exemplo 3.2.8 possui $M = 4$ elementos e é binário, pois $q = 2$. Então, temos que $k = \dim C = \log_2 4 = 2$. Logo, podemos denotar tal código C por código linear $(5, 2)$.

A seguir, mostraremos um tipo de matriz que é utilizada para gerar um código linear.

Definição 3.2.13. Sejam $C \in K^n$ um código linear e $\mathcal{B} = \{v_1, \ldots, v_k\}$ uma base ordenada de C. A matriz G, cujas linhas são os vetores $v_i = (v_{i1}, \ldots, v_{in})$, $i = 1, \ldots, k$ é denominada **matriz geradora** do código C associada à base $\mathcal{B}$. Isto é,

$$G = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ \vdots & \vdots & & \vdots \\ v_{k1} & v_{k2} & \cdots & v_{kn} \end{bmatrix}.$$

A matriz G não é a única matriz geradora de C, pois ela depende da escolha da base. Isto é, para cada base diferente de C obtemos uma matriz geradora diferente.

Consideremos uma transformação linear $T : K^k \to K^n$ definida por $T(x) = x \cdot G$, com $x \in K^k$. Se $x = (x_1, \ldots, x_k)$, temos que

$$T(x) = xG = x_1 v_1 + \cdots + x_k v_k,$$

logo $T(K^k) = C$. Sendo assim, temos K^k como o código da fonte e C como o código de canal. A transformação T é considerada uma codificação, que transforma o código da fonte no código de canal.

Sabemos que uma base de um espaço vetorial pode ser obtida a partir de outra por meio de sequências de operações elementares. Analogamente, também obtemos uma matriz geradora a partir de outra matriz geradora, efetuando operações elementares.

Vamos supor que seja dada uma palavra $y \in K^n$ de um código linear C, que foi gerado por uma matriz G. Desejamos decodificar a palavra y. Então, precisamos determinar a palavra $x \in K^n$ da qual y se origina, por meio da transformação T. Logo, temos que resolver o sistema

$$x \cdot G = y$$

.

Definição 3.2.14. Uma matriz geradora G de um código C está na **forma padrão** se

$$G = [Id_k|A],$$

onde Id_k é a matriz identidade de ordem k e A é uma matriz $k \times (n-k)$.

Ao efetuarmos sequências de operações como: permutação de duas colunas ou multiplicação de uma coluna por um escalar não nulo, sobre a matriz geradora G de um código linear C, obtemos uma matriz G' de um código C' linearmente equivalente a C.

Teorema 3.2.15. Dado um código C, existe um código equivalente C' com matriz geradora na forma padrão.

Demonstração: A demonstração deste teorema encontra-se em (HEFEZ; VILLELA, 2017, p. 92). □

Exemplo 3.2.16. Consideremos o código linear (5,2) do Exemplo 3.2.8, que tem dimensão $k = 2$ sobre $\mathbb{Z}_2^5$. Sejam $b = \{10110, 01011\}$ uma base e

$$G = \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{array}\right]$$

uma matriz geradora associada à base b, ambos desse código linear. Observemos que a matriz $G = [\mathrm{Id}_k|A]$ está na forma padrão. Logo, temos que

$$A = \left[\begin{array}{ccc} 1 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right].$$

Seja a transformção linear $T : \mathbb{Z}_2^2 \to \mathbb{Z}_2^5$ definida por $T(x) = x \cdot G$, com $x \in \mathbb{Z}_2^2$. Aplicando

T para $x = (10)$, temos

$$T(10) = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \end{bmatrix},$$

que pertence ao código linear (5,2) acima. Portanto, a palavra $x = (10)$ do código da fonte foi codificada como $T(10) = (10110)$ do código do canal.

Definição 3.2.17. Seja $C \subset K^n$ um código linear, o **complemento ortogonal** de C, denotado por $C^\perp$, é definido como

$$C^\perp = \{v \in K^n;\ \langle u, v\rangle = 0,\ \forall u \in C\}.$$

sendo $\langle u, v\rangle$ o produto interno dos elementos u e v pertencentes a K^n.

A partir das propriedades de produto interno, temos que $C^\perp$ é um subespaço vetorial de K^n. Logo, $C^\perp$ também é um código linear. Além disso, $(C^\perp)^\perp = C$.

Definição 3.2.18. Seja $C \subset K^n$ um código linear. O código linear $C^\perp$ chama-se **código dual** de C.

Proposição 3.2.19. Seja $C \subset K^n$ um código linear, com matriz geradora G. Dado $x \in K^n$, teremos $x \in C^\perp$ se, e somente se,

$$G \cdot x^t = 0.$$

Demonstração: Seja $x \in K^n$. Teremos $x \in C^\perp$ se, e somente se, x for ortogonal a todos os elementos de C se, e somente se, x for ortogonal a todos os elementos de uma base de C, o que é equivalente a dizer que $G \cdot x^t = 0$, pois as linhas de G são uma base de C. □

Proposição 3.2.20. Seja $C \subset K^n$ um código de dimensão k com matriz geradora $G = [Id_k|A]$, na forma padrão. Então
i) $\dim C^\perp = n - k$;
ii) $H = [-A^t|Id_{n-k}]$ é uma matriz geradora de $C^\perp$.

Demonstração: (i) Pela Proposição 3.2.19, $x = (x_1, \ldots, x_n) \in C^\perp \iff G \cdot x^t = 0$. A matriz G está na forma padrão e a matriz A é de ordem $k \times (n-k)$, temos então que:

$$[Id_k|A]_{n\times n} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1} = 0 \iff \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}_{k\times 1} = -A_{k\times(n-k)} \begin{bmatrix} x_{k+1} \\ \vdots \\ x_n \end{bmatrix}_{(n-k)\times 1}.$$

Portanto, $C^\perp$ possui q^{n-k} elementos, que são as possíveis escolhas arbitrárias de $x_{k+1}, \ldots, x_n$. Logo, $C^\perp$ tem dimensão $n - k$.

(ii) Por causa de Id_{n-k}, as linhas de H são linearmente independentes, portanto, geram um subespaço vetorial de dimensão $n-k$. Como as linhas de H são ortogonais às linhas de G, temos que o espaço gerado pelas linhas de H está contido em $C^{\perp}$; e como esses dois subespaços vetoriais têm a mesma dimensão, eles coincidem, provando assim que $H = [-A^t|Id_{n-k}]$ é uma matriz geradora de $C^{\perp}$. □

A matriz H, que é a geradora do código linear $C^{\perp}$, pode ser considerada também uma matriz na forma padrão, quando estiver escrita na forma

$$H = [-A^t|Id_{n-k}]$$

.

Lema 3.2.21. Seja $C \in K^n$ um código linear de dimensão k, com matriz geradora G. Uma matriz H de ordem $(n-k) \times n$, com coeficientes em K^n e com linhas linearmente independentes, é uma matriz geradora de $C^{\perp}$ se, e somente se,

$$G \cdot H^t = 0.$$

Demonstração: ($\Rightarrow$) É imediata.
($\Leftarrow$) Como as linhas de H são linearmente independentes, elas geram um subespaço vetorial de K^n de dimensão $n-k$, que é igual a dimensão de $C^{\perp}$. Se representarmos por $h_1, \ldots, h_{n-k}$ e por $g_1, \ldots, g_k$, respectivamente, as linhas de H e de G, temos que

$$[G \cdot H^t]_{i,j} = \langle g_i, h_j \rangle.$$

Por hipótese, $G \cdot H^t = 0$, logo todos os vetores do subespaço vetorial gerado pelas linhas de H estão em $C^{\perp}$. Porém, como esse subespaço tem a mesma dimensão de $C^{\perp}$, concluimos que são o mesmo subespaço vetorial. □

Proposição 3.2.22. Sejam C um código linear e H uma matriz geradora de $C^{\perp}$. Temos então que

$$v \in C \text{ se, e somente se, } H \cdot v^t = 0.$$

Demonstração: É consequência imediata do Lema 3.2.21 e do fato de que $(C^{\perp})^{\perp} = C$. □

Observação 3.2.23. A matriz geradora H de $C^{\perp}$ é chamada de **matriz teste de paridade** de C. A proposição acima nos permite utilizar essa matriz para verificar se um dado vetor v pertence ou não a um determinado código.

Observação 3.2.24. Sejam $C \in K^n$ um código linear com matriz teste de paridade H e um vetor $v \in K^n$, chamamos o vetor $H \cdot v^t$ de **síndrome** de v .

Exemplo 3.2.25. Consideremos o código linear (5,2) do Exemplo 3.2.8. A matriz teste de paridade deste código é

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Utilizaremos essa matriz H para verificarmos se a palavra $v = (11101)$ pertence ao código mencionado acima. Efetuando o produto $H \cdot v^t$, obtemos

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

que é o vetor nulo. Portanto, $v = (11101)$ pertence ao código linear (5,2) em questão. Além disso, a síndrome de v é o vetor $H \cdot v^t = 0$.

A proposição a seguir nos mostrará que a matriz teste de paridade de um código contém informações sobre o valor do peso do código, ou seja, sua própria distância mínima.

Proposição 3.2.26. Sejam $C \in K^n$ um código linear e H a matriz teste de paridade de C. Temos que

$$w(C) \geq s \iff \text{ quaisquer } s-1 \text{ colunas de H são linearmente independentes.}$$

Demonstração: Sejam $x = (x_1, \ldots, x_n)$ uma palavra não nula de C e $h^1, \ldots, h^n$ as colunas de H. Suponhamos que cada conjunto de $s-1$ colunas de H é linearmente independente. Como $Hx^t = 0$, temos que

$$H \cdot x^t = x_1 h^1 + x_2 h^2 + \ldots + x_n h^n = 0. \tag{17}$$

O peso $w(x)$ é o número de componentes não nulas de x, então se $w(x) \leq s-1$, teríamos por (17) uma combinação nula de t colunas de H, com $1 \leq t \leq s-1$, o que é uma contradição. Logo, $w(x) \geq s$ e, portanto, $w(C) \geq s$.

Por outro lado, suponhamos que $w(C) \geq s$. Suponhamos também, por absurdo, que H tenha $s-1$ colunas linearmente dependentes, digamos $h^{i_1}, h^{i_2}, \ldots, h^{i_{s-1}}$. Logo, existiriam $x_{i_1}, \ldots, x_{i_{s-1}}$, nem todos nulos, tais que

$$x_{i_1} h^{i_1} + \cdots + x_{i_{s-1}} h^{i_{s-1}} = 0.$$

Portanto, $x = (0, \ldots, x_{i_1}, 0, \ldots, x_{i_{s-1}}, 0, \ldots, 0) \in C$ e consequentemente,

$w(x) \leq s-1 < s$, o que seria um absurdo, pois a hipótese inicial é a de que $w(C) \geq s$. □

Teorema 3.2.27. Seja H a matriz teste de paridade de um código C. Temos que o peso de C é igual a s se, e somente se, quaisquer $s-1$ colunas de H são linearmente independentes e existem s colunas de H linearmente dependentes.

Demonstração: Suponhamos que $w(C) = s$, logo da Proposição 3.2.26, todo conjunto de $s-1$ colunas de H é linearmente independente. Por outro lado, existem s colunas de H linearmente dependentes, pois, caso contrário, teríamos $w(C) \geq s+1$.

Reciprocamente, suponhamos que todo conjunto de $s-1$ vetores colunas de H é linearmente independente e existem s colunas linearmente dependentes. Logo, da Proposição 3.2.26, temos que $w(C) \geq s$. Caso $w(C)$ seja maior do que s, novamente da Proposição 3.2.26, todo conjunto com s colunas de H é linearmente independente, gerando uma contradição. □

Corolário 3.2.28. (Cota de Singleton) Os parâmetros (n, k, d) de um código linear satisfazem à desigualdade

$$d \leq n - k + 1.$$

Demonstração: Sejam C um código linear de distância mínima d e H uma matriz teste de paridade de C. Pelo Teorema 3.2.27, temos que $d-1$ é menor ou igual ao posto de H, que é $n-k$. Portanto, temos: $d-1 \leq n-k \Longrightarrow d \leq n-k+1$. □

Caso se verifique a igualdade $d = n-k+1$ em um código, este será chamado de **MDS (Maximum Distance Separable)**. Isto é, um código que tem máxima distância mínima.

O processo de decodificação consiste na detecção e correção de erros num determinado código. A seguir, introduziremos um método de decodificação que, segundo (HEFEZ; VILLELA, 2017), é um aperfeiçoamento do que foi inventado por D. Slepian na década de 60. Além disso, consideraremos que um vetor é uma palavra do código.

Definição 3.2.29. O **vetor erro $\boldsymbol{v_e}$** é a diferença entre o vetor recebido $\boldsymbol{v_r}$ e o vetor transmitido $\boldsymbol{v_t}$, isto é,

$$v_e = v_r - v_t.$$

Podemos citar duas características importantes em relação ao vetor erro v_e. A primeira é que o peso de v_e corresponde, exatamente, ao número de erros cometidos num vetor durante sua transmissão. A segunda é que v_e tem a mesma síndrome que o vetor recebido v_r. De fato, se considerarmos H a matriz teste de paridade de um código C, como v_t pertence a C, temos que $Hv_t^t = 0$. Portanto:

$$Hv_e^t = H(v_r^t - v_t^t) = Hv_r^t - Hv_t^t = Hv_r^t.$$

De outra forma, se denotarmos por h^i a i-ésima coluna de H e considerarmos $v_e = (\alpha_1, \ldots, \alpha_n)$, teremos então que

$$\sum_{i=1}^{n} \alpha_i h^i = H v_e^t = H v_r^t.$$

Lema 3.2.30. Seja $C \in K^n$ um código linear com capacidade de correção k. Se $v_r \in K^n$ e $v_t \in C$ são tais que $d(v_t, v_r) \leq k$, então existe um único vetor v_e com $\omega(v_e) \leq k$, cuja síndrome é igual à de v_r e

$$v_t = v_r - v_e$$

.

Demonstração: Inicialmente, para provarmos a existência do vetor v_e, basta considerarmos $v_e = v_r - v_t$, já que $w(v_e) = d(v_t, v_r) \leq k$. Para provarmos a unicidade, vamos supor dois vetores $v_{e_1} = (x_1 \ldots x_n)$ e $v_{e_2} = (y_1 \ldots y_n)$ tais que $w(v_{e_1}) \leq k$ e $w(v_{e_2}) \leq k$ e tenham mesma síndrome que v_r. Então, se H é uma matriz teste de paridade de C, temos

$$H \cdot v_{e_1}^t = H \cdot v_{e_2}^t \Longrightarrow \sum_{i=1}^{n} x_i h^i = \sum_{i=1}^{n} y_i h^i,$$

A igualdade acima nos fornece uma relação de dependência linear entre $2k\,(\leq d-1)$ colunas de H. De acordo com o Teorema 3.2.15, quaisquer $d-1$ colunas de H são linearmente independentes. Então, temos que $x_i = y_i$ para todo i, portanto $v_{e_1} = v_{e_2}$. □

Consideradas as hipóteses do Lema 3.2.30, o problema que surge, então, é como determinar esse único vetor erro v_e a partir de $H \cdot v_r^t$. Vejamos como v_e poderá ser determinado a seguir:

Vamos supor um código C com distância mínima $d \geq 3$, matriz teste de paridade H e que o vetor erro v_e, introduzido entre o vetor transmitido v_t e o vetor recebido v_r, seja tal que $w(v_e) \leq 1$. (No máximo, apenas um erro foi cometido durante a transmissão).

Observamos que se $H \cdot v_e^t = 0$, então $v_r \in C$ e, neste caso, tomamos $v_t = v_r$. Isto é, nenhum erro foi introduzido durante a transmissão. Por outro lado, se $H \cdot v_e^t \neq 0$; $w(v_e) = 1$ e, portanto, v_e tem apenas uma coordenada não nula. Vamos considerar que $v_e = (0, \ldots, \alpha, \ldots, 0)$ com $\alpha \neq 0$ na i-ésima posição. Logo,

$$H \cdot v_e^t = \alpha h^i,$$

onde h^i é a i-ésima coluna de H. Como v_r e v_e têm a mesma síndrome, então,

$$H \cdot v_e^t = H \cdot v_r^t = \alpha h^i.$$

Podemos determinar v_e como sendo o vetor com todas as componentes nulas exceto a i-ésima componente que é α. Note que i acima é bem determinado, pois $d \geq 3$.

A seguir apresentamos o algoritmo de decodificação em códigos corretores de um erro, que foi extraído de (HEFEZ; VILLELA, 2017).

Algoritmo 1 - Algoritmo de Decodificação em Códigos Corretores de Um Erro

Seja H a matriz teste de paridade do código C e seja v_r um vetor recebido. (Suponha $d \geq 3$.) Passo 1: Calcule $H \cdot v_r^t$. Passo 2: Se $H \cdot v_r^t = 0$, aceite v_r como sendo o vetor transmitido. Passo 3: Se $H \cdot v_r^t = s^t \neq 0$, compare s^t com as colunas de H. Passo 4: Se existirem i e α tais que $s^t = \alpha h^i$, para $\alpha \in K$, então v_e é a n-upla com α na posição i e zeros nas outras posições. Corrija v_r colocando $v_t = v_r - v_e$. Passo 5: Se ocorrer o contrário do Passo 4, então foi cometido mais de um erro.

Exemplo 3.2.31. Consideremos o código linear (5,2) do Exemplo 3.2.8 com matriz teste de paridade

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Suponhamos que este código seja utilizado por um monitor de TV, que possui 4 canais. Cada um desses canais é selecionado através de um comando enviado por um controle remoto para o monitor. E cada comando é um elemento do código. Suponhamos também que o controle remoto enviou um comando para o monitor e, por causa de alguma interferência, o comando recebido foi $v_r = (01010)$. Utilizando a matriz teste de paridade

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix},$$

verificamos que $v_r = (01010)$ não pertence ao código, pois o resultado do produto acima não foi o vetor nulo. Vemos que a síndrome de v_r corresponde a quinta coluna de H, ou seja,

$$H \cdot v_r^t = 1 \cdot h^5.$$

Então, pelo algoritmo acima, temos que $v_e = (00001)$. Logo, poderemos corrigir v_t fazendo

$$v_t = v_r - v_e = (01010) - (00001) = (01011).$$

Portanto, o comando enviado pelo controle remoto ao monitor foi $v_t = (01011)$.

Em algumas situações podem ser introduzidos mais de um erro durante a transmissão de um vetor. Para corrigir um vetor v_r, que foi recebido com mais de um erro, iremos apresentar algumas definições e outro algoritmo de decodificação mais geral.

Seja $C \subset K^n$ um código corretor de erros com matriz teste de paridade H. Sejam d a distância mínima de C e $k = [\frac{d-1}{2}]$. Como v_e e v_r têm a mesma síndrome e, se $w(v_e) = d(v_r, v_t) \leq k$, então v_e é univocamente determinado por v_r.

Definição 3.2.32. Seja $v \in K^n$. Chama-se **classe lateral de v determinada por C** ao conjunto

$$v + C = \{v + u;\ u \in C\}.$$

Note que

$$v + C = C \iff v \in C.$$

Lema 3.2.33. Os vetores u e v de K^n têm a mesma síndrome se, e somente se, $u \in v + C$.

Demonstração: $H \cdot u^t = H \cdot v^t \iff H \cdot (u - v)^t = 0 \iff u - v \in C \iff u \in v + C$. □

A seguir, apenas enunciaremos as propriedades das classes laterais determinadas por C.

Proposição 3.2.34. Seja $C \in K^n$ um código linear de dimensão k com q elementos. Temos que
i) $v + C = v' + C \Leftrightarrow v - v' \in C$;
ii) $(v + C) \cap (v' + C) \neq \emptyset \Longrightarrow v + C = v' + C$;
iii) $\bigcup_{v \in K^n} (v + C) = K^n$;
iv) $|(v + C)| = |C| = q^k$.

Segue imediatamente de (ii)-(iv) da Proposição 3.2.34 que o número de classes laterais segundo C é

$$\frac{q^n}{q^k} = q^{n-k}.$$

Notemos que o Lema 3.2.33 estabelece uma correspondência 1 a 1 entre classes laterais e síndromes. Todos os elementos de uma classe lateral determinada por um código têm a mesma síndrome, e elementos de classes distintas possuem síndromes distintas.

Definição 3.2.35. Um vetor de peso mínimo numa classe lateral é chamado de **elemento líder** dessa classe.

Proposição 3.2.36. Seja C um código linear em K^n com distância mínima d. Se $u \in K^n$ é tal que

$$w(u) \leq \left[\frac{d-1}{2}\right] = k,$$

então u é o único elemento líder de sua classe.

Demonstração: Suponhamos que $u, v \in K^n$ tais que

$$w(u) \leq [\frac{d-1}{2}] \quad \text{e} \quad w(v) \leq [\frac{d-1}{2}].$$

Se $u - v \in C$, então

$$w(u-v) \leq w(u) + w(v) \leq \left[\frac{d-1}{2}\right] + \left[\frac{d-1}{2}\right] \leq d-1;$$

logo, $u - v = 0$ e, portanto, $u = v$. □

A seguir, apresentaremos um algoritmo que permite corrigir vetores que tenham sido recebidos com até k erros. Devemos lembrar que $k = [\frac{d-1}{2}]$ é a capacidade de correção do código.

Inicialmente, precisamos determinar os líderes de classes laterais, sendo cada um deles o único líder em sua classe. Faremos isso encontrando todos os elementos u de K^n, tal que $w(u) \leq k$.

Em seguida, devemos calcular as síndromes desses elementos e inserir esses vetores com suas respectivas síndromes numa tabela. Após esses passos, é só seguir o algoritmo abaixo, que foi extraído de (HEFEZ; VILLELA, 2017):

Algoritmo 2 - Algoritmo de Decodificação em Códigos Corretores de k Erros

	Seja v_r um vetor recebido.
Passo 1:	Calcule a síndrome $s^t = H \cdot v_r^t$.
Passo 2:	Se s está na tabela, seja l o elemento líder da classe determinada por s; troque v_r por $v_r - l$.
Passo 3:	Se s não está na tabela, então foram cometidos mais do que k erros no vetor recebido.

Se v_r, v_t e v_e forem respectivamente os vetores recebido, transmitido e erro, como $H \cdot v_e^t = H \cdot v_r^t$, temos que a classe lateral onde v_e se encontra está determinada pela síndrome de v_r. Se $w(v_e) \leq k$, então v_e é o único elemento líder l de sua classe e, portanto, é conhecido e se encontra na tabela. Consequentemente, pelo Lema 3.2.30, $v_t = v_r - v_e = v_r - l$ é determinado.

3.3 Códigos Cíclicos Binários

Os códigos cíclicos formam uma subclasse importante dos códigos lineares e possuem propriedades algébricas que simplificam a sua implementação. Nesta seção, aborda-

remos tais propriedades dos códigos cíclicos binários, em particular trataremos os códigos BCH binários, que também são cíclicos.

Definição 3.3.1. Sejam o corpo $\mathbb{Z}_2$, um código linear $C \subset \mathbb{Z}_2^n$ e $v = (v_0, v_1, \ldots, v_{n-1})$ um vetor de C. Se cada uma das $n-1$ primeiras componentes de v forem deslocadas uma posição para a direita e a última componente v_{n-1} for deslocada para a primeira posição à esquerda, obteremos o vetor

$$v_{d_1} = (v_{n-1}, v_0, v_1, \ldots, v_{n-2}),$$

que é chamado de **vetor deslocamento cíclico** de v. Neste caso, o deslocamento foi de uma posição. Se o deslocamento cíclico for de $i \in \mathbb{N}$ $(1 \leq i \leq n-1)$ posições para a direita no vetor v, o vetor deslocamento cíclico obtido será o seguinte

$$v_{d_i} = (v_{n-i}, \ldots, v_0, \ldots, v_{n-1-i}).$$

Observação 3.3.2. O deslocamento cíclico de $i = n$ posições para a direita gera um vetor deslocamento cíclico igual ao vetor original. E o deslocamento cíclico de uma posição para a esquerda é o mesmo que um deslocamento cíclico de $i = n-1$ posições para a direita.

Definição 3.3.3. Um código linear $C \subset \mathbb{Z}_2^n$ será chamado de **código cíclico binário *C*** se, para todo vetor $v = (v_0, \ldots, v_{n-1})$ de C, então o vetor deslocamento $v_{d_i} = (v_{n-i}, \ldots, v_{n-1-i})$ também pertence a C, $\forall i = 1, \ldots, n-1$.

Observação 3.3.4. Em algumas situações utilizaremos a notação $C(n,k)$, para nos referirmos a um código cíclico binário de comprimento n, distância mínima d e dimensão k de C sobre $\mathbb{Z}_2$.

Sabemos que um polinômio $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ pode ser representado como uma n-upla $p = (a_0, a_1, \ldots, a_{n-1})$, onde as componentes da n-upla p são os coeficientes de $p(x)$ e vice-versa. Portanto, cada vetor $v = (a_0, a_1, \ldots, a_{n-1})$ de um código cíclico C está associado a um polinômio $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ tal que $gr(p(x)) \leq n-1$. A representação vetorial de um polinômio é útil no desenvolvimento de propriedades algébricas dos códigos cíclicos binários, e será usada em algumas situações daqui em diante neste trabalho.

Observação 3.3.5. Sejam $v = (v_0, v_1, \ldots, v_{n-1})$ um vetor de um código cíclico C e

$$v(x) = v_0 + v_1x + \cdots + v_{n-1}x^{n-1} \tag{18}$$

o polinômio associado a v. Como C é cíclico, o vetor $v_{d_i} = (v_{n-i}, v_{n-i+1}, \ldots, v_{n-i-1})$, com $0 \leq i \leq n-1$, também pertence a C.

Podemos escrever a representação polinomial de v_{d_i} da seguinte forma:

$$v_{d_i}(x) = v_{n-i} + v_{n-i+1}x + \ldots + v_{n-1}x^{i-1} + v_0x^i + v_1x^{i+1} + \ldots + v_{n-i-1}x^{n-1}.$$

Existe uma relação entre $v(x)$ e $v_{d_i}(x)$ que veremos a seguir.
Multiplicaremos $v(x)$, dado em (18), por x^i

$$x^iv(x) = v_0x^i + v_1x^{i+1} + \ldots + v_{n-i-1}x^{n-1} + \ldots + v_{n-1}x^{n+i-1}.$$

A expressão acima pode ser reescrita como:

$$\begin{aligned} x^iv(x) &= v_{n-i} + v_{n-i+1}x + \ldots + v_{n-1}x^{i-1} + v_0x^i + \ldots + v_{n-i-1}x^{n-1} + \\ &\quad + v_{n-i}(x^n+1) + v_{n-i+1}x(x^n+1) + \ldots + v_{n-1}x^{i-1}(x^n+1). \end{aligned}$$

E obtemos

$$x^iv(x) = [v_{n-i} + v_{n-i+1}x + \cdots + v_{n-1}x^{i-1}](x^n+1) + v_{d_i}(x) \tag{19}$$

De (19), temos que $v_{d_i}(x)$ é o resto da divisão de $x^iv(x)$ por (x^n+1).

Exemplo 3.3.6. Consideremos o polinômio $v(x) = v_0 + v_1x + v_2x^2 + v_3x^3 + v_4x^4 + v_5x^5$. Vamos efetuar um deslocamento cíclico de três posições para a direita em $v(x)$. Então, multiplicando $v(x)$ por x^3, obtemos

$$\begin{aligned} x^3v(x) &= v_0x^3 + v_1x^4 + v_2x^5 + v_3x^6 + v_4x^7 + v_5x^8 \\ &= (v_5x^2 + v_4x + v_3)(x^6+1) + v_{d_3}(x). \end{aligned}$$

De (19), temos que $v_{d_3}(x)$ é o resto da divisão de $x^3v(x)$ por (x^6+1). Logo,

$$v_{d_3}(x) = v_2x^5 + v_1x^4 + v_0x^3 + v_5x^2 + v_4x + v_3.$$

Observação 3.3.7. Ao longo deste texto, usaremos o vetor $v = (v_0, v_1, \ldots, v_{n-1})$ e o polinômio $v(x) = v_0 + v_1x + \cdots + v_{n-1}x^{n-1}$ associado a ele indistintamente como elementos de um código cíclico $C(n,k)$. O polinômio $v(x)$ poderá ser chamado de **polinômio código**.

A seguir, mostraremos algumas propriedades dos códigos cíclicos binários que servirão de base para a apresentação dos códigos BCH binários.

Teorema 3.3.8. O polinômio não nulo de grau mínimo de um código cíclico binário $C(n,k)$ é único.

Demonstração: Seja $g(x) = g_0 + g_1x + \cdots + g_{r-1}x^{r-1} + x^r$ um polinômio não nulo de grau mínimo r, de um código cíclico binário $C(n,k)$. Suponhamos que $g(x)$ não seja único

e que exista outro polinômio de grau r,

$$h(x) = h_0 + h_1x + \cdots + h_{r-1}x^{r-1} + x^r.$$

Como $C(n,k)$ é linear, a soma de grau $r-1$,

$$g(x) + h(x) = (g_0 + h_0) + (g_1 + h_1)x + \cdots + (g_{r-1} + h_{r-1})x^{r-1}$$

pertence a $C(n,k)$. Se $g(x) + h(x) \neq 0$, então $g(x) + h(x)$ é um polinômio não nulo de grau menor do que r. Isto é um absurdo, pois r é o grau mínimo. Então, devemos ter

$$g(x) + h(x) = 0.$$

Logo, $g(x) = h(x)$. Portanto, $g(x)$ é único. □

Teorema 3.3.9. Seja $g(x) = g_0 + g_1x + \cdots + g_{r-1}x^{r-1} + x^r$ um polinômio não nulo de grau mínimo r de um código cíclico binário $C(n,k)$. Então, o termo independente g_0 de $g(x)$ é igual a 1.

Demonstração: Suponhamos que $g_0 = 0$. Então

$$g(x) = g_1x + g_2x^2 + \cdots + g_{r-1}x^{r-1} + x^r.$$

Efetuando um deslocamento cíclico de $n-1$ posições para a direita, obtemos

$$g(x) = g_1 + g_2x + \cdots + g_{r-1}x^{r-2} + x^{r-2} + x^{r-1},$$

que tem grau menor que r, que é o grau mínimo. Isto é um absurdo. Pela hipótese, $g(x)$ é o polinômio não nulo de grau mínimo r. Logo, $g_0 = 1$. □

Teorema 3.3.10. Seja $g(x) = 1 + g_1x + \cdots + g_{r-1}x^{r-1} + x^r$ um polinômio não nulo de grau mínimo r de um código cíclico binário $C(n,k)$. Um polinômio $p(x)$ de grau $s \leq n-1$ pertence a $C(n,k)$ se, e somente se, $p(x)$ for múltiplo de $g(x)$.

Demonstração: ($\Rightarrow$) Seja $p(x)$ um polinômio de grau $s \leq n-1$. Suponhamos que $p(x)$ seja múltiplo de $g(x)$. Então

$$p(x) = (p_0 + p_1x + \cdots + p_sx^s) \cdot g(x) = p_0g(x) + p_1xg(x) + \cdots + p_sx^sg(x).$$

De acordo com a Observação 3.3.5, os polinômios $g(x), xg(x), \ldots, x_sg(x)$ são deslocamentos cíclicos de $g(x)$ e, portanto, pertencem a $C(n,k)$. Se $p(x)$ é uma combinação linear dos polinômios $g(x), xg(x), \ldots, x_sg(x)$ e como $C(n,k)$ é linear, então $p(x)$ pertence a $C(n,k)$.

($\Leftarrow$) Seja $p(x)$ um polinômio de um código cíclico binário $C(n,k)$. Dividindo $p(x)$ por $g(x)$, obtemos

$$p(x) = g(x) \cdot q(x) + r(x),$$

onde $r(x) = 0$ ou $gr(r(x)) < gr(g(x))$. A equação acima pode ser escrita da seguinte forma

$$r(x) = p(x) + g(x) \cdot q(x).$$

Segue da prova da primeira parte do teorema, que $g(x)q(x)$ pertence a $C(n,k)$. Como $p(x)$ também pertence a $C(n,k)$, então $r(x)$ deve pertencer a $C(n,k)$. Se $r(x) \neq 0$, então $r(x)$ é um polinômio não nulo de grau menor do que o grau de $g(x)$. Isto é um absurdo. Pela hipótese do teorema, $g(x)$ é um polinômio não nulo de grau mínimo. Então, devemos ter $r(x) = 0$. Logo,

$$p(x) + g(x)q(x) = 0 \quad \text{e} \quad p(x) = g(x)q(x).$$

Ou seja, $p(x)$ é múltiplo de $g(x)$. □

Observação 3.3.11. De acordo com o Teorema 3.3.8, $g(x)$ é único. Como $C(n,k)$ é binário e possui dimensão k, existem 2^k polinômios nesse código. Por outro lado, existem 2^{n-r} polinômios de grau $s \leq n-1$, que são múltiplos de $g(x)$ de grau mínimo r. Esses polinômios formam todos os polinômios de um código cíclico binário $C(n,k)$, conforme o Teorema 3.3.10. Dessa forma, devemos ter: $2^{n-r} = 2^k$. E, consequentemente, $r = n - k$. Isto é,

$$gr(g(x)) = r = n - k.$$

Teorema 3.3.12. Seja $g(x) = 1 + g_1x + g_2x^2 + \cdots + g_{r-1}x^{r-1} + g_rx^r$ um polinômio não nulo de grau mínimo r de um código cíclico binário $C(n,k)$. Existe um e somente um polinômio de grau $n-k$ em $C(n,k)$ e $r = n - k$. Então,

$$g(x) = 1 + g_1x + g_2x^2 + \cdots + g_{n-k-1}x^{n-k-1} + g_{n-k}x^{n-k}. \tag{20}$$

Definição 3.3.13. Seja $g(x) = 1+g_1x+g_2x^2+\cdots+g_{n-k-1}x^{n-k-1}+g_{n-k}x^{n-k}$ um polinômio não nulo de grau mínimo $n-k$ de um código cíclico binário $C(n,k)$. O polinômio $g(x)$ é chamado de **polinômio gerador** de $C(n,k)$.

Observação 3.3.14. Seja $g(x)$ um polinômio não nulo de grau mínimo $n-k$ de um código cíclico binário $C(n,k)$. O número de dígitos de verificação de paridade de $C(n,k)$ é igual a $n-k$.

Exemplo 3.3.15. O código cíclico binário $C(7,4)$, que está representado na tabela 6, é gerado pelo polinômio $g(x) = 1 + x + x^3$. Notemos que neste código, $n = 7$ e $k = 4$. Além disso, $gr(g(x)) = 7 - 4 = 3$ e todos os polinômios pertencentes a C são múltiplos de $g(x)$.

Tabela 6 - Código cíclico binário $C(7,4)$ gerado por $g(x) = 1 + x + x^3$

Mensagem	Palavra código	Polinômios
(0,0,0,0)	(0,0,0,0,0,0,0)	$0 = 0 \cdot g(x)$
(1,0,0,0)	(1,1,0,1,0,0,0)	$1 + x + x^3 = g(x)$
(0,1,0,0)	(0,1,1,0,1,0,0)	$x + x^2 + x^4 = xg(x)$
(1,1,0,0)	(1,0,1,1,1,0,0)	$1 + x^2 + x^3 + x^4 = (1 + x)g(x)$
(0,0,1,0)	(1,1,1,0,0,1,0)	$1 + x + x^2 + x^5 = (1 + x^2)g(x)$
(1,0,1,0)	(0,0,1,1,0,1,0)	$x^2 + x^3 + x^5 = x^2g(x)$
(0,1,1,0)	(1,0,0,0,1,1,0)	$1 + x^4 + x^5 = (1 + x + x^2)g(x)$
(1,1,1,0)	(0,1,0,1,1,1,0)	$x + x^3 + x^4 + x^5 = (x + x^2)g(x)$
(0,0,0,1)	(1,0,1,0,0,0,1)	$1 + x^2 + x^6 = (1 + x + x^3)g(x)$
(1,0,0,1)	(0,1,1,1,0,0,1)	$x + x^2 + x^3 + x^6 = (x + x^3)g(x)$
(0,1,0,1)	(1,1,0,0,1,0,1)	$1 + x + x^4 + x^6 = (1 + x^3)g(x)$
(1,1,0,1)	(0,0,0,1,1,0,1)	$x^3 + x^4 + x^6 = x^3g(x)$
(0,0,1,1)	(0,1,0,0,0,1,1)	$x + x^5 + x^6 = (x + x^2 + x^3)g(x)$
(1,0,1,1)	(1,0,0,1,0,1,1)	$1 + x^3 + x^5 + x^6 = (1 + x + x^2 + x^3)g(x)$
(0,1,1,1)	(0,0,1,0,1,1,1)	$x^2 + x^4 + x^5 + x^6 = (x^2 + x^3)g(x)$
(1,1,1,1)	(1,1,1,1,1,1,1)	$1 + x + x^2 + x^3 + x^4 + x^5 + x^6 = (1 + x^2 + x^5)g(x)$

Fonte: (LIN; DANIEL, 1983)

Teorema 3.3.16. O polinômio gerador $g(x)$ de um código cíclico binário $C(n,k)$ é um fator de $x^n + 1$.

Demonstração: Multiplicando $g(x)$ por x^k obtemos o polinômio $x^k g(x)$, que tem grau n. Dividindo $x^k g(x)$ por $x^n + 1$ obtemos

$$x^k g(x) = (x^n + 1) + g_{d_k}(x), \tag{21}$$

onde $g_{d_k}(x)$ é o resto da divisão. De acordo com a Definição 3.3.3, $g_{d_k}(x)$ representa o vetor deslocamento cíclico de k posições para direita do vetor $g(x)$. Portanto, $g_{d_k}(x)$ é um múltiplo de $g(x)$. Isto é, $g_{d_k}(x) = g(x)p(x)$. Substituindo essa igualdade em (21), obtemos

$$x^k g(x) = (x^n + 1) + g(x)p(x).$$

Reorganizando essa igualdade, temos

$$x^n + 1 = [x^k + p(x)]g(x).$$

Portanto, $g(x)$ é um fator de $x^n + 1$. □

Teorema 3.3.17. Se $g(x)$ é um fator de $x^n + 1$ e $gr(g(x)) = n - k$, então $g(x)$ gera um código cíclico binário $C(n,k)$.

Demonstração: Sejam os k polinômios $g(x), xg(x), \ldots, x^{k-1}g(x)$, todos com grau menor ou igual a $n-1$ e $g(x)$ um polinômio de grau $n-k$. Seja $v(x)$ uma combinação linear desses k polinômios,

$$v(x) = a_0 g(x) + a_1 x g(x) + \cdots + a_{k-1} x^{k-1} g(x) = (a_0 + a_1 x + \cdots + a_{k-1} x^{k-1}) g(x).$$

O polinômio $v(x)$ é um polinômio de grau menor ou igual a $n-1$ e é múltiplo de $g(x)$. Existe um total de 2^k polinômios que formam um código cíclico binário $C(n,k)$.

Seja $v(x) = v_0 + v_1 x + \cdots + v_{n-1} x^{n-1}$ um polinômio de $C(n,k)$. Multiplicando $v(x)$ por x, obtemos

$$\begin{aligned} xv(x) &= v_0 x + v_1 x^2 + \cdots + v_{n-2} x^{n-1} + v_{n-1} x^n \\ &= v_{n-1}(x^n + 1) + (v_{n-1} + v_0 x + \cdots + v_{n-2} x^{n-1}) \\ &= v_{n-1}(x^n + 1) + v_{d_1}(x), \end{aligned}$$

onde $v_{d_1}(x)$ é um deslocamento cíclico de $v(x)$. Se os polinômios $xv(x)$ e $x^n + 1$ são múltiplos de $g(x)$, então $v_{d_1}(x)$ também deve ser múltiplo de $g(x)$. Logo, $v_{d_1}(x)$ é um múltiplo de $g(x)$ e é uma combinação linear dos k polinômios $g(x), xg(x), \ldots, x^{k-1}g(x)$. Portanto, $v_{d_1}(x)$ também pertence a $C(n,k)$. De acordo com a Definição 3.3.3, o código linear gerado pelos polinômios $g(x), xg(x), \ldots, x^{k-1}g(x)$ é um código cíclico binário $C(n,k)$. □

Exemplo 3.3.18. O polinômio $x^7 + 1$ pode ser fatorado da seguinte forma:

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

Na fatoração acima há dois fatores de grau 3: $1 + x^2 + x^3$ e $1 + x + x^3$. Cada um deles gera um código cíclico binário $C(7,4)$. O polinômio $g(x) = 1 + x + x^3$ gera o código cíclico binário $C(7,4)$ do Exemplo 3.3.15, que está representado na tabela 6 (página 58). Este código tem distância mínima $d = 3$ e é um código corretor de 1 erro.

Dada uma mensagem $u(x)$ que foi transmitida por um canal de comunicação. A codificação de $u(x)$ é feita multiplicando-se $u(x)$ por $g(x)$, gerando o vetor $v(x) = v_t(x)g(x)$, pertencente ao código cíclico binário $C(n,k)$.

Suponhamos que a mensagem $u = (u_0, u_1, \ldots, u_{k-1})$ foi transmitida e precisa ser codificada. O polinômio correspondente é

$$u(x) = u_0 + u_1 x + \cdots + u_{k-1} x^{k-1}.$$

multiplicando $u(x)$ por x^{n-k}, obtemos um polinômio de grau menor ou igual a $n-1$,

$$x^{n-k} u(x) = u_0 x^{n-k} + u_1 x^{n-k+1} + \cdots + u_{k-1} x^{n-1}.$$

Dividindo $x^{n-k}u(x)$ pelo polinômio gerador $g(x)$, temos

$$x^{n-k}u(x) = g(x)q(x) + r(x) \tag{22}$$

onde $q(x)$ e $r(x)$ são respectivamente o quociente e o resto da divisão. Como o grau de $g(x)$ é $n-k$, então o grau de $r(x)$ tem que ser menor ou igual a $n-k-1$. Isto é,

$$r(x) = r_0 + r_1x + \cdots + r_{n-k-1}x^{n-k-1}.$$

Reorganizando a igualdade (22), obtemos o seguinte polinômio de grau menor ou igual a $n-k$:

$$r(x) + x^{n-k}u(x) = g(x)q(x). \tag{23}$$

O polinômio $r(x) + x^{n-k}u(x)$ é múltiplo do polinômio gerador $g(x)$ e, portanto, pertence ao código cíclico binário gerado por $g(x)$. Escrevendo $r(x) + x^{n-k}u(x)$ em função de seus coeficientes, temos

$$r(x) + x^{n-k}u(x) = r_0 + r_1x + \cdots + r_{n-k-1}x^{n-k-1} + u_0x^{n-k} + \cdots + u_{k-1}x^{n-1}, \tag{24}$$

cujo vetor correspondente é

$$(r_0, r_1, \ldots, r_{n-k-1}, u_0, \ldots, u_{k-1}).$$

Notemos que o vetor acima consiste de k dígitos de informação $(r_0, r_1, \ldots, r_{n-k-1})$, seguidos de $n-k$ dígitos de verificação de paridade $(u_0, \ldots, u_{k-1})$. O processo descrito acima gera um código cíclico binário $C(n,k)$ na forma sistemática. Dessa maneira, considerando o polinômio $r(x) + x^{n-k}u(x)$, os coeficientes de $1, x, \ldots, x^{n-k-1}$ são os $n-k$ dígitos de verificação de paridade, e os coeficientes de $x^{n-k}, x\ n-k+1, \ldots, x^{n-1}$ são os k dígitos da informação, que correspondem a mensagem transmitida.

Exemplo 3.3.19. Consideremos o código cíclico binário $C(7,4)$ gerado por $g(x) = 1 + x + x^3$. Seja $u = (1,0,0,1)$ a mensagem transmitida que será codificada. O polinômio correspondente de u é $u(x) = 1 + x^3$. Dividindo $x^3u(x) = x^3 + x^6$ por $g(x)$, obtemos

$$x^3 + x^6 = (1 + x + x^3)(x + x^3) + (x + x^2),$$

onde $r(x) = x + x^2$ é o resto da divisão. Então, o polinômio resultante da codificação é

$$v(x) = r(x) + x^3u(x) = x + x^2 + x^3 + x^6,$$

cujo vetor correspondente é

$$v = (0, 1, 1, 1, 0, 0, 1),$$

onde os quatro ($k = 4$) últimos dígitos da direita são os dígitos da mensagem $u = (1, 0, 0, 1)$. E os três ($n-k = 3$) primeiros dígitos da esquerda são os dígitos de verificação de paridade.

Na próxima seção, veremos uma classe de códigos corretores de erros chamados códigos BCH binários. Tais códigos foram primeiramente descobertos por A. Hocquenghem em 1959 e, de forma independente, por R.C. Bose e D.K. Ray-Chaudhuri em 1960.

Embora os códigos BCH sejam cíclicos, não nos aprofundaremos a respeito disso, pois nossa intenção é exemplificar a utilização de polinômios vinculando-os aos códigos corretores de erros. Muitas das definições e resultados que apresentaremos encontram-se mais detalhados em (LIN; DANIEL, 1983).

3.4 Códigos BCH Binários

Inicialmente, apresentaremos algumas propriedades básicas dos códigos BCH binários.

Vimos na Definição 3.2.10 que os parâmetros de um código linear C são: a dimensão k de C sobre o corpo finito K, a distância mínima d de C e o comprimento n de C.

Definição 3.4.1. Chamamos de **código BCH binário com capacidade de correção de t erros** ao código cíclico binário de distância mínima $d \geq 2t + 1$, com $n - k \leq mt$ dígitos de verificação de paridade e comprimento $n = 2^m - 1$, para quaisquer inteiros positivos $m \geq 3$ e $t < 2^m - 1$.

O código BCH binário definido acima é gerado por um polinômio que é especificado em termos de suas raízes no corpo finito $\mathbb{F}_{2^m}$. Isto nos conduz à seguinte definição.

Definição 3.4.2. Sejam o corpo finito $\mathbb{F}_{2^m}$ e α um elemento primitivo de $\mathbb{F}_{2^m}$. Seja $min_{\alpha^i}(x)$ o polinômio mínimo de α^i. O **polinômio gerador $g(x)$** do código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros é o polinômio de menor grau de $\mathbb{Z}_2[x]$, que tem como raízes os elementos

$$\alpha, \alpha^2, \ldots, \alpha^i, \ldots, \alpha^{2t} \tag{25}$$

Então, $g(x)$ deve ser o mínimo múltiplo comum de $min_{\alpha^1}(x), min_{\alpha^2}(x), \ldots, min_{\alpha^{2t}}$, ou seja,

$$g(x) = \text{MMC}(min_{\alpha^1}(x), min_{\alpha^2}(x), \ldots, min_{\alpha^{2t}}(x)). \tag{26}$$

Em (25), se $i \in \mathbb{Z}$ é par, então pode ser expresso como um produto da seguinte

forma: $i = j \cdot 2^l$, onde $j \in \mathbb{Z}$ é ímpar e $l \geq 1 \in \mathbb{Z}$. Logo, $\alpha^i = (\alpha^j)^{2^l}$ é um conjugado de α^j e, portanto, α^i e α^j têm o mesmo polinômio mínimo, ou seja,

$$min_{\alpha^i} = min_{\alpha^j}.$$

Como resultado, cada potência par de α tem o mesmo polinômio mínimo de alguma potência ímpar de α precedente na mesma sequência. Consequentemente, o polinômio gerador $g(x)$ de um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros dado por (26) pode ser reduzido para

$$g(x) = mmc(min_{\alpha^1}(x), min_{\alpha^3}(x), \ldots, min_{\alpha^{2t-1}}(x)). \tag{27}$$

Como o grau de cada polinômio mínimo é menor ou igual a m, então temos que

$$gr(g(x)) = n - k \leq mt.$$

De acordo com (LIN; DANIEL, 1983), não existe uma fórmula simples para a enumeração de $n-k$, mas para valores pequenos de t, $n-k$ é exatamente igual a mt. Na tabela 7 estão indicados os parâmetros de todos os códigos BCH binários de comprimento $n = 2^m - 1$, com $m \leq 5$.

Tabela 7 - Parâmetros dos códigos BCH(n, k) binários para $3 \leq m \leq 5$

Código	Parâmetros			Código	Parâmetros		
	$n = 2^m - 1$	k	t		$n = 2^m - 1$	k	t
BCH(7,4)	7	4	1	BCH(31,26)	31	26	1
BCH(15,11)	15	11	1	BCH(31,21)		21	2
BCH(15,7)		7	2	BCH(31,16)		16	3
BCH(15,5)		5	3	BCH(31,11)		11	5
XXXXX	X	X	X	BCH(31,6)		6	6

Fonte: O autor, com informações extraídas de (LIN; DANIEL, 1983), 2019

Observação 3.4.3. Em algumas situações, utilizaremos a notação BCH(n, k) para nos referirmos a um código BCH de comprimento $n = 2^m - 1$ com capacidade de correção de t erros.

De (27) podemos observar que um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de um erro é gerado pelo polinômio

$$g(x) = min_{\alpha}(x),$$

e

$$min_{\alpha}(x) = min_{\alpha^{2t-1}}(x).$$

Observação 3.4.4. De acordo com a Definição 3.4.2, α é um elemento primitivo de $\mathbb{F}_{2^m}$. Logo, $min_{\alpha}(x)$ é um polinômio primitivo de grau m. Portanto, um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de um erro é um código de Hamming.

Exemplo 3.4.5. Consideremos o corpo finito $\mathbb{F}_{2^3}$ gerado por $p(x) = 1 + x + x^3$ e dado pela tabela 4 (página 31). Seja α um elemento primitivo de $\mathbb{F}_{2^3}$. Da tabela 5 (página 35), obtemos o polinômio mínimo de α que é

$$min_{\alpha}(x) = 1 + x + x^3.$$

De (27) verificamos que o código BCH binário de comprimento $n = 2^3 - 1 = 7$ com capacidade de correção de um erro é gerado por

$$g(x) = mmc(min_{\alpha}(x)),$$

ou seja

$$g(x) = min_{\alpha}(x).$$

Então, temos que

$$g(x) = 1 + x + x^3.$$

Portanto, de acordo com a tabela 7, o código acima é um código BCH(7,4) binário, que tem distância mínima maior ou igual a 3. Como o polinômio gerador $g(x)$ tem peso 3, a distância mínima deste código BCH binário é exatamente 3. Além disso, o código BCH(7,4) binário é o código cíclico C(7,4) binário dado pela tabela 6 (página 58). Isto porque, este código cíclico atende às condições da Definição 3.4.1 e ambos os códigos citados possuem o mesmo polinômio gerador $g(x) = 1 + x + x^3$.

De acordo com as definições 3.4.1 e 3.4.2, o polinômio gerador $g(x)$ de um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros possui como raízes, os elementos $\alpha, \alpha^2, \ldots, \alpha^{2t}$ e seus conjugados. Com isso, podemos verificar se um determinado polinômio pertence a um código BCH binário. Para tanto, consideremos o polinômio $v(x) = v_0 + v_1x + \cdots + v_{n-1}x^{n-1} \in \mathbb{Z}_2[x]$. Se os elementos $\alpha, \alpha^2, \ldots, \alpha^{2t}$ são as raízes de $v(x)$, então, pelo Teorema 2.4.9, $v(x)$ é divisível pelos polinômios mínimos $min_{\alpha}(x), min_{\alpha^2}(x), \ldots, min_{\alpha^{2t}}(x)$. E, consequentemente, $v(x)$ é divisível pelo MMC desses polinômios, que é polinômio gerador $g(x)$ do código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros, dado por (26). Portanto, $v(x)$ pertence ao código BCH binário. Isso nos conduz às seguintes definições.

Definição 3.4.6. Sejam o corpo finito $\mathbb{F}_{2^m}$ e o polinômio $v(x) = v_0+v_1x+\cdots+v_{n-1}x^{n-1} \in \mathbb{Z}_2[x]$. Se o polinômio $v(x)$ pertence a um código BCH binário de comprimento $2^m - 1$ com capacidade de correção de t erros, então $v(x)$ é um **polinômio código**.

Conforme mencionado no capítulo 1, um polinômio $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ pode ser expresso como uma n-upla, da seguinte forma: $p = (a_0, a_1, a_2, \ldots, a_n)$. Esta representação vetorial de um polinômio será utilizada na próxima definição e em outras partes desta seção.

Definição 3.4.7. Sejam o corpo finito $\mathbb{F}_{2^m}$, o polinômio $v(x) = v_0+v_1x+\cdots+v_{n-1}x^{n-1} \in \mathbb{Z}_2[x]$ e os elementos $\alpha, \alpha^2, \ldots, \alpha^{2t} \in \mathbb{F}_{2^m}$. A n-upla binária $v = (v_0, v_1, \ldots, v_{n-1})$ é uma **palavra código** se, e somente se, os elementos $\alpha, \alpha^2, \ldots, \alpha^{2t}$ são as raízes do polinômio $v(x)$.

Seja $v(x)$ um polinômio código de um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros. Seja α^i uma raiz de $v(x)$. Então, para $i = 1, \ldots, 2t$, temos que

$$v(\alpha^i) = v_0 + v_1\alpha^i + v_2\alpha^{2i} + \cdots + v_{n-1}\alpha^{(n-1)i} = 0 \tag{28}$$

Esta igualdade pode ser escrita como o seguinte produto de matrizes

$$[v_0, v_1, \ldots, v_{n-1}] \cdot \begin{bmatrix} 1 \\ \alpha^i \\ \alpha^{2i} \\ \vdots \\ \alpha^{(n-1)i} \end{bmatrix} = 0 \tag{29}$$

para $i = 1, \ldots, 2t$, e onde $(v_0, v_1, \ldots, v_{n-1})$ é a representação vetorial de $v(x)$. Em (29), temos a condição de que o produto interno de $(v_0, v_1, \ldots, v_{n-1})$ por $\alpha, \alpha^2, \ldots, \alpha^{2t}$ é igual a zero.

Consideremos a seguinte matriz de ordem $2t \times n$

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \cdots & \alpha^{n-1} \\ 1 & (\alpha^2) & (\alpha^2)^2 & (\alpha^2)^3 & \cdots & (\alpha^2)^{n-1} \\ 1 & (\alpha^3) & (\alpha^3)^2 & (\alpha^3)^3 & \cdots & (\alpha^3)^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & (\alpha^{2t}) & (\alpha^{2t})^2 & (\alpha^{2t})^3 & \cdots & (\alpha^{2t})^{n-1} \end{bmatrix}. \tag{30}$$

De acordo com (29), se $v = (v_0, v_1, \ldots, v_{n-1})$ é uma palavra código de um código BCH

binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros, então

$$v \cdot H^T = 0. \tag{31}$$

Se uma n-upla $v = (v_0, v_1, \ldots, v_{n-1})$ satisfaz a condição da igualdade (31), segue de (28) e (29), que α^i é uma raiz do polinômio $v(x)$, para $i = 1, \ldots, 2t$. Neste caso, v deve ser uma palavra código de um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros. Consequentemente, o produto em (31) é o vetor nulo do código. E a matriz H é a matriz teste de paridade do código. Se para algum i e j, temos que α^i é um conjugado de α^j, então, de acordo com o Teorema 2.4.2, $v(\alpha^j) = 0$ se, e somente se, $v(\alpha^i) = 0$. Assim sendo, se o produto interno de $v = (v_0, v_1, \ldots, v_{n-1})$ e a i-ésima linha da matriz H é zero, então o produto interno de $v = (v_0, v_1, \ldots, v_{n-1})$ e a j-ésima linha da matriz H também é zero. Logo, a j-ésima linha da matriz H pode ser omitida e a matriz H dada em (30) pode ser reduzida a uma matriz de ordem $t \times n$, da seguinte forma:

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \cdots & \alpha^{n-1} \\ 1 & (\alpha^3) & (\alpha^3)^2 & (\alpha^3)^3 & \cdots & (\alpha^3)^{n-1} \\ 1 & (\alpha^5) & (\alpha^5)^2 & (\alpha^5)^3 & \cdots & (\alpha^5)^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & (\alpha^{2t-1}) & (\alpha^{2t-1})^2 & (\alpha^{2t-1})^3 & \cdots & (\alpha^{2t-1})^{n-1} \end{bmatrix}. \tag{32}$$

Notemos que as entradas da matriz H são elementos do corpo finito $\mathbb{F}_{2^m}$. Cada elemento de $\mathbb{F}_{2^m}$ pode ser representado por uma m-upla sobre $\mathbb{Z}_2$. Se cada entrada de H é substituída por sua correspondente m-upla sobre $\mathbb{Z}_2$ e colocada na forma de coluna, obtemos uma matriz binária de teste de paridade do código.

Exemplo 3.4.8. Consideremos o código BCH(7,4) binário com capacidade de correção de 1 erro, dado no Exemplo 3.4.5. Seja α um elemento primitivo do corpo finito $\mathbb{F}_{2^3}$. Então, a matriz teste de paridade deste código BCH binário é

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 \end{bmatrix}.$$

Agora, vamos utilizar essa matriz H para verificarmos se a palavra $v = (1, 1, 1, 1, 1, 1, 1)$ é uma palavra código. Isto é, se v pertence ao código BCH(7,4) binário com capacidade de correção de 1 erro. Para tanto, usaremos os valores da tabela 4, que se encontra na

página 31. Efetuando o produto de v e H^t, temos

$$[1,1,1,1,1,1,1] \cdot \begin{bmatrix} 1 \\ \alpha \\ \alpha^2 \\ \alpha^3 \\ \alpha^4 \\ \alpha^5 \\ \alpha^6 \end{bmatrix} = 1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4 + \alpha^5 + \alpha^6 = \alpha^3 + \alpha^4 + \alpha^6 = 0.$$

Como o resultado foi o vetor nulo, então a identidade (31) foi verificada. Portanto, $v = (1,1,1,1,1,1,1)$ pertence ao código BCH(7,4) binário.

Observação 3.4.9. A distância mínima d de um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros é maior ou igual a $2t + 1$. A prova desta afirmação pode ser verificada em (LIN; DANIEL, 1983, pp. 148–150).

A observação acima nos conduz a definição de um novo parâmetro, conforme a seguir.

Definição 3.4.10. Seja um código BCH binário de comprimento $n = 2^m - 1$ com capacidade de correção de t erros. O parâmetro $2t + 1$ é chamado de **distância projetada d_0** de um código BCH binário.

Na realidade, a distância mínima de um código BCH binário pode ser ou não igual a distância projetada.

Veremos agora, uma definição mais geral de código BCH binário.

Definição 3.4.11. Sejam o corpo finito $\mathbb{F}_{2^m}$, o elemento $\beta \in \mathbb{F}_{2^m}$ e $l_0 \geq 0 \in \mathbb{Z}$. Um **código BCH binário com distância projetada d_0** é gerado pelo polinômio binário $g(x)$ de grau mínimo que tem como raízes as seguintes potências consecutivas de β:

$$\beta^{(l_0)}, \beta^{(l_0+1)}, \ldots, \beta^{(l_0+d_0-2)}$$

Para $i = 0, \ldots, d_0 - 1$, sejam o polinômio mínimo $min_{\beta^{(l_0+i)}}(x)$ de $\beta^{(l_0+i)}$ e n_i a ordem de $\beta^{(l_0+i)}$. Então

$$g(x) = mmc(min_{\beta^{(l_0)}}(x), min_{\beta^{(l_0+1)}}(x), \ldots, min_{\beta^{(l_0+d_0-2)}}(x))$$

e o comprimento do código é

$$n = mmc(n_0, n_1, \ldots, n_{d_0-2}).$$

Observação 3.4.12. O código BCH binário definido acima tem distância mínima maior ou igual a d_0 e, no máximo, $m(d_0 - 1)$ dígitos de verificação de paridade. Então, esse código BCH binário possui capacidade de corrigir até $[(d_0 - 1)/2]$ erros. Caso $l_0 = 1$, $d_0 = 2t + 1$ e se β é um elemento primitivo de um corpo finito $\mathbb{F}_{2^m}$, temos um **código BCH primitivo binário** de comprimento $n = 2^m - 1$ com capacidade de correção de t erros . Se fizermos: $l_0 = 1$, $d_0 = 2t + 1$ e se β não é um elemento primitivo de um corpo finito $\mathbb{F}_{2^m}$, temos um **código BCH não primitivo binário** com capacidade de correção de t erros e de tamanho n que é a ordem do elemento β. Neste texto, trabalharemos apenas com códigos BCH primitivos binários.

Observemos que, na Definição 3.4.11, foi exigido que o polinômio gerador $g(x)$ do código tenha $d_0 - 1$ potências consecutivas do elemento β como raízes. Essa exigência garante que a distância mínima do código BCH binário assim definido seja maior ou igual a d_0. Desta forma, d_0 é um limite inferior para a distância mínima do código BCH binário e é denominado de *BCH bound*.

Apresentaremos a seguir o processo de decodificação dos códigos BCH binários, que é similar ao da decodificação de códigos lineares mostrado na seção 3.3.

Suponhamos que uma palavra código $v_t = (v_0, v_1, \ldots, v_{n-1})$ seja transmitida, e que os erros introduzidos pelo canal durante a sua transmissão tenham gerado a seguinte palavra recebida

$$v_r = (r_0, r_1, \ldots, r_{n-1}).$$

Consideremos as representações polinomiais da palavra código transmitida v_t e da palavra recebida v_r como sendo, respectivamente, os polinômios $v_t(x) = v_{t_0} + v_{t_1}x + \cdots + v_{t_{n-1}}x^{n-1}$ e $v_r(x) = v_{r_0} + v_{r_1}x + \cdots + v_{r_{n-1}}x^{n-1}$. Seja v_e o padrão de erro introduzido pelo canal, cuja representação polinomial é $v_e(x)$, então temos que

$$v_r(x) = v_t(x) + v_e(x). \tag{33}$$

O primeiro passo do processo de decodificação de um código BCH binário com capacidade de correção de t erros é o calculo da síndrome do vetor recebido v_r (ou $v_r(x)$). E nesse processo de decodificação, essa síndrome é uma $2t$-upla como a seguir

$$S = (S_1, S_2, \ldots, S_{2t}) = v_r \cdot H^T, \tag{34}$$

onde H é a matriz teste de paridade dada por (30). De (30) e (34), temos que a síndrome S_i do i-ésimo componente é dada por

$$S_i = v_r(\alpha^i) = v_{r_0} + v_{r_1}\alpha^i + v_{r_2}\alpha^{2i} + \cdots + v_{r_{n-1}}\alpha^{(n-1)i} \tag{35}$$

para $i = 1, \ldots, 2t$. Podemos observar que as componentes da síndrome são elementos

do corpo finito $\mathbb{F}_{2^m}$. Essas componentes podem ser calculadas a partir de $v_r(x)$, como mostraremos a seguir.

Dividindo $v_r(x)$ pelo polinômio mínimo $min_{\alpha^i}(x)$ de α^i, obtemos

$$v_r(x) = a_i(x) \cdot min_{\alpha^i}(x) + b_i(x),$$

onde $b_i(x)$ é o resto dessa divisão e possui grau menor que $min_{\alpha^i}(x)$. Como $min_{\alpha^i}(\alpha^i) = 0$, temos que

$$S_i = v_r(\alpha^i) = b_i(\alpha^i). \tag{36}$$

Logo, a síndrome S_i do i-ésimo componente é obtida substituindo-se x por α^i em $b_i(x)$.

Exemplo 3.4.13. Consideremos o código BCH(7,4) binário com capacidade de correção de um erro dado no Exemplo 3.4.5. Suponhamos que a palavra código $v_t = (0,0,0,0,0,0,0)$ foi transmitida por um canal de comunicação. E que, por causa de algum ruído durante essa transmissão, foi recebida a palavra $v_r = (0,1,0,0,0,0,0)$. O polinômio correspondente a v_r é

$$v_r(x) = x.$$

A síndrome consiste de 2 componentes, pois para $t = 1$, $2t = 2$. Então, temos que

$$S = (S_1, S_2).$$

De acordo com a tabela 5 (página 35), os polinômios mínimos de α e α^2 são iguais e, portanto, temos $min_\alpha(x) = min_{\alpha^2}(x) = 1 + x + x^3$. Dividindo $v_r(x) = x$ por $min_\alpha(x) = 1 + x + x^3$, obtemos o resto

$$b_1(x) = x.$$

Usando os valores do corpo finito $\mathbb{F}_{2^3}$ dados pela tabela 4 (página 31) e substituindo α e α^2 em $b_1(x)$, obtemos

$$S_1 = b_1(\alpha) = \alpha \qquad \text{e} \qquad S_2 = b_1(\alpha^2) = \alpha^2.$$

Portanto, a síndrome procurada é

$$S = (\alpha, \alpha^2).$$

Os componentes dessa síndrome também podem ser obtidos substituindo-se α e α^2 em $v_r(x) = x$, de acordo com (35). Então, dessa forma, temos que

$$S_1 = v_r(\alpha) = \alpha \qquad \text{e} \qquad S_2 = v_r(\alpha^2) = \alpha^2.$$

Logo, a síndrome procurada é

$$S = (\alpha, \alpha^2).$$

Como $\alpha, \alpha^2, \ldots, \alpha^{2t}$ são as raízes de cada polinômio código, então $v_t(\alpha^i) = 0$ para $i = 1, \ldots, 2t$. De (33) e (35) obtemos a seguinte relação entre os componentes S_i da síndrome e o padrão de erro:

$$S_i = v_e(\alpha^i) \tag{37}$$

para $i = 1, \ldots, 2t$. Notemos que, de acordo com (37), a síndrome S depende apenas do padrão de erro. Suponhamos que o padrão de erro $v_e(x)$ tenha w erros nas posições $x^{j_1}, x^{j_2}, \ldots, x^{j_w}$. Isto é,

$$e(x) = x^{j_1} + x^{j_2} + \cdots + x^{j_w}, \tag{38}$$

onde $0 \leq j_1 < j_2 < \cdots < j_w < n$. De (37) e (38) obtemos o seguinte conjunto de equações:

$$\begin{cases} S_1 &= \alpha^{j_1} + \alpha^{j_2} + \cdots + \alpha^{j_w} \\ S_2 &= (\alpha^{j_1})^2 + (\alpha^{j_2})^2 + \cdots + (\alpha^{j_w})^2 \\ S_3 &= (\alpha^{j_1})^3 + (\alpha^{j_2})^3 + \cdots + (\alpha^{j_w})^3 \\ \vdots & \\ S_{2t} &= (\alpha^{j_1})^{2t} + (\alpha^{j_2})^{2t} + \cdots + (\alpha^{j_w})^{2t}, \end{cases} \tag{39}$$

onde $\alpha^{j_1}, \alpha^{j_2}, \ldots, \alpha^{j_w}$ são desconhecidos.

Observação 3.4.14. Qualquer método para a solução do sistema de equações (39) é um algoritmo de decodificação dos códigos BCH.

As potências $j_1, j_2, \ldots, j_w$ indicam as posições dos erros no padrão de erro $e(x)$ como em (38), ao serem encontrados os valores de $\alpha^{j_1}, \alpha^{j_2}, \ldots, \alpha^{j_w}$. De um modo geral, o sistema de equações (39) tem muitas soluções possíveis. E cada solução gera um padrão de erro diferente. Então, entre as soluções cujo padrão de erro $e(x)$ possui um número de erros menor ou igual a t, a solução correta é a que tem o menor número de erros. Ou seja, o padrão de erro correspondente à esta solução é o mais provável padrão de erro $e(x)$ causado pelo ruído do canal de comunicação. Para valores grandes de t, calcular uma solução direta para o sistema de equações (39) é difícil e ineficaz.

Apresentaremos a seguir, o algoritmo de Peterson (3) (citado em (LIN; DANIEL, 1983)), que é um procedimento eficaz para a determinação de α^{j_l}, para $l = 1, \ldots, w$, dos componentes S_l da síndrome.

Nosso objetivo será apenas apresentar o algoritmo sem entrarmos em detalhes sobre sua fundamentação matemática, a qual envolve conhecimentos que fogem do escopo deste

trabalho. Para mais informações, consultar <http://www.professores.uff.br/nmedeiros/wp-content/uploads/sites/88/2017/08/Algebra-III-2016_2-galois.pdf>.

Algoritmo 3 - Algoritmo de Peterson

Para $l = 1, \ldots, w$, seja

$$\theta_l = \alpha^{j_l} \tag{40}$$

Estes elementos são chamados de **números de localização de erros**, pois eles indicam as posições dos erros. Então, o sistema de equações (39) pode ser escrito da seguinte forma:

$$\begin{cases} S_1 &= \theta_1 + \theta_2 + \cdots + \theta_w \\ S_2 &= \theta_1^2 + \theta_2^2 + \cdots + \theta_w^2 \\ \vdots & \\ S_{2t} &= \theta_1^{2t} + \theta_2^{2t} + \cdots + \theta_w^{2t}. \end{cases} \tag{41}$$

Consideremos o seguinte polinômio

$$\begin{aligned} \sigma(x) &= (1 + \theta_1 x)(1 + \theta_2 x) \cdot \ldots \cdot (1 + \theta_w x) \\ &= \sigma_0 + \sigma_1 x + \sigma_2 x^2 + \cdots + \sigma_w x^w. \end{aligned} \tag{42}$$

As raízes de $\sigma(x)$ são os inversos dos números de localização de erros: $\theta_1^{-1}, \theta_2^{-1}, \ldots, \theta_w^{-1}$. Por esta razão, o polinômio $\sigma(x)$ é chamado de **polinômio localizador de erros**. Observemos que $\sigma(x)$ é um polinômio desconhecido com coeficientes que precisam ser determinados. Os coeficientes de $\sigma(x)$ e os números de localização de erros estão relacionados pelo seguinte sistema de equações:

$$\begin{cases} \sigma_0 &= 1 \\ \sigma_1 &= \theta_1 + \theta_2 + \cdots + \theta_w \\ \sigma_2 &= \theta_1\theta_2 + \theta_2\theta_3 + \cdots + \theta_{w-1}\theta_w \\ \vdots & \\ \sigma_w &= \theta_1\theta_2 \cdot \ldots \cdot \theta_w. \end{cases} \tag{43}$$

De (41) e (43) vemos que os σ_i's se referem às componentes S_i da síndrome. Essa relação

está indicada nas seguintes identidades:

$$\begin{aligned}
S_1 + \sigma_1 &= 0 \\
S_2 + \sigma_1 S_1 + 2\sigma_2 &= 0 \\
S_3 + \sigma_1 S_2 + \sigma_2 S_1 + 3\sigma_3 &= 0 \\
&\vdots \\
S_w + \sigma_1 S_{w-1} + \cdots + \sigma_{w-1} S_1 + w\sigma_w &= 0 \\
S_{w+1} + \sigma_1 S_w + \cdots + \sigma_{w-1} S_2 + \sigma_w S_1 &= 0
\end{aligned} \tag{44}$$

Para o caso binário, como $1 + 1 = 2 = 0$, temos que

$$i\sigma_i = \begin{cases} \sigma_i & \text{para } i \text{ ímpar} \\ 0 & \text{para } i \text{ par.} \end{cases} \tag{45}$$

Portanto, ao aplicarmos as alterações de (45) nas identidades (44), temos que as identidades dos componentes de ordem par não estabelecem relações entre os σ_i's com os componentes S_i's da síndrome. Logo, o sistema (44) resulta em:

$$\begin{aligned}
S_1 + \sigma_1 &= 0 \\
S_3 + \sigma_1 S_2 + \sigma_2 S_1 + \sigma_3 &= 0 \\
S_5 + \sigma_1 S_4 + \sigma_2 S_3 + \sigma_3 S_2 + \sigma_4 S_1 + \sigma_5 &= 0 \\
&\vdots
\end{aligned} \tag{46}$$

que é um sistema de $\frac{2t}{2} = t$ identidades.

Suponhamos que seja possível determinar as funções elementares $\sigma_1, \sigma_2, \ldots, \sigma_w$ das equações (44). Então, os números de localização de erros $\theta_1, \theta_2, \ldots, \theta_w$ podem ser encontrados pela determinação das raízes do polinômio localizador de erros $\sigma(x)$. As equações (44) possuem muitas soluções possíveis. Entretanto, a solução será aquela que gerar um $\sigma(x)$ de grau mínimo. Este polinômio $\sigma(x)$ produzirá um padrão de erro com um número mínimo de erros. Se $w \leq t$, o polinômio $\sigma(x)$ fornecerá o verdadeiro padrão de erro.

Basicamente, o algoritmo de Peterson (3) pode ser resumido em 3 passos principais:
1) Calcular a síndrome $S = (S_1, S_2, \ldots, S_{2t})$ a partir do polinômio recebido $v_r(x)$;
2) Determinar o polinômio localizador de erros $\sigma(x)$ dos componentes $S_1, S_2, \ldots, S_w$ da síndrome;
3) Determinar os números localizadores de erros $\theta_1, \theta_2, \ldots, \theta_w$, através das raízes de $\sigma(x)$ e corrigir os erros em $v_r(x)$.

Observação 3.4.15. Os passos 1 e 3 são simples. Por outro lado, o passo 2 é o mais complicado da decodificação de códigos BCH binários.

Exemplo 3.4.16. Consideremos o código BCH(7,4) binário com capacidade de correção de 1 erro dado no Exemplo 3.4.5. Suponhamos que a palavra código $v_t = (0,0,0,0,0,0,0)$ foi transmitida por um canal de comunicação. Mas, por causa de algum ruído no canal, foi recebida a palavra $v_r = (0,0,0,1,0,0,0)$. O polinômio correspondente a palavra v_r é

$$v_r(x) = x^3.$$

Como o código é binário usaremos as identidades do sistema (46). Logo, da primeira identidade desse sistema, obtemos a igualdade

$$S_1 = \sigma_1.$$

Além disso, pelo fato de o código BCH(7,4) binário ter capacidade de correção de 1 erro, a síndrome tem apenas um componente, então, temos que

$$S = (S_1),$$

e o polinômio localizador de erros, dado por (42), fica

$$\sigma(x) = 1 + \sigma_1 x.$$

Então, calculando S_1 temos que

$$S_1 = v_r(\alpha) = \alpha^3.$$

E o polinômio localizador de erros para esse caso é

$$\sigma(x) = 1 + \sigma_1 x = 1 + \alpha^3 x.$$

Calculando as raízes de $\sigma(x)$ temos que

$$\begin{aligned}
\sigma(1) &= 1 + \alpha^3 = \alpha \\
\sigma(\alpha) &= 1 + \alpha^3 \cdot \alpha = 1 + \alpha^4 = \alpha^5 \\
\sigma(\alpha^2) &= 1 + \alpha^3 \cdot \alpha^2 = 1 + \alpha^5 = \alpha^4 \\
\sigma(\alpha^3) &= 1 + \alpha^3 \cdot \alpha^3 = 1 + \alpha^6 = \alpha^2 \\
\sigma(\alpha^4) &= 1 + \alpha^3 \cdot \alpha^4 = 1 + \alpha^7 = 1 + 1 = 0 \\
\sigma(\alpha^5) &= 1 + \alpha^3 \cdot \alpha^5 = 1 + \alpha^8 = 1 + \alpha = \alpha^3 \\
\sigma(\alpha^6) &= 1 + \alpha^3 \cdot \alpha^6 = 1 + \alpha^9 = 1 + \alpha^2 = \alpha^6
\end{aligned}$$

Observamos que α^4 é raiz do polinômio localizador de erros $\sigma(x) = 1 + \sigma_1 x = 1 + \alpha^3 x$.

A posição do erro é o inverso da raiz α^4. Então, temos que

$$\theta_1 = \frac{1}{\alpha^4} = \alpha^{-4} = \alpha^3.$$

Portanto, o padrão de erro é $v_e(x) = x^3$ e temos que a palavra transmitida foi

$$v_t(x) = v_e(x) + v_r(x) = x^3 + x^3 = 0,$$

que é o polinômio identicamente nulo correspondente a palavra $v_t = (0, 0, 0, 0, 0, 0, 0)$.

No Exemplo 3.4.16 foi utilizado o código BCH(7,4) binário com capacidade de correção de 1 erro. Porém, existem situações em que ocorrem mais de um erro. Veremos a seguir um exemplo que ilustra como é feita a correção de 2 erros, usando um código BCH binário.

Exemplo 3.4.17. Consideremos o corpo finito $\mathbb{F}_{2^4}$, gerado por $p(x) = 1 + x + x^4$, dado pela tabela 8.

Tabela 8 - Representações dos elementos de $\mathbb{F}_{2^4}$, gerado por $p(x) = 1 + x + x^4$

Potência	Polinômio	Vetor
0	0	(0 0 0 0)
1	1	(1 0 0 0)
α	α	(0 1 0 0)
α^2	α^2	(0 0 1 0)
α^3	α^3	(0 0 0 1)
α^4	$1 + \alpha$	(1 1 0 0)
α^5	$\alpha + \alpha^2$	(0 1 1 0)
α^6	$\alpha^2 + \alpha^3$	(0 0 1 1)
α^7	$1 + \alpha + \alpha^3$	(1 1 0 1)
α^8	$1 + \alpha^2$	(1 0 1 0)
α^9	$\alpha + \alpha^3$	(0 1 0 1)
α^{10}	$1 + \alpha + \alpha^2$	(1 1 1 0)
α^{11}	$\alpha + \alpha^2 + \alpha^3$	(0 1 1 1)
α^{12}	$1 + \alpha + \alpha^2 + \alpha^3$	(1 1 1 1)
α^{13}	$1 + \alpha^2 + \alpha^3$	(1 0 1 1)
α^{14}	$1 + \alpha^3$	(1 0 0 1)

Fonte: (LIN; DANIEL, 1983)

Consideremos agora os polinômios mínimos de $\mathbb{F}_{2^4}$, gerado por $p(x) = 1 + x + x^4$, dados pela tabela 9.

Tabela 9 - Polinômios mínimos dos elementos de $\mathbb{F}_{2^4}$ gerado por $p(x) = 1 + x + x^4$

Raízes conjugadas	Polinômios mínimos
0	x
1	$x + 1$
$\alpha, \alpha^2, \alpha^4, \alpha^8$	$x^4 + x + 1$
$\alpha^3, \alpha^6, \alpha^9, \alpha^{12}$	$x^4 + x^3 + x^2 + x + 1$
α^5, α^{10}	$x^2 + x + 1$
$\alpha^7, \alpha^{11}, \alpha^{13}, \alpha^{14}$	$x^4 + x^3 + 1$

Fonte: (LIN; DANIEL, 1983)

Sejam α e α^3 dois elementos primitivos de $\mathbb{F}_{2^4}$. Da tabela 9, obtemos o polinômio mínimo desses elementos, que são os seguintes:

$$min_{\alpha}(x) = 1 + x + x^4 \quad \text{e} \quad min_{\alpha^3}(x) = 1 + x + x^2 + x^3 + x^4.$$

De (27) verificamos que o código BCH com capacidade de correção de dois erros e comprimento $n = 2^4 - 1 = 15$ é gerado por

$$g(x) = \text{MMC}(min_{\alpha}(x), min_{\alpha^3}(x)),$$

ou seja

$$g(x) = (1 + x + x^4)(1 + x + x^2 + x^3 + x^4).$$

Então, temos que

$$g(x) = 1 + x^4 + x^6 + x^7 + x^8.$$

Logo, de acordo com a tabela 7 (página 62), o código acima é o código BCH(15,7) binário, que tem distância mínima maior ou igual a 5. Como o polinômio gerador $g(x)$ tem peso 5, a distância mínima deste código é 5.

Suponhamos que a palavra código $v_t = (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)$ tenha sido transmitida por um canal de comunicação e que, por causa de alguma interferência neste canal, a palavra código recebida foi $v_r = (1,1,0,0,0,0,0,0,0,0,0,0,0,0,0)$. O polinômio correspondente desta palavra código é

$$v_r(x) = 1 + x.$$

Como estamos trabalhando com o código BCH(15,7) binário, devemos calcular as seguin-

tes síndromes, usando os valores da tabela 8 e considerando que $\alpha^{15} = 1$:

$$S_1 = v_r(\alpha) = 1 + \alpha \quad \text{e} \quad S_3 = v_r(\alpha^3) = 1 + \alpha^3.$$

Usando as identidades do sistema (46), obtemos

$$\sigma_1 = 1 + \alpha \quad \text{e} \quad \sigma_2 = \alpha.$$

E o polinômio localizador de erros é

$$\sigma(x) = 1 + (1 + \alpha)x + \alpha x^2.$$

Efetuando os cálculos

$$\sigma(1) = 1 + (1 + \alpha) \cdot 1 + \alpha \cdot 1^2 = 0$$

e

$$\sigma(\alpha^{14}) = 1 + (1 + \alpha) \cdot \alpha^{14} + \alpha \cdot (\alpha^{14})^2 = 1 + \alpha^{14} + \alpha^{15} + \alpha^{29} = 0,$$

verificamos que as raízes de $\sigma(x)$ são os elementos 1 e α^{14}. Como as posições dos erros são os inversos dessas raízes, temos que

$$\theta_1 = \frac{1}{1} = 1 \quad \text{e} \quad \theta_2 = \frac{1}{\alpha^{14}} = \alpha^{-14} = \alpha,$$

e os erros estão nas posições 1 e x. Portanto, o padrão de erro é $v_e(x) = 1 + x$ e a palavra transmitida foi

$$v_t(x) = v_e(x) + v_r(x) = 1 + x + 1 + x = 0,$$

que corresponde ao vetor $v_t = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$.

CONCLUSÃO

Em nossa prática docente como professores de Matemática, vemos que é comum os alunos questionarem sobre a necessidade de estudarem esta disciplina e quais são as aplicações dos conteúdos matemáticos vistos por eles. Isso também é observado quando ensinamos polinômios e funções polinomiais na Escola Básica. Os exemplos que apresentamos em sala de aula, tais como: trajetórias de projéteis, expressões de áreas ou alguns modelos de fenômenos naturais, entre outros, envolvem expressões polinomiais cujos coeficientes são números reais ou complexos. No entanto, conforme comentamos na introdução deste texto, podemos ter polinômios que não usam tais conjuntos, como é o caso daqueles que ocorrem em Criptografia, mais precisamente em alguns tipos de códigos corretores de erros.

Estes tipos de códigos foram criados com o propósito de tratarem da segurança da informação, tanto no armazenamento como na transmissão de dados e são cada vez mais usados no nosso dia-a-dia. Neste texto, escolhemos abordar códigos corretores de erros BCH que trabalham sobre Corpos Finitos. Mostramos de maneira direta e objetiva e por meio de vários exemplos, como os polinômios são utilizados nesses códigos. Fizemos isso de uma forma que não fosse muito aprofundada, mas nem por isso, superficial.

Além de polinômios, abordamos, de forma direta ou indireta, outros conteúdos matemáticos essenciais para o estudo desses códigos corretores de erros, tais como: anéis, corpos, transformação linear, matrizes, vetores e aritmética modular. Dessa maneira, trouxemos uma aplicação de polinômios vinculada a um tema da atualidade (segurança da informação) e provavelmente do futuro; um assunto estudado nas aulas de Matemática e que é usado diariamente por todos que acessam à internet, assistem TV digital, usam celular, entre outras tecnologias.

REFERÊNCIAS

ARAGÃO, Canuto Ruan Santos. Códigos cíclicos: uma introdução aos códigos corretores de erros. Universidade Federal de Sergipe, 2017.

BERLEKAMP, Elwyn R. Algebraic coding theory-revised edition. World Scientific Publishing Co., Inc., 2015.

BIAZZI, Ricardo Neves. Polinômios irredutíveis: critérios e aplicações. Universidade Estadual Paulista (UNESP), 2014.

BOLDRINI, José Luiz et al. *Álgebra linear*. [S.l.]: Harper & Row, 1980.

CARVALHO, Sézani Morais de. Matrizes, determinantes e polinômios: aplicações em códigos em corretores de erros, como estratégias motivacional para o ensino de matemática. 2014.

COUTINHO, M. Corpos finitos e códigos corretores de erros. *Juiz de Fora: Universidade Federal de Juiz de Fora*, 2014.

HEFEZ, Abramo. Curso de algebra, vol. 1. *Coleção Matemática Universitária, IMPA/CNPq, RJ*, 1993.

HEFEZ, Abramo; VILLELA, Maria Lúcia. *Códigos Corretores de Erros*. 2. ed. Rio de Janeiro: IMPA, 2017.

IEZZI, Gelson. *Fundamentos de matemática elementar: complexos, polinômios, equações*. [S.l.]: Atual, 2005.

LIN, Shu; DANIEL, Costello. *COSTELLO. Error Control Coding: Fundamentals and Applications*. [S.l.]: Prentice-Hall, Englewood Cliffs, NJ, 1983.

MILIES, César Polcino. Breve introdução à teoria dos códigos corretores de erros. *Colóquio de Matemática da Região Centro-Oeste*, 2009. Disponível em: <https://www.sbm.org.br/docs/coloquios/NE-1.04.pdf>. Acesso em: 10 mar 2018.

NICOLETTI, Everton Rodrigo. Aplicações de ágebra linear aos códigos corretos de erros e ao ensino médio. Universidade Estadual Paulista (UNESP), 2015.

STALLINGS, William. *Criptografia e segurança de redes, tradução Vieira, Daniel*. [S.l.]: São Paulo. Pearson Prentice Hall, 2008.