UNIVERSIDADE FEDERAL DE OURO PRETO

Pedro Leonardo Pinto de Souza

Códigos Corretores de Erros

Ouro Preto, Brasil

2020

UNIVERSIDADE FEDERAL DE OURO PRETO

Pedro Leonardo Pinto de Souza

Códigos Corretores de Erros

Monografia submetida ao Colegiado do Curso de Matemática do Instituto de Ciências Exatas e Biológicas da Universidade Federal de Ouro Preto como requisito parcial para a conclusão do curso de Bacharelado em Matemática.

Orientador: Prof. Dr. Edney Augusto Jesus de Oliveira

Coorientador: Prof. Me. Vinícius Vivaldino Pires de Almeida

Universidade Federal de Ouro Preto – UFOP

Instituto de Ciências Exatas e Biológicas

Departamento de Matemática

Ouro Preto, Brasil

2020

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S729c Souza, Pedro Leonardo Pinto de .
Códigos corretores de erros. [manuscrito] / Pedro Leonardo Pinto de Souza. - 2020.
145 f.: il.: color., tab..

Orientador: Prof. Dr. Edney Augusto Jesus de Oliveira.
Coorientador: Prof. Me. Vinícius Vivaldino Pires de Almeida.
Monografia (Bacharelado). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Graduação em Matemática .

1. Códigos corretores de erros (Teoria da informação) . 2. Códigos . 3. Teoria da codificação. 4. Decodificação. I. Almeida, Vinícius Vivaldino Pires de. II. Oliveira, Edney Augusto Jesus de. III. Universidade Federal de Ouro Preto. IV. Título.

CDU 512

Bibliotecário(a) Responsável: Celina Brasil Luiz - CRB6-1589

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIENCIAS EXATAS E BIOLOGICAS
COLEGIADO DO CURSO DE BACHARELADO EM MATEMATICA

FOLHA DE APROVAÇÃO

Pedro Leonardo Pinto de Souza

Códigos corretores de erros

Monografia apresentada ao Curso de Bacharelado em Matemática da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de bacharel em Matemática

Aprovada em 28 de setembro de 2020

Membros da banca

Dr. Edney Augusto Jesus de Oliveira - Orientador (Universidade Federal de Ouro Preto)
Ms. Vinícius Vivaldino Pires de Almeida - Coorientador (Universidade Federal de Ouro Preto)
Dr. Juliano Soares Dias - (Universidade Federal de Ouro Preto)
Dra. Mariana Garabini Cornelissen Hoyos - (Universidade Federal de São João Del-Rei)

Edney Augusto Jesus de Oliveira, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 30/03/2021

Documento assinado eletronicamente por **Edney Augusto Jesus de Oliveira**, **PROFESSOR DE MAGISTERIO SUPERIOR**, em 30/03/2021, às 21:31, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0 , informando o código verificador **0149426** e o código CRC **42C367A5**.

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35400-000
Telefone: - www.ufop.br

Agradecimentos

Agradeço, primeiramente, aos meus pais, por serem presentes e me apoiarem sempre, com muito afeto e cuidado, especialmente minha mãe, que com toda sua generosidade e sabedoria ensinou-me as coisas mais importantes da vida. Agradeço à minha irmã, por sempre me apoiar com muito carinho. Agradeço à minha família, por todo o suporte, carinho e por ser meu porto seguro. Agradeço a minha amiga e companheira por todo carinho e apoio, além do incentivo incansável. Agradeço ao meu orientador e coorientador pelo excelente trabalho desenvolvido comigo. Agradeço também por toda a paciência, dedicação e principalmente por acreditarem no meu trabalho. Foi um caminho longo e difícil, que só foi possível pela confiança depositada. Sou imensamente grato por isso. Agradeço aos professores que se dispuseram a participar da banca de defesa da monografia e contribuir para a melhoria e o enriquecimento do trabalho. Agradeço a todos os professores que tive durante a graduação, por todo o conhecimento compartilhado com grande profissionalismo. Agradeço ao PETMAT-UFOP, por contribuir intensamente na minha formação e por ser o meu agente de transformação na universidade. Agradeço aos meus amigos e colegas pelo apoio e torcida, especialmente os mais íntimos, por saberem ouvir, serem pacientes e presentes sempre que precisei. Por fim, agradeço a todos que fizeram parte dessa trajetória e contribuíram direta ou indiretamente para minha formação.

"Nada a temer senão o correr da luta."
(Luiz Carlos Sá e Sérgio Magrão)

Resumo

Os códigos corretores de erros são ferramentas essenciais na comunicação eletrônica, em virtude da sua confiabilidade ao garantirem a integridade da informação transmitida. Nosso objetivo nesse trabalho é apresentarmos, com um viés matemático, a construção dos códigos lineares, códigos cíclicos e códigos BCH, e mostrarmos como ocorre o processo de detecção e correção de erros de cada um. Em um primeiro momento, a principal motivação para esse estudo é entendermos como a álgebra linear está relacionada aos códigos lineares através de conceitos como espaços vetoriais, transformações lineares e suas respectivas matrizes, além de observarmos como alguns resultados desses códigos, herdados de espaços vetoriais, facilitam vários dos cálculos necessários para codificarmos e decodificarmos as suas palavras. Em um segundo momento, nossa motivação é entendermos como alguns resultados da álgebra abstrata, relacionados a corpos finitos, anéis, ideais e anéis de polinômios, possibilitam definirmos os códigos cíclicos e BCH com a estrutura adicional de ideal, a qual permite o desenvolvimento de algoritmos de codificação e decodificação mais eficientes, do ponto de vista matemático, do que para os códigos lineares. Além disso, como os códigos BCH se destacam dos outros dois quando utilizamos conceitos como extensões de corpos e de raízes da unidade em sua construção. Para isso, exemplificamos cada um dos códigos corretores de erros estudados.

Palavras-chave: Códigos Corretores de Erros, Códigos Lineares, Códigos Cíclicos, Códigos BCH, Codificação, Decodificação.

Sumário

1 Introdução

A comunicação eletrônica como conhecemos hoje foi primordialmente desenvolvida pelo matemático estadunidense e cientista da computação Claude Elwood Shannon (1926 - 2001). A descoberta de Shannon, de que era possível transmitir informações de maneira segura através de canais de comunicação imperfeitos, como linhas telefônicas, foi publicada em 1948 no artigo mais importante de sua carreira, intitulado *"A mathematical theory of communication"*, dando início a Teoria da Informação e, consequentemente, o desenvolvimento dos códigos corretores de erros (IEEE, 2020).

Shannon trabalhava na empresa Bell Telephone Laboratories, Inc.[1], situado em Nova Jersey/EUA, quando publicou seu artigo. No entanto, antes mesmo de seu artigo ser difundido fora do Laboratório, um de seus colegas, o também estadunidense Richard Wesley Hamming (1915 - 1998) (LEE, 2019), projetou um conjunto de códigos binários lineares, capaz de corrigir (durante uma transmissão de dados) apenas um único bit em cada bloco de bits (BERLEKAMP, 2015, página vii, prefácio). Era de se esperar que os pesquisadores da época estivessem interessados em desenvolver códigos mais eficientes, o que se tornou realidade poucos anos depois, com destaque para os códigos de Reed-Muller, cuja maior vantagem é a simplicidade de codificar e decodificar mensagens, desenvolvidos pelos americanos D. E. Muller e I. S. Reed em 1954, (RAAPHORST, 2003, página 2, introdução); os códigos BCH (Bose-Chaudhuri-Hocquenhem), que são capazes de corrigir múltiplos erros, sendo desenvolvidos pelo francês Alexis Hocquenghem em 1959 e de maneira independente pelos indianos Raj Chandra Bose e D. K. Chaudhuri em 1960 (REED; CHEN, 2001, página 189, capítulo 5); e os códigos de Reed-Solomon (1960), desenvolvidos pelos americanos Gustave Solomon e I. S. Reed, capazes de corrigirem erros não binários (GEISEL, 1990, páginas 1-8, introdução).

Desde a descoberta de Shannon e o início do desenvolvimento dos códigos corretores de erros, a humanidade fez grandes progressos na comunicação eletrônica e ainda o faz. Este fato deixa evidente a relevância desses estudos tanto no aspecto científico, quanto social.

Neste trabalho abordamos os códigos corretores de erros, mais especificamente, os códigos lineares, os códigos cíclicos e os códigos BCH. Nosso objetivo é apresentarmos a construção desses códigos com um viés matemático, destacando suas principais diferenças na prática, como o seu comprimento, sua capacidade de detecção e correção de erros e o quanto a

[1] Bell Telephone Laboratories, Inc. foi fundada em 1925, sendo uma empresa marcada pela pesquisa científica e desenvolvimento tecnológico de alto nível ao longo de sua história. Em 2016 o Bell Labs e a Nokia se uniram para criar um Bell Labs maior e possui o nome de Nokia Bell Labs, (LABS, 2019).

base teórica se torna complexa quando exigimos uma maior praticidade e eficiência, do ponto de vista matemático, de tais códigos. Esse estudo foi baseado em (HEFEZ; VILLELA, 2008).

No capítulo 2 apresentamos os principais resultados teóricos utilizados ao longo do trabalho por meio de definições, teoremas e proposições da álgebra linear e álgebra abstrata, como espaços e subespaços vetoriais, ideais, anéis de polinômios, extensão de corpos e adjunção de raízes. Alguns resultados foram demonstrados e exemplificados quando julgados necessários, e outros optamos por apenas referenciá-los, a fim de proporcionar uma melhor compreensão durante a leitura. Os estudos relacionados à álgebra linear foram baseados em (COELHO; LOURENCO, 2013), já os estudos relativos à álgebra abstrata foram baseados em (BHATTACHARYA; JAIN; NAGPAUL, 1994), (HEFEZ, 2002), (HEFEZ; VILLELA, 2008) e (GONÇALVES, 2017).

No capítulo 3, apresentamos os códigos corretores de erros e falamos sobre a motivação de estudá-los. Além disso, notamos que no processo de envio de uma mensagem entre o remetente e o destinatário, a mensagem atravessa um ou mais canais de comunicação. Por essa razão, podem ocorrer interferências nesses canais de comunicação que podem afetar a integridade da mensagem. Vale ressaltar que nosso interesse é garantirmos a integridade dessa mensagem e não o sigilo, sendo esse último, um assunto da teoria de Criptografia. Para discutirmos a construção de um código corretor de erros e como é o seu processo de correção, demos um exemplo de uma lâmpada com leds coloridos que altera a cor da luz ambiente por meio de um controle manual, cuja transmissão ocorre do controle para a lâmpada, ou seja, o controle é o remetente e a lâmpada o destinatário. Posteriormente, apresentamos qual o ponto de partida para a criação de um código corretor de erros e alguns resultados acerca da metrização de um código, como a métrica de Hamming, conceitos de disco e esfera, e distância mínima de um código. Para os estudos desse capítulo utilizamos a bibliografia (HEFEZ; VILLELA, 2008).

No capítulo 4 abordamos os códigos lineares com foco nos conceitos matemáticos fundamentais para a sua construção. Observamos que um código linear pode ser visto como subespaço vetorial de um espaço vetorial dado, o que possibilita considerarmos seus elementos como vetores. Além disso, mostramos como algumas propriedades de espaços vetoriais facilitam na manipulação das palavras do código. Destacamos também as transformações lineares, pois a partir delas podemos enxergar um código linear como imagem ou núcleo dessa transformação e, mais ainda, a matriz dessa transformação é uma matriz que gera todos os elementos do código, a qual chamamos de matriz geradora. Mostramos que tal matriz possibilita gerar uma outra matriz, chamada de matriz teste de paridade, sendo esta última imprescindível no processo de detecção e correção de erros, o qual descrevemos como algoritmo da decodificação para códigos lineares. Todo o estudo feito nesse capítulo foi baseado em (HEFEZ; VILLELA, 2008).

No capítulo 5 apresentamos os códigos cíclicos, cuja principal diferença para os códigos

lineares é que este pode ser visto como ideal de um anel quociente. Com isso, além da estrutura de espaço vetorial, um código cíclico possui a estrutura de um ideal. Deduzimos que este ideal é principal e que através do polinômio gerador deste ideal podemos construir a matriz geradora do código cíclico. Além disso, a partir desse mesmo polinômio, conseguimos encontrar um outro polinômio, que chamamos de polinômio recíproco, o qual nos fornece uma maneira de encontrar a matriz teste de paridade do código. A partir daí, estabelecemos o algoritmo da codificação de um código cíclico, já o algoritmo da decodificação que utilizamos é o mesmo dos códigos lineares. Apesar de conseguirmos codificar e decodificar os elementos de um código cíclico de maneira mais simples, não conseguimos determinar a sua distância mínima, o que justifica o estudo dos códigos BCH do capítulo 6. Esse capítulo foi inteiramente baseado em (HEFEZ; VILLELA, 2008).

No capítulo 6 abordamos os códigos BCH, os quais são um caso particular dos códigos cíclicos, cuja condição extra está em seu polinômio gerador, que agora é um polinômio minimal de raízes n-ésimas da unidade. Veremos que essa restrição sobre os polinômios geradores nos permitirá determinar uma cota inferior para a sua distância mínima. Além disso, utilizamos os mesmos algoritmos da codificação e decodificação dos códigos cíclicos e, como sabemos a cota para a sua distância mínima, esses algoritmos tornam-se altamente eficientes, do ponto de vista matemático, em virtude da facilidade de construir esses códigos e da sua capacidade de corrigir múltiplos erros. Os estudos desse capítulo foram todos baseados em (HEFEZ; VILLELA, 2008).

2 Base Teórica

Apresentaremos neste capítulo alguns conceitos relativos à álgebra linear e à teoria de anéis que serão relevantes para o desenvolvimento desse trabalho. Introduziremos tais conceitos através de definições, proposições, teoremas, entre outros resultados, a fim de elucidar quais os princípios teóricos fundamentam a Teoria dos Códigos, da qual os códigos corretores de erros fazem parte. Assumiremos ao longo de todo o texto as definições de *Anel* e *Corpo* estabelecidas em (HEFEZ, 2002).

2.1 Tópicos de Álgebra Linear

Nesta seção, apresentaremos alguns conceitos da álgebra linear, relativos a espaços vetoriais sobre um corpo finito K e transformações lineares entre dois espaços vetoriais, uma vez que os códigos corretores de erros são desenvolvidos utilizando espaços vetoriais sobre corpos finitos.

Definição 1. *Seja V um espaço vetorial sobre K.*

(i) Dizemos que um vetor $\mathbf{v} \in V$ é uma combinação linear *dos vetores $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, quando existem escalares $\alpha_1, \ldots, \alpha_n \in K$ tais que*

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \sum_{i=1}^{n} \alpha_i\mathbf{v}_i.$$

(ii) Seja $\mathcal{B}$ um subconjunto não vazio de V. Dizemos que $\mathcal{B}$ é um conjunto gerador *de V quando todo elemento de V é uma combinação linear de um número finito de elementos de $\mathcal{B}$.*

Definição 2. *Sejam V um espaço vetorial sobre K e $\mathcal{B}$ um subconjunto de V.*

(i) Dizemos que $\mathcal{B}$ é linearmente independente quando $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}$, para $\mathbf{v}_i \in \mathcal{B}$ e $\alpha_i \in K, i = 1, \ldots, n$, implica em $\alpha_1 = \cdots = \alpha_n = 0$.

(ii) O conjunto $\mathcal{B}$ é chamado de linearmente dependente quando não é linearmente independente.

Combinando as Definições 1 e 2, definimos o conceito de base.

Definição 3. *Seja V um espaço vetorial sobre um corpo K. Dizemos que um subconjunto $\mathcal{B}$ de V é uma base de V quando:*

(i) $\mathcal{B}$ *é um conjunto gerador de* V*;*

(ii) $\mathcal{B}$ *é linearmente independente.*

Definição 4. *Dizemos que um espaço vetorial* V *sobre* K *é finitamente gerado se possui um conjunto gerador finito.*

Antes de prosseguirmos adotaremos que o espaço vetorial nulo possui como base o conjunto vazio, pois convencionamos que o conjunto vazio é linearmente independente e gera o espaço vetorial nulo. Além disso, convém observarmos que todo espaço vetorial finitamente gerado não nulo possui uma base e, mais ainda, que quaisquer duas bases desse espaço vetorial têm a mesma cardinalidade. Os resultados que contêm essas afirmações podem ser encontrados em (COELHO; LOURENCO, 2013, páginas 45, 48, 51-52 e 54-55, capítulo 2).

Definição 5. *Seja* V *um espaço vetorial sobre* K*. Se* V *admite uma base finita, então chamamos de dimensão de* V *e denotamos por* $dim_K V$*, o número de elementos de tal base. Caso contrário dizemos que a dimensão de* V *é infinita.*

A próxima definição será imprescindível quando estivermos tratando de códigos lineares no Capítulo 4.

Definição 6. *Seja* V *um espaço vetorial sobre um corpo* K*. Um subconjunto* W *de* V *é um subespaço vetorial de* V *quando a restrição das operações de* V *a* W *torna esse conjunto um espaço vetorial sobre* K*.*

Proposição 1. *Sejam* V *um espaço vetorial sobre* K *e* $W \subseteq V$ *um subconjunto. Então* W *é um subespaço de* V *se, e somente se, satisfaz as seguintes propriedades:*

(i) $\mathbf{0} \in W$*;*

(ii) Se $\mathbf{v}_1, \mathbf{v}_2 \in W$*, então* $\mathbf{v}_1 + \mathbf{v}_2 \in W$*;*

(iii) Se $\lambda \in K$ *e* $\mathbf{v} \in W$*, então* $\lambda \cdot \mathbf{v} \in W$*.*

Demonstração. Suponha que W é um subespaço vetorial de V, assim segue que W, com as operações de V, é um espaço vetorial sobre K. Logo as três propriedades são imediatamente satisfeitas.

Reciprocamente, se as três propriedades são satisfeitas, então pelo item (i) o elemento neutro da adição de V também está em W. Pelos itens (ii) e (iii), W é fechado pela adição e multiplicação por escalar e, por W ser um subconjunto de V, as demais propriedades necessárias

para que um conjunto seja um espaço vetorial também são satisfeitas. Consequentemente, W é um espaço vetorial sobre K com as operações de V e portanto é um subespaço vetorial de V. □

Observe que a proposição acima nos fornece uma maneira prática de verificar que um conjunto é um subespaço vetorial, sem a necessidade de verificarmos todas as propriedades que um conjunto deve satisfazer para ser espaço vetorial.

Definição 7. *Sejam W_1 e W_2 K-subespaços vetoriais do K-espaço vetorial V. Definimos a soma desses subespaços como*

$$W_1 + W_2 = \{\mathbf{w}_1 + \mathbf{w}_2;\ \mathbf{w}_1 \in W_1 \text{ e } \mathbf{w}_2 \in W_2\}.$$

Exemplo 1. *Se W_1 e W_2 são K-subespaços vetoriais do K-espaço vetorial V, então $W_1 + W_2$ é também um K-subespaço vetorial.*

De fato, temos que como W_1 e W_2 são subespaços vetoriais, então $\mathbf{0} \in W_1$ e $\mathbf{0} \in W_2$, logo $\mathbf{0} = \mathbf{0} + \mathbf{0} \in W_1 + W_2$. Além disso, para todo $\mathbf{v}_1 = \mathbf{w}_1 + \mathbf{w}_2 \in W_1 + W_2$ e $\mathbf{v}_2 = \mathbf{u}_1 + \mathbf{u}_2 \in W_1 + W_2$, temos

$$\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{u}_1 + \mathbf{u}_2 = \underbrace{\mathbf{w}_1 + \mathbf{u}_1}_{\in W_1} + \underbrace{\mathbf{w}_2 + \mathbf{u}_2}_{\in W_2},$$

logo $\mathbf{v}_1 + \mathbf{v}_2 \in W_1 + W_2$. Observe também que, para todo $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2 \in W_1 + W_2$ e para todo $\lambda \in K$, temos

$$\lambda\mathbf{v} = \underbrace{\lambda\mathbf{w}_1}_{\in W_1} + \underbrace{\lambda\mathbf{w}_2}_{\in W_2},$$

consequentemente, $\lambda\mathbf{v} \in W_1 + W_2$. Portanto, pela Proposição 1, $W_1 + W_2$ é um K-subespaço vetorial de V.

Definição 8. *Sejam W_1 e W_2 dois subespaços vetoriais de um espaço vetorial V. Dizemos que a soma $W_1 + W_2$ é direta quando $W_1 \cap W_2 = \{\mathbf{0}\}$ e a chamamos de soma direta de W_1 com W_2.*

Trataremos agora de uma classe específica de funções entre espaços vetoriais que obedecem as operações, já definidas, de espaços vetoriais.

Definição 9. *Sejam U e V espaços vetoriais sobre um corpo K. Uma função $T : U \to V$ é uma* transformação linear *quando*

(i) $T(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2)$, *para todos* $\mathbf{u}_1, \mathbf{u}_2 \in U$;

(ii) $T(\lambda\mathbf{u}) = \lambda T(\mathbf{u})$, *para todo* $\lambda \in K$ *e* $\mathbf{u} \in U$.

Note que, quando uma função T é uma transformação linear que leva de U em V, temos $T(\mathbf{0}_U) = \mathbf{0}_V$. De fato,

$$\mathbf{0}_V + T(\mathbf{0}_U) = T(\mathbf{0}_U) = T(\mathbf{0}_U + \mathbf{0}_U) = T(\mathbf{0}_U) + T(\mathbf{0}_U)$$

e, portanto, $T(\mathbf{0}_U) = \mathbf{0}_V$. Esta observação é de suma importância, pois se tivermos uma aplicação em que $T(\mathbf{0}) \neq \mathbf{0}$, então T não é uma transformação linear.

Outra observação importante a ser feita é quando uma transformação linear T de U em V é bijetiva. Neste caso, temos que T é um isomorfismo linear e dizemos que U é isomorfo a V. Denotamos esse isomorfismo por $U \cong V$.

Lema 1. *Sejam U e V espaços vetoriais sobre K. Então uma função $T : U \to V$ é uma transformação linear se, e somente se,*

$$T(\lambda\mathbf{u}_1 + \mathbf{u}_2) = \lambda T(\mathbf{u}_1) + T(\mathbf{u}_2), \quad \forall\, \mathbf{u}_1, \mathbf{u}_2 \in U,\ \forall\, \lambda \in K. \tag{2.1}$$

Demonstração. Suponha que a função T é uma transformação linear, logo valem as propriedades (i) e (ii) da Definição 9 e, portanto,

$$\begin{aligned} T(\lambda\mathbf{u}_1 + \mathbf{u}_2) &\overset{(i)}{=} T(\lambda\mathbf{u}_1) + T(\mathbf{u}_2) \\ &\overset{(ii)}{=} \lambda T(\mathbf{u}_1) + T(\mathbf{u}_2). \end{aligned}$$

Reciprocamente, suponha válida a igualdade (2.1). Como ela é válida para todo $\lambda \in K$, em particular vale para $\lambda = 1$, assim temos

$$T(\mathbf{u}_1 + \mathbf{u}_2) = T(1 \cdot \mathbf{u}_1 + \mathbf{u}_2) = 1 \cdot T(\mathbf{u}_1) + T(\mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2),$$

isto é, vale a propriedade (i) da Definição 9. Temos também que a igualdade é válida para todo $\mathbf{u}_2$ de V, em particular para $\mathbf{u}_2 = \mathbf{0}$, temos

$$T(\lambda\mathbf{u}_1) = T(\lambda\mathbf{u}_1 + \mathbf{0}) = \lambda T(\mathbf{u}_1) + T(\mathbf{0}) = \lambda T(\mathbf{u}_1),$$

ou seja, também vale a propriedade (ii) da Definição 9. Portanto, a função T é uma transformação linear. □

Lema 2. *Sejam U e V espaços vetoriais sobre K e $T : U \to V$ uma transformação linear. Então*

(i) $T(-\mathbf{u}) = -T(\mathbf{u})$, *para cada* $\mathbf{u} \in U$;

(ii) $T\left(\sum_{i=1}^{m} \alpha_i \mathbf{u}_i\right) = \sum_{i=1}^{m} \alpha_i T(\mathbf{u}_i)$, *em que* $\alpha_i \in K$ *e* $\mathbf{u}_i \in U$, *para* $i = 1, \ldots, m$.

Demonstração. (i) De fato,

$$T(-\mathbf{u}) = T(-\mathbf{u} + \mathbf{0}) = T((-1)\mathbf{u} + \mathbf{0}) = (-1)T(\mathbf{u}) + T(\mathbf{0}) = -T(\mathbf{u}) + \mathbf{0} = -T(\mathbf{u}).$$

(ii) Para demonstrar esta igualdade, basta usar sucessivamente as propriedades da definição de transformação linear.

□

Definição 10. *Sejam U e V dois espaços vetoriais sobre um corpo K e $T : U \to V$ uma transformação linear.*

(i) O conjunto $\{\mathbf{u} \in U : T(\mathbf{u}) = \mathbf{0}\}$ é chamado de núcleo de T e denotado por $Ker(T)$.

(ii) O conjunto $\{\mathbf{v} \in V : \exists\, \mathbf{u} \in U \text{ com } T(\mathbf{u}) = \mathbf{v}\}$ é chamado de imagem de T e denotado por $Im(T)$.

Os conjuntos do núcleo e da imagem de uma transformação linear $T : U \to V$ são subespaços vetoriais de U e V, respectivamente. A demonstração desse fato segue da Proposição 1 e o Lema 2.

Definição 11. *Seja M uma matriz $M_{m\times n}(K)$. O posto coluna (linha) de M é o número máximo de colunas (linhas) linearmente independentes de M.*

A partir da definição acima é possível determinar qual a dimensão do espaço gerado pelas linhas (espaço linha) e do espaço gerado pelas colunas (espaço coluna) de uma matriz M.

Lema 3. *Seja M uma matriz $M_{m\times n}(K)$, com $m \leq n$, tal que suas m linhas são linearmente independentes. Então m colunas de M são linearmente independentes.*

A demonstração desse Lema pode ser encontrada em (OLIVEIRA, 2018, páginas 1-2).

Teorema 1. (Teorema do Posto) *Sejam M uma matriz $M_{m\times n}(K)$ e p (o posto linha (coluna) de M) o número máximo de linhas (colunas) linearmente independentes de M. O número máximo de colunas (linhas) linearmente independentes de M é p.*

O Teorema do Posto desempenha um papel fundamental quando estamos buscando informações sobre as dimensões dos espaços linha e coluna de uma matriz M. Esse Teorema será essencial no Capítulo 6 e sua prova pode ser vista em (OLIVEIRA, 2018, página 2).

Teorema 2. *Sejam U e V dois espaços vetoriais sobre K com $dim_K U$ finita e $T : U \to V$ uma transformação linear. Então*

$$dim_K U = dim_K Ker(T) + dim_K Im(T). \tag{2.2}$$

A demonstração desse resultado pode ser encontrada em (COELHO; LOURENCO, 2013, páginas 87-88, capítulo 3).

Uma importante representação de uma transformação linear entre espaços vetoriais que nos será bastante útil nos próximos capítulos, é a sua representação matricial. Para entender melhor essa ideia, façamos a seguinte definição.

Definição 12. *Sejam U e V espaços vetoriais finitos sobre K de bases $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ e $\mathcal{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, respectivamente. Então a representação matricial de T com relação as bases $\mathcal{B}$ e $\mathcal{B}'$ e denotada por $[T]_{\mathcal{B},\mathcal{B}'}$, é dada por*

$$[T]_{\mathcal{B},\mathcal{B}'} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \ldots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \ldots & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m1} & \lambda_{m2} & \ldots & \lambda_{mn} \end{bmatrix},$$

em que $T(\mathbf{u}_j) = \lambda_{1j}\mathbf{v}_1 + \lambda_{2j}\mathbf{v}_2 + \cdots + \lambda_{mj}\mathbf{v}_m$, $j = 1, \ldots, n$ e $\lambda_{ij} \in K$, para todo $i = 1, \ldots, m$, $j = 1, \ldots, n$.

Agora iremos estudar uma importante ferramenta de álgebra linear, o *produto interno* de vetores de K^n, que nos permitirá definir o *complemento ortogonal* de um subespaço vetorial, estrutura essencial na construção do algoritmo de decodificação para códigos lineares que será visto no Capítulo 4. Para tal, sejam $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in K^n$. Consideremos o produto de $\mathbf{u}$ por $\mathbf{v}$, definido como

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + \ldots + u_nv_n,$$

e chamado de produto interno entre $\mathbf{u}$ e $\mathbf{v}$.

Proposição 2. *Sejam K um corpo finito e K^n o K-espaço vetorial. O produto interno definido acima tem as seguintes propriedades*

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$, *para todos $\mathbf{u}, \mathbf{v} \in K^n$ (simetria);*

(ii) $\langle \mathbf{u} + \lambda\mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \lambda\langle \mathbf{w}, \mathbf{v} \rangle$, *para todos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in K^n$ e para todo $\lambda \in K$ (bilinear).*

Demonstração. Dados $\mathbf{u}, \mathbf{v}, \mathbf{w} \in K^n$ e $\lambda \in K$, então

(i) $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + \ldots + u_nv_n = v_1u_1 + \ldots + v_nu_n = \langle \mathbf{v}, \mathbf{u} \rangle$.

(ii) $\langle \mathbf{u} + \lambda\mathbf{w}, \mathbf{v} \rangle = \langle (u_1 + \lambda w_1, \ldots, u_n + \lambda w_n), (v_1, \ldots, v_n) \rangle =$
$= u_1v_1 + \ldots u_nv_n + \lambda(w_1v_1 + \ldots + w_nv_n) = \langle \mathbf{u}, \mathbf{v} \rangle + \lambda\langle \mathbf{w}, \mathbf{v} \rangle$.

□

Convém observar que no item (ii) da Proposição 2 vale também a linearidade na segunda coordenada, isto é, $\langle \mathbf{u}, \mathbf{w} + \lambda\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \lambda\langle \mathbf{u}, \mathbf{v} \rangle$.

Definição 13. *Seja $V \subset K^n$ um K-subespaço vetorial. Definimos o* complemento ortogonal *de V como o subconjunto de K^n, dado por $V^\perp = \{\mathbf{v} \in V;\ \langle \mathbf{u}, \mathbf{v} \rangle = 0,\ \forall\, \mathbf{u} \in V\}$, em que $\langle \cdot, \cdot \rangle$ representa o produto interno usual de K^n.*

Lema 4. *Se $V \subset K^n$ é um K-subespaço vetorial, com uma base $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ e $V^\perp$ o complemento ortogonal de V, então $V^\perp$ é também um K-subespaço vetorial de K^n.*

Demonstração. Dados $\mathbf{u}, \mathbf{v} \in C^\perp$ e $\lambda \in K$, temos que

(i) $\mathbf{0} \in C^\perp$. De fato, se $\mathbf{v} = \mathbf{0}$, então $\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{0}, \mathbf{x} \rangle = 0$ para todo $\mathbf{x} \in C$, logo $\mathbf{0} \in C^\perp$.

(ii) $\mathbf{u} + \mathbf{v} \in C^\perp$. Com efeito, $\langle \mathbf{u} + \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{u}, \mathbf{x} \rangle + \langle \mathbf{v}, \mathbf{x} \rangle = 0 + 0 = 0$ para todo $\mathbf{x} \in C$, logo $\mathbf{u} + \mathbf{v} \in C^\perp$.

(iii) $\lambda\mathbf{v} \in C^\perp$, pois $\langle \lambda\mathbf{v}, \mathbf{x} \rangle = \lambda\langle \mathbf{v}, \mathbf{x} \rangle = \lambda \cdot 0 = 0$ para todo $\mathbf{x} \in C$, logo $\lambda\mathbf{v} \in C^\perp$.

Portanto, $C^\perp$ é um subespaço vetorial de K^n. □

Proposição 3. *Seja V um K-espaço vetorial de dimensão $n \geq 1$ com produto interno, e seja $W \subsetneq V$ um subespaço próprio de V. Então $V = W \oplus W^\perp$ e*

$$dim_K V = dim_K W + dim_K W^\perp.$$

Vale observar que a Proposição 3, cuja demonstração pode ser encontrada em (COELHO; LOURENCO, 2013, páginas 191-192, capítulo 6), será utilizada para provar a próxima proposição e necessária para estabelecer algumas relações entre os códigos lineares no Capítulo 4.

Proposição 4. *Sejam V um K-espaço vetorial de dimensão $n \geq 1$ com produto interno e $W \subset V$ um K-subespaço vetorial dimensão k, então $W^\perp$ é um K-subespaço vetorial de dimensão $n - k$.*

Demonstração. De fato, temos pela Proposição 3 que $V = W \oplus W^\perp$ e $dim_K V = dim_K W + dim_K W^\perp$, logo

$$dim_K W^\perp = dim_K V - dim_K W = n - k.$$

□

2.2 Tópicos da Teoria de Anéis

Nesta seção apresentaremos os resultados mais relevantes para esse trabalho a respeito da teoria de anéis. Focaremos aqui em ideais, corpos finitos e anéis quocientes (principalmente o anel de polinômios).

A partir da definição de anel dada em (HEFEZ, 2002), em que consideramos que todo anel é comutativo e com unidade, podemos pensar nos seguintes exemplos de anéis: o conjunto dos números inteiros $\mathbb{Z}$, dos números racionais $\mathbb{Q}$, dos números reais $\mathbb{R}$ e dos números complexos $\mathbb{C}$, com as operações binárias usuais + (adição) e $\cdot$ (multiplicação). No entanto, o conjunto dos números naturais $\mathbb{N}$ não é um anel, pois não possui a propriedade de elemento simétrico, já que $\mathbb{N}$ não possui um elemento neutro aditivo.

Temos também, a partir da definição de corpo dada em (HEFEZ, 2002), que os conjuntos $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ são exemplos de corpos. Porém, o conjunto dos inteiros $\mathbb{Z}$ não é um corpo, pois não possui a propriedade de que todos os seus elementos não nulos são invertíveis.

Quando estivermos nos referindo à multiplicação entre dois elementos a e b de um anel, adotaremos também a escrita por *justaposição* ab ao invés de $a \cdot b$.

Faremos a seguir uma importante definição sobre elementos específicos de um anel, que utilizaremos bastante ao longo desse texto.

Definição 14. *Seja A um anel. Dizemos que $a \in A \setminus \{0\}$, com a não invertível, é um elemento primo quando*

$$\forall\, b, c \in A, \quad a \mid b \cdot c \Rightarrow a \mid b \text{ ou } a \mid c.$$

Definição 15. *Seja K um corpo e L um subconjunto não vazio de K. Então L é chamado de subcorpo de K quando L é também um corpo com as operações de K.*

Proposição 5. *Sejam K um corpo e L um subconjunto de K. Temos que L é um subcorpo de K se, e somente se, as seguintes condições são satisfeitas:*

(i) $a - b \in L$ *para todos* $a, b \in L$*;*

(ii) $ab^{-1} \in L$ *para todos* $a, b \in L$*, com* $b \neq 0$*.*

Demonstração. Suponha que L seja um subcorpo de K, logo L é um corpo com as operações de K. Portanto as propriedades (i) e (ii) são satisfeitas.

Reciprocamente, se valem as propriedades (i) e (ii), então pelo item (i) L é fechado pela adição de K e o elemento neutro da adição de K também está em L. Pelo item (ii), L é fechado

pela multiplicação de K e, por L ser um subconjunto de K, as demais propriedades da definição de corpo também são satisfeitas. Portanto L é um subcorpo de K. □

Definição 16. *Seja I um subconjunto do anel A. Dizemos que I é um ideal de A quando as condições abaixo são satisfeitas:*

(i) $I \neq \emptyset$;

(ii) $a - b \in I$, *para todos* $a, b \in I$;

(iii) $ca \in I$, *para todo* $a \in I$ *e para todo* $c \in A$.

Definição 17. *Seja I um ideal de A, com $I \neq A$. Dizemos que I é um ideal maximal de A sempre que existir um J, ideal de A, tal que $I \subset J \subset A$ implicar em $I = J$ ou $J = A$.*

Definição 18. *Se $a \in A$, então o conjunto*

$$I = I(a) = \{ca;\ c \in A\}, \tag{2.3}$$

é um ideal de A, chamado de ideal principal *gerado por a.*

Podemos generalizar a Definição 18 da seguinte forma: se $a_1, a_2, \ldots, a_n \in A$, então o conjunto

$$I = I(a_1, a_2, \ldots, a_n) = \{c_1a_1 + \cdots + c_na_n;\ c_1, \ldots, c_n \in A\}$$

é um ideal de A. Os elementos $a_1, a_2, \ldots, a_n \in A$ são chamados de geradores de I.

Definição 19. *Sejam A um anel e I um ideal de A. Dados elementos $a, b \in A$, dizemos que a é* congruente *a b módulo I se, e somente se, $a - b \in I$. Nesse caso, escrevemos*

$$a \equiv b\ (mod\ I).$$

A relação de congruência possui as seguintes propriedades:

Proposição 6. *Sejam a, b, c elementos quaisquer de A e I um ideal de A.*

(i) $a \equiv a\ (mod\ I)$;

(ii) Se $a \equiv b\ (mod\ I)$, *então* $b \equiv a\ (mod\ I)$;

(iii) Se $a \equiv b\ (mod\ I)$ *e* $b \equiv c\ (mod\ I)$, *então* $a \equiv c\ (mod\ I)$.

Demonstração. (i) De fato, temos que $a - a = 0 \in I$. Logo, pela Definição 19 $a \equiv a\ (mod\ I)$.

(ii) Pela Definição 19, $a \equiv b\,(mod\,I)$ implica em $a - b \in I$. Logo, pela Definição 16 $b - a = -(a - b) \in I$ e, portanto, $b \equiv a\,(mod\,I)$.

(iii) Para provar esse item devemos mostrar que $a - c \in I$. Como $a - b \in I$ e $b - c \in I$, temos que $a - b + b - c \in I$. Logo $a - c \in I$ e, portanto, $a \equiv c\,(mod\,I)$.

□

Uma relação que satisfaz os itens (i), (ii), e (iii) é chamada de relação de equivalência. Deste modo a relação de congruência modulo m é uma relação de equivalência sobre I.

Definição 20. *Sejam A um anel, I um ideal de A e $a \in A$. Definimos a* classe residual *de a módulo I por*

$$[a] = a + I = \{a + x;\ x \in I\}.$$

Proposição 7. *Seja I um ideal de A. O conjunto*

$$A/I = \{[a] = a + I;\ a \in A\}.$$

é um anel quando munido das operações: $[a] + [b] = [a + b]$, $[a] \cdot [b] = [a \cdot b] \in A/I$, para todos $[a], [b] \in A/I$.

A prova deste fato pode ser vista na bibliografia (BHATTACHARYA; JAIN; NAGPAUL, 1994, páginas 183-184, capítulo 10).

Observação 1. Sejam $a, b \in A$, temos que $[a] = [b]$ se, e somente se $a \equiv b\,(mod\,I)$.

Definição 21. *Seja I um ideal de A. Então o anel A/I é chamado de anel quociente de A módulo I.*

Exemplo 2. *Seja $I(n) = \{an;\ a \in \mathbb{Z}\}$ um ideal de $\mathbb{Z}$. Se $n \neq 0$, então o anel quociente $\mathbb{Z}/I(n) = \{0, 1, \ldots, n-1\} \cong \mathbb{Z}_n$, ou seja, é isomorfo ao conjunto das classes residuais módulo n.*

Definição 22. *Sejam A e B anéis. Uma função $f : A \to B$ será chamada de homomorfismo quando, para todos $a, a' \in A$ valem:*

(i) $f(a + a') = f(a) + f(a')$;

(ii) $f(a \cdot a') = f(a) \cdot f(a')$.

Proposição 8. *Seja $f : A \to B$ um homomorfismo entre os anéis A e B e sejam $a, a' \in A$. Valem*

(i) $f(0) = 0$;

(ii) $f(-a) = -f(a)$;

(iii) $f(1) = 1$;

(iv) *Se* $a \in A$ *é invertível e* $f(1) = 1$, *então* $f(a)$ *é invertível e* $f(a^{-1}) = (f(a))^{-1}$;

(v) *Se* f *é bijetora, então a função* f^{-1}, *inversa de* f, *é um homomorfismo;*

(vi) *Se* A *e* B *são corpos e* $f(1) = 1$, *então* f *é injetora e* $Im(f)$ *é um subcorpo de* B.

Demonstração. (i) De fato,

$$f(0) = f(0+0) = f(0) + f(0) \Rightarrow 0 = f(0).$$

(ii) Temos que

$$0 = f(0) = f(a+(-a)) = f(a)+f(-a) \Rightarrow -f(a) = -f(a)+f(a)+f(-a) \Rightarrow -f(a) = f(-a).$$

(iii) Temos que existe $f(a) \in B$ tal que

$$1 \cdot f(a) = f(a) = f(a \cdot 1) = f(a) \cdot f(1) \Rightarrow 1 = f(1).$$

(iv) Se a é invertível, então

$$1 \overset{(iii)}{=} f(1) = f(aa^{-1}) = f(a)f(a^{-1}),$$

logo $f(a)$ é invertível e $f(a^{-1})$ é o inverso de $f(a)$, isto é, $f(a^{-1}) = (f(a))^{-1}$.

(v) Devemos mostrar que f^{-1} atende às duas condições da Definição 22. Para demonstrar a primeira e segunda condições, considere $b, b' \in B$, tais que $a = f^{-1}(b)$ e $a' = f^{-1}(b')$. Deste modo, $f(a) = f(f^{-1}(b)) = b$ e $f(a') = f(f^{-1}(b')) = b'$. Daí, segue que

$$f^{-1}(b+b') = f^{-1}(f(a)+f(a')) = f^{-1}(f(a+a')) = a+a' = f^{-1}(b)+f^{-1}(b');$$

$$f^{-1}(bb') = f^{-1}(f(a)f(a')) = f^{-1}(f(aa')) = aa' = f^{-1}(b)f^{-1}(b').$$

Portanto, f^{-1} é um homomorfismo.

(vi) Suponha que A e B são corpos. Suponha também, que $f(a) = f(a')$, daí pelo item (ii) e por f ser um homomorfismo,

$$f(a-a') = f(a) - f(a') = 0.$$

No entanto, se $a \neq a'$, então $a - a' \neq 0$, isto é, $a - a'$ e invertível e, pelo item (iv), segue que $f(a-a')$ é também invertível, contradizendo $f(a) = f(a')$. Logo $a = a'$ e, portanto,

f é uma função injetora. Por outro lado, para demonstrar que $Im(f)$ é um subcorpo, devemos provar que $Im(f)$ satisfaz as condições da Proposição 5. Deste modo, suponha que $b = f(a), b' = f(a') \in Im(f)$. Daí,

$$b - b' = f(a) - f(a') = f(a - a') \in Im(f);$$

$$b(b')^{-1} = f(a)f((a')^{-1}) = f(a(a')^{-1}) \in Im(f).$$

Portanto, $Im(f)$ é subcorpo de B.

□

2.3 O Anel de Polinômios

Nesta seção, introduziremos os polinômios em uma incógnita e também alguns resultados de polinômios de várias incógnitas. O estudo que faremos aqui serve como base para construção dos códigos cíclicos e BCH. Focaremos em apresentar os principais resultados necessários para este trabalho, relativos a polinômios sobre corpos finitos, utilizando como base a bibliografia (HEFEZ; VILLELA, 2008).

Seja A um anel e X uma incógnita. Chamaremos de um polinômio com coeficientes em A e incógnita X a expressão formal

$$p(X) = \sum_{i=0}^{n} a_i X^i = a_0 + a_1 X + \ldots + a_n X^n, \tag{2.4}$$

em que a_i, para todo $i = 1, \ldots, n$, são elementos do anel A, chamados de coeficientes de $p(X)$ e n é um inteiro não negativo. Denotaremos o conjunto de todos os polinômios em uma incógnita X com coeficientes no anel A, por $A[X]$.

Definição 23. *Dois polinômios $p(X), q(X) \in A[X]$ são iguais se, e somente se, seus coeficientes correspondentes são todos iguais em A.*

Definição 24. *Quando $p(X) = a_0 + a_1 X + \ldots + a_n X^n \in A[X]$ é tal que $a_0 = a_1 = \ldots = a_n = 0$, dizemos que $p(X)$ é o polinômio nulo em $A[X]$ e o denotamos por 0.*

Definição 25. *Dizemos que o grau de um polinômio não nulo $p(X) \in A[X]$ é igual a n, quando $p(X) = a_0 + a_1 X + \ldots + a_n X^n$ é tal que $a_n \neq 0$ e $a_j = 0$ para todo $j > n$ e o denotamos por $gr(p(X))$.*

Definição 26. *Quando $p(X) = a_0 + a_1 X + \ldots + a_n X^n \in A[X]$, em que $a_i = 0$ para todo $i \geq 1$, isto é, $p(X) = a_0$, dizemos que $p(X)$ é um polinômio constante.*

Definição 27. *Sejam* $p(X) = \sum_{i=0}^{n} a_i X^i$, $q(X) = \sum_{j=0}^{m} b_j X^j \in A[X]$, *definimos as operações:*

$$\begin{array}{rccc} +: & A[X]\times A[X] & \to & A[X] \\ & (p(X),q(X)) & \mapsto & p(X)+q(X) \end{array} \qquad \begin{array}{rccc} \cdot: & A[X]\times A[X] & \to & A[X] \\ & (p(X),q(X)) & \mapsto & p(X)\cdot q(X) \end{array},$$

por

$$p(X)+q(X) = \sum_{k=0}^{\max\{n,m\}} (a_k + b_k)X^k$$

e

$$p(X)\cdot q(X) = \sum_{i=0}^{n+m} c_i X^i,$$

em que $c_i = \sum_{j=0}^{i} a_j \cdot b_{i-j}$.

Teorema 3. *O conjunto* $A[X]$ *é um anel com as operações definidas acima.*

A demonstração segue ao verificar as propriedades da definição de anel e pode ser encontrada em (HEFEZ, 2002, páginas 23-24, capítulo 2).

Proposição 9. *O anel* A *é isomorfo ao conjunto dos polinômios constantes de* $A[X]$.

Demonstração. Defina uma função Φ por

$$\begin{array}{rccc} \Phi: & A & \to & \tilde{A} \subset A[X] \\ & a & \mapsto & a + 0X + \ldots + 0X^n \end{array},$$

em que $\tilde{A}$ é conjunto dos polinômios constantes de $A[X]$. Sejam $a, b \in A$ tais que $a + b = c$ e $a \cdot b = d$, então

$$\begin{aligned} \Phi(a+b) &= \Phi(c) \\ &= c + 0X + \ldots + 0X^n \\ &= a + b + 0X + \ldots + 0X^n \\ &= a + 0X + \ldots + 0X^n + b + 0X + \ldots + 0X^n \\ &= \Phi(a) + \Phi(b) \end{aligned}$$

e

$$\begin{aligned} \Phi(a\cdot b) &= \Phi(d) \\ &= d + 0X + \ldots + 0X^n \\ &= a\cdot b + 0X + \ldots + 0X^n \\ &= (a + 0X + \ldots + 0X^n)\cdot(b + 0X + \ldots + 0X^n) \\ &= \Phi(a)\cdot\Phi(b). \end{aligned}$$

Logo, a função Φ é um homomorfismo.

Agora, sendo $a, b \in A$, temos que

$$\Phi(a) = \Phi(b) \Leftrightarrow a + 0X + \ldots = b + 0X + \ldots \Leftrightarrow a = b,$$

o que prova a injetividade de Φ. Além disso, dado $a + 0X + \ldots \in \tilde{A}$, temos que $a \in A$ tal que $\Phi(a) = a + 0X + \ldots$, implica que Φ é também sobrejetiva e, portanto, bijetiva. Consequentemente, Φ é um isomorfismo de A em $\tilde{A}$. □

Observe que, como o anel de polinômios $A[X]$ contém uma cópia isomorfa a A, usaremos o abuso de linguagem que $A[X]$ contém A. Mais ainda, se dado $k \in A$ e $p(X) \in A[X]$ o seu produto terá como resultado o mesmo que o produto do polinômio constante $k \in A[X]$ pelo polinômio $p(X)$. No entanto, esse tipo de produto nos permite formalizar a seguinte operação (que por abuso notacional, usamos o mesmo símbolo do produto entre polinômios) denominada produto por escalar

$$\begin{array}{rcl} \cdot : A \times A[X] & \to & A[X] \\ (k,\ p(X)) & \mapsto & k \cdot p(X) \text{ (o mesmo como produto de polinômios).} \end{array}$$

E com isso, no caso de A ser um corpo, o anel de polinômios munidos da sua adição usual dada na Definição 27 e a operação de produto por escalar, constituem um A-espaço vetorial (mais geralmente, se A não é um corpo, tal estrutura é um A-módulo). Vale observar que o A-espaço vetorial $A[X]$ não tem dimensão finita.

Proposição 10. *Seja A um domínio de integridade.*

(i) Se $p(X), q(X) \in A[X] \setminus \{0\}$, então $gr(p(X) \cdot q(X)) = gr(p(X)) + gr(q(X))$;

(ii) $A[X]$ é um domínio de integridade;

(iii) Dado $p(X) \in A[X]$, $p(X)$ é invertível em $A[X]$ se, e somente se, $p(X)$ é invertível em A.

Demonstração. (i) Sejam $p(X) = \sum_{i=0}^{n} a_i X^i$, $q(X) = \sum_{j=0}^{m} b_j X^j \in A[X]$ dois polinômios de graus n e m, respectivamente. Temos que

$$p(X) \cdot q(X) = \sum_{i=0}^{n+m} c_i X^i.$$

Desde que A seja um domínio de integridade, segue que $c_{n+m} = a_n \cdot b_m \neq 0$, pois $a_n \neq 0$ e $b_m \neq 0$, logo

$$gr(p(X) \cdot q(X)) = gr\left(\sum_{i=0}^{n+m} c_i X^i\right) = n + m = gr(p(X)) + gr(q(X)).$$

(ii) Sejam $p(X), q(X) \in A[X]$ e suponha que

$$p(X) \cdot q(X) = 0,$$

então, $p(X) \cdot q(X)$ não possui grau, pois é o polinômio nulo. Agora suponha por absurdo que $p(X)$ e $q(X)$ são polinômios não nulos. Logo, segue do item (i) que

$$gr(p(X) \cdot q(X)) = gr(p(X)) + gr(q(X)) \geq 0,$$

o que é uma contradição, a menos que $p(X) = 0$ ou $q(X) = 0$. Portanto $A[X]$ é um domínio de integridade.

(iii) Sendo $p(X) \in A[X]$ um elemento invertível em $A[X]$, temos que existe $q(X) \in A[X]$ tal que $p(X) \cdot q(X) = 1$. Como o elemento nulo não possui inverso, segue que $p(X) \neq 0$ e $q(X) \neq 0$, assim, pelo item (i)

$$0 = gr(1) = gr(p(X) \cdot q(X)) = gr(p(X)) + gr(q(X)).$$

Logo $gr(p(X)) = 0$ e $gr(q(X)) = 0$, o que implica em $p(X) = a$ e $q(X) = b$, com a e b constantes não nulas. Além disso, pela Proposição 9 temos que $p(X), q(X) \in A$ com $p(X) \cdot q(X) = 1$ em A. Portanto $p(X)$ é invertível em A.

Reciprocamente, se $a \in A$ é um elemento invertível, então existe $b \in A$ tal que $a \cdot b = 1$, logo $a \neq 0, b \neq 0$. Pela Proposição 9, podemos escrever $a = p(X)$, $b = q(X)$ sendo dois polinômios constantes, ou seja, $a, b \in A[X]$. Com isso,

$$a \cdot b = 1 \Leftrightarrow p(X) \cdot q(X) = 1$$

em $A[X]$. Portanto $p(X) = a$ é invertível em $A[X]$.

□

Se no lugar do anel A, na Proposição 10, tomarmos um corpo K, temos que $K[X]$ também é um domínio de integridade.

Definição 28. *Sejam $p(X), q(X) \in A[X]$, dizemos que $p(X)$ e $q(X)$ são associados quando existe uma constante invertível $c \in A$, tal que $p(X) = cq(X)$.*

Vale ressaltar que, quando a Definição 28 é dada sobre um corpo K, basta exigirmos que a constante $c \in K$ seja não nula, uma vez que todo elemento não nulo de K é invertível.

Definição 29. *Dizemos que um polinômio $p(X) \in K[X]$ de grau n é um* polinômio mônico *(ou simplesmente* mônico*) quando $a_n = 1$.*

Proposição 11. *Todo polinômio não nulo é associado a apenas um polinômio mônico.*

Demonstração. (Existência) Seja $p(X) = a_0 + a_1X + \ldots + a_nX^n \in K[X]$, com $a_n \neq 0$, temos que $p(X) = a_n(a_n^{-1} \cdot a_0 + a_n^{-1} \cdot a_1X + \ldots + X^n)$, em que $a_n^{-1} \cdot a_0 + a_n^{-1} \cdot a_1X + \ldots + X^n$ é um polinômio mônico. Com isso, garantimos a existência do polinômio mônico associado a $p(X)$. (Unicidade) Seja $p(X) \in K[X]$ não nulo e suponha que $p(X)$ é associado aos polinômios mônicos $f(X)$ e $g(X)$, com $f(X) \neq g(X)$, isto é, existem $a, b \in K$ não nulos tais que

$$p(X) = af(X) \quad \text{e} \quad p(X) = bg(X).$$

Assim,

$$f(X) = a^{-1}bg(X).$$

Como $f(X) \neq g(X)$, segue que $a^{-1}b \neq 1$. Logo $f(X)$ não é mônico, o que é um absurdo. Portanto $p(X)$ é associado a um único polinômio mônico. □

Observe que esse resultado é apenas válido para anéis de polinômios com coeficientes em um corpo. Por exemplo, o polinômio $2X + 1$ não é associado a nenhum mônico em $\mathbb{Z}[X]$, pois como vimos no resultado anterior, o único candidato a ser o polinômio mônico associado a $2X+1$ é o polinômio $X+2^{-1}$ que, por sua vez, não possui coeficientes em $\mathbb{Z}$ e, consequentemente, não é um polinômio de $\mathbb{Z}[X]$.

Um importante resultado para esse capítulo é o *algoritmo da divisão polinomial*, em virtude de ser uma ferramenta de bastante utilidade na divisão de polinômios como veremos ao longo de todo o texto.

Teorema 4. (Algoritmo da Divisão Polinomial) *Sejam $f(X), g(X) \in K[X]$ com $g(X) \neq 0$. Então existem únicos polinômios $q(X), r(X) \in K[X]$ tais que*

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

com $r(X) = 0$ ou $gr(r(X)) < gr(g(X))$.

Demonstração. (Existência) Sejam $f(X) = \sum_{i=0}^{n} a_iX^i, g(X) = \sum_{j=0}^{m} b_jX^j \in K[X]$ com $a_n \neq 0$ e $b_m \neq 0$. Primeiramente vamos mostrar a existência dos polinômios $q(X)$ e $r(X)$. Se $f(X)$ for identicamente nulo, então basta tomar $q(X) = 0$ e $r(X) = 0$ para obtermos $f(X) = q(X)g(X) + r(X) = 0$. Se $f(X)$ não é o polinômio nulo, então $gr(f(X)) = n \geq 0$. Suponha que $n < m$, neste caso basta tomar $q(X) = 0$ e $r(X) = f(X)$, pois do contrário

$$n = gr(f(X)) = gr(q(X)g(X) + r(X)) \geq m,$$

o que seria um absurdo. Considere então $n \geq m$, vamos mostrar por indução finita sobre o grau de $f(X)$ a existência de $q(X)$ e $r(X)$. Defina

$$f_1(X) = f(X) - a_n b_m^{-1} X^{n-m} g(X) \in K[X]. \tag{2.5}$$

Podemos reescrever (2.5) como

$$f(X) = a_n b_m^{-1} X^{n-m} g(X) + f_1(X), \tag{2.6}$$

em que $gr(f_1(X)) < gr(f(X))$. Se $n = 0$, então $m = 0$ e assim

$$f(X) = a_0 b_0^{-1} g(X) + f_1(X).$$

Daí, pelo item (i) da Proposição 10, $g(X)$ é constante, donde $f_1(X) = 0$. Logo, $f(X) = a_0 b_0^{-1} g(X)$ e, portanto, basta tomar $q(X) = a_0 b_0^{-1}$ e $r(X) = f_1(X) = 0$. Suponha agora que o teorema seja válido para todo polinômio de grau menor que n. Segue da hipótese de indução que existem $q_1(X)$ e $r_1(X)$, tais que

$$f_1(X) = q_1(X)g(X) + r_1(X), \tag{2.7}$$

com $r_1(X) = 0$ ou $gr(r_1(X)) < gr(g(X))$. Assim, pelas igualdades (2.5) e (2.7)

$$f(X) = (a_n b_m^{-1} X^{n-m} + q_1(X))g(X) + r_1(X), \tag{2.8}$$

com $r_1(X) = 0$ ou $gr(r_1(X)) < gr(g(X))$. Portanto, tomando $q(X) = a_n b_m^{-1} X^{n-m} + q_1(X)$ e $r(X) = r_1(X)$ garantimos a existência de $q(X)$ e $r(X)$ de modo que $f(X) = q(X) \cdot g(X) + r(X)$, com $r(X) = 0$ ou $gr(r(X)) < gr(g(X))$.
(Unicidade) Suponha que existem $q_1(X), q_2(X), r_1(X), r_2(X) \in K[X]$ tais que

$$f(X) = q_1(X)g(X) + r_1(X) = q_2(X)g(X) + r_2(X), \tag{2.9}$$

com $r_1(X) = 0$ ou $gr(r_1(X)) < gr(g(X))$ e $r_2(X) = 0$ ou $gr(r_2(X)) < gr(g(X))$. Deste modo,

$$(q_1(X) - q_2(X))g(X) + (r_1(X) - r_2(X)) = 0. \tag{2.10}$$

Com isso,

$$gr((q_1(X) - q_2(X))g(X)) = gr((r_1(X) - r_2(X))),$$

logo $gr((r_1(X) - r_2(X))) \geq gr(g(X))$, o que é uma contradição, a menos que os polinômios $(q_1(X) - q_2(X))g(X)$ e $r_1(X) - r_2(X)$ sejam nulos, e como $g(X)$ é não nulo, temos

$$q_1(X) = q_2(X) \quad \text{e} \quad r_1(X) = r_2(X).$$

□

Definição 30. *Sejam $f(X), g(X) \in K[X]$ não simultaneamente nulos. Dizemos que $m(X) \in K[X]$ é o* mínimo múltiplo comum (mmc) *de $f(X)$ e $g(X)$ quando se verificam as seguintes condições:*

(i) $f(X) \mid m(X)$ *e* $g(X) \mid m(X)$*;*

(ii) Se existe $m'(X) \in K[X]$ *tal que* $f(X) \mid m'(X)$ *e* $g(X) \mid m'(X)$*, então* $m(X) \mid m'(X)$*.*

Definição 31. *Sejam $f(X), g(X) \in K[X]$ não simultaneamente nulos. Dizemos que $d(X) \in K[X]$ é o* máximo divisor comum (mdc) *de $f(X)$ e $g(X)$ quando se verificam as seguintes condições:*

(i) $d(X) \mid f(X)$ *e* $d(X) \mid g(X)$*;*

(ii) Se existe $d'(X) \in K[X]$ *tal que* $d'(X) \mid f(X)$ *e* $d'(X) \mid g(X)$*, então* $d'(X) \mid d(X)$*.*

Teorema 5. (Existênca do mdc) *Sejam $f(X), g(X) \in K[X]$ não simultaneamente nulos. Então existe o mdc de $f(X)$ e $g(X)$.*

Demonstração. Temos que se $f(X) = g(X) = 0$, então $mdc(f(X), g(X)) = 0$. Sejam $f(X), g(X) \in K[X]$ com $g(X) \neq 0$, então pelo Teorema 4 existem únicos $q_1(X)$ e $r_1(X)$ tais que

$$f(X) = q_1(X)g(X) + r_1(X), \quad \text{com } r_1(X) = 0 \text{ ou } gr(r_1(X)) < gr(g(X)). \tag{2.11}$$

Se $r_1(X) = 0$, então $f(X) = q_1(X)g(X)$, ou seja, $g(X) \mid f(X)$, logo $g(X)$ atende a condição (i). Por outro lado, se existe $d'(X) \in K[X]$ tal que $d'(X) \mid f(X)$ e $d'(X) \mid g(X)$, segue imediatamente que $g(X)$ satisfaz a condição (ii).
Se $r_1(X) \neq 0$, então pelo Teorema 4 existem únicos $q_2(X)$ e $r_2(X)$ tais que

$$g(X) = q_2(X)r_1(X) + r_2(X), \quad \text{com } r_2(X) = 0 \text{ ou } gr(r_2(X)) < gr(r_1(X)). \tag{2.12}$$

Se $r_2(X) = 0$, então $r_1(X) \mid g(X)$ e, como

$$f(X) = q_1(X)g(X) + r_1(X) = q_1(X)(q_2(X)r_1(X)) + r_1(X) = (q_1(X)q_2(X) + 1)r_1(X),$$

segue que $r_1(X) \mid f(X)$, logo $r_1(X)$ atende a condição (i). Em contrapartida, se existe $d'(X) \in K[X]$ tal que $d'(X) \mid g(X)$ e $d'(X) \mid f(X)$, então $d'(X) \mid (f(X) - g(X))$ e, como $r_1(X) = f(X) - q_1(X)g(X)$, segue que $d'(X) \mid r_1(X)$. Desse modo, $r_1(X)$ satisfaz a condição (ii). Se $r_2(X) \neq 0$, novamente pelo Teorema 4, existem únicos $q_3(X)$ e $r_3(X)$ tais que

$$r_1(X) = q_3(X)r_2(X) + r_3(X), \quad \text{com } r_3(X) = 0 \text{ ou } gr(r_3(X)) < gr(r_2(X)). \tag{2.13}$$

Esse processo será repetido da mesma maneira até encontrarmos um polinômio $r_{n+1}(X) \in K[X]$, com $n \in \mathbb{N}$, de modo que $r_{n+1}(X) = 0$, ou seja,

$$\vdots$$

$$r_2(X) = q_4(X)r_3(X) + r_4(X), \text{ com } r_4(X) = 0 \text{ ou } gr(r_4(X)) < gr(r_3(X)) \quad (2.14)$$
$$r_3(X) = q_5(X)r_4(X) + r_5(X), \text{ com } r_5(X) = 0 \text{ ou } gr(r_5(X)) < gr(r_4(X)) \quad (2.15)$$

$$\vdots$$

$$r_{n-2}(X) = q_n(X)r_{n-1}(X) + r_n(X), \text{ com } r_n(X) = 0 \text{ ou } gr(r_n(X)) < gr(r_{n-1}(X)) \quad (2.16)$$
$$r_{n-1}(X) = q_{n+1}(X)r_n(X). \quad (2.17)$$

Ora, se $r_{n+1}(X) = 0$, de maneira análoga aos passos anteriores, obtemos que $r_n(X) \mid g(X)$ e $r_n(X) \mid f(X)$, assim segue que $r_n(X)$ atende a condição (i) e além disso, se existe $d'(X) \in K[X]$ tal que $d'(X) \mid g(X)$ e $d'(X) \mid f(X)$, então $d'(X) \mid r_n(X)$, atendendo a condição (ii). Portanto $r_n(X) = mdc(f(X), g(X))$.

□

Vale ressaltar que sempre existirá um n natural tal que $r_{n+1}(X) = 0$, pois caso contrário teríamos uma sequência infinita de números inteiros não negativos, tais que

$$gr(g(X)) > gr(r_1(X)) > gr(r_2(X)) > \ldots \geq 0,$$

o que é um absurdo.

Corolário 1. *Dados quaisquer $f(X)$ e $g(X)$ em $K[X]$ não simultaneamente nulos, existe um único mdc mônico de $f(X)$ e $g(X)$.*

Demonstração. (Existência) O Teorema 5 assegura a existência do mdc entre $f(X)$ e $g(X)$, digamos $d(X)$. Como pela Proposição 11, um polinômio não nulo é associado a um único polinômio mônico, então este é um mdc mônico de $f(X)$ e $g(X)$.
(Unicidade) Suponha que existem $d_1(X), d_2(X) \in K[X]$ mônicos, tais que

$$d_1(X) = mdc(f(X), g(X)) \quad \text{e} \quad d_2(X) = mdc(f(X), g(X)). \quad (2.18)$$

Temos pelo item (ii) do Teorema 5, que $d_1(X) \mid d_2(X)$ e $d_2(X) \mid d_1(X)$, daí

$$d_1(X) = p_1(X)d_2(X), \quad \text{com } p_1(X) \in K[X]$$

e

$$d_2(X) = p_2(X)d_1(X), \quad \text{com } p_2(X) \in K[X],$$

ou seja,

$$d_1(X) = p_1(X)p_2(X)d_1(X),$$

em que $gr(d_1(X)) = gr(p_1(X)p_2(X)d_1(X))$. Pelo item (i) da Proposição 10,

$$gr(d_1(X)) = gr(p_1(X)p_2(X)d_1(X)) = gr(p_1(X)p_2(X)) + gr(d_1(X)),$$

o que implica em $gr(p_1(X)p_2(X)) = 0$. Assim,

$$gr(p_1(X)) = gr(p_2(X)) = 0,$$

isto é, $p_1(X)$ é constante, digamos $p_1(X) = c \in K$. Logo $d_1(X) = cd_2(X)$ e, como $d_1(X)$ e $d_2(X)$ são mônicos, segue que $c = 1$ e portanto $d_1(X) = d_2(X)$. □

De modo análogo ao *mdc*, temos os resultados que garantem a existência e unicidade do *mmc* mônico. Com isso em mente, quando falarmos do *mdc* ou *mmc* de polinômios, estaremos nos referindo ao *mdc* e *mmc* mônicos.

Definição 32. *Um polinômio não nulo $p(X)$, com grau maior ou igual a 1, é dito* irredutível *em $K[X]$ (ou* irredutível *sobre K) quando toda vez que $p(X) = f(X)g(X)$, com $f(X), g(X) \in K[X]$, implicar em $f(X)$ ou $g(X)$ invertível em K. Se $p(X)$ não é irredutível em $K[X]$ dizemos que ele redutível em $K[X]$.*

Proposição 12. *Todo polinômio de grau 1 é irredutível em $K[X]$.*

Demonstração. De fato, se $p(X) = aX + b$ é redutível em $K[X]$, então existem $f(X), g(X)$ em $K[X]$ tais que

$$aX + b = f(X)g(X).$$

Temos pela Definição 25 que

$$1 = gr(aX + b) = gr(f(X)g(X)) = gr(f(X)) + gr(g(X)).$$

Note que $gr(f(X)) + gr(g(X)) = 1$ implica necessariamente em uma das duas situações:

- $gr(f(X)) = 1 \quad \text{e} \quad gr(g(X)) = 0$, donde $g(X)$ invertível, ou;
- $gr(g(X)) = 1 \quad \text{e} \quad gr(f(X)) = 0$, donde $f(X)$ invertível.

Portanto, pela Definição 32, segue que $p(X) = aX + b$ é irredutível em $K[X]$. □

Definição 33. *Dizemos que dois polinômios são primos entre si quando seu mdc é igual a* 1.

Teorema 6. *Sejam* $f(X), g(X)$ *e* $d(X) \in K[X]$ *tais que* $d(X) = mdc(f(X), g(X))$*, então existem* $t(X), s(X) \in K[X]$ *tais que*

$$d(X) = s(X)f(X) + t(X)g(X). \tag{2.19}$$

Demonstração. Primeiramente, mostremos por indução finita sobre n que os restos não nulos das divisões sucessivas do Teorema 5 podem ser escritos como

$$r_n(X) = s(X)f(X) + t(X)g(X). \tag{2.20}$$

Temos pelo Teorema 5 que, se $n = 2$, então pela equação (2.12)

$$\begin{aligned} r_2(X) &= g(X) - q_2(X)r_1(X) \\ &= g(X) - q_2(X)[f(X) - q_1(X)g(X)] \\ &= \underbrace{(-q_2(X))}_{s_1(X)\in K[X]} f(X) + \underbrace{(1 - q_1(X)q_2(X))}_{t_1(X)\in K[X]} g(X). \end{aligned}$$

Suponha que o resultado seja válido para todo $n \leq k-1$ e provemos que também vale para $n = k$. Temos pela equação (2.16) do Teorema 5 que

$$r_k(X) = r_{k-2}(X) - q_k(X)r_{k-1}(X).$$

Logo, como o resultado vale para todo $n \leq k-1$, segue que existem $s'(X), s''(X), t'(X)$ e $t''(X)$ em $K[X]$ tais que

$$\begin{aligned} r_k(X) &= s''(X)f(X) + t''(X)g(X) - q_k(X)[s'(X)f(X) + t'(X)g(X)] \\ &= \underbrace{[s''(X) - q_k(X)s'(X)]}_{s(X)\in K[X]} f(X) + \underbrace{[t''(X) - q_k(X)t'(X)]}_{t(X)\in K[X]} g(X), \end{aligned}$$

ou seja, $r_k(X) = s(X)f(X) + t(X)g(X)$.

Observe que, pelo Teorema 5, o último resto não nulo das divisões sucessivas realizadas no Teorema 5 é o mdc de $f(X)$ e $g(X)$, ou seja, $r_n(X) = d(X) = mdc(f(X), g(X))$. Portanto, pela prova de indução que vimos acima, temos

$$d(X) = r_n(X) = s(X)f(X) + t(X)g(X).$$

□

Proposição 13. *Sejam* $f(X), g(X), h(X) \in K[X]$*. Se* $f(X) \mid g(X)h(X)$ *e* $mdc(f(X), g(X)) = 1$*, então* $f(X) \mid h(X)$*.*

Demonstração. Se $f(X) \mid g(X)h(X)$, então existe $q(X) \in K[X]$ tal que

$$g(X)h(X) = f(X)q(X)$$

e, como $mdc(f(X), g(X)) = 1$, segue pelo Teorema 6 que existem $t(X), s(X) \in K[X]$ tais que

$$t(X)f(X) + s(X)g(X) = 1.$$

Logo,

$$\begin{aligned} h(X) &= (t(X)f(X) + s(X)g(X))h(X) \\ &= t(X)f(X)h(X) + s(X)g(X)h(X) \\ &= t(X)f(X)h(X) + s(X)f(X)q(X). \end{aligned}$$

Portanto, $f(X) \mid h(X)$. □

Proposição 14. *Todo elemento não invertível e irredutível no anel $K[X]$ é primo.*

Demonstração. Suponha $p(X) \in K[X]$ não inversível e irredutível, de modo que existem $f(X), g(X) \in K[X]$ tais que $p(X) \mid f(X)g(X)$. Sabemos pela Definição 14 que $p(X)$ é primo quando $p(X) \mid f(X)g(X)$ implica em $p(X) \mid f(X)$ ou $p(X) \mid g(X)$. Assim, se $p(X) \mid f(X)$ segue imediatamente que $p(X)$ é primo. Por outro lado, suponha que $p(X) \nmid f(X)$. Temos pelo Corolário 1 que existe um único $d(X) \in K[X]$ mônico, tal que $mdc(p(X), f(X)) = d(X)$, donde devemos ter, $d(X) \mid p(X)$. Desde que $p(X)$ é irredutível, segue que $d(X) = 1$ ou $d(X) = p(X)$. Se $d(X) = p(X)$, então $p(X) \mid f(X)$, absurdo. Logo devemos ter $d(X) = 1$ e, pela Proposição 13, $p(X) \mid g(X)$, provando deste modo que $p(X)$ é primo. □

Teorema 7. (Teorema da Fatoração Única) *Seja $f(X) \in K[X] \setminus K$ de grau n, em que $K[X] \setminus K$ é o conjunto dos polinômios não constantes de $K[X]$, então $f(X)$ é irredutível ou $f(X)$ se escreve de maneira única, a menos pela ordem dos fatores, como produto de polinômios irredutíveis de $K[X]$.*

Demonstração. Mostremos que o resultado é válido utilizando indução finita sobre o grau de $f(X)$. Para tanto, seja $f(X) \in K[X]$ com $gr(f(X)) = n \geq 1$. Supondo $n = 1$, temos pela Proposição 12 que $f(X)$ é irredutível em $K[X]$. Agora suponha que todo polinômio de grau menor do que n seja irredutível ou produto de irredutíveis em $K[X]\setminus K$. Seja $f(X) \in K[X]$ com $gr(f(X)) = n$. Se $f(X)$ é irredutível, o teorema está provado. Se f(X) é redutível, então existem $p(X)$ e $q(X)$ em $K[X]$ tais que

$$f(X) = p(X)q(X),$$

com $gr(p(X)) > 0$ e $gr(q(X)) > 0$ e, mais ainda,

$$n = gr(f(X)) = gr(p(X)q(X)) = gr(p(X)) + gr(q(X)),$$

logo $gr(p(X)) < n$ e $gr(q(X)) < n$. Portanto, pela hipótese de indução aplicada a $p(X)$ e $q(X)$ o resultado segue. Para provar a unicidade, suponha que existem $p_1(X), \ldots, p_s(X), p'_1(X), \ldots, p'_r(X)$ em $K[X]$ irredutíveis tais que

$$f(X) = p_1(X)p_2(X)\ldots p_s(X) \tag{2.21}$$

e

$$f(X) = p'_1(X)p'_2(X)\ldots p'_r(X). \tag{2.22}$$

Daí, segue que

$$p_1(X)p_2(X)\ldots p_s(X) = p'_1(X)p'_2(X)\ldots p'_r(X). \tag{2.23}$$

Agora, assuma sem perda de generalidade que $r \geq s$ (o caso em que $r \leq s$ é análogo). Como por (2.21)

$$p_j(X) \mid p_1(X)p_2(X)\ldots p_s(X),$$

para algum $j = 1, \ldots, s$, então, por (2.23)

$$p_j(X) \mid p'_1(X)p'_2(X)\ldots p'_r(X),$$

para algum $j = 1, \ldots, s$. Além disso, como $p_j(X)$ é primo, segue que

$$\underbrace{p_j(X) \mid p'_1(X)}_{(I)} \quad \text{ou} \quad \underbrace{p_j(X) \mid p'_2(X)\ldots p'_r(X)}_{(II)}.$$

Se vale (I), como ambos são irredutíveis, existe $c' \in K \setminus \{0\}$ tal que $p'_1(X) = c'p_j(X)$, ou seja, $p'_1(X)$ é associado a $p_j(X)$ em $K[X]$. Caso contrário, vale (II), e assim

$$p_j(X) \mid p'_2(X)p'_3(X)\ldots p'_r(X).$$

Logo

$$\underbrace{p_j(X) \mid p'_2(X)}_{(III)} \quad \text{ou} \quad \underbrace{p_j(X) \mid p'_3(X)\ldots p'_r(X)}_{(IV)}.$$

De modo análogo ao passo anterior, obtemos que ou $p'_2(X)$ é associado a $p_j(X)$ em $K[X]$, ou

$$p_j(X) \mid p'_3(X)p'_4(X)\ldots p'_r(X).$$

Procedendo recursivamente, de modo análogo aos passos anteriores, temos que

$$p_j(X) \mid p'_{r-1}(X) \quad \text{ou} \quad p_j(X) \mid p'_r(X).$$

Assim, $p'_{r-1}(X)$ é associado a $p_j(X)$ ou $p'_r(X)$ é associado a $p_j(X)$ em $K[X]$. Logo $p'_i(X)$ é associado a $p_j(X)$ em $K[X]$ para algum $i = 1, \ldots, r$ e $j = 1, \ldots, s$, ou seja, existe $u' \in K[X]$, tal que $p'_i(X) = u'p_j(X)$.

Agora, suponha que $r > s$. Como qualquer que seja o fator $p_j(X)$, obtemos um fator $p'_i(X)$ associado a ele, então restam $r - s$ fatores da decomposição de $f(X)$ dada em (2.22), que não são associados a nenhum fator $p_j(X)$ de $f(X)$. Assim, sejam $p'_{k_{s+1}}, \ldots, p'_{k_r}$ os $r - s$ fatores de $f(X)$ que não são associados a um fator $p_j(X)$, então

$$p'_{k_{s+1}}(X) \ldots p'_{k_r}(X) = 1,$$

pois caso contrário não teríamos a igualdade (2.23), o que é uma contradição, visto que $p'_{k_l}(X)$ é primo, para todo $l = s+1, \ldots, r$. Logo $r = s$, pois supomos anteriormente que $r \geq s$ e acabamos de ver que não vale a desigualdade estrita. Portanto a decomposição de $f(X)$ em polinômios irredutíveis é única, a menos da ordem dos fatores.

□

Corolário 2. *Todo polinômio $f(X)$ em $K[X]$ de grau maior ou igual a 1 pode ser escrito de modo único como $f(X) = u(p_1(X))^{\alpha_1} \ldots (p_k(X))^{\alpha_k}$ com cada $p_i(X)$ irredutível e mônico, em que $\alpha_i \in \mathbb{N}, 1 \leq i \leq k$ e $u \neq 0 \in K$.*

Demonstração. A existência do polinômio $f(X)$ segue de maneira análoga ao Teorema 7, ao considerarmos como hipótese de indução, que qualquer polinômio $f(X)$ de grau menor do que n, ou é irredutível, ou se fatora como produto de polinômios irredutíveis mônicos e de multiplicidade maior ou igual a 1. Para provar a unicidade, suponha que existem $p_1(X), \ldots, p_k(X), p'_1(X), \ldots, p'_r(X)$ mônicos irredutíveis e inteiros não negativos $\alpha_1, \ldots, \alpha_k, \alpha'_1, \ldots, \alpha'_r$ tais que

$$(p_1(X))^{\alpha_1} \ldots (p_k(X))^{\alpha_k} = f(X) = (p_1(X)')^{\alpha'_1} \ldots (p_r(X)')^{\alpha'_r}.$$

Como

$$(p_i(X))^{\alpha_i} = \underbrace{p_i(X) \ldots p_i(X)}_{\alpha_i \text{ vezes}} \quad \text{e} \quad (p_j(X)')^{\alpha'_j} = \underbrace{p_j(X)' \ldots p_j(X)'}_{\alpha'_j \text{ vezes}},$$

para $i = 1, \ldots, k$ e $j = 1, \ldots, r$. Então

$$\underbrace{p_1(X) \ldots p_1(X)}_{\alpha_1 \text{ vezes}} \ldots \underbrace{p_k(X) \ldots p_k(X)}_{\alpha_k \text{ vezes}} = f(X) = \underbrace{p_1(X)' \ldots p_1(X)'}_{\alpha'_1 \text{ vezes}} \ldots \underbrace{p_r(X)' \ldots p_r(X)'}_{\alpha'_r \text{ vezes}}.$$

Portanto, de maneira análoga ao Teorema 7, segue que todo polinômio $f(X) \in K[X]$ pode ser escrito de maneira única, a menos da ordem dos fatores, como

$$f(X) = u \underbrace{p_1(X) \ldots p_1(X)}_{\alpha_1 \text{ vezes}} \ldots \underbrace{p_k(X) \ldots p_k(X)}_{\alpha_k \text{ vezes}} = u(p_1(X))^{\alpha_1} \ldots (p_k(X))^{\alpha_k}.$$

□

Sejam $f(X), g(X) \in K[X] \setminus K$, pelo Corolário 2, existem $p_1(X), p_2(X), \ldots, p_r(X) \in K[X] \setminus K$ mônicos irredutíveis distintos, inteiros não negativos $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_r$ e $a, b \in K \setminus \{0\}$ tais que

$$f(X) = a(p_1(X))^{\alpha_1} \ldots (p_r(X))^{\alpha_r}$$

e

$$g(X) = b(p_1(X))^{\beta_1} \ldots (p_r(X))^{\beta_r}.$$

Temos que

$$mdc(f(X), g(X)) = (p_1(X))^{\delta_1} \ldots (p_r(X))^{\delta_r}, \tag{2.24}$$

em que $\delta_i = \min\{\alpha_i, \beta_i\}$ para cada $1 \leq i \leq r$. De fato, considerando $d(X) = (p_1(X))^{\delta_1} \ldots (p_r(X))^{\delta_r}$, com $\delta_i = \min\{\alpha_i, \beta_i\}$ para cada $1 \leq i \leq r$, podemos reescrever $f(X)$ e $g(X)$ como

$$f(X) = (a(p_1(X))^{\alpha_1 - \delta_1} \ldots (p_r(X))^{\alpha_r - \delta_r}) d(X)$$

e

$$g(X) = (b(p_1(X))^{\beta_1 - \delta_1} \ldots (p_r(X))^{\beta_r - \delta_r}) d(X),$$

em que $\alpha_i - \delta_i > 0$ e $\beta_i - \delta_i > 0$ para todo $i = 1, \ldots, r$. Logo $d(X) \mid f(X)$ e $d(X) \mid g(X)$. Além disso, se existe $d'(X) = (p_1(X))^{\xi_1} \ldots (p_r(X))^{\xi_r} \in K[X]$, com $\xi_i \leq \min\{\alpha_i, \beta_i\}$ para cada $1 \leq i \leq r$, tal que $d'(X) \mid f(X)$ e $d'(X) \mid g(X)$, então $d'(X) \mid d(X)$, pois como $\delta_i - \xi_i \geq 0$ para todo $i = 1 \ldots, r$, segue que

$$d(X) = (p_1(X))^{\delta_1} \ldots (p_r(X))^{\delta_r} = ((p_1(X))^{\delta_1 - \xi_1} \ldots (p_r(X))^{\delta_r - \xi_r}) d'(X).$$

Portanto, pela Definição 31,

$$mdc(f(X), g(X)) = d(X) = (p_1(X))^{\delta_1} \ldots (p_r(X))^{\delta_r},$$

em que $\delta_i = \min\{\alpha_i, \beta_i\}$ para algum $1 \leq i \leq r$.

Temos também, que o mínimo múltiplo comum de $f(X)$ e $g(X)$ é dado por

$$mmc(f(X), g(X)) = (p_1(X))^{\gamma_1} \ldots (p_r(X))^{\gamma_r}, \tag{2.25}$$

em que $\gamma_i = \max\{\alpha_i, \beta_i\}$ para cada $1 \leq i \leq r$. A demonstração desse fato é análoga a que foi feita para o máximo divisor comum.

Lema 5. *Sejam m e n dois inteiros positivos. Então*

(i) $mdc(X^n - 1, X^m - 1) = X^{mdc(n,m)} - 1$;

(ii) $mdc(X^n - X, X^m - X) = X^{q^{mdc(n,m)}} - X$.

A demonstração desse lema pode ser encontrada em (HEFEZ; VILLELA, 2008, páginas 44-45, capítulo 3).

Definição 34. *Seja K um corpo. Se K é um subcorpo de um corpo F, então dizemos que F é uma extensão de corpo de K. Podemos olhar F como um espaço vetorial sobre K e, nesse caso, dizemos que F é uma extensão finita ou infinita de K de acordo com a dimensão deste espaço vetorial. Denotaremos F extensão de K por $F \supset K$.*

Convém observarmos que utilizaremos nesse texto somente extensão finita de corpos finitos.

Definição 35. *Seja F uma extensão de K, ou seja, $F \supset K$ e seja $f(X) \in K[X]$. Um elemento α de F é dito raiz de $f(X)$ ou um zero de $f(X)$ quando $f(\alpha) = 0$.*

Proposição 15. *Sejam K um corpo, $p(X) \in K[X]$ e $\alpha \in K$. Temos que α é uma raiz de $f(X)$ se, e somente se, $X - \alpha$ é um divisor de $f(X)$ sobre K.*

Demonstração. Suponha que α é uma raiz de $f(X)$. Pela Definição 35, temos $f(\alpha) = 0$, além disso, pelo Teorema 4, existem únicos $q(X), r(X) \in K[X]$ tais que

$$f(X) = q(X)(X - \alpha) + r(X), \quad \text{com } r(X) = 0 \text{ ou } gr(r(X)) < gr(X - \alpha) = 1.$$

Como ou $r(X) = 0$ ou $gr(r(X)) < gr(X - a) = 1$, temos que $r(X) = r \in K$ e sendo assim,

$$0 = f(\alpha) = q(\alpha)(\alpha - \alpha) + r = r.$$

Logo,

$$f(X) = q(X)(X - \alpha),$$

consequentemente, $(X - \alpha)$ é um divisor de $f(X)$.

Reciprocamente, se $X - \alpha$ é um divisor de $f(X)$, então existe $q(X) \in K[X]$ tal que $f(X) = q(X)(X - \alpha)$, logo $f(\alpha) = q(\alpha)(\alpha - \alpha) = 0$. Portanto, pela Definição 35, α é uma raiz de $f(X)$. □

Teorema 8. *Sejam K um corpo e $f(X) \in K[X]$ um polinômio de grau $n \geq 0$. Então $f(X)$ tem no máximo n raízes distintas em K.*

Demonstração. Vamos provar por indução finita sobre n que o polinômio $f(X)$, de grau n, possui no máximo n raízes distintas em K. Primeiramente, observamos que se $f(X)$ não possui raízes

em K, o teorema está provado. Por outro lado, suponha que $f(X)$ possui raízes em K. Temos que se $n = 1$, então $f(X)$ possui apenas uma raiz em K e, portanto, o resultado vale. Agora suponha que o resultado vale para n, ou seja, o polinômio $f(X)$ de grau n possui no máximo n raízes em K. Vamos mostrar que também vale para $n+1$. De fato, seja $f(X) \in K[X]$ de grau $n+1$ e $\alpha \in K$ uma raiz de $f(X)$, temos pela Proposição 15 que

$$f(X) = (X - \alpha)g(X),$$

com $gr(g(X)) = n$. Como o resultado vale para n, segue que $g(X)$ possui no máximo n raízes distintas em K. Portanto $f(X)$ possui no máximo $n+1$ raízes distintas em K. □

Definição 36. *Sejam K um corpo e $\alpha \in K$ uma raiz de $p(X) \in K[X]$. Dizemos que α é uma raiz de $p(X)$ de* multiplicidade *$m \geq 1$ quando m é a maior potência de $(X - \alpha)$ que divide $p(X)$. Quando $m = 1$ dizemos que α é uma raiz* simples *de $p(X)$ e que $(X - \alpha)$ é um* fator múltiplo constante *em $K[X]$ e, quando $m > 1$, dizemos que α é uma raiz* múltipla *de $p(X)$ e que $(X - \alpha)$ é um* fator múltiplo não constante *em $K[X]$.*

Definição 37. *Seja K um corpo e $p(X) = \sum_{i=0}^{n} a_i X^i \in K[X]$. A derivada de $p(X)$ é definida como*

$$p'(X) = \sum_{i=1}^{n} i a_i X^{i-1}. \tag{2.26}$$

Note que se $p(X) = a$, constante, então $p'(X) = 0$.

Os resultados das regras operatórias da derivada formal de polinômios aqui definida são os mesmos vistos no Cálculo. Tais resultados podem ser demonstrados por meio de manipulações diretas das definições de derivada, soma e produto de polinômios.

Proposição 16. *Seja $f(X) \in K[X]$ com um fator múltiplo não constante. Então $mdc(f(X), f'(X)) \neq 1$.*

Demonstração. Seja $h(X) \in K[X]$ um fator múltiplo não constante de $f(X)$, então existem $r \in \mathbb{N}$ e $g(X) \in K[X]$, com $r > 1$, tais que

$$f(X) = h(X)^r g(X).$$

Daí,

$$\begin{aligned} f'(X) &= rh(X)^{r-1}h'(X)g(X) + h(X)^r g'(X) \\ &= h(X)(rh(X)^{r-2}h'(X)g(X) + h(X)^{r-1}g'(X)). \end{aligned}$$

Logo, $h(X)$ é um divisor de $f'(X)$ e, portanto, $mdc(f(X), f'(X)) \neq 1$.

□

Proposição 17. *Um ideal de $K[X]$ é da forma $I = I(g(X))$, em que $g(X) \in K[X]$.*

Demonstração. Seja I um ideal de $K[X]$. Se $I = \{0\}$, então $I = I(0)$. Agora, se $I \neq \{0\}$, considere $g(X) \in I$ não nulo, de modo que $g(X)$ tenha o menor grau possível em I, (o que é possível, pois o conjunto formado pelos graus dos polinômios de I é não vazio, logo pelo princípio da boa ordenação, este possui um menor elemento). Vamos provar que $I = I(g(X))$ e, para tanto, basta mostrarmos que $I \subset I(g(X))$.

Seja $f(X) \in I$, temos pelo Teorema 4 que existem únicos $q(X), r(X) \in K[X]$ tais que

$$f(X) = g(X)q(X) + r(X), \quad \text{com } r(X) = 0 \text{ ou } gr(r(X)) < gr(g(X))$$

e, como $g(X), f(X) \in I$, segue que

$$r(X) = f(X) + g(X)(-q(X)) \in I.$$

Daí, se $r(X) \neq 0$, então teríamos um polinômio de I com grau menor do que o grau de $g(X)$. Logo $r(X) = 0$, assim,

$$f(X) = g(X)q(X), \text{ ou seja, } f(X) \in I(g(X)).$$

Portanto, $I \subset I(g(X))$ e, consequentemente, $I = I(g(X))$. □

Corolário 3. *Seja $I \neq \{0\}$ um ideal de $K[X]$. Então, existe um único polinômio mônico $g(X)$ em I (de grau mínimo), tal que $I = I(g(X))$.*

Demonstração. Observe que, como $I \neq \{0\}$, segue da Proposição 17 que $I = I(f(X))$, com $f(X) \in K[X]$. Pela Proposição 11 temos que todo polinômio é associado a apenas um polinômio mônico, ou seja, existem únicos $g(X) \in K[X]$ e $a \in K \setminus \{0\}$, com $g(X)$ mônico, tais que

$$f(X) = ag(X) \Leftrightarrow f(X) \in I(g(X)) \Leftrightarrow I(f(X)) \subset I(g(X)).$$

Por outro lado, temos

$$g(X) = a^{-1}f(X) \Leftrightarrow g(X) \in I(f(X)) \Leftrightarrow I(g(X)) \subset I(f(X)).$$

Portanto $I = (f(X)) = I(g(X))$. □

Seja $p(X) \in K[X]$, temos que o anel quociente do anel $K[X]$ pelo ideal gerado por $p(X)$, em que $p(X)$ é um polinômio mônico de grau $n > 0$, é dado por

$$\frac{K[X]}{I(p(X))} = \{[r(X)];\ r(X) \in K[X], \text{ com } r(X) = 0, \text{ ou } gr(r(X)) < n\}. \tag{2.27}$$

Observação 2. Se $r_1(X), r_2(X) \in K[X]$, com $gr(r_1(X)) < n$ e $gr(r_2(X)) < n$, são tais que $r_1(X) \neq r_2(X)$, então $[r_1(X)] \neq [r_2(X)]$. De fato, suponha que $r_1(X) \neq r_2(X)$ e suponha por absurdo que $[r_1(X)] = [r_2(X)]$. Temos que

$$\begin{aligned}[r_1(X)] = [r_2(X)] &\Rightarrow [r_1(X)] - [r_2(X)] = [0] \Rightarrow [r_1(X) - r_2(X)] = [p(X)] \Rightarrow \\ &\Rightarrow r_1(X) - r_2(X) = p(X)q(X) \Rightarrow \\ &\Rightarrow n > gr(r_1(X) - r_2(X)) = gr(p(X)q(X)) \geq n.\end{aligned}$$

O que é um absurdo. Logo $r_1(X) = r_2(X)$. Além disso, $[r_1(X)] = [r_2(X)]$ se, e somente se, $r_1(X) \equiv r_2(X)\ (mod\, I(p(X)))$.

Observação 3. Temos que a soma e o produto em $\dfrac{K[X]}{I(p(X))}$ são dados por

$$[f(X)] + [g(X)] = [f(X) + g(X)], [f(X)] \cdot [g(X)] = [f(X) \cdot g(X)],$$

para todos $[f(X)], [g(X)] \in \dfrac{K[X]}{I(p(X))}$. Além disso, dados $\lambda \in K$ e $[f(X)] \in \dfrac{K[X]}{I(p(X))}$, temos $\lambda[f(X)] = [\lambda f(X)] \in \dfrac{K[X]}{I(p(X))}$. Com essas operações, $\dfrac{K[X]}{I(p(X))}$ é um espaço vetorial de dimensão $n = gr(p(X))$ sobre o corpo K, sendo $\{[1], [X], \ldots, [X^{n-1}]\}$ uma base para este espaço vetorial.

Teorema 9. *Todo ideal de* $\dfrac{K[X]}{I(p(X))}$ *é da forma* $I = I([f(X)])$*, em que* $f(X)$ *é um divisor de* $p(X)$.

Demonstração. Seja I um ideal de $\dfrac{K[X]}{I(p(X))}$. Defina o conjunto

$$J = \{g(X) \in K[X];\ [g(X)] \in I\}.$$

Note que $p(X) \in J$, pois $[p(X)] \in I$, consequentemente, $J \neq \emptyset$. Sejam $g_1(X), g_2(X) \in J$, mostremos que $g_1(X) - g_2(X) \in J$. De fato, temos que

$$g_1(X), g_2(X) \in J \Leftrightarrow [g_1(X)], [g_2(X)] \in I.$$

Logo,

$$[g_1(X) - g_2(X)] = [g_1(X)] - [g_2(X)] \in I,$$

consequentemente, $g_1(X) - g_2(X) \in J$.

Mostremos agora que para $g(X) \in J$ e $h(X) \in K[X]$, temos $g(X)h(X) \in J$. Com efeito,

$$\begin{aligned}g(X) \in J, h(X) \in K[X] &\Rightarrow [g(X)] \in I \quad \text{e} \quad [h(X)] \in \frac{K[X]}{I(p(X))} \\ &\Rightarrow [g(X)h(X)] = [g(X)][h(X)] \in I \\ &\Rightarrow g(X)h(X) \in J.\end{aligned}$$

Portanto, J é um ideal de $K[X]$.

Falta mostrarmos que I é um ideal principal. Note que, pela da Proposição 17, existe $f(X) \in K[X] \setminus \{0\}$ tal que $J = I(f(X))$ e, como $p(X) \in J$, existe $q(X) \in K[X]$ tal que $p(X) = f(X)q(X)$, isto é, $f(X)$ é um divisor de $p(X)$. Agora, observe que é imediato a verificação de que $I = \{[g(X)];\ g(X) \in J\}$. Como $J = I(f(X))$, segue que

$$[g(X)] = [h(X)f(X)] = [f(X)][h(X)].$$

Logo,

$$I = \left\{[f(X)][h(X)]; [h(X)] \in \frac{K[X]}{I(p(X))}\right\} = I([f(X)]),$$

como queríamos. □

Proposição 18. *Seja* $[f(X)] \in \dfrac{K[X]}{I(p(X))}$. *Temos que* $[f(X)]$ *é invertível se, e somente se,* $mdc(f(X), p(X)) = 1$.

Demonstração. Suponha que $[f(X)]$ seja invertível. Daí, existe $[g(X)] \in \dfrac{K[X]}{I(p(X))}$ tal que

$$\begin{aligned}[f(X)g(X)] = [f(X)][g(X)] = [1] &\Leftrightarrow f(X)g(X) \equiv 1\ (mod\ I(p(X))) \\ &\Leftrightarrow f(X)g(X) - 1 \equiv 0\ (mod\ I(p(X))).\end{aligned}$$

Logo, existe $q(X) \in K[X]$ tal que

$$f(X)g(X) - 1 = q(X)p(X) \Leftrightarrow f(X)g(X) - q(X)p(X) = 1.$$

Portanto, segue do Teorema 6 que $mdc(f(X), p(X)) = 1$.

Reciprocamente, suponha que $mdc(f(X), p(X)) = 1$. Então, pelo Teorema 6, existem $g(X), q(X) \in K[X]$ tais que

$$f(X)g(X) + q(X)p(X) = 1$$

e assim,

$$f(X)g(X) - 1 = (-q(X))p(X).$$

Tomando a classe em ambos os lados da igualdade acima, temos

$$\begin{aligned}[f(X)g(X) - 1] = [(-q(X))p(X)] &\Leftrightarrow [f(X)][g(X)] - [1] = [0] \\ &\Leftrightarrow [f(X)][g(X)] = [1].\end{aligned}$$

Logo $[f(X)]$ é invertível. □

Teorema 10. *O anel* $\dfrac{K[X]}{I(p(X))}$ *é um corpo se, e somente se, o polinômio* $p(X)$ *é irredutível.*

Demonstração. Suponha que o anel $\frac{K[X]}{I(p(X))}$ é corpo e suponha por absurdo que $p(X)$ é redutível. Temos pela Definição 32 que existem $f(X), g(X) \in K[X]$ não nulos e não invertíveis, tais que

$$p(X) = f(X)g(X), \text{ com } gr(f(X)) < gr(p(X)) \text{ e } gr(g(X)) < gr(p(X)).$$

Daí, $[f(X)], [g(X)] \in \frac{K[X]}{I(p(X))}$ e assim,

$$[0] = [p(X)] = [f(X)g(X)] = [f(X)][g(X)].$$

Logo $\frac{K[X]}{I(p(X))}$ não é um domínio de integridade, contradizendo o fato dele ser corpo. Portanto $p(X)$ é irredutível.

Reciprocamente, suponha que $p(X)$ é irredutível. Devemos mostrar que o anel $\frac{K[X]}{I(p(X))}$ é corpo. Como ele é um anel comutativo com unidade, para que seja corpo falta mostrarmos que é também um anel de divisão, ou seja, que todos os seus elementos não nulos são invertíveis. De fato, seja $[r(X)] \in \frac{K[X]}{I(p(X))}$ qualquer e não nulo, temos que $[r(X)] \neq [p(X)]$, logo $p(X) \nmid r(X)$. Note que $d(X) = mdc(r(X), p(X))$ é tal que $d(X)$ é invertível ou $d(X) = p(X)$, pois $p(X)$ é irredutível. Desta forma, $p(X) \nmid r(X)$ implica que $d(X) = d \in K \setminus \{0\}$ invertível, logo pelo Teorema 6, existem $s(X), t(X) \in K[X]$ tais que

$$s(X)r(X) + t(X)p(X) = d.$$

Daí, como d é invertível,

$$r(X)\left(\frac{1}{d} \cdot s(x)\right) + p(X)\left(\frac{1}{d} \cdot t(X)\right) = 1$$

Assim,

$$\left[r(X)\left(\frac{1}{d} \cdot s(x)\right)\right] + \left[p(X)\left(\frac{1}{d} \cdot t(X)\right)\right] = [1] \Rightarrow [r(X)]\left[\left(\frac{1}{d} \cdot s(x)\right)\right] = [1].$$

Logo $[r(X)]$ é invertível em $\frac{K[X]}{I(p(X))}$ e portanto esse quociente é corpo. □

Até o presente momento focamos no estudo de polinômios em uma incógnita. A partir de agora apresentaremos alguns resultados referentes a polinômios de várias incógnitas, que serão utilizados no Capítulo 6.

Definição 38. *Sejam $Y_1, \ldots, Y_n$ incógnitas. Definimos os* polinômios simétricos elementares *nas incógnitas $Y_1, \ldots, Y_n$, por*

$$\begin{cases} S_k(Y_1, \ldots, Y_n) = \displaystyle\sum_{1 \leq i_1 < \ldots < i_k \leq n} Y_{i_1} \ldots Y_{i_k}, & se \quad 1 \leq k \leq n; \\ S_0(Y_1, \ldots, Y_n) = 1. \end{cases}$$

Observação 4. Temos que, ao expandirmos os somatórios da Definição 38 os polinômios S_k, para cada $k = 0, \ldots, n$, são:

$$\begin{aligned}
&S_0(Y_1, \ldots, Y_n) = 1;\\
&S_1(Y_1, \ldots, Y_n) = Y_1 + \ldots + Y_n;\\
&S_2(Y_1, \ldots, Y_n) = Y_1Y_2 + \ldots + Y_{n-1}Y_n;\\
&\vdots\\
&S_k(Y_1, \ldots, Y_n) = Y_1Y_2 \ldots Y_k + \ldots + Y_{n-k+1}Y_{n-k+2} \ldots Y_n\\
&\vdots\\
&S_{n-1}(Y_1, \ldots, Y_n) = Y_1 \ldots Y_{n-1} + \ldots + Y_2 \ldots + Y_n;\\
&S_n(Y_1, \ldots, Y_n) = Y_1 \ldots Y_n,
\end{aligned}$$

em que S_k é uma soma de $\binom{n}{k}$ parcelas, sendo que cada parcela é um produto de k elementos distintos do conjunto $\{Y_1, \ldots, Y_n\}$.

Esses polinômios nos serão úteis, pois eles fornecem relações entre os coeficientes e raízes de um polinômio, como veremos a seguir

Lema 6. *Sejam $S_k = S_k(Y_1, \ldots, Y_n)$ os polinômios simétricos para $1 \leq k \leq n$. Temos que*

$$(X - Y_1) \ldots (X - Y_n) = X^n - S_1X^{n-1} + S_2X^{n-2} - \ldots + (-1)^nS_n. \tag{2.28}$$

A demonstração desse resultado pode ser encontrada em (HEFEZ; VILLELA, 2008, páginas 56-57, capítulo 3)

Proposição 19. *Dados $S_i = S_i(Y_1, \ldots, Y_n)$, polinômios simétricos elementares nas incógnitas $Y_1, \ldots, Y_n$, e $P_j = P_j(Y_1, \ldots, Y_n) = Y_1^j + \ldots + Y_n^j$, com $j \in \mathbb{N}$. Então valem as seguintes relações:*

$$\begin{aligned}
&P_1 - S_1 = 0\\
&P_2 - S_1P_1 + 2S_2 = 0\\
&\vdots\\
&P_n - S_1P_{n-1} + \cdots + (-1)^{n-1}S_{n-1}P_1 + (-1)^n nS_n = 0\\
&P_i - S_1P_{i-1} + \cdots + (-1)^nS_nP_{i-n} = 0 \, se \; i > n.
\end{aligned} \tag{2.29}$$

A prova desse resultado pode ser encontrada em (HEFEZ; VILLELA, 2008, página 59, capítulo 3).

2.4 Corpos Finitos

Apresentaremos nesta seção alguns resultados de corpos finitos, como característica de um corpo, existência de corpos finitos e elementos primitivos. Tais resultados serão cruciais nos capítulos 5 e 6.

Antes de definirmos a característica de um corpo K, convém observarmos que dados m inteiro e a um elemento de K, temos

$$ma = \begin{cases} \underbrace{a + a + \cdots + a}_{\text{m parcelas}}, & \text{se } m > 0; \\ 0, & \text{se } m = 0; \\ \underbrace{(-a) + (-a) + \cdots + (-a)}_{\text{m parcelas}}, & \text{se } m < 0. \end{cases}$$

Com isso, temos que $ma = m(1_K \cdot a) = (m \cdot 1_K)a$.

Lema 7. *Se $m, n \in \mathbb{Z}$ e $a, b \in K$, valem:*

(i) $(mn)a = m(na)$;

(ii) $(m+n)a = (ma) + (na)$;

(iii) $m(a+b) = (ma) + (mb)$.

Demonstração.

(i) Por indução finita sobre n, temos que para $n = 1$, vale $(m \cdot 1)a = ma = m(1 \cdot a)$. Agora, suponha que o resultado vale para n, ou seja,

$$(mn)a = m(na). \tag{2.30}$$

Vamos mostrar que também vale para $n+1$. De fato, temos que

$$\begin{aligned} (m(n+1))a &= \underbrace{a + \ldots + a}_{m(n+1) \text{ vezes}} = \underbrace{(a + \ldots + a)}_{mn \text{ vezes}} + \underbrace{(a + \ldots + a)}_{m \text{ vezes}} = (mn)a + ma \overset{(2.30)}{=} m(na) + ma = \\ &= m(na + a) = m((n+1)a). \end{aligned}$$

(ii) De modo similar ao item (i), temos que para $n = 1$, vale

$$(m+1)a = \underbrace{a + \ldots + a}_{m+1 \text{ vezes}} = \underbrace{(a + \ldots + a)}_{m \text{ vezes}} + a = ma + 1 \cdot a.$$

Suponha que o resultado vale para n, ou seja,

$$(m+n)a = (ma) + (na). \tag{2.31}$$

Vamos mostrar que também vale para $n+1$. Com efeito,

$$\begin{aligned}(m+(n+1))a &= \underbrace{a+\ldots+a}_{m+(n+1)\text{ vezes}} = \underbrace{(a+\ldots+a)}_{m+n\text{ vezes}} + a = (m+n)a + a \overset{(2.31)}{=} (ma)+(na)+a = \\ &= (ma) + ((n+1)a).\end{aligned}$$

(iii) Para este item, façamos a prova por indução finita sobre m. Se $m = 1$, então vale

$$1(a+b) = a+b = 1\cdot a + 1\cdot b.$$

Supondo que o resultado vale para m, temos que

$$m(a+b) = (ma) + (mb). \tag{2.32}$$

vamos mostrar que também vale para $m+1$. De fato,

$$\begin{aligned}(m+1)(a+b) &= \underbrace{(a+b)+\ldots+(a+b)}_{(m+1)\text{ vezes}} = \underbrace{(a+b)+\ldots+(a+b)}_{m\text{ vezes}} + (a+b) = \\ &= m(a+b) + (a+b) \overset{(2.32)}{=} (ma)+(mb)+(a+b) = ((m+1)a) + ((m+1)b).\end{aligned}$$

□

Definição 39. *Seja Λ_K o conjunto definido como*

$$\Lambda_K = \{n \in \mathbb{N}^*; n\cdot 1_K = 0\} \subset \mathbb{N}. \tag{2.33}$$

Definimos por característica de um corpo K, o inteiro positivo

$$car(K) = \min \Lambda_K = \min\{n \in \mathbb{N}^*; n\cdot 1_K = 0\}. \tag{2.34}$$

Se $\Lambda_K = \emptyset$, definimos a característica de K como 0. Temos como exemplos de característica 0 os corpos $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Por outro lado, quando $\Lambda_K \neq \emptyset$, prova-se utilizando o princípio da boa ordenação e a divisão euclidiana, que existe um menor elemento $m \in \Lambda_K$ tal que m é a característica de K.

Note que, se K é um corpo finito, então o conjunto $\{n\cdot 1_K\}_{n\in\mathbb{N}}$ é finito. Daí, segue que existem $n_1, n_2 \in \mathbb{N}$ tais que

$$n_1\cdot 1_K = n_2\cdot 1_K \Rightarrow (n_1-n_2)\cdot 1_K = 0 \Rightarrow n_1 - n_2 \in \Lambda_K \Rightarrow \Lambda_K \neq \emptyset,$$

isto é, se K é um corpo finito, então K possui característica diferente de 0. Veremos que a próxima proposição é um resultado ainda mais forte a respeito da característica de um corpo finito K.

Proposição 20. *Seja K um corpo finito, então $car(K) = p$, com p primo.*

Demonstração. Seja $m = car(K)$ e suponha por absurdo que m não é um número primo. Temos que existem $n_1, n_2 < m \in \mathbb{N}$ tais que $m = n_1 n_2$. Daí,

$$m \cdot 1_K = 0 \Rightarrow (n_1 n_2) \cdot 1_K = 0 \Rightarrow (n_1 \cdot 1_K)(n_2 \cdot 1_K) = 0.$$

Como K é um domínio de integridade, segue que $(n_1 \cdot 1_K) = 0$ ou $(n_2 \cdot 1_K) = 0$, o que contradiz a minimalidade de m. Portanto m é primo. □

Proposição 21. *Seja K um corpo finito com $car(K) = p$. Se $ma = 0$ para algum m inteiro e a um elemento de K, então m é um múltiplo de p ou $a = 0$.*

Demonstração. Suponha que $ma = 0$, com $m \in \mathbb{Z}$ e $a \in K$. Se $a = 0$, nada temos a provar. Suponha então $a \neq 0$. Daí, como $ma = (m \cdot 1_K)a$, então $(m \cdot 1_K)a = 0$ e, dado que K é corpo, segue que $m \cdot 1_K = 0$ ou $a = 0$. Como existem únicos $q, r \in \mathbb{Z}$ tais que

$$m = qp + r, \text{ com } 0 \leq r < q,$$

então

$$0 = m \cdot 1_K = (qp+r) \cdot 1_K \overset{\text{Lema } 7}{=} (qp) \cdot 1_K + r \cdot 1_K = q(p \cdot 1_K) + r \cdot 1_K \overset{car(K)=p}{=} q \cdot 0 + r \cdot 1_K = r \cdot 1_K.$$

Por definição de característica, p é o menor inteiro positivo tal que $p \cdot 1_K = 0$, então $r \cdot 1_K = 0$ se, e somente se, $r = 0$. Logo $m = qp$, isto é, m é múltiplo de p. □

Teorema 11. *Seja K um corpo finito com $car(K) = p$, em que p é um número primo. Então, K contém um subcorpo isomorfo a $\mathbb{Z}_p$. Em particular, K tem p^r elementos, para algum r natural.*

Demonstração. Considere a aplicação

$$\begin{array}{rccc} \varphi: & \mathbb{Z}_p & \to & K \\ & [n] & \mapsto & n \cdot 1_K. \end{array}$$

A função φ está bem definida. De fato, sejam $m, n \in \mathbb{Z}$ tais que $[n] = [m]$, então existe $q \in \mathbb{Z}$ tal que $n = m + qp$. Assim,

$$n \cdot 1_K = (m + qp) \cdot 1_K \overset{\text{Lema } 7}{=} m \cdot 1_K + (qp) \cdot 1_K = m \cdot 1_K + q(p \cdot 1_K) = m \cdot 1_K.$$

Além disso, φ é um homomorfismo, pois

$$\varphi([n+m]) = (n+m) \cdot 1_K = n \cdot 1_K + m \cdot 1_K = \varphi([n]) + \varphi([m])$$

e

$$\varphi([nm]) = (nm) \cdot 1_K = (n \cdot 1_K)(m \cdot 1_K) = \varphi([n])\varphi([m]).$$

Com isso, temos pela Proposição 8, item (vi), que φ é injetora e $\varphi(\mathbb{Z}_p)$ é um subcorpo de K, logo φ restrita a imagem de $\mathbb{Z}_p$ é bijetora. Portanto, K possui um subcorpo isomorfo a $\mathbb{Z}_p$, a saber, $\varphi(\mathbb{Z}_p)$.

Podemos então olhar K como um espaço vetorial sobre $\mathbb{Z}_p$ e, como K é um corpo finito, segue que K é um espaço vetorial de dimensão finita sobre $\mathbb{Z}_p$, assim existe uma base $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de K sobre $\mathbb{Z}_p$. Logo, qualquer que seja $\mathbf{v} \in K$, temos

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n,$$

com $\alpha_1, \ldots, \alpha_n \in \mathbb{Z}_p$. Portanto, $|K| = p^n$. □

Note que $\mathbb{Z}_p$ é um corpo com p elementos que estão carregados de notação, pois quando se trata de $\mathbb{Z}_p$, pensamos nos conjuntos de classes módulo p e as suas propriedades aritméticas, como restos de divisões, equações diofantinas e etc. Para evitar, de certo modo, embarcar todo esse conhecimento prévio a respeito do conjunto $\mathbb{Z}_p$, consideraremos um corpo arbitrário, que chamaremos de $\mathbb{F}_p$, com p elementos, os quais denotamos por $0, 1, \ldots, p-1$. Além disso, para denotar um corpo finito com qualquer quantidade q de elementos, utilizaremos a notação $\mathbb{F}_q$.

Proposição 22. *Seja K um corpo finito de característica p e seja $q = p^r$, para algum inteiro positivo r. Se $a, b \in K$, temos que*

$$(a \pm b)^q = a^q + b^q.$$

Demonstração. Temos, pelo Binômio de Newton, que

$$(a \pm b)^p = a^p \pm \binom{p}{1}a^{p-1}b + \cdots + (\pm 1)^i\binom{p}{i}a^{p-i}b^i + \cdots \pm b^p.$$

Note que $p \mid \binom{p}{i}$; $i = 1, \ldots, p-1$. Com efeito,

$$\binom{p}{i} = \underbrace{\frac{p!}{i!(p-i)!}}_{\in \mathbb{Z}} = p \cdot \underbrace{\frac{(p-1)!}{i!(p-i)!}}_{\in \mathbb{Z}}.$$

Logo $p \mid \binom{p}{i}$ e, portanto,

$$(a \pm b)^p = a^p \pm \binom{p}{1}a^{p-1}b + \cdots + (\pm 1)^i\binom{p}{i}a^{p-i}b^i + \cdots \pm b^p = a^p \pm b^p.$$

Agora, vamos mostrar que vale para $(a \pm b)^q$, com $q = p^r$, fazendo indução sobre r. O caso em que $r = 1$ foi provado acima. Suponha por indução que a igualdade seja válida para $r - 1$, ou seja,

$$(a \pm b)^{p^{r-1}} = a^{p^r-1} \pm b^{p^{r-1}}.$$

Logo,

$$(a \pm b)^q = (a \pm b)^{p^r} = ((a \pm b)^{p^{r-1}})^p = (a^{p^{r-1}} \pm b^{p^{r-1}})^p = a^{p^r} \pm b^{p^r} = a^q \pm b^q.$$

□

Observação 5. É importante notar que, se $a_1, \ldots, a_n$ são elementos de um corpo finito K, de característica p, e sendo $q = p^r$ para algum natural r, então aplicando a Proposição 22 sucessivas vezes, temos

$$(a_1 + a_2 + \cdots + a_n)^q = a_1^q + a_2^q + \cdots + a_n^q.$$

Observe também que, se $p(X) = a_0 + a_1X + \cdots + a_nX^n \in K[X]$, então

$$(p(X))^q = (a_0 + a_1X + \cdots + a_nX^n)^q \overset{\text{Prop. } 22}{=} a_0^q + a_1^qX^q + \cdots + a_n^qX^{qn}. \tag{2.35}$$

Corolário 4. *Seja K um corpo finito com $car(K) = p$. Se $q = p^r$, com $r \in \mathbb{Z}$, então a função*

$$\begin{array}{rccc} f_q: & K & \to & K \\ & x & \mapsto & x^q \end{array}$$

é um isomorfismo.

Demonstração. Para provar esse resultado, mostremos que f_q é um homomorfismo bijetor. Para tanto, sendo $x, y \in K$ e q uma potência da característica de K, temos que

$$f_q(x + y) = (x + y)^q = x^q + y^q = f_q(x) + f_q(y)$$

e

$$f_q(xy) = (xy)^q = x^qy^q = f_q(x)f_q(y).$$

Logo, f_q é um homomorfismo.

Mostremos agora que f_q é bijetiva. De fato, como $f_q(1) = 1$ e K é corpo, segue do item (vi) da Proposição 8 que a função f_q é injetiva e, como é uma função de K em K, é também sobrejetiva. □

Corolário 5. *Sejam F um corpo de característica $p > 0$ e $q = p^r$, com $r \in \mathbb{Z}$. Então o conjunto $K = \{\alpha \in F; \alpha^q - \alpha = 0\}$ é um subcorpo de F.*

Demonstração. Para provarmos que K é um subcorpo de F, de acordo com a Proposição 5, é necessário e suficiente mostrarmos que dados $\alpha, \beta \in K$, com $\beta \neq 0$, implica em $\alpha - \beta \in K$ e $\alpha\beta^{-1} \in K$. Desta forma, temos

$$\begin{aligned}(\alpha-\beta)^q - (\alpha-\beta) &= \alpha^q - \beta^q - \alpha + \beta \\ &= (\alpha^q - \alpha) - (\beta^q - \beta),\end{aligned}$$

logo $\alpha - \beta \in K$. Além disso,

$$(\alpha\beta^{-1})^q - \alpha\beta^{-1} = \alpha^q\beta^{-q} - \alpha\beta^{-1} = \alpha^q(\beta^q)^{-1} - \alpha\beta^{-1} = \alpha\beta^{-1} - \alpha\beta^{-1} = 0.$$

□

Proposição 23. *Sejam K um corpo finito com característica p e $f(X) \in K[X]$. Temos que $f'(X) = 0$ se, e somente se, existe um polinômio $g(X) \in K[X]$ tal que $f(X) = g(X)^p$.*

Demonstração. Suponha que $f(X) = a_0 + a_1X + \cdots + a_nX^n \in K[X]$ é tal que $f'(X) = 0$, ou seja,

$$\begin{aligned}f'(X) = a_1 + 2a_2X + \cdots + na_nX^{n-1} &= 0 \\ \Leftrightarrow a_1 = 2a_2 = \cdots = na_n &= 0,\end{aligned}$$

o que é equivalente, pela Proposição 21, a i ser múltiplo de p ou $a_i = 0$. Logo, podemos reescrever o polinômio $f(X)$ como potência da característica, como segue

$$f(X) = a_0 + a_pX^p + a_{2p}X^{2p} + a_{3p}X^{3p} \ldots;$$

pois todos os coeficientes, cujos índices não são múltiplos da característica, são iguais a zero. Assim, pelo Corolário 4 podemos escolher $b_i \in K$ de modo que $b_i^p = a_{ip}$ e considerar

$$g(X) = b_0 + b_1X + b_2X^2 + b_3X^3 + \ldots.$$

Daí,

$$\begin{aligned}(g(X))^p &= (b_0 + b_1X + b_2X^2 + b_3X^3 + \ldots)^p \\ &= b_0^p + b_1^pX^p + b_2^pX^{2p} + b_3^pX^{3p} + \ldots \\ &= a_0 + a_pX^p + a_{2p}X^{2p} + a_{3p}X^{3p} + \ldots \\ &= f(X).\end{aligned}$$

Reciprocamente, supondo $f(X) = g(X)^p$, temos que $f'(X) = pg(X)^{p-1}g'(X) = 0$. □

Proposição 24. *Seja K um corpo finito de característica p e seja $q = p^r$, para algum inteiro positivo r. O polinômio $f(X) = X^q - X$ não possui fatores irredutíveis múltiplos em $K[X]$.*

Demonstração. Sendo $f(X) = X^q - X$, temos que $f'(X) = qX^{q-1} - 1 = -1$. Como $mdc(X^q - X, qX^{q-1} - 1) = 1$, segue da Proposição 16 que $f(X) = X^q - X$ não possui fatores irredutíveis múltiplos não constantes em $K[X]$. □

Lema 8. *Seja K um corpo finito com q elementos. Para todo $\alpha \in K^*$, em que $K^* = K \setminus \{0\}$, temos que $\alpha^{q-1} = 1$.*

Demonstração. Seja $\alpha \in K^*$. Defina a aplicação

$$\begin{array}{rccc} f_\alpha : & K^* & \to & K^* \\ & a & \mapsto & \alpha a \end{array}.$$

Temos que f_α é injetiva. De fato, dados $a, b \in K^*$,

$$f_\alpha(a) = f_\alpha(b) \Leftrightarrow \alpha a = \alpha b \Leftrightarrow a = b.$$

Além disso, como K^* é um corpo finito e f_α está definida de K^* em K^*, segue que f_α é também sobrejetiva e, portanto, bijetiva.

Considerando $K^* = \{\alpha a_1, \ldots, \alpha a_{q-1}\} = \{a_1, \ldots, a_{q-1}\}$, temos que

$$(\alpha a_1) \ldots (\alpha a_{q-1}) = a_1 \ldots a_{q-1} \Rightarrow \alpha^{q-1} = 1.$$

□

Proposição 25. *Seja K um corpo finito com q elementos. Para todo $\alpha \in K$ e para todo $i \in \mathbb{N}$, temos que $\alpha^{q^i} = \alpha$.*

Demonstração. Provemos por indução finita sobre i a igualdade $\alpha^{q^i} = \alpha$. Temos que para $i = 1$,

$$\alpha^q = \alpha^{q-1}\alpha \overset{\text{Lema } 8}{=} 1 \cdot \alpha = \alpha.$$

Suponha que a igualdade seja válida para $i = k - 1$, isto é, $\alpha^{q^{k-1}} = \alpha$. Logo, para $i = k$, temos

$$\alpha^{q^k} = \left(\alpha^{q^{k-1}}\right)^q = \alpha^q = \alpha.$$

Portanto, $\alpha^{q^i} = \alpha$ para todo i natural. □

Corolário 6. *Sejam K um corpo finito de característica p com q elementos e F uma extensão de K. Então os elementos de K são os elementos de F que são raízes de $X^q - X$, enquanto que os elementos do subcorpo $\mathbb{Z}_p$ de F são as raízes do polinômio $X^p - X$.*

Demonstração. Da Proposição 25, segue que $\alpha^q = \alpha$ para todo $\alpha \in K$. Logo, todo elemento α de K é raiz do polinômio $X^q - X \in F[X]$. Como pelo Teorema 8 este polinômio possui no máximo q raízes distintas, segue que suas raízes são todos os elementos de K. A prova de que todas as raízes do polinômio $X^p - X$ de $F[X]$ são elementos de $\mathbb{Z}_p$ segue de modo análogo ao anterior, quando consideramos $K = \mathbb{Z}_p$. □

Exemplo 3. *Seja K um corpo com q elementos e seja F um corpo finito, extensão de K. Sejam $\beta_0, \ldots, \beta_{r-1}$ elementos de F linearmente independentes sobre K. Vamos mostrar que o determinante da matriz M é não nulo, em que*

$$M = \begin{bmatrix} \beta_0 & \beta_0^q & \beta_0^{q^2} & \ldots & \beta_0^{q^{r-1}} \\ \beta_1 & \beta_1^q & \beta_1^{q^2} & \ldots & \beta_1^{q^{r-1}} \\ \vdots & \vdots & \vdots & & \vdots \\ \beta_{r-1} & \beta_{r-1}^q & \beta_{r-1}^{q^2} & \ldots & \beta_{r-1}^{q^{r-1}} \end{bmatrix}.$$

Seja $car(F) = car(K) = p$, com p primo, temos pelo Teorema 11 que $q = p^n$ para algum natural n. Observe também que, como os elementos $\beta_0, \ldots, \beta_{r-1}$ de F são linearmente independentes sobre K, estes geram um subespaço vetorial V r-dimensional de F sobre K, consequentemente, $|V| = q^r$. Suponhamos por absurdo que $det\ M = 0$, equivalentemente, que as colunas de M são linearmente dependentes em F. Assim, denotando as colunas de M por $\bar{\beta}_j$, com $j = 0, \ldots, r-1$, existem escalares $c_0, \ldots, c_{r-1} \in F$, nem todos nulos, tais que

$$\begin{aligned} & c_0\bar{\beta}_0 + \ldots + c_{r-1}\bar{\beta}_{r-1} = 0 \\ & \Leftrightarrow c_0\beta_j + c_1\beta_j^q + \cdots + c_{r-1}\beta_j^{q^{r-1}} = 0, \quad j = 0, \ldots r-1. \end{aligned}$$

Logo, $\beta_0, \ldots, \beta_{r-1}$ são raízes do polinômio de grau no máximo q^{r-1}

$$f(X) = c_0X + c_1X^q + \ldots + c_{r-1}X^{q^{r-1}} \in F[X].$$

Além disso, desde que β_k, β_ℓ são raízes de $f(X)$, então para todo $a \in K$ o elemento $a\beta_k + \beta_\ell$ também é raiz de $f(X)$, pois

$$\begin{aligned} f(a\beta_k + \beta_\ell) &= c_0(a\beta_k + \beta_\ell) + c_1(a\beta_k + \beta_\ell)^q + \ldots + c_{r-1}(a\beta_k + \beta_\ell)^{q^{r-1}} \\ &= c_0a\beta_k + c_0\beta_\ell + c_1a\beta_k^q + c_1\beta_\ell^q + \ldots + c_{r-1}a\beta_k^q + c_{r-1}\beta_\ell^{q^{r-1}} \\ &= a(c_0\beta_k + c_1\beta_k^q + \ldots + c_{r-1}\beta_k^{q^{r-1}}) + (c_0\beta_\ell + c_1\beta_\ell^q + \ldots + c_{r-1}\beta_\ell^{q^{r-1}}) \\ &= 0. \end{aligned}$$

Desta forma, como uma combinação linear arbitrária de elementos de V é também uma raiz de $f(X)$, então $f(X)$ possui mais do que q^{r-1} raízes, o que é um absurdo, visto que pelo Teorema 8 esse polinômio possui no máximo q^{r-1} raízes distintas. Portanto, $det\ M \neq 0$, como queríamos.

Observação 6. É importante observarmos que sob as condições do Lema 8, o conjunto

$$\{n \in \mathbb{N}^*;\ \alpha^n = 1\}$$

é não vazio. Tal observação possibilita darmos a definição a seguir.

Definição 40. *Seja K um corpo finito e seja $\alpha \in K^*$. Definimos a ordem de α o inteiro positivo*

$$ord(\alpha) = min\{n \in \mathbb{N}^*; \alpha^n = 1\}.$$

Proposição 26. *Seja K um corpo finito com q elementos e seja $\alpha \in K^*$. Se para algum inteiro positivo m temos que $\alpha^m = 1$, então $ord(\alpha)$ é um divisor de m. Em particular, $ord(\alpha)$ divide $(q-1)$.*

Demonstração. Denotando $\text{ord}(\alpha) = k$ e sendo m algum inteiro positivo tal que $\alpha^m = 1$, temos que existe um único inteiro positivo s tal que

$$m = ks + r, \quad \text{com } 0 \leq r < k.$$

Logo,

$$\alpha^m = 1 \Rightarrow (\alpha^k)^s \alpha^r = 1 \Rightarrow 1 \cdot \alpha^r = 1 \Rightarrow \alpha^r = 1.$$

Mas, como $r < k$ e k é, por definição, o menor inteiro tal que $\alpha^k = 1$, segue que o único valor de r que satisfaz a igualdade acima é $r = 0$. Portanto, m divide $k = \text{ord}(\alpha)$. Note que, pelo Lema 8 temos $\alpha^{q-1} = 1$, logo $\text{ord}(\alpha)$ divide $(q-1)$. □

Proposição 27. *Seja K um corpo finito. Sejam α e β elementos de K tais que $mdc(ord(\alpha), ord(\beta)) = 1$. Então $ord(\alpha\beta) = ord(\alpha)ord(\beta)$.*

Demonstração. Suponha que $\text{ord}(\alpha) = m$, $\text{ord}(\beta) = n$ e $\text{ord}(\alpha\beta) = t$. Vamos mostrar que $t = mn$. Primeiramente observamos que

$$(\alpha\beta)^{mn} = (\alpha^m)^n (\beta^m)^n = 1 \Rightarrow t = \text{ord}(\alpha\beta) \mid mn. \tag{2.36}$$

Como $\text{ord}(\alpha\beta) = t$, então

$$((\alpha\beta)^t)^m = 1 \Rightarrow \alpha^{tm}\beta^{tm} = 1 \Rightarrow 1 \cdot \beta^{tm} = 1 \Rightarrow \beta^{tm} = 1$$

e

$$((\alpha\beta)^t)^n = 1 \Rightarrow \alpha^{tn}\beta^{tn} = 1 \Rightarrow 1 \cdot \alpha^{tn} = 1 \Rightarrow \alpha^{tn} = 1.$$

Pela Proposição 26, temos que $n \mid tm$ e $m \mid tn$. Além disso, como $mdc(m,n) = 1$, então necessariamente $m \mid t$ e $n \mid t$, mais ainda, $mn \mid t$. Portanto $\text{ord}(\alpha\beta) = mn = \text{ord}(\alpha)\text{ord}(\beta)$. □

Proposição 28. *Seja K um corpo finito e sejam $\alpha \in K^*$ e $i \in \mathbb{N}^*$. Suponhamos que $ord(\alpha) = m$, então*

$$ord(\alpha^i) = \frac{m}{mdc(m,i)}.$$

Demonstração. Suponha que $\text{ord}(\alpha^i) = t$, daí

$$(\alpha^i)^t = 1 \Leftrightarrow \alpha^{it} = 1,$$

em que t é o menor inteiro positivo que satisfaz as igualdades acima. Como $m = \text{ord}(\alpha)$, pela Proposição 26 temos que $m \mid it$, logo $it = mmc(m,i)$, pois caso contrário teríamos $is < it$, com s inteiro positivo, tal que $is = mmc(m,i)$ e, nesse caso, teríamos $m \mid is$, o que contradiz $\text{ord}(\alpha^i) = t$, uma vez que teríamos $(\alpha^i)^s = 1$ com $s < t$. Além disso, como

$$mi = mmc(m,i)mdc(m,i),$$

segue que

$$it = mmc(m,i) = \frac{mi}{mdc(m,i)} \Rightarrow t = \frac{mmc(m,i)}{i} = \frac{m}{mdc(m,i)}.$$

Portanto,

$$\text{ord}(\alpha^i) = \frac{m}{mdc(m,i)}.$$

□

Proposição 29. *Seja K um corpo finito com q elementos e seja $f(X)$ um polinômio irredutível em $K[X]$, de grau d. Considere o corpo $F = \dfrac{K[X]}{I(f(X))}$. Temos que*

(i) o conjunto $\mathcal{B} = \{1, [X], [X^2], \ldots, [X^{d-1}]\}$ forma uma base de F sobre K;

(ii) $[X]^{q^d} = [X]$ em F;

(iii) $f(X) \mid (X^{q^d} - X)$ em $K[X]$;

(iv) Os elementos $[X], [X]^q, \ldots, [X^{q^{d-1}}]$ de F são distintos e são as raízes de $f(X)$.

Demonstração. (i) Seja $p(X)$ um polinômio de $K[X]$, temos pelo Teorema 4 que existem únicos $q(X), r(X) \in K[X]$ tais que

$$p(X) = f(X)q(X) + r(X), \quad \text{com } r(X) = 0 \text{ ou } gr(r(X)) < gr(f(X)) = d,$$

em que $r(X)$ é um polinômio do tipo $a_0 + a_1X + \cdots + a_{d-1}X^{d-1}$, com $a_0, \ldots, a_{d-1} \in K$, assim

$$p(X) = f(X)q(X) + (a_0 + a_1X + \cdots + a_{d-1}X^{d-1}).$$

Agora, suponha que $[p(X)] \in F$, daí

$$\begin{aligned}[p(X)] &= [f(X)q(X) + (a_0 + a_1X + \cdots + a_{d-1}X^{d-1})] \\ &= [f(X)][q(X)] + [a_0 + a_1X + \cdots + a_{d-1}X^{d-1}] \\ &= [0] + [a_0 + a_1X + \cdots + a_{d-1}X^{d-1}] \\ &= [a_0] + [a_1][X] + \cdots + [a_{d-1}][X^{d-1}].\end{aligned}$$

Note que, pela Proposição 9, $K \cong \tilde{K}$, em que $\tilde{K} \subset F$ é o conjunto das classes residuais dos polinômios constantes em $K[X]$, logo cometeremos o abuso de linguagem em escrever que F contém K. Dessa forma, temos

$$\begin{aligned}[p(X)] &= [a_0] + [a_1][X] + \cdots + [a_{d-1}][X^{d-1}] \\ &= a_0 + a_1[X] + \cdots + a_{d-1}[X^{d-1}].\end{aligned}$$

Logo,

$$F = \{[p(X)]; p(X) \in K[X]\} = \{a_0 + a_1[X] + \cdots + a_{d-1}[X]^{d-1}; a_0, \ldots, a_{d-1} \in K\},$$

ou seja, o conjunto $\mathcal{B} = \{1, [X], [X^2], \ldots, [X^{d-1}]\}$ gera F.

Agora, suponha $[b_0+b_1X+\cdots+b_{d-1}X^{d-1}] = [0]$. Temos que $b_0+b_1X+\cdots+b_{d-1}X^{d-1} \in I(f(X))$, consequentemente,

$$\underbrace{b_0 + b_1X + \ldots + b_{d-1}X^{d-1}}_{\substack{gr(b_0+b_1X+\ldots+b_{d-1}X^{d-1}) \leq d-1 \\ \text{ou polinômio nulo}}} = \underbrace{f(X)h(X).}_{\substack{gr(f(X)h(X)) \geq d \\ \text{ou polinômio nulo}}}$$

Assim, se $b_0 + b_1X + \cdots + b_{d-1}X^{d-1} \neq 0$, então $f(X)h(X) \neq 0$, logo

$$d - 1 \geq gr(b_0 + b_1X + \cdots + b_{d-1}X^{d-1}) = gr(f(X)h(X)) \geq d,$$

o que é um absurdo. Assim, $b_0+b_1X+\cdots+b_{d-1}X^{d-1} = 0$ e daí, $b_0 = b_1 = \cdots = b_{d-1} = 0$. Portanto, $\mathcal{B}$ é uma base de F.

(ii) Temos que F é um espaço vetorial de dimensão d sobre K, logo F tem q^d elementos e, pela Proposição 25, vale $\alpha^{q^d} = \alpha$ para todo $\alpha \in F$. Em particular, para $[X] \in F$, portanto $[X]^{q^d} = [X]$.

(iii) Temos que $[f(X)] = [0]$, logo

$$\begin{aligned}[X^{q^d}] - [X] = 0 &\Leftrightarrow [X^{q^d} - X] = [0] \\ &\Leftrightarrow X^{q^d} - X \equiv 0 \,(mod\, I(f(X)) \\ &\Leftrightarrow X^{q^d} - X = f(X)q(X) \\ &\Leftrightarrow f(X) \mid (X^{q^d} - X).\end{aligned}$$

(iv) A demonstração dessa afirmação será feita em seis etapas. Primeiramente, considere $g(Y) = (Y - [X])(Y - [X]^q) \dots (Y - [X]^{q^{d-1}}) \in F[Y]$.

(Afirmação 1): $g(Y^q) = (g(Y))^q$.

$$\begin{aligned} g(Y^q) &= (Y - [X])(Y - [X]^q) \dots (Y - [X]^{q^{d-1}}) \\ &\overset{(ii)}{=} (Y - [X]^{q^d})(Y - [X]^q) \dots (Y - [X]^{q^{d-1}}) \\ &= (Y - [X]^{q^{d-1}})^q (Y - [X])^q \dots (Y - [X]^{q^{d-2}})^q \\ &= \left((Y - [X]^{q^{d-1}})(Y - [X]) \dots (Y - [X]^{q^{d-2}})\right)^q \\ &\overset{\text{F é corpo}}{=} \left((Y - [X])(Y - [X]^{q^{d-1}}) \dots (Y - [X]^{q^{d-2}})\right)^q \\ &= (g(Y))^q. \end{aligned}$$

(Afirmação 2): $g(Y) \in K[Y]$.

Desde que $g(Y) \in F[Y]$ é um polinômio de grau d, podemos escrevê-lo como

$$g(Y) = [b_0(X)] + [b_1(X)]Y + \dots + [b_{d-1}(X)]Y^{d-1} + Y^d.$$

Daí,

$$g(Y^q) = [b_0^q(X)] + [b_1^q(X)]Y^q + \dots + [b_{d-1}^q(X)]Y^{q(d-1)} + Y^{dq}$$

e

$$(g(Y))^q = \left([b_0 X] + [b_1 X]Y + \dots + [b_{d-1}(X)]Y^{d-1} + Y^d\right)^q.$$

Como pela Afirmação 1 $g(Y^q) = (g(Y))^q$, segue que $[b_i(X)] = [b_i^q(X)]$ para todo $i = 0, \dots, d-1$. Assim, $[b_i(X)] - [b_i^q(X)] = [0]$, logo $[b_i(X)]$ é raiz de $[Y^q] - [Y]$ para todo $i = 0, \dots, d-1$. Pela Proposição 6, $[b_i(X)] \in \tilde{K}$ para todo $i = 0, \dots, d-1$, logo $g(Y) \in \tilde{K}[Y]$. Portanto,

$$g(Y) \in \tilde{K}[X] \overset{K \cong \tilde{K}}{\Rightarrow} g(Y) \in K[Y].$$

(Afirmação 3): $mdc(f(Y), g(Y)) \neq 1$ em $K[Y]$.

Note que

$$\begin{aligned} [f(X)] &= [a_0 + a_1 X \dots + a_{d-1} X^{d-1}] \\ &= [a_0] + [a_1][X] \dots + [a_{d-1}][X^{d-1}] \\ &= a_0 + a_1[X] + \dots + a_{d-1}[X^{d-1}] \\ &= f([X]), \end{aligned}$$

então

$$[f(X)] = [0] \Leftrightarrow f([X]) = 0,$$

isto é, a classe de $f(X)$ é igual a classe do zero se, e somente se, $[X]$ é raiz de $f(Y)$. Como por hipótese $[f(X)] = [0]$, então $[X]$ é uma raiz de $f(Y)$, além disso, $g([X]) = 0$. Logo, $f(Y)$ e $g(Y)$ possuem uma raiz não nula em comum, isto é, possuem um fator não constante em comum. Portanto, $mdc(f(Y), g(Y)) \neq 1$ em $F[Y]$ e, consequentemente, $mdc(f(Y), g(Y)) \neq 1$ em $K[Y]$.

(Afirmação 4): $f(Y) = g(Y)$.

Pela Afirmação 3, temos que $mdc(f(Y), g(Y)) \neq 1$ e, como $f(Y)$ é irredutível e mônico, segue que $mdc(f(Y), g(Y)) = f(Y)$, logo $f(Y) \mid g(Y)$, ou seja, existe $q(Y) \in K[Y]$ tal que $g(Y) = f(Y)q(Y)$. Além disso, desde que $gr(g(Y)) = gr(f(Y))$, segue que $gr(q(Y)) = 0$, ou seja, $q(Y)$ é um polinômio constante, logo $f(Y)$ é associado a $g(Y)$. Por outro lado, $g(Y)$ é um polinômio mônico, portanto $f(Y) = g(Y)$.

(Afirmação 5): $g(Y) \mid (Y^{q^d} - Y)$.

Pelo item (iii) dessa proposição, temos que $f(Y) \mid (Y^{q^d} - Y)$, e pela afirmação anterior, $f(Y) = g(Y)$, logo $g(Y) \mid (Y^{q^d} - Y)$.

(Afirmação 6): $g(Y)$ não possui fatores múltiplos em uma extensão F de K.

Pela Afirmação 5 e a Proposição 24, temos que $g(Y)$ não possui fatores múltiplos irredutíveis em F extensão de K. Logo as raízes $[X], [X^q], \ldots, [X^{q^{d-1}}]$ são raízes de $f(Y)$ duas a duas distintas.

□

Observação 7. Note que, pela Proposição 29, o número d é tal que

$$d = \min\{j;\ [X^{q^j}] = [X]\}, \tag{2.37}$$

pois caso contrário, teríamos uma contradição com item (iv) dessa mesma proposição.

Proposição 30. *Seja K um corpo finito com q elementos e seja n um inteiro positivo. Temos a seguinte igualdade em $K[X]$:*

$$X^{q^n} - X = \prod_{d \mid n} G_d(X),$$

em que $G_d(X)$ é o produto de todos os polinômios mônicos irredutíveis de grau d em $K[X]$, e o produto na fórmula acima é efetuado sobre todos os inteiros positivos d que dividem n.

Demonstração. Provemos que $X^{q^n} - X = \prod_{d \mid n} G_d(X)$, para tanto, mostremos que $X^{q^n} - X \mid \prod_{d \mid n} G_d(X)$ e $\prod_{d \mid n} G_d(X) \mid X^{q^n} - X$.

Seja $f(X)$ um polinômio mônico e irredutível em $K[X]$ de grau m. Note que, por definição, $f(X)$ é um fator de $G_m(X)$. Mostremos que $G_m(X)$ é um fator de $\prod_{d\,|\,n} G_d(X)$. Temos pela Proposição 16 que

$$mdc(X^{q^n} - X, q^n X^{q^n-1} - 1) = 1,$$

logo $X^{q^n} - X$ não possui fatores múltiplos não constantes. Agora suponha que $f(X) \mid (X^{q^n} - X)$, ou seja, $f(X)$ é um fator de $X^{q^n} - X$. Pelo item (iv) da Proposição 29 $f(X)$ divide $X^{q^n} - X$, temos que se $f(X)$ divide $X^{q^m} - X$, então $f(X)$ divide $mdc(X^{q^n} - X, X^{q^m} - X)$. Daí, segue do item (ii) do Lema 5 que $f(X)$ divide $(X^{q^{mdc(m,n)}} - X)$, em que $mdc(m,n) \leq m$, mais ainda, que

$$[X^{q^{mdc(m,n)}} - X] = [f(X)q(X)] = [f(X)][q(X)] = [0] \Rightarrow [X]^{q^{mdc(m,n)}} = [X],$$

o que pelo item (iv) da Proposição 29 e por (2.37), só é possível se $mdc(m,n) \geq m$. Deste modo, $mdc(m,n) = m$, o que implica em $m \,|\, n$. Logo, $G_m(X)$ é um fator de $\prod_{d\,|\,n} G_d(X)$ e, portanto, $f(X)$ é um fator de $\prod_{d\,|\,n} G_d(X)$, consequentemente, $X^{q^n} - X \mid \prod_{d\,|\,n} G_d(X)$.

Reciprocamente, suponha que $f(X)$ é um fator de $\prod_{d\,|\,n} G_d(X)$. Como $f(X)$ é irredutível e mônico, segue que $f(X)$ é um fator de $G_{d'}(X)$ para algum inteiro positivo d' tal que $d' \,|\, n$. Assim,

$$m = gr(f(X)) = d' \Rightarrow m \,|\, n.$$

Deste modo, $(X^{q^m} - X) \mid (X^{q^n} - X)$ e, como $f(X) \mid (X^{q^m} - X)$ (item (iii) da Proposição 29), segue que $f(X) \mid (X^{q^n} - X)$. Como cada $G_d(X)$ é o produto de todos os fatores mônicos irredutíveis de grau d em $F[X]$, cujos fatores aparecem uma única vez, então $\prod_{d\,|\,n} G_d(X)$ é formado por fatores mônicos irredutíveis de grau d, com $d \,|\, n$, em que cada fator tem multiplicidade 1. Logo, $\prod_{d\,|\,n} G_d(X) \mid X^{q^n} - X$. Portanto $X^{q^n} - X = \prod_{d\,|\,n} G_d(X)$. □

Corolário 7. *Seja $I(n)$ o número de polinômios mônicos irredutíveis de grau n em $K[X]$, em que K é um corpo finito com q elementos. Temos que*

$$q^n = \sum_{d\,|\,n} dI(d).$$

Demonstração. Temos que para cada $G_j(X)$, o grau de $G_j(X)$ é $jI(j)$. Deste modo, para cada divisor d_j de n,

$$gr(G_{d_j}(X)) = d_j I(d_j).$$

Logo,

$$q^n = gr(X^{q^n} - X) \overset{\text{Prop. } 30}{=} gr\left(\prod_{d\,|\,n} G_d(X)\right) = \sum_{d\,|\,n} gr(G_d(X)) = \sum_{d\,|\,n} dI(d).$$

□

Teorema 12. *Seja K um corpo finito qualquer. Para cada número natural n, existe pelo menos um polinômio irredutível de grau n em K[X].*

Demonstração. Se $n = 1$, temos que $p(X) = X$ é um polinômio irredutível em $K[X]$. Seja $n > 1$ e sejam $1 = d_1 < d_2 < \ldots < d_s < n$, os divisores de n, com $s \geq 1$. Tomando $|K| = q$, temos pelo Corolário 7 que

$$\begin{aligned} q^n &= \sum_{d \mid n} dI(d) = \sum_{j=1}^{s} d_j I(d_j) + nI(n) \\ &= d_1 I(d_1) + \cdots + d_s I(d_s) + nI(n) \\ &\leq \sum_{d \mid d_1} dI(d) + \cdots + \sum_{d \mid d_s} dI(d) + nI(n) \\ &= \sum_{j=1}^{s} \left(\sum_{d \mid d_j} dI(d) \right) + nI(n) \\ &= \sum_{j=1}^{s} q^{d_j} + nI(n). \end{aligned}$$

Daí, escrevendo em sequência todos os naturais do intervalo $[d_j, d_{j+1}]$, temos

$$d_j < a_{j1} < \ldots < a_{jm_j} < d_{j+1},$$

incluindo aqueles que não são divisores de n, a fim de obtermos uma progressão geométrica de razão q e primeiro termo igual a 1, do seguinte modo:

$$q^0 + q^{a_{11}} + \cdots + q^{a_{1m_1}} + q^{d_1} + \cdots + q^{a_{s1}} + \cdots + q^{a_{sm_s}} + q^{d_s} = \sum_{j=0}^{d_s} q^j.$$

Assim, seque que

$$q^n \leq \sum_{j=1}^{s} q^{d_j} + nI(n) \quad < \quad \sum_{j=0}^{d_s} q^j + nI(n) = \frac{q^{d_s+1} - 1}{q - 1} + nI(n) < q^{d_s+1} + nI(n).$$

Desde que $d_s \mid n$, com $d_s < n$, temos que $n = \lambda d_s$ com $\lambda > 1$. Com isso,

$$\begin{aligned} d_s = \frac{n}{\lambda} \leq \frac{n}{2} &\Rightarrow q^{d_s+1} \leq q^{\frac{n}{2}+1} \\ &\Rightarrow -q^{d_s+1} \geq -q^{\frac{n}{2}+1} \\ &\Rightarrow -q^{d_s+1} + q^n \geq -q^{\frac{n}{2}+1} + q^n \\ &\Rightarrow nI(n) > -q^{d_s+1} + q^n \geq -q^{\frac{n}{2}+1} + q^n \\ &\Rightarrow nI(n) > q^n(1 - q^{-\frac{n}{2}+1}). \end{aligned}$$

Como estamos analisando o caso em que $n > 1$, temos que $1 - \frac{n}{2} \leq 0$, donde

$$(1 - q^{-\frac{n}{2}+1}) \geq (1 - q^0) = 0.$$

Portanto,

$$nI(n) > 0 \overset{n>1}{\Rightarrow} I(n) > 0.$$

□

Teorema 13. *Dois corpos finitos com o mesmo número de elementos são isomorfos.*

A demonstração desse resultado por ser vista em (HEFEZ; VILLELA, 2008, páginas 76-77, capítulo 4).

Definição 41. *Um elemento α de um corpo finito $\mathbb{F}_q$ é chamado de elemento primitivo quando*

$$\mathbb{F}_q^* = \{1, \alpha, \alpha_2, \ldots, \alpha^{q-2}\},$$

isto é, se $ord(\alpha) = q - 1$.

Teorema 14. *Todo corpo finito possui elementos primitivos.*

A demonstração desse resultado encontra-se em (HEFEZ; VILLELA, 2008, página 78, capítulo 4).

O Teorema 14 garante que todo corpo finito com q elementos possui elementos primitivos, o que significa que deve existir um elemento α de tal modo que $\alpha^{q-1} = 1$. Note que para calcular o produto entre dois elementos α^i e α^j de $\mathbb{F}_q^*$, fazemos

$$\alpha^i \cdot \alpha^j = \alpha^{[i+j]},$$

sendo $[i + j]$ o resto da divisão de $i + j$ por $q - 1$. No entanto surge uma pergunta: como fazemos para somar esses elementos? Observe que, se $i < j$, então

$$\alpha^i + \alpha^j = \alpha^i(1 + \alpha^{j-i})$$

e, como sabemos multiplicar elementos de $\mathbb{F}_q^*$, para obtermos o resultado dessa soma é necessário que determinemos o inteiro $z(j - i)$ tal que $1 + \alpha^{j-i} = \alpha^{z(j-i)}$. Assim,

$$\alpha^i + \alpha^j = \alpha^i \cdot \alpha^{z(j-i)}.$$

De modo geral, para somarmos potências de um elemento primitivo α de $\mathbb{F}_q^*$, calculamos os valores dos $z(r)$ tais que $1 + \alpha^r = \alpha^{z(r)}$, para $r = 1, \ldots q - 2$, e os organizamos em uma tabela chamada de tabela logarítmica de Zech. O exemplo a seguir deixa evidente como construir essa tabela e como efetuar a soma de potências de um dado elemento primitivo de $\mathbb{F}_q^*$ e será imprescindível no exemplo de código BCH que construiremos no Capítulo 6.

Exemplo 4. *Determinaremos a tabela de Zech do corpo $\mathbb{F}_{16}$, construído a partir do polinômio irredutível $X^4 + X + 1$ sobre $\mathbb{F}_2[X]$ e consideraremos o elemento primitivo $\alpha = [X]$ de $\mathbb{F}_{16}^*$.*

Primeiramente, determinamos as representações dos elementos de $\mathbb{F}_{16}^*$ em termos do elemento primitivo $\alpha = [X]$ e da base $\{1, \alpha, \alpha^2, \alpha^3\}$ de $\mathbb{F}_{16} = \dfrac{\mathbb{F}_2[X]}{X^4 + x + 1}$ sobre $\mathbb{F}_2$ da seguinte maneira:

$$
\begin{array}{ll}
1 & \alpha^8 = \alpha^4 + \alpha^2 + \alpha = \alpha^2 + 1 \\
\alpha & \alpha^9 = \alpha^3 + \alpha \\
\alpha^2 & \alpha^{10} = \alpha^4 + \alpha^2 = \alpha^2 + \alpha + 1 \\
\alpha^3 & \alpha^{11} = \alpha^3 + \alpha^2 + \alpha \\
\alpha^4 = \alpha + 1 & \alpha^{12} = \alpha^4 + \alpha^3 + \alpha^2 = \alpha^3 + \alpha^2 + \alpha + 1 \\
\alpha^5 = \alpha^2 + \alpha & \alpha^{13} = \alpha^4 + \alpha^3 + \alpha^2 + \alpha = \alpha^3 + \alpha^2 + 1 \\
\alpha^6 = \alpha^3 + \alpha^2 & \alpha^{14} = \alpha^4 + \alpha^3 + \alpha = \alpha^3 + 1 \\
\alpha^7 = \alpha^4 + \alpha^3 = \alpha^3 + \alpha + 1 & \alpha^{15} = \alpha^4 + \alpha = 1
\end{array}
$$

Agora observe que conseguimos determinar o inteiro $z(r)$ tal que $1 + \alpha^r = \alpha^{z(r)}$ para cada $r = 1, \ldots, 14$. Com isso, obtemos a seguinte tabela de Zech:

r	$z(r)$	r	$z(r)$
1	4	8	2
2	8	9	7
3	14	10	5
4	1	11	12
5	10	12	11
6	13	13	6
7	9	14	3

Portanto, conseguimos agora somar potências do elemento primitivo $\alpha = [X]$ de $\mathbb{F}_{16}^*$, como por exemplo,

$$\alpha^5 + \alpha^9 = \alpha^5 \cdot \alpha^{z(4)} = \alpha^5 \cdot \alpha = \alpha^6.$$

2.4.1 Extensões de Corpos Finitos

Nesta seção apresentaremos alguns resultados referentes a extensões de corpos finitos, como polinômios minimais e corpos de frações, que serão muito utilizados no Capítulo 6.

Proposição 31. *Seja p um número primo positivo. Um corpo finito F com p^n elementos contém um subcorpo K com p^m elementos se, e somente se, $m \mid n$. Nesse caso, existe um único subcorpo com a referida propriedade, e seus elementos são as raízes em F da equação $X^{p^m} - X = 0$.*

Demonstração. Suponha que F e K sejam corpos finitos tais que $|F| = p^n$ e $|K| = p^m$, para $n, m \in \mathbb{N}$. Temos pelo Corolário 6 que os elementos de F são as raízes de $X^{p^n} - X$ e os elementos de K são as raízes de $X^{p^m} - X$. Se K é um subcorpo de F, então novamente pelo Corolário 6, os elementos de K são os elementos de F que são raízes de $X^{p^n} - X$ e, como todos os elementos de K são as raízes de $X^{p^m} - X$, segue que esse polinômio é um fator de $X^{p^n} - X$, logo $(X^{p^m} - X) \mid (X^{p^n} - X)$ e portanto $m \mid n$.
Reciprocamente, se $m \mid n$, então $(X^{p^m} - X) \mid (X^{p^n} - X)$, logo as raízes de $X^{p^m} - X$ são também raízes de $X^{p^n} - X$. Como os elementos de K são as raízes de $X^{p^m} - X$, segue que os elementos de K são também as raízes de $X^{p^n} - X$, consequentemente, são elementos de F. Portanto K é um subcorpo de F.

A unicidade do corpo K é uma consequência imediata do Teorema 13.

□

Corolário 8. *Seja q uma potência inteira de um número primo p e sejam m e n inteiros positivos. O corpo $\mathbb{F}_{q^n}$ contém um subcorpo isomorfo a $\mathbb{F}_{q^m}$ se, e somente se, $m \mid n$. Nesse caso, tal subcorpo é único, e seus elementos são as raízes de $X^{q^m} - X$ em $\mathbb{F}_{q^n}$.*

Demonstração. Seja $q = p^r$, então $q^n = p^{rn}$ e $q^m = p^{rm}$. Como $rm \mid rn$ se, e somente se, $m \mid n$, segue da Proposição 31 que $\mathbb{F}_{q^m}$ é um subcorpo de $\mathbb{F}_{q^n}$ se, e somente se, $m \mid n$. Portanto $\mathbb{F}_{q^m}$ é um subcorpo de $\mathbb{F}_{q^n}$. □

Antes do próximo resultado, convém observarmos que, sendo F um corpo finito e uma extensão de K, temos que F é um espaço vetorial de dimensão finita n sobre K. Além disso, seja $\beta \in F$, então o conjunto $\{1, \beta, \beta^2, \ldots, \beta^n\}$ é linearmente dependente. Logo existem $a_i \in K$, com $0 \leq i \leq n$, nem todos nulos, tais que

$$a_0 + a_1\beta + \ldots + a_n\beta^n = 0.$$

Deste modo, β é uma raiz do polinômio $f(X) = a_0 + a_1X + \ldots + a_nX^n \in K[X] \setminus \{0\}$, ou seja, $f(\beta) = 0$. Por consequência disso, o conjunto

$$J_\beta = \{p(X) \in K[X];\ p(\beta) = 0\} \neq \{0\}.$$

Proposição 32. *O conjunto definido como*

$$J_\beta = \{p(X) \in K[X];\ p(\beta) = 0\}$$

é um ideal de $K[X]$.

Demonstração. De fato, sejam $p(X), q(X) \in J_\beta$, temos que

$$(p-q)(\beta) = (p-q)(\beta) = p(\beta) - q(\beta) = 0, \quad \forall\ p(X), q(X) \in J_\beta.$$

Logo, $(p-q)(X) \in J_\beta$ para todo $p(X), q(X) \in J_\beta$. Além disso, sejam $p(X) \in J_\beta$ e $h(X) \in K[X]$, temos que

$$(hp)(\beta) = h(\beta)p(\beta) = h(\beta) \cdot 0 = 0, \quad \forall\ p(X) \in J_\beta,\ h(X) \in K[X].$$

Logo, $(hp)(X) \in J_\beta$ para todo $p(X) \in J_\beta$, $h(X) \in K[X]$. Portanto J_β é um ideal de $K[X]$. □

Segue do Corolário 3 que $J_\beta = I(m_\beta(X))$ para um único polinômio mônico $m_\beta(X)$, que chamaremos de *polinômio mínimo* ou *polinômio minimal* de β sobre K.

Proposição 33. *Sejam K e F corpos finitos tais que F é uma extensão de K. Seja $p(X) \in K[X]$ um polinômio mônico e β um elemento de F. As seguintes condições são equivalentes:*

(i) $p(X)$ *é o polinômio mínimo de* β*;*

(ii) $p(X)$ *é um polinômio de menor grau em* $K[X]$ *tal que* $p(\beta) = 0$*;*

(iii) $p(X)$ *é irredutível e* $p(\beta) = 0$*.*

Demonstração. $(i) \Rightarrow (ii)$: Seja $f(X) \in K[X]$ tal que $f(\beta) = 0$, temos que $f(X) \in J_\beta = I(p(X))$, ou seja, existe $q(X) \in K[X]$ tal que $f(X) = q(X)p(X)$, consequentemente, $gr(f(X)) \geq gr(p(X))$. Portanto, não existe polinômio de grau menor que $p(X)$ que se anula em β, ou seja, $p(X)$ é um polinômio de menor grau em $K[X]$ tal que $p(\beta) = 0$.
$(ii) \Rightarrow (iii)$: Suponha por absurdo que $p(X)$ é um polinômio redutível em $K[X]$. Assim, segue que existem $q(X), h(X) \in K[X]$ não invertíveis, ambos com grau menor que $gr(p(X))$, tais que $p(X) = q(X)h(X)$. Logo

$$0 = p(\beta) = q(\beta)h(\beta).$$

Como K é um domínio de integridade, então $q(\beta) = 0$ ou $h(\beta) = 0$. Contradição com o fato de $p(X)$ ser um polinômio de menor grau que se anula em β. Portanto $p(X)$ é irredutível em $K[X]$.
$(iii) \Rightarrow (i)$: Seja $p(X) \in K[X]$, irredutível e mônico, tal que $p(\beta) = 0$. Então, $p(X) \in J_\beta = I(m_\beta(X))$. Logo existe $q(X) \in K[X]$ tal que $p(X) = q(X)m_\beta(X)$, consequentemente, $m_\beta(X) \mid p(X)$. Como $p(X)$ é irredutível, então pela Definição 32 $q(X)$ é invertível em K, ou seja, $p(X) = qm_\beta(X)$ com $q \in K \setminus \{0\}$. Além disso, como $p(X)$ é mônico, segue que $q = 1$, isto é, $p(X) = m_\beta(X)$. Portanto, $p(X)$ é o polinômio mínimo de β. □

Concluímos com a Proposição 33 que existe um único polinômio mônico $m_\beta(X)$ que satisfaz todas as suas condições equivalentes.

Proposição 34. *Sejam F uma extensão de K e $\beta \in F$. O conjunto*

$$K(\beta) = \left\{ \frac{p(\beta)}{q(\beta)};\ p(X), q(X) \in K[X], q(\beta) \neq 0 \right\}$$

é um subcorpo de F contendo β e K, contido em qualquer subcorpo de F que tem essa propriedade.

Demonstração. Sejam $\frac{p_1(\beta)}{q_1(\beta)}, \frac{p_2(\beta)}{q_2(\beta)} \in K(\beta)$, com $\frac{p_2(\beta)}{q_2(\beta)} \neq 0$. Temos que

$$\frac{p_1(\beta)}{q_1(\beta)} - \frac{p_2(\beta)}{q_2(\beta)} = \frac{p_1(\beta)q_2(\beta) - p_2(\beta)q_1(\beta)}{q_1(\beta)q_2(\beta)} \in K(\beta),$$

pois $p_1(X)q_2(X) - p_2(X)q_1(X)$ e $q_1(X)q_2(X)$ são elementos de $K[X]$, com $q_1(X)q_2(X) \neq 0$. Além disso,

$$\frac{p_1(\beta)}{q_1(\beta)} \cdot \left(\frac{p_2(\beta)}{q_2(\beta)}\right)^{-1} = \frac{p_1(\beta)}{q_1(\beta)} \cdot \frac{q_2(\beta)}{p_2(\beta)} = \frac{p_1(\beta)q_2(\beta)}{q_1(\beta)p_2(\beta)} \in K(\beta),$$

pois $p_1(X)q_2(X), q_1(X)p_2(X) \in K[X]$, com $q_1(X)p_2(X) \neq 0$. Logo $K(\beta)$ é um subcorpo de F contendo β e K.

Agora mostremos que se existir um subcorpo E de F que contém β e K, então $K(\beta)$ é um subcorpo de E. Primeiramente, mostremos que $\frac{p(\beta)}{q(\beta)} \in E$ para todo $p(X), q(X) \in K[X]$. De fato, seja $p(\beta) = a_0 + a_1\beta + \ldots + a_r\beta^r \in K[X]$. Como $a_0, \ldots, a_r \in K \subset E$ e $\beta^r \in E$, para todo $r \in K$ (pois $\beta \in E$), então $p(\beta) \in E$. Similarmente, $q(\beta) \in E$ e, como $q(\beta) \neq 0$, então ele é invertível em E, ou seja, $\frac{1}{q(\beta)} \in E$. Logo $\frac{p(\beta)}{q(\beta)} \in E$, consequentemente, $K(\beta) \subset E$ para todo $p(X), q(X) \in K[X]$. Portanto, $K(\beta)$ é o menor subcorpo de F que contém β e K. □

Proposição 35. *Sejam F um corpo finito, K um subcorpo de F e $\beta \in F$.*

(i) Se $m = gr(m_\beta(X))$, então $\{1, \beta, \ldots, \beta^{m-1}\}$ é uma base de $K(\beta)$ sobre K. Em particular, $dim_K K(\beta) = gr(m_\beta(X))$.

(ii) Se $q = |K|$, então $\beta^{q^m} = \beta$ e $\beta^{q^i} \neq \beta^{q^j}$ para $i \neq j$, $i, j = 0, \ldots, m-1$. Além disso,

$$m_\beta(X) = (X - \beta)(X - \beta^q) \ldots (X - \beta^{q^{m-1}}).$$

(iii) Existe $\alpha \in F$ tal que $F = K(\alpha)$.

Demonstração. (i) Seja $m_\beta(X)$ o polinômio minimal de β sobre K. Se $gr(m_\beta(X)) = m \geq 1$, então $m_\beta(X) = a_0 + a_1X + \ldots + a_mX^m \in K[X]$, com $a_m \neq 0$. Daí,

$$0 = m_\beta(\beta) = a_0 + a_1\beta + \ldots + a_m\beta^m,$$

logo $\{1, \beta, \ldots, \beta^m\}$ é linearmente dependente. Por outro lado, $\{1, \beta, \ldots, \beta^{m-1}\}$ é linearmente independente, pois caso contrário, teríamos $f(X) = b_0 + b_1X + \ldots + b_{m-1}X^{m-1} \in K[X]$, com $b_{m-1} \neq 0$, tal que

$$0 = f(\beta) = b_0 + b_1\beta + \ldots + b_{m-1}\beta^{m-1},$$

o que contradiz a minimalidade do grau de $m_\beta(X)$. Portanto $\{1, \beta, \ldots, \beta^{m-1}\}$ é linearmente independente.

Mostremos agora que esse conjunto gera $K(\beta)$. Para isso, dado $\dfrac{p(\beta)}{q(\beta)} \in K(\beta)$, temos que $q(\beta) \neq 0$. Como $m_\beta(X)$ é um polinômio mônico e irredutível, então

$$mdc(q(X), m_\beta(X)) = 1 \quad \text{ou} \quad mdc(q(X), m_\beta(X)) = m_\beta(X).$$

Se $mdc(q(X), m_\beta(X)) = m_\beta(X)$, então $q(X) = m_\beta(X)g(X)$, consequentemente, $q(\beta) = m_\beta(\beta)g(\beta) = 0$. Contradição com fato de $\dfrac{p(\beta)}{q(\beta)} \in K(\beta)$, com $q(\beta) \neq 0$. Logo, $mdc(q(X), m_\beta(X)) = 1$. Além disso, temos pela Proposição 6, que existem $s(X), t(X) \in K[X]$ tais que

$$s(X)q(X) + t(X)m_\beta(X) = 1,$$

logo

$$s(\beta)q(\beta) + t(\beta)m_\beta(\beta) = 1 \Leftrightarrow s(\beta)q(\beta) = 1 \Leftrightarrow s(\beta) = \frac{1}{q(\beta)}.$$

Deste modo, temos $\dfrac{p(\beta)}{q(\beta)} = p(\beta)s(\beta)$ e, pelo Teorema 4, existem únicos $h(X), r(X) \in K[X]$ tais que

$$p(X)s(X) = h(X)m_\beta(X) + r(X), \quad \text{com } r(X) = 0 \text{ ou } gr(r(X)) < gr(m_\beta(X)) = m.$$

Assim,

$$p(\beta)s(\beta) = h(\beta)m_\beta(\beta) + r(\beta) = r(\beta).$$

Consequentemente,

$$\frac{p(\beta)}{q(\beta)} = p(\beta)s(\beta) = r(\beta) = c_0 + c_1\beta + \ldots + c_{m-1}\beta^{m-1},$$

em que $c_0, c_1, \ldots, c_{m-1} \in K$. Logo, $\{1, \beta, \ldots, \beta^{m-1}\}$ é um conjunto gerador de $K(\beta)$ e, portanto, é uma base de $K(\beta)$.

(ii) Seja φ a aplicação

$$\begin{array}{rccc} \varphi: & \dfrac{K[X]}{I(m_\beta(X))} & \to & K(\beta) \\ & [a_0 + a_1X + \ldots + a_{m-1}X^{m-1}] & \mapsto & a_0 + a_1\beta + \ldots + a_{m-1}\beta^{m-1} \end{array}.$$

Temos que, pelo item (i) da Proposição 29, o conjunto $\{1, [X], \ldots, [X^{m-1}]\}$ é uma base para o espaço vetorial $\dfrac{K[X]}{I(m_\beta(X))}$, que possui, como consequência, dimensão $m-1$. Por outro lado, vimos no item anterior que, o conjunto $\{1, \beta, \ldots, \beta^{m-1}\}$ é uma base para o espaço vetorial $K(\beta)$ e este, por sua vez, possui dimensão $m-1$. Logo, a aplicação φ é um isomorfismo linear. Além disso, φ é um homomorfismo, pois

$$\varphi([p+q](X)) = \big(p+q\big)(\beta) = p(\beta) + q(\beta) = \\ = \varphi([p(X)]) + \varphi([q(X)]), \quad \forall\ [p(X)], [q(X)] \in \frac{K[X]}{I(m_\beta(X))}$$

e

$$\varphi([pq](X)) = \big(pq\big)(\beta) = p(X)\cdot q(X) = \\ = \varphi([p(X)])\cdot\varphi([q(X)]), \quad \forall\ [p(X)], [q(X)] \in \frac{K[X]}{I(m_\beta(X))}.$$

Portanto, a aplicação φ é um isomorfismo de corpos e, consequentemente, $\varphi([X]^i) = \beta^i$ para todo i. Observe que, pelo item (iv) da Proposição 29, temos que $\dfrac{K[X]}{I(m_\beta(X))}$ é um corpo, cujas raízes $[X], [X]^q, \ldots, [X]^{q^{m-1}}$ são duas a duas distintas e, mais ainda, pelo item (ii) da Proposição 29, temos $[X]^{q^m} = [X]$. Portanto, $\beta^{q^m} = \beta$ e $\beta^{q^i} \neq \beta^{q^j}$ quando $i \neq j$, para todo $i, j = 0, \ldots, m-1$.

Mostremos agora que $m_\beta(X) = (X-\beta)(X-\beta^q)\ldots(X-\beta^{q^{m-1}})$. Para isso, seja

$$p(X) = (X-\beta)(X-\beta^q)\ldots(X-\beta^{q^{m-1}}) \in K(\beta),$$

tal que $p(\beta) = 0$ e $p(X)$ mônico.

Temos, pelo Lema 6, que

$$p(X) = \prod_{j=0}^{m-1}(X-\beta^{q^j}) = \sum_{j=0}^{m-1}(-1)^j S_j(1, \beta, \ldots, \beta^{q^{m-1}})X^{m-j-1},$$

em que $(-1)^j S_j(1, \beta, \ldots, \beta^{q^{m-1}}) = a_j \in K(\beta)$, sendo a_j os coeficientes do polinômio $p(X)$. Como

$$S_j(\beta, \ldots, \beta^{q^{m-1}})^q = S_j(\beta^q, \ldots, \beta) = S_j(\beta, \ldots, \beta^{q^{m-1}}),$$

então $a_j^q = a_j$, isto é, a_j é raiz de $X^q - X$, assim $a_j \in K$, logo $p(X) \in K[X]$. Por consequência disso, $m_\beta(X) \mid p(X)$ e, como $gr(p(X)) = gr(m_\beta(X)) = m$, então $p(X)$ é associado a $m_\beta(X)$. Além disso, $p(X)$ e $m_\beta(X)$ são mônicos, Portanto $p(X) = m_\beta(X)$.

(iii) Tome $\alpha \in F$ um elemento primitivo, o que é possível em vista do Teorema 14. Como $\alpha \in F$ e $F \supset K$, segue que $K(\alpha) \subset F$. Por outro lado, como α é um elemento primitivo de F, então para todo $b \neq 0 \in F$ temos que existe $j \in \mathbb{N}$ tal que $b = \alpha^j$, logo $b \in K(\alpha)$, pois $\alpha^j \in K(\alpha)$ para todo $j \in \mathbb{N}$, isto é, $F \subset K(\alpha)$. Portanto $F = K(\alpha)$.

□

2.4.2 Raízes da Unidade

Nesta seção veremos alguns resultados relativos às raízes da unidade, que serão essenciais no Capítulo 6, visto que definiremos um código cíclico através das raízes de um polinômio $g(X)$ em alguma extensão F de um corpo K.

Teorema 15. *Sejam F um corpo finito e, n, um inteiro positivo que divide $|F| - 1$. Então, existe um elemento $\gamma \in F$ tal que*

$$X^n - 1 = (X - \gamma^0)(X - \gamma)(X - \gamma^2) \dots (X - \gamma^{n-1}),$$

com $1, \gamma, \gamma^2, \dots, \gamma^{n-1}$ dois a dois distintos.

Demonstração. Seja $\alpha \in F$ um elemento primitivo, cuja existência é garantida pelo Teorema 14. Então

$$\alpha^{|F|-1} = 1 \quad \text{e} \quad \alpha^m \neq 1;\ 0 \leq m \leq |F| - 1.$$

O caso em que $n = 1$ é trivial, sendo assim, suponha $n \geq 2$. Definindo $\gamma = \alpha^{\frac{|F|-1}{n}}$, temos que $\gamma^0, \gamma^1\gamma^2, \dots, \gamma^{n-1}$ são raízes de $X^n - 1$. Com efeito, considerando $\gamma^i = \left(\alpha^{\frac{|F|-1}{n}}\right)^i$, temos que

$$(\gamma^i)^n - 1 = \gamma^{in} - 1 = \alpha^{\left(\frac{|F|-1}{n}\right)in} - 1 = \alpha^{(|F|-1)i} - 1 = 0,$$

para todo $i = 1, \dots, n-1$. Falta mostrarmos que $1, \gamma, \gamma^2, \dots, \gamma^{n-1}$ são dois a dois distintos e, para tanto, suponhamos por absurdo que $\gamma^i = \gamma^j$, com $i \neq j$. Sem perda de generalidade, podemos supor $j > i$. Com isso, temos

$$\alpha^{\left(\frac{|F|-1}{n}\right)j} = \alpha^{\left(\frac{|F|-1}{n}\right)i} \Rightarrow \alpha^{\left(\frac{|F|-1}{n}\right)(j-i)} = 1.$$

Como $|F| - 1$ é o menor valor tal que $\alpha^{|F|-1} = 1$, então

$$\left(\frac{|F|-1}{n}\right)(j-i) \geq |F| - 1 \Leftrightarrow (|F|-1)(j-i) \geq n(|F|-1) \Leftrightarrow (j-i) \geq n.$$

Absurdo, pois $0 \leq (j-i) \leq n$. Logo $1, \gamma, \gamma^2, \dots, \gamma^{n-1}$ são dois a dois distintos. □

Corolário 9. *Sejam K um corpo finito com q elementos e, n, um inteiro co-primo com q. Então, existem uma extensão F de K e um elemento $\gamma \in F$ tais que*

$$X^n - 1 = (X - \gamma^0)(X - \gamma)(X - \gamma^2)\dots(X - \gamma^{n-1}),$$

com $1, \gamma, \gamma^2, \dots, \gamma^{n-1}$ dois a dois distintos.

Demonstração. Temos que, se o grau da extensão F de K é $n = 1$, então $F = K$ e $\gamma = 1$. Suponha $n \geq 2$, com n e q co-primos. Daí, a classe residual de q módulo n, ou seja, $[q]$, é invertível em $\mathbb{Z}_n$, além disso, como $\mathbb{Z}_n$ é um conjunto finito, o conjunto $\mathcal{A} = \{[q], [q]^2, [q]^3, \dots\} \subset \mathbb{Z}_n$ é também um conjunto finito, logo possui repetições. Como $[q] \in \mathbb{Z}_n$, então $[q]^i \in \mathbb{Z}_n$ e, pelo fato de $[q]$ ser um elemento invertível em $\mathbb{Z}_n$, segue que $[q]^{-1} \in \mathbb{Z}_n$, assim

$$[q]^{-i} \in \mathbb{Z}_n \Rightarrow ([q]^{-1})^i \in \mathbb{Z}_n.$$

Como o conjunto $\mathcal{A}$ possui repetições, então existem i e j distintos (sem perda de generalidade, $j > i$), tais que

$$[q]^j = [q]^i \Leftrightarrow [q]^{(j-i)} = 1,$$

logo existe $m = j - i$ tal que $[q]^m = 1$. Seja F uma extensão do corpo K com q^m elementos. Como $[q]^m = 1$ em $\mathbb{Z}_n$, segue que

$$q^m \equiv 1\,(mod\, n) \Leftrightarrow n \mid (q^m - 1) \Leftrightarrow n \mid (\mid F \mid - 1).$$

Logo, temos as hipóteses do Teorema 15 e, portanto, vale o resultado. □

Observe que, se F é uma extensão de K com q^k elementos e $k < m$, então pelo Corolário 9, $[q]^k = 1$ em $\mathbb{Z}_n$. Daí,

$$q^k \equiv 1\,(mod\, n) \Leftrightarrow n \mid (q^k - 1) \Leftrightarrow n \mid (\mid F \mid - 1),$$

o que implica pelo Teorema 15 que $X^n - 1$ se fatora como produto de fatores lineares em um corpo F com q^k elementos. Com isso, podemos escolher m o menor elemento tal que $[q]^m = 1$ e, consequentemente, o corpo F com q^m elementos é o menor corpo onde o polinômio $X^n - 1$ se fatora como produto de fatores lineares.

2.4.3 Polinômios q-lineares

Definição 42. *Um polinômio q-linear sobre $F = \mathbb{F}_{q^m}$ é um polinômio mônico da forma*

$$L(X) = \sum_{i=0}^{r} l_i X^{q^i} = l_0 X^{q^0} + l_1 X^{q^1} + l_2 X^{q^2} + \ldots + l_r X^{q^r} \in F[X]. \tag{2.38}$$

Exemplo 5. *O polinômio* $L(X) = X^8 + X^4 + X^2 + X$ *é 2-linear sobre* $\mathbb{F}_2$ *e o polinômio* $T(X) = X^9 + X$ *é 3-linear sobre* $\mathbb{F}_3$.

Um polinômio q-linear pode ser visto como uma aplicação linear L, definida como

$$\begin{array}{rccc} L: & \mathbb{F}_{q^m} & \to & \mathbb{F}_{q^m} \\ & x & \mapsto & L(x) = \sum_{i=0}^{r} l_i x^{q^i} \end{array}, \tag{2.39}$$

que associa um elemento de $\mathbb{F}_{q^m}$ com um polinômio q-linear em $\mathbb{F}_{q^m}$. Com efeito,

$$\begin{aligned} L(x+\lambda y) &= \sum_{i=0}^{r} l_i (x+\lambda y)^{q^i} \\ &= \sum_{i=0}^{r} l_i (x^{q^i} + \lambda^{q^i} y^{q^i}) \\ &\overset{\text{Prop 25}}{=} \sum_{i=0}^{r} (l_i x^{q^i} + l_i \lambda y^{q^i}) \\ &= \sum_{i=0}^{r} l_i x^{q^i} + \lambda \sum_{i=0}^{r} l_i y^{q^i} \\ &= L(x) + \lambda L(y). \end{aligned}$$

Logo L é linear. Além disso, se β_j com $j = 1, \ldots, q^r$ são raízes do polinômio $L(X) = \sum_{i=0}^{r} l_i X^{q^i}$, então $L(\beta_j) = 0$ para todo $j = 1, \ldots, q^r$. Como L é uma aplicação linear, segue que $\beta_j \in Ker(L)$ para todo $j = 0, \ldots, q^r$. Por outro lado, para todo $\beta \in Ker(L)$ temos que $L(\beta) = 0$, logo β é também uma raiz de $L(X)$ ou seja $\beta = \beta_j$ para algum $j = 1, \ldots, q^r$. Portanto, as raízes do polinômio $L(X)$ formam o subespaço $Ker(L)$.

Lema 9. *Seja* V *um* $\mathbb{F}_q$*-subespaço vetorial de* $\mathbb{F}_{q^m}$*, logo*

$$L(X) = \prod_{\beta \in V} (X - \beta) \tag{2.40}$$

é um polinômio mônico q*-linear sobre* $\mathbb{F}_q$*, de grau* q^h*, em que* $h = dim_{\mathbb{F}_q} V$.

Demonstração. Considere $\mathcal{B} = \{\beta_0, \ldots, \beta_{h-1}\}$ uma base de V sobre $\mathbb{F}_q$. Temos que

$$det \begin{bmatrix} \beta_0 & \beta_0^q & \beta_0^{q^2} & \ldots & \beta_0^{q^{h-1}} \\ \beta_1 & \beta_1^q & \beta_1^{q^2} & \ldots & \beta_1^{q^{h-1}} \\ \vdots & \vdots & \vdots & & \vdots \\ \beta_{h-1} & \beta_{h-1}^q & \beta_{h-1}^{q^2} & \ldots & \beta_{h-1}^{q^{h-1}} \end{bmatrix} \neq 0,$$

de acordo com o Exemplo 3. Então, existe uma única solução $l_0, \ldots, l_{h-1} \in \mathbb{F}_{q^h}$ para o sistema, como se segue

$$\begin{bmatrix} \beta_0 & \beta_0^q & \beta_0^{q^2} & \ldots & \beta_0^{q^{h-1}} \\ \beta_1 & \beta_1^q & \beta_1^{q^2} & \ldots & \beta_1^{q^{h-1}} \\ \vdots & \vdots & \vdots & & \vdots \\ \beta_{h-1} & \beta_{h-1}^q & \beta_{h-1}^{q^2} & \ldots & \beta_{h-1}^{q^{h-1}} \end{bmatrix} \cdot \begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \\ \vdots \\ Y_{h-1} \end{bmatrix} = \begin{bmatrix} -\beta_0^{q^h} \\ -\beta_1^{q^h} \\ -\beta_2^{q^h} \\ \vdots \\ -\beta_{h-1}^{q^h} \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \beta_0 Y_0 + \beta_0^q Y_1 + \beta_0^{q^2} Y_2 + \cdots + \beta_0^{q^{h-1}} Y_{h-1} &= -\beta_0^{q^h} \\ \beta_1 Y_0 + \beta_1^q Y_1 + \beta_1^{q^2} Y_2 + \cdots + \beta_1^{q^{h-1}} Y_{h-1} &= -\beta_1^{q^h} \\ \vdots & \vdots \quad \vdots \\ \beta_{h-1} Y_0 + \beta_{h-1}^q Y_1 + \beta_{h-1}^{q^2} Y_2 + \cdots + \beta_{h-1}^{q^{h-1}} Y_{h-1} &= -\beta_{h-1}^{q^h} \end{cases} \Leftrightarrow$$

$$\Leftrightarrow Y_0 = l_0, Y_1 = l_1, Y_2 = l_2, \ldots, Y_{h-1} = l_{h-1},$$

ou seja,

$$\sum_{j=0}^{h-1} \beta_i^{q^j} l_j = -\beta_i^{q^h}, \ i = 0, \ldots, h-1.$$

Daí, segue que

$$\sum_{j=0}^{h-1} \beta_i^{q^j} l_j = -\beta_i^{q^h} \Leftrightarrow \beta_i^{q^h} + \sum_{j=0}^{h-1} \beta_i^{q^j} l_j = 0, \quad i = 0, \ldots, h-1.$$

Logo, $\beta_0, \ldots, \beta_{h-1}$ são raízes do polinômio q-linear

$$L(X) = X^{q^h} + \sum_{j=0}^{h-1} l_j X^{q^j}.$$

Além disso, temos que toda combinação linear de $\beta_0, \ldots, \beta_{h-1}$ com coeficientes em $\mathbb{F}_q$ é também raiz de $L(X)$. Com efeito, considere a combinação linear $a_0\beta_0 + \ldots + a_{h-1}\beta_{h-1}$, com $a_0, \ldots, a_{h-1} \in \mathbb{F}_q$. Temos que

$$\begin{aligned} L(a_0\beta_0 + \ldots + a_{h-1}\beta_{h-1}) &= (a_0\beta_0 + \ldots + a_{h-1}\beta_{h-1})^{q^h} + \sum_{j=0}^{h-1} l_j (a_0\beta_0 + \ldots + a_{h-1}\beta_{h-1})^{q^j} \\ &= (a_0\beta_0^{q^h} + \ldots + a_{h-1}\beta_{h-1}^{q^h}) + (a_0 l_0 \beta_0 + \ldots + a_{h-1} l_0 \beta_{h-1}) + \ldots \\ &\quad \ldots + (a_0 l_{h-1} \beta_0^{q^{h-1}} + \ldots + a_{h-1} l_{h-1} \beta_{h-1}^{q^{h-1}}) \\ &= a_0 \underbrace{\left(\beta_0^{q^h} + \sum_{j=0}^{h-1} l_j \beta_0^{q^j}\right)}_{0} + \ldots + a_{h-1} \underbrace{\left(\beta_{h-1}^{q^h} + \sum_{j=0}^{h-1} l_j \beta_{h-1}^{q^j}\right)}_{0} \\ &= 0. \end{aligned}$$

Logo todos os elementos de V são raízes de $L(X)$. Portanto,

$$L(X) = \prod_{\beta \in V} (X - \beta).$$

□

3 Códigos Corretores de Erros

O nosso objeto de estudo nesse trabalho é o que chamamos de *Códigos Corretores de Erros*, sendo nossa grande motivação entender como a matemática desempenha um papel fundamental na teoria que envolve esses códigos.

Os Códigos Corretores de Erros estão intimamente ligados ao nosso cotidiano quando fazemos uso de diversos tipos de canais de comunicação como o rádio, televisão, celular, aplicativos de troca de mensagens, dentre outros. Tais códigos tem sido estudados e desenvolvidos há décadas e tiveram grande relevância na corrida espacial, sendo essenciais no sucesso das transmissões de imagens de Marte realizadas pelas naves espaciais Mariner 4 (1965) e Mariner 9 (1972) (HEFEZ; VILLELA, 2008, página 3, capítulo 1). Mas afinal, o que são os códigos corretores de erros?

> "Um código corretor de erros é, em essência, um modo organizado de acrescentar algum dado adicional a cada informação que se queira transmitir ou armazenar, que permita, ao recuperar a informação, detectar e corrigir erros." (HEFEZ; VILLELA, 2008, página i, prefácio).

Agora vejamos um exemplo para entender os princípios dos códigos corretores de erros. Consideremos uma lâmpada com quatro leds distintos nas cores vermelho, branco, azul e verde, e que conseguimos alterar a cor emitida por ela utilizando um controle remoto. Para tornar esse exemplo mais simples e objetivo faremos as seguintes convenções: não há possibilidade de acionarmos dois ou mais leds simultaneamente; a intensidade da luz emitida pelos leds é sempre a mesma e, ao pressionarmos um botão do controle, se houver alguma outra luz acesa, o seu led é instantaneamente desativado, enquanto o led correspondente ao botão pressionado é instantaneamente acionado. Dito isso, para ilustrarmos esse exemplo, o qual chamaremos de *Código da Luz Ambiente*, considere a Figura 1.

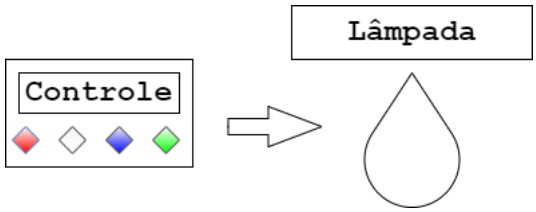

Figura 1 – Ilustração do controlador de luz ambiente.

Como o processo de mudança da luz ambiente ocorre com o acionamento do botão do controle correspondente à cor desejada, estabeleceremos códigos compostos por elementos de

$\{0,1\} \times \{0,1\}$ para o acionamento de cada botão.

$$\begin{array}{ccc} \text{Vermelho} & \mapsto & 00 \\ \text{Branco} & \mapsto & 01 \\ \text{Azul} & \mapsto & 10 \\ \text{Verde} & \mapsto & 11 \end{array}$$

Chamaremos esses elementos de palavras do código, entretanto, do ponto de vista da álgebra linear, trataremos essas palavras como vetores ou matrizes linhas e os denotaremos por, respectivamente, 00; $(0,0)$; $[0\ \ 0]$, quando conveniente.

Agora observe que se quisermos, por exemplo, um ambiente com a luz azul, acionamos o botão correspondente ao azul, o que fará com que o controle gere o código 10 e envie uma mensagem com esse código à lâmpada. Este código é chamado de *código fonte*. Note que, se a mensagem for transmitida por um canal via rádio, por exemplo, podem ocorrer erros durante a transmissão e a mensagem recebida ser diferente daquela que foi transmitida.

Suponhamos, por exemplo, que queríamos um ambiente com a cor branca e a mensagem recebida foi 00. Neste caso foi acesa a luz de cor vermelha, isto prova que ocorreu um erro na segunda coordenada da palavra transmitida e consequentemente a luz ambiente não corresponde à desejada. Uma maneira de lidar com esse tipo de problema é adicionar redundâncias ao código, de modo que mesmo em caso de ocorrência de erros, seja possível identificá-los e corrigí-los. Para tanto, é necessário fazermos uma recodificação do código da fonte, que chamaremos de *código de canal*. Podemos considerar, por exemplo, a recodificação abaixo:

$$\begin{array}{ccc} 00 & \mapsto & 00000 \\ 01 & \mapsto & 01111 \\ 10 & \mapsto & 10110 \\ 11 & \mapsto & 11101 \end{array}$$

Observe que nessa recodificação, o código da fonte corresponde aos primeiros dois dígitos do código de canal e os outros dígitos são as redundâncias do código de canal (o que não é via de regra). Com o código de canal, se quiséssemos um ambiente com a luz de mesma cor da citada anteriormente e durante a transmissão ocorreu o mesmo erro, ou seja, ocorreu um erro no segundo bit da palavra, então esta seria 00111, diferente da transmitida 01111. Por outro lado, a palavra recebida não pertence ao código, o que nos leva a suspeitar que ocorreram erros durante a transmissão. Todavia, podemos interpretá-la como uma palavra relativamente próxima da palavra transmitida 01111, pois tem a menor quantidade de dígitos diferentes, o que nos induz a pensar que a palavra recebida é supostamente a palavra 01111. Trataremos dessa proximidade com o conceito de distância entre palavras, que será abordado na próxima seção.

A Figura 2 nos fornece uma maneira bastante clara sobre o processo de transmissão de mensagens (dados), em que o objeto de transmissão pode ser um canal de rádio, canal wireless, ou qualquer outro capaz de transportar o código de canal.

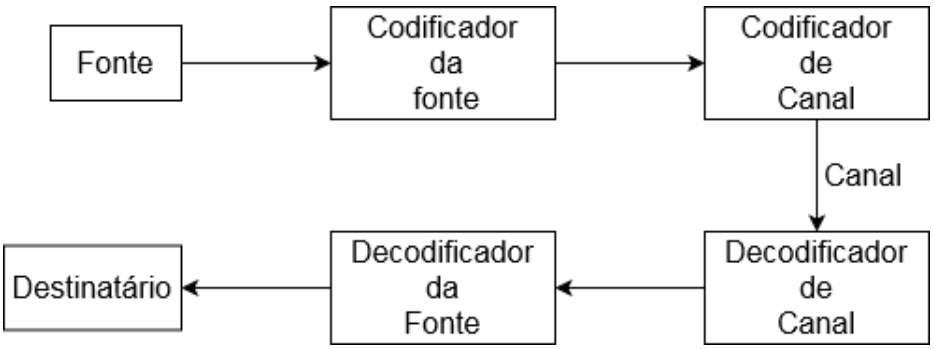

Figura 2 – Fluxograma do processo de transmissão de dados entre fonte e destinatário.

O Fluxograma descreve o seguinte processo: a mensagem que queremos enviar, chamada de código fonte, é codificada pelo codificador da fonte, de modo que o sistema onde ela será armazenada, antes de ser enviada, seja capaz de interpretá-la. Após esse passo, a nova mensagem é recodificada pelo codificador de canal, onde são adicionadas redundâncias às suas palavras. Com isso, a mensagem estará pronta para ser transmitida pelo canal de transmissão. Entretanto, como podem ocorrer erros durante o seu envio, a mensagem é corrigida, se possível, e decodificada pelo decodificador de canal e em seguida pelo decodificador da fonte, para então ser entregue ao destinatário, (HEFEZ; VILLELA, 2008).

3.1 Metrização de um código

Para iniciar a construção de um código corretor de erros, é necessário dar um conjunto finito $\mathcal{A}$ (alfabeto). Um código corretor de erros é um subconjunto próprio qualquer de $\mathcal{A}^n$, para algum $n \in \mathbb{N}$, em que n é o comprimento dos elementos do código corretor de erros, isto é, o vetor $\mathbf{u} = (u_1, u_2, \ldots, u_n) \in \mathcal{A}^n$, tal que $u_1, u_2, \ldots, u_n \in \mathcal{A}$ é um elemento do código, cujo comprimento é n.

A fim de estabelecer uma maneira precisa de identificar o quão próximas as palavras do código estão, apresentaremos a seguir uma maneira de medir distâncias entre as palavras em $\mathcal{A}^n$.

Definição 43. (Métrica de *Hamming*) *Dados dois elementos* $\mathbf{u}$, $\mathbf{v} \in \mathcal{A}^n$, *definimos a distância de Hamming entre* $\mathbf{u}$ *e* $\mathbf{v}$, *denotada por* $d(\mathbf{u}, \mathbf{v})$, *sendo a quantidade de componentes diferentes, isto é,*

$$d(\mathbf{u}, \mathbf{v}) = |\{i;\ u_i \neq v_i, 1 \leq i \leq n\}|.$$

Exemplo 6 (Código da Luz Ambiente)**.** *A distância entre quaisquer duas palavras do código da luz ambiente, em que* $\mathcal{A} = \{0, 1\}$ *e* $n = 5$*, são:*

$$\begin{aligned} d(00000, 01111) &= 4 \\ d(00000, 10110) &= 3 \\ d(00000, 11101) &= 4 \\ d(01111, 10110) &= 3 \\ d(01111, 11101) &= 2 \\ d(10110, 11101) &= 2 \end{aligned}$$

Proposição 36. *A distância de Hamming é uma métrica, ou seja, dados* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{A}^n$*, valem as seguintes propriedades:*

(i) $d(\mathbf{u}, \mathbf{v}) \geq 0$*, em que* $d(\mathbf{u}, \mathbf{v}) = 0$ *se, e somente se,* $\mathbf{u} = \mathbf{v}$*;*

(ii) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$*;*

(iii) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$.

Demonstração. Sejam $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{A}^n$.

(i) Note que $d(\mathbf{u}, \mathbf{v}) = |\{i;\ u_i \neq v_i, 1 \leq i \leq n\}| \geq 0$. Além disso,

$$\begin{aligned} d(\mathbf{u}, \mathbf{v}) = 0 &\Leftrightarrow |\{i;\ u_i \neq v_i, 1 \leq i \leq n\}| = 0 \\ &\Leftrightarrow u_i = v_i, \forall\, i = 1, \ldots, n \\ &\Leftrightarrow \mathbf{u} = \mathbf{v}. \end{aligned}$$

(ii) Basta observarmos que:

$$d(\mathbf{u}, \mathbf{v}) = |\{i;\ u_i \neq v_i, 1 \leq i \leq n\}| = |\{i;\ v_i \neq u_i, 1 \leq i \leq n\}| = d(\mathbf{v}, \mathbf{u}).$$

(iii) Temos que, se $\mathbf{u}$ e $\mathbf{w}$ têm α coordenadas distintas e $\mathbf{v}$ e $\mathbf{w}$ possuem β coordenadas distintas, então observe que $\mathbf{u}$ e $\mathbf{v}$ têm no máximo $\alpha + \beta$ coordenadas distintas, de modo que temos $d(\mathbf{u}, \mathbf{v}) \leq \alpha + \beta = d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$.

□

Definição 44. *Sejam* $\mathbf{a} \in \mathcal{A}^n$ *e* $t \geq 0$*, definimos o disco e a esfera de centro* $\mathbf{a}$ *e raio* t *por* $\mathcal{D}(\mathbf{a}, t) = \{\mathbf{u} \in \mathcal{A}^n; d(\mathbf{u}, \mathbf{a}) \leq t\}$ *e* $\mathcal{S}(\mathbf{a}, t) = \{\mathbf{u} \in \mathcal{A}^n; d(\mathbf{u}, \mathbf{a}) = t\}$*, respectivamente.*

Lema 10. *Para todo* $\mathbf{a} \in \mathcal{A}^n$ *e todo número natural* $r > 0$*, temos que*

$$| \mathcal{D}(\mathbf{a}, r) | = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i.$$

A demonstração desse resultado pode ser vista em (HEFEZ; VILLELA, 2008, página 6, capítulo 1).

Definição 45. *Seja* C *um código. A distância mínima de* C *é o número* $d = \min\{d(\mathbf{u}, \mathbf{v}); \mathbf{u}, \mathbf{v} \in C \text{ e } \mathbf{u} \neq \mathbf{v}\}$*, em que* d *representa a distância de Hamming.*

Observação 8. Note que a distância mínima no Exemplo 6 é igual a 2, pois $d = \min\{2, 3, 4\} = 2$.

Observação 9. Vale ressaltar que conseguimos determinar qual a distância mínima do código da luz ambiente, pois calculamos todas as distâncias possíveis entre quaisquer duas palavras dele e, para isso, verificamos duas a duas qual a distância entre todas as 4 palavras do código, ou seja, fizemos uma quantidade de cálculos igual a combinação $\binom{4}{2}$. Isto nos induz a pensar que para calcular a distância mínima de um código é necessário efetuar $\binom{M}{2}$ cálculos (sendo M é a quantidade de palavras do código), pois precisamos comparar tais distâncias e concluir qual a menor delas, o que é bastante trabalhoso quando se trata de códigos de muitas palavras.

Lema 11. *Seja* C *um código com distância mínima* d*. Se* $\mathbf{c}, \mathbf{c}' \in C$*, então*

$$\mathcal{D}(\mathbf{c}, l) \cap \mathcal{D}(\mathbf{c}', l) = \emptyset,$$

em que $l = \left\lfloor \frac{d-1}{2} \right\rfloor$.

Demonstração. Suponha por absurdo que a interseção é não vazia, ou seja, existe $\mathbf{x} \in \mathcal{D}(\mathbf{c}, l) \cap \mathcal{D}(\mathbf{c}', l)$. Então $d(\mathbf{x}, \mathbf{c}) \leq l$ e $d(\mathbf{x}, \mathbf{c}') \leq l$. Pelas propriedades (ii) e (iii) da Proposição 36, temos que

$$d(\mathbf{c}, \mathbf{c}') \leq d(\mathbf{c}, \mathbf{x}) + d(\mathbf{x}, \mathbf{c}') \leq 2l \leq d - 1,$$

o que é uma contradição com o fato da distância mínima ser igual a d. □

Teorema 16. *Seja* C *um código cuja distância mínima é* d*. Então* C *pode corrigir até* $\left\lfloor \frac{d-1}{2} \right\rfloor$ *erros e detectar até* $d - 1$ *erros.*

Demonstração. Se durante a transmissão de uma palavra $\mathbf{c}$ do código ocorreram t erros, com $t \leq \left\lfloor \frac{d-1}{2} \right\rfloor$, e a palavra recebida foi $\mathbf{r}$, então $d(\mathbf{r}, \mathbf{c}) = t \leq \left\lfloor \frac{d-1}{2} \right\rfloor$. O Lema 11 assegura que a distância entre quaisquer duas palavras do código é maior do que l, pois d é a distância mínima

do código. Logo **c** é determinada de maneira única a partir de **r**, visto que a distância entre **r** e **c** é menor do que a distância entre **r** e qualquer outra palavra do código C. Por outro lado, dada uma palavra **c** qualquer do código, podemos introduzir até $d-1$ erros em **c** sem encontrar nenhuma outra palavra $\mathbf{c}'$ do código, pois **c** está no mínimo a uma distância d de qualquer outra palavra do código. Portanto é possível detectar até $d-1$ erros. □

Definição 46. *Dado um código C, definimos por κ a capacidade de correção de erros do código C, em que κ é o número*

$$\kappa = \left\lfloor \frac{d-1}{2} \right\rfloor . \tag{3.1}$$

Exemplo 7. Suponha que foi transmitido à lâmpada do exemplo do *Código da Luz Ambiente* o comando para acender a luz vermelha, ou seja, foi transmitida a palavra 00000. Entretanto, houve alguma falha na comunicação e a palavra recebida foi 11110 (não pertencente ao código). Assim, foi acesa a luz de cor azul. Note que $d(11110, 00000) = 4$ e, portanto, maior do que a distância mínima $d = 3$. Por outro lado $d(11110, 10110) = 1 = \kappa$, logo, pelo Teorema 16 a palavra do código 10110 é univocamente determinada pela palavra 11110, ou seja, a partir dessa correção a lâmpada irá acender a luz de cor azul. Isso é uma contradição com o fato da palavra transmitida ser 00000 (luz de cor vermelha).

O Teorema 16 é um resultado de grande importância para os códigos corretores de erros, pois nos fornece uma maneira prática de detectar e corrigir erros quando uma palavra **r** é recebida por meio da transmissão de dados. Sendo assim, quando um receptor recebe uma palavra **r**, temos as seguintes possibilidades:

1. A palavra **r** encontra-se num disco de raio κ e pode ser substituída pela palavra **c** do código, em caso de erro na transmissão.
2. A palavra **r** não está em nenhum disco em torno da palavra **c** e, neste caso, não é possível corrigí-la.

Note que nem sempre podemos garantir que no caso 1 a palavra substituída de fato é a palavra transmitida, pois como mostra o Exemplo 7, a palavra recebida aproximou-se de uma outra palavra do código e não da palavra transmitida, visto que foram cometidos mais do que κ erros.

Definição 47. *Um código $C \subset \mathcal{A}^n$ possui três parâmetros principais:*

$$(n, M, d), \tag{3.2}$$

que são, respectivamente, o comprimento de suas palavras, a sua quantidade de elementos e a sua distância mínima.

A Definição 47 e as relações entre esses parâmetros, são de suma importância na construção dos códigos corretores de erros, pois nos fornecem diversas informações acerca do código e serão mencionadas posteriormente.

Uma importante definição que daremos a respeito dos códigos corretores de erros é a de equivalência entre códigos, pois, como veremos no Capítulo 4, às vezes trabalhar com um código equivalente a um dado código C é mais simples do que trabalharmos com o próprio código C. Antes disso, daremos a seguinte definição:

Definição 48. *Sejam $\mathcal{A}$ um alfabeto e n um número natural. Dizemos que uma função $F : \mathcal{A}^n \to \mathcal{A}^n$ é uma* isometria *quando ela preserva distâncias de Hamming, ou seja,*

$$d(F(\mathbf{x}), F(\mathbf{y})) = d(\mathbf{x}, \mathbf{y}), \quad \forall\, \mathbf{x}, \mathbf{y} \in \mathcal{A}^n.$$

Definição 49. *Dados dois códigos C e $\tilde{C}$ em $\mathcal{A}^n$, dizemos que $\tilde{C}$ é* equivalente *a C quando existe uma isometria F de $\mathcal{A}^n$ tal que $F(C) = \tilde{C}$.*

4 Códigos Lineares

Discutiremos nesse capítulo uma das classes de códigos corretores de erros mais utilizados na prática: *os códigos lineares*. Apresentaremos aqui um método mais eficaz para o cálculo da distância mínima de um código do que o visto no Capítulo 3, no qual calculamos todas as distâncias entre as palavras do código para determinar qual a menor delas, além de apresentarmos alguns conceitos teóricos que embasam a construção de um código linear, tais como matriz geradora, matriz teste de paridade e, mais importante, algoritmos de codificação e decodificação para esses códigos. Os estudos feitos aqui foram baseados em (HEFEZ; VILLELA, 2008) e nos conceitos teóricos relativos à álgebra linear, corpos finitos e códigos corretores de erros, apresentados nos capítulos 2 e 3.

Para construir um código linear, o primeiro passo, como vimos no Capítulo 3, é considerarmos como alfabeto um conjunto de símbolos que seja possível muni-lo de uma estrutura algébrica. Matematicamente, a melhor estrutura para o alfabeto é um corpo finito K. Dito isso, um código linear será um subconjunto próprio do espaço vetorial K^n sobre K e, para que as propriedades de espaços vetoriais valham também nesse subconjunto, ele deve ser um subespaço vetorial de K^n. Com isso em mente, daremos a definição de códigos lineares a seguir.

Definição 50. *Sejam K um corpo finito com q elementos e K^n o espaço vetorial sobre K. Um código $C \subset K^n$ será chamado de código linear se for um subespaço vetorial de K^n sobre K.*

Exemplo 8. Considere o código da luz ambiente, visto no Capítulo 3. Como o alfabeto adotado é $\{0, 1\}$, é natural substituí-lo pelo corpo $K = \mathbb{F}_2$. Considere a seguinte transformação linear

$$\begin{array}{rccc} T: & \mathbb{F}_2^2 & \to & \mathbb{F}_2^5 \\ & (x_1, x_2) & \mapsto & (x_1, x_2, x_1 + x_2, x_1 + x_2, x_2) \end{array}.$$

Temos que

$$\begin{aligned} T(0,0) &= (0,0,0,0,0) = \mathbf{v}_1; \\ T(0,1) &= (0,1,1,1,1) = \mathbf{v}_2; \\ T(1,0) &= (1,0,1,1,0) = \mathbf{v}_3; \\ T(1,1) &= (1,1,0,0,1) = \mathbf{v}_4. \end{aligned}$$

Assim, $C = Im(T)$, além disso,

$$(x_1, x_2, x_1 + x_2, x_1 + x_2, x_2) = x_1(1,0,1,1,0) + x_2(0,1,1,1,1) = x_1\mathbf{v}_3 + x_2\mathbf{v}_2.$$

Deste modo, segue que C é um subespaço vetorial de $\mathbb{F}_2^5$ de dimensão 2 e $\mathcal{B}_C = \{\mathbf{v}_2, \mathbf{v}_3\}$ é uma base de C.

Esse exemplo pode ser generalizado para o caso em que C é um subespaço vetorial de K^n de dimensão k, com base $\mathcal{B}_C = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$. Logo qualquer palavra $\mathbf{v}$ de C pode ser escrita, de maneira única, como combinação linear dos elementos da base, isto é,

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_k\mathbf{v}_k; \quad \lambda_i \in K \; \forall \, i = 1, \ldots, k.$$

Observe que o corpo K possui q elementos, logo para cada λ_i, com $i = 1, \ldots, k$, temos q possibilidades de escalares. Portanto $M = |C| = q^k$.

A estrutura de subespaço vetorial de um código linear C possibilita descrevê-lo como imagem de uma transformação linear. De fato, seja $\mathcal{B}_C = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ uma base para o código C, então para toda palavra $\mathbf{x}$ de C, existem $x_1, \ldots, x_k \in K$ (ordenados) tais que

$$\mathbf{x} = x_1\mathbf{v}_1 + \ldots + x_k\mathbf{v}_k = \sum_{i=1}^{k} x_i\mathbf{v}_i.$$

Considerando a base canônica de K^k e a transformação linear dada por $T_I(\mathbf{e}_i) = \mathbf{v}_i$, temos que

$$T_I(x_1, \ldots, x_k) = x_1\mathbf{v}_1 + \ldots + x_k\mathbf{v}_k = \mathbf{x} \tag{4.1}$$

e T_I construída deste modo é injetiva, pois leva a base canônica de K^k na base de C. Assim,

$$C = Im(T_I).$$

Por outro lado, verificar que dado $\mathbf{v} \in K^n$ é um elemento do código C não é uma tarefa fácil, visto que é necessário resolver o sistema de n equações nas k incógnitas $x_1, \ldots, x_k$ abaixo

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_k\mathbf{v}_k = \mathbf{v}.$$

Vejamos o seguinte exemplo. Seja $C \subset \mathbb{F}_2^8$ de dimensão 4 e uma base de C dada por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, em que $\mathbf{v}_i = (v_{1i}, \ldots, v_{8i})$, com $i = 1, \ldots, 4$. Para verificar que a palavra 11101001 de $\mathbb{F}_2^8$ pertence ao código, é necessário resolvermos o seguinte sistema

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 = (1,1,1,0,1,0,0,1) \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x_1v_{11} + x_2v_{12} + x_3v_{13} + x_4v_{14} = 1 \\ x_1v_{21} + x_2v_{22} + x_3v_{23} + x_4v_{24} = 1 \\ x_1v_{31} + x_2v_{32} + x_3v_{33} + x_4v_{34} = 1 \\ x_1v_{41} + x_2v_{42} + x_3v_{43} + x_4v_{44} = 0 \\ x_1v_{51} + x_2v_{52} + x_3v_{53} + x_4v_{54} = 1 \\ x_1v_{61} + x_2v_{62} + x_3v_{63} + x_4v_{64} = 0 \\ x_1v_{71} + x_2v_{72} + x_3v_{73} + x_4v_{74} = 0 \\ x_1v_{81} + x_2v_{82} + x_3v_{83} + x_4v_{84} = 1. \end{cases}$$

Note que, apesar de ser prático do ponto de vista teórico, ver um código linear como a imagem de uma transformação linear, tal opção pode nos conduzir a processos que possuem diversos cálculos. A seguir iremos estudar uma outra forma de conceber um código linear que, como veremos, é mais eficiente do que o anterior em alguns aspectos. Por esse motivo iremos identificar um código como o núcleo de uma transformação linear T_N, uma vez que a verificação de que um elemento $\mathbf{x}$ de K^n é um elemento do código C se resumirá a verificar se $T_N(\mathbf{x})$ é o vetor nulo. Para tal, escreveremos

$$K^n = C \oplus C',$$

em que C' é o subespaço vetorial de K^n complementar de C. Assim, dado $\mathcal{B}_C = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ uma base para C e $\mathcal{B}_{C'} = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n-k}\}$ uma base para C'. Podemos definir $T_N : C \oplus C' \to K^{n-k}$ da seguinte maneira:

$$T_N(\mathbf{v}_i) = \mathbf{0} \text{ para } i = 1, \ldots, k \quad \text{e} \quad T_N(\mathbf{u}_j) = \mathbf{e}_j \text{ para } j = 1, \ldots, n-k,$$

em que os vetores $\mathbf{e}_1, \ldots, \mathbf{e}_{n-k}$ formam uma base canônica para C'. Com isso, vemos que C é o núcleo de T_N. De fato, se $\mathbf{c} \in C$, então existem $\lambda_1, \ldots, \lambda_k \in K$ tais que $\mathbf{c} = \lambda_1\mathbf{v}_1 + \ldots + \lambda_k\mathbf{v}_k$. Logo,

$$T_N(\mathbf{c}) = T_N(\lambda_1\mathbf{v}_1 + \ldots + \lambda_k\mathbf{v}_k) = \lambda_1 T_N(\mathbf{v}_1) + \ldots + \lambda_k T_N(\mathbf{v}_k) = \mathbf{0} \Rightarrow \mathbf{c} \in Ker(T_N) \Rightarrow C \subset Ker(T_N).$$

Por outro lado, sendo $\mathbf{c}$ não nulo, um elemento de $Ker(T_N)$, então

$$\begin{aligned} T_N(\mathbf{c}) = \mathbf{0} &\Leftrightarrow T_N(\lambda_1\mathbf{v}_1 + \ldots + \lambda_k\mathbf{v}_k + \beta_1\mathbf{u}_1 + \ldots + \beta_{n-k}\mathbf{u}_{n-k}) = \mathbf{0} \\ &\Leftrightarrow \beta_1 T_N(\mathbf{u}_1) + \ldots + \beta_{n-k} T_N(\mathbf{u}_{n-k}) = \mathbf{0} \\ &\Leftrightarrow \beta_1 T_N(\mathbf{e}_1) + \ldots + \beta_{n-k} T_N(\mathbf{e}_{n-k}) = \mathbf{0} \\ &\Leftrightarrow \beta_1 = \ldots = \beta_{n-k} = 0 \Leftrightarrow \mathbf{c} \in C \Leftrightarrow Ker(T_N) \subset C. \end{aligned}$$

Portanto $C = Ker(T_N)$. Nesse caso, é muito mais simples determinar se um elemento $\mathbf{v} \in K^n$ pertence ao código C, uma vez que não precisamos resolver um sistema de várias equações com múltiplas variáveis.

Vamos agora aplicar o resultado acima para o código da luz ambiente que introduzimos no Capítulo 3.

Exemplo 9. Vimos no Exemplo 8 que o conjunto $\mathcal{B}_C = \{\mathbf{v}_2, \mathbf{v}_3\} = \{(0,1,1,1,1), (1,0,1,1,0)\}$ é uma base de C. Como queremos descrever esse código como núcleo de uma transformação linear $T : C \oplus C' = \mathbb{F}_2^5 \to \mathbb{F}_2^3$, vamos construí-la de modo que $Ker(T)$ seja gerado por $(0,1,1,1,1)$ e $(1,0,1,1,0)$, ou seja, que

$$T(0,1,1,1,1) = (0,0,0) \quad \text{e} \quad T(1,0,1,1,0) = (0,0,0).$$

Como $(0,1,1,1,1)$ e $(1,0,1,1,0)$ são linearmente independentes, formam uma base de $Ker(T)$, a qual pode ser completada a

$$\{(0,1,1,1,1),(1,0,1,1,0),(0,1,0,0,0),(0,0,1,0,0),(0,0,0,0,1)\},$$

uma base de $\mathbb{F}_2^5$ e por fim, fazemos

$$\begin{aligned} T(0,1,1,1,1) &= (0,0,0),\\ T(1,0,1,1,0) &= (0,0,0),\\ T(0,1,0,0,0) &= (1,0,0),\\ T(0,0,1,0,0) &= (0,1,0),\\ T(0,0,0,0,1) &= (0,0,1). \end{aligned}$$

Assim,

$$T(x_1,x_2,x_3,x_4,x_5) = (x_1+x_2+x_4, x_3+x_4, x_1+x_4+x_5).$$

Seja $\mathbf{u} = (1,0,0,0,1) \in \mathbb{F}_2^5$. Temos que

$$T(\mathbf{u}) = T(1,0,0,0,1) = (1,0,0).$$

Logo, $\mathbf{u} \notin Ker(T)$ e, portanto, a palavra 10001 não pertence ao código C. Por outro lado, $\mathbf{v}_4 = (1,1,0,0,1)$ é tal que

$$T(\mathbf{v}_4) = T(1,1,0,0,1) = (0,0,0).$$

Logo $\mathbf{v}_4 \in Ker(T)$, isto é, a palavra 11001 é uma palavra do código C.

Definição 51. *Dado $\mathbf{x} \in K^n$, definimos o peso de $\mathbf{x}$ por $\omega(\mathbf{x})$, em que $\omega(\mathbf{x})$ é o número inteiro obtido pela contagem de letras não nulas da palavra $\mathbf{x}$, ou seja,*

$$\omega(\mathbf{x}) := |\{i;\ x_i \neq 0\}|.$$

Note que, dessa forma temos que $\omega(\mathbf{x}) = d(\mathbf{x},0)$, em que d representa a distância de *Hamming*.

Definição 52. *O peso de um código linear C é o inteiro*

$$\omega(C) := \min\{\omega(\mathbf{x});\ \mathbf{x} \in C \setminus \{\mathbf{0}\}\}.$$

Proposição 37. *Seja $C \subset K^n$ um código linear com distância mínima d. Temos que*

(i) $d(\mathbf{x},\mathbf{y}) = \omega(\mathbf{x}-\mathbf{y}),\ \forall\ \mathbf{x},\mathbf{y} \in K^n$;

(ii) $d = \omega(C)$.

Demonstração. (i) Pelas definições de distância de *Hamming* e de peso, temos que

$$d(\mathbf{x}, \mathbf{y}) = |\{i;\, x_i \neq y_i,\ 1 \leq i \leq n\}| = |\{i;\, x_i - y_i \neq 0,\ 1 \leq i \leq n\}| = \omega(\mathbf{x} - \mathbf{y}).$$

(ii) Note que, para todo $\mathbf{x}, \mathbf{y} \in C$ com $\mathbf{x} \neq \mathbf{y}$, temos $\mathbf{x} - \mathbf{y} \in C \setminus \{\mathbf{0}\}$, logo, pelo item anterior $d(\mathbf{x}, \mathbf{y}) = \omega(\mathbf{x} - \mathbf{y})$ para todo $\mathbf{x}, \mathbf{y} \in C$. Em particular, vale para $\mathbf{c}, \mathbf{c}' \in C$, cuja distância é a menor possível entre dois vetores (palavras) de C, ou seja, $\mathbf{c}, \mathbf{c}'$ possuem a menor quantidade de coordenadas diferentes, consequentemente, $\omega(\mathbf{c} - \mathbf{c}')$ tem a menor quantidade possível de coordenadas não nulas, uma vez que quando as coordenadas de $\mathbf{c}$ e $\mathbf{c}'$ são iguais, as coordenadas do vetor $\mathbf{c} - \mathbf{c}'$ são nulas, e quando as coordenadas de $\mathbf{c}$ e $\mathbf{c}'$ são diferentes, as coordenadas de $\mathbf{c} - \mathbf{c}'$ são não nulas. Portanto, $d = d(\mathbf{c}, \mathbf{c}') = \omega(\mathbf{c} - \mathbf{c}') = \omega(C)$.

□

Uma importante consequência da Proposição 37 é que para obtermos a distância mínima de um código linear basta determinarmos o seu peso, o que pode ser feito determinando o peso de suas palavras. Vejamos o exemplo a seguir.

Exemplo 10. Retomando o exemplo do código da luz ambiente, considere $\mathbf{c}_1 = (0, 1, 1, 1, 1)$, $\mathbf{c}_2 = (1, 0, 1, 1, 0)$, $\mathbf{c}_3 = (1, 1, 0, 0, 1) \in C \setminus \{(0, 0, 0, 0, 0)\}$, temos que $\omega(\mathbf{c}_1) = 4$, $\omega(\mathbf{c}_2) = 3$ e $\omega(\mathbf{c}_3) = 3$, o que implica em $\omega(C) = 3 = d$.

Decorre da Proposição 37 que, em códigos lineares com M elementos, podemos encontrar a distância mínima a partir de $M - 1$ cálculos, uma vez que para encontrá-la, basta calcular o peso de suas palavras não nulas e tomar a distância mínima como o menor dos pesos calculados. Porém, apesar de mais simples do que quando comparamos com a combinação $\binom{M}{2}$ do Capítulo 3, ainda assim demandará um número considerável de cálculos se o código tiver muitas palavras.

Para exemplificar essa situação, consideremos novamente o exemplo do código da luz ambiente. Como ele tem quatro palavras, é simples encontrar a distância mínima, pois basta calcular o peso de suas três palavras não nulas. Entretanto, considerando por exemplo um código que tenha 2^{10} palavras, para encontrar sua distância mínima seria necessário o cálculo do peso de 1023 palavras, para então compará-los a fim de determinar qual o menor deles. Com essa observação fica evidente que quanto mais palavras o código possui, maior a dificuldade de encontrar sua distância mínima por esse método. Deste modo, buscaremos desenvolver outras maneiras para determinar a distância mínima de um código.

4.1 Matriz geradora de um código

No Capítulo 2 definimos por parâmetros de um código C a terna dada em (3.2). Para os códigos lineares os parâmetros são praticamente os mesmos, (n, k, d), a diferença está no parâmetro k, que é a dimensão do código, pois a partir dela é possível determinar quantas palavras o código possui.

Além disso, dado K um corpo finito com q elementos e sendo $C \subset K^n$ um código linear, conseguimos uma base ordenada de C dada por $\mathcal{B}_C = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, com $\mathbf{v}_i = (v_{i1}, v_{i2}, \ldots, v_{in})$, para $i = 1, \ldots, k$. Considere a seguinte matriz

$$G = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_k \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{k1} & v_{k2} & \ldots & v_{kn} \end{bmatrix}. \tag{4.2}$$

Note que, se $\mathbf{x} \in K^k$, então

$$\mathbf{x}G = \begin{bmatrix} x_1 & \ldots & x_k \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_k \end{bmatrix} = \begin{bmatrix} x_1\mathbf{v}_1 + \ldots + x_k\mathbf{v}_k \end{bmatrix} \Leftrightarrow x_1\mathbf{v}_1 + \ldots + x_k\mathbf{v}_k \in C,$$

ou seja, a matriz G é a matriz da transformação linear dada em (4.1), cuja imagem é o código C. Tal matriz será chamada de matriz geradora de C associada à base $\mathcal{B}_C$. Com isso, diremos que a transformação linear

$$\begin{array}{rccc} T: & K^k & \to & K^n \\ & \mathbf{x} & \mapsto & \mathbf{x}G \end{array} \tag{4.3}$$

é uma codificação do código C, em que $Im(T) = C$.

Note que a matriz geradora G não é única, pois depende da escolha da base de C. Dessa forma, fazendo uma mudança de base do espaço vetorial C, mudamos a matriz da transformação linear que nos dá o código C. Essa nova matriz, digamos $\tilde{G}$, obtida através da mudança de base de C, pode ser obtida ao efetuarmos operações matriciais elementares sobre as linhas da matriz G,

(L1) Permutação de duas linhas.

(L2) Multiplicação de uma linha por um escalar não nulo.

(L3) Adição de um múltiplo escalar de uma linha a outra.

Inversamente, dada uma matriz G, cujas linhas são linearmente independentes, podemos definir um código C a partir dessa matriz, como a imagem da transformação linear dada em (4.3).

Exemplo 11. *Sejam $K = \mathbb{F}_2$ e*

$$G = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Temos que, se $\alpha, \beta, \gamma \in K$, então

$$\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 + \gamma\mathbf{v}_3 = \mathbf{0} \Leftrightarrow \alpha(1,0,1,0,1) + \beta(1,1,0,1,0) + \gamma(1,1,1,1,1) = (0,0,0,0,0).$$

Isso implica em

$$\begin{cases} \alpha + \beta + \gamma = 0 \\ \beta + \gamma = 0 \\ \alpha + \gamma = 0 \\ \beta + \alpha = 0 \\ \alpha + \gamma = 0 \end{cases}.$$

Daí, temos $\alpha = \beta = \gamma = 0$, ou seja, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ são linearmente independentes.

Considerando

$$\begin{array}{rccc} T: & \mathbb{F}_2^3 & \to & \mathbb{F}_2^5 \\ & \mathbf{x} & \mapsto & \mathbf{x}G \end{array},$$

obtemos um código linear $C \subset \mathbb{F}_2^5$ através de T, sendo $C = Im(T)$.

Observe que a codificação da palavra do código da fonte 101, através da transformação linear T, é a palavra do código de canal 01010. Por outro lado, para encontrarmos a palavra do código da fonte a partir de 01010, resolvemos o sistema

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} G = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \end{bmatrix},$$

cuja única solução é $x_1 = 1, x_2 = 0, x_3 = 1$ e portanto 01010 é a palavra do código de canal decodificada como 101.

Observe que realizando operações matriciais elementares sobre as linhas da matriz G, obtemos a matriz $\tilde{G}$ abaixo, geradora do código C.

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

Neste caso, obtemos mais facilmente que

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \end{bmatrix} \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 1 \\ x_3 = 0 \\ x_2 = 1 \\ x_3 = 0 \end{cases},$$

ou seja, a palavra 01010 do código de canal é decodificada como 010.

Definição 53. *Dizemos que uma matriz geradora G de um código C está na forma padrão quando temos*

$$G = (Id_k \mid A),$$

em que Id_k é a matriz identidade $k \times k$ e A, uma matriz $k \times (n-k)$.

Considere

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

a matriz geradora de um código C, em que $C \subset \mathbb{F}_2^5$. Note que as duas primeiras colunas dessa matriz são nulas, por esse motivo não é possível escrever $G = (Id_{2x2} \mid A_{3x2})$ utilizando as operações elementares sobre as linhas, logo não conseguimos encontrar uma matriz geradora do código C na forma padrão. Entretanto, se efetuarmos permutações das colunas de G, podemos obter a matriz

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

geradora de um código $\tilde{C}$ que está na forma padrão e, conforme (HEFEZ; VILLELA, 2008, página 91, capítulo 5), o código $\tilde{C}$ é equivalente ao código C.

Além da permutação de colunas, podemos multiplicar uma coluna por um escalar não nulo e ainda obter um código $\tilde{C}$ equivalente a C.

Observe que, ao efetuarmos operações nas linhas da matriz geradora de C, não alteramos em nada as palavras do código fonte, apenas alteramos as coordenadas da codificação dessas palavras, pois mudamos a base da imagem da transformação linear T. No entanto, ao efetuarmos operações em suas colunas, alteramos a base do domínio da transformação T, o que implica em alterar todas as palavra do código C, em virtude de alterarmos o código fonte.

Teorema 17. *Dado um código C, existe um código equivalente $\tilde{C}$ com matriz geradora na forma padrão.*

A demonstração desse resultado pode ser encontrada em (HEFEZ; VILLELA, 2008, páginas 92-93, capítulo 5).

4.2 Códigos Duais

Daremos nessa seção algumas definições importantes para os códigos lineares, como código dual a um código C, matriz teste de paridade e a síndrome de um elemento de K^n. As definições e resultados que traremos aqui serão de grande utilidade no processo de decodificação dos códigos lineares.

Lema 12. *Seja $C \subset K^n$ um código linear, com uma base $\mathcal{B}_C = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ e matriz geradora G. Temos que $\mathbf{x} \in C^\perp$ se, e somente se, $G\mathbf{x}^t = \left[0\right]_{k\times 1}$.*

Demonstração. De fato, temos que $\mathbf{x} \in C^\perp$ se, e somente se, $\langle \mathbf{x}, \mathbf{v}\rangle = 0$ para todo $\mathbf{v} \in C$. Como

$$G = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_k \end{bmatrix} = \begin{bmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{k1} & \ldots & v_{kn} \end{bmatrix},$$

em que $\mathbf{v}_i = (v_{i1}, \ldots, v_{in}) \in C$, segue que

$$G\mathbf{x}^t = \begin{bmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{k1} & \ldots & v_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} v_{11}x_1 + \ldots + v_{1n}x_n \\ \vdots \\ v_{k1}x_1 + \ldots + v_{kn}x_n \end{bmatrix} = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{x}\rangle \\ \vdots \\ \langle \mathbf{v}_k, \mathbf{x}\rangle \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

logo $G\mathbf{x}^t = \left[0\right]_{k\times 1}$. Por outro lado, se $G\mathbf{x}^t = \left[0\right]_{k\times 1}$, então

$$\begin{bmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{k1} & \ldots & v_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Assim,

$$\begin{cases} v_{11}x_1 + \ldots + v_{1n}x_n = 0 \\ \vdots \quad\quad\quad\quad\quad\quad \vdots \\ v_{k1}x_1 + \ldots + v_{kn}x_n = 0 \end{cases} \Rightarrow \begin{cases} \langle \mathbf{v}_1, \mathbf{x}\rangle = 0 \\ \vdots \\ \langle \mathbf{v}_k, \mathbf{x}\rangle = 0, \end{cases}$$

para todo $\mathbf{v}_i \in C$, com $i = 1, \ldots, k$ e, como qualquer elemento $\mathbf{v}$ de C é uma combinação linear do tipo

$$\mathbf{v} = \alpha_1\mathbf{v}_1, \ldots, \alpha_k\mathbf{v}_k, \quad \text{com } \alpha_1, \ldots, \alpha_k \in K,$$

segue que $\langle \mathbf{v}, \mathbf{x}\rangle = 0$ para todo $\mathbf{v} \in C$. Logo $\mathbf{x} \in C^\perp$. □

Definição 54. *Sejam $C \subset K^n$ um código linear e $C^\perp$ o complemento ortogonal a C. Dizemos que $C^\perp$ é o código dual a C.*

Observe que, se $C \subset K^n$ é um código linear de dimensão k, então o seu dual $C^\perp$, pela Proposição 4, tem dimensão $n - k$.

Proposição 38. *Seja $C \subset K^n$ um código linear de dimensão k com matriz geradora $G = (Id_k \mid A)$, na forma padrão. Então, $H = (-A^t \mid Id_{n-k})$ é uma matriz geradora de $C^\perp$.*

Demonstração. Observe que a matriz H possui um bloco identidade Id_{n-k} e, por esse motivo, esta possui $n - k$ linhas linearmente independentes. Consequentemente, as linhas de H geram um subespaço vetorial de dimensão $n - k$. Além disso, as linhas de H são ortogonais às linhas de G. De fato, sejam

$$\mathbf{h}_i = (-a_{1i}, \ldots, -a_{ki}, 0, \ldots, 1, \ldots, 0)$$

e

$$\mathbf{g}_j = (0, \ldots, 1, \ldots, 0, a_{j1}, \ldots, a_{j(n-k)})$$

a i-ésima linha da matriz H e a j-ésima linha da matriz G, respectivamente. Temos que

$$\begin{aligned} \langle \mathbf{h}_i, \mathbf{g}_j \rangle &= \langle (-a_{1i}, \ldots, -a_{ki}, 0, \ldots, 1, \ldots, 0), (0, \ldots, 1, \ldots, 0, a_{j1}, \ldots, a_{j(n-k)}) \rangle \\ &= -a_{ji} + a_{ji} = 0, \ \forall\, i = 1, \ldots, n-k; j = 1, \ldots, k. \end{aligned}$$

Com isso, as linhas de H geram um subespaço vetorial de $C^\perp$ de dimensão $n - k$ e, como pela Proposição 4, a dimensão de $C^\perp$ é $n - k$, segue que H gera $C^\perp$. □

Lema 13. *Seja $C \subset K^n$ um código de dimensão k com matriz geradora G. Uma matriz H de ordem $(n - k) \times n$, com coeficientes em K e linhas linearmente independentes, é uma matriz geradora de $C^\perp$ se, e somente se, $G \cdot H^t = \left[0\right]_{k\times(n-k)}$.*

Demonstração. Se uma matriz H de ordem $(n - k) \times n$ é uma matriz geradora de $C^\perp$, então as linhas de H, que denotaremos por $h_1, h_2, \ldots, h_{n-k}$, são linearmente independentes e portanto formam uma base para $C^\perp$. Então, qualquer elemento $\mathbf{u}$ de $C^\perp$ se escreve de maneira única como

$$\mathbf{u} = a_1 h_1 + a_2 h_2 + \cdots + a_{n-k} h_{n-k}, \quad \text{com } a_1, \ldots, a_{n-k} \in K. \tag{4.4}$$

Como $G \cdot h_j^t = \left[0\right]_{k\times 1}$ para todo $h_j \in C^\perp$, com $1 \leq j \leq n - k$, decorre da equação (4.4), $G \cdot \mathbf{u}^t = \left[0\right]_{k\times 1}$ para todo $\mathbf{u} \in C^\perp$. Logo $G \cdot H^t = \left[0\right]_{k\times(n-k)}$.

Por outro lado, se $G \cdot H^t = \left[0\right]_{k\times(n-k)}$, então $\langle g_i, h_j \rangle = 0$, com $1 \leq i \leq k$, $1 \leq j \leq n - k$ e, consequentemente, $\langle g_i, \mathbf{u} \rangle = 0$ para todo $\mathbf{u} \in C^\perp$. Logo o conjunto $\{h_1, h_2, \ldots, h_{n-k}\} \subset C^\perp$ é gerador de um subespaço vetorial de $C^\perp$, de dimensão $n - k$, que coincide com a dimensão de $C^\perp$. Portanto, H é uma matriz geradora de $C^\perp$. □

Corolário 10. *Se C é um subespaço vetorial de K^n, então $(C^\perp)^\perp = C$.*

Demonstração. Sejam G e H as matrizes geradoras de C e $C^{\perp}$, respectivamente. Temos, pelo Lema 13, que $G \cdot H^t = \left[0\right]_{k\times(n-k)}$ e daí temos que $(G \cdot H^t)^t = H \cdot G^t = \left[0\right]_{(n-k)\times k}$. Logo, G é a matriz geradora de $(C^{\perp})^{\perp}$ e, como por hipótese G gera C, segue que $(C^{\perp})^{\perp} = C$. □

O resultado a seguir nos fornece uma maneira bastante útil de verificar se uma palavra pertence ou não ao código.

Proposição 39. *Sejam C um código linear e H uma matriz geradora de $C^{\perp}$. Então*

$$\mathbf{v} \in C \Leftrightarrow H\mathbf{v}^t = \left[0\right]_{k\times 1}.$$

Demonstração. Suponha que $\mathbf{v} \in C$ e denote por $h_1, \ldots, h_{n-k}$, as linhas da matriz H. Temos pelo Corolário 10 que $\mathbf{v} \in (C^{\perp})^{\perp}$ e, como a matriz H gera $C^{\perp}$, segue que $\langle h_i, \mathbf{v}\rangle = 0$, para todo $h_i \in \{h_1, h_2, \ldots, h_{n-k}\}$, mais ainda, $\langle \mathbf{u}, \mathbf{v}\rangle = 0$ para todo $\mathbf{u} \in C^{\perp}$. Logo $H\mathbf{v}^t = \left[0\right]_{k\times 1}$.
Reciprocamente, se $H\mathbf{v}^t = \left[0\right]_{k\times 1}$, então $\langle h_i, \mathbf{v}\rangle = 0$ para todo $h_i \in \{h_1, h_2, \ldots, h_{n-k}\}$, com $1 \leq i \leq n-k$. Assim, tomando $\mathbf{w} \in C^{\perp}$ qualquer, temos que $\mathbf{w} = \sum_{i=1}^{m} \alpha_i h_i$, $\alpha_i \in K$ para $i = 1, \ldots, m$, logo

$$\begin{aligned}\langle \mathbf{w}, \mathbf{v}\rangle &= \left\langle \sum_{i=1}^{m} \alpha_i h_i, \mathbf{v} \right\rangle = \langle \alpha_1 h_1 + \alpha_2 h_2 + \cdots + \alpha_m h_m, \mathbf{v}\rangle \\ &= \alpha_1\langle h_1, \mathbf{v}\rangle + \alpha_2\langle h_2, \mathbf{v}\rangle + \ldots + \alpha_m\langle h_m, \mathbf{v}\rangle \\ &= \sum_{i=1}^{m} \alpha_i\langle h_i, \mathbf{v}\rangle = 0.\end{aligned}$$

Portanto $\mathbf{v} \in (C^{\perp})^{\perp}$ e, pelo Corolário 10, segue que $\mathbf{v} \in C$. □

Definição 55. *Sejam H uma matriz geradora de $C^{\perp}$ e $\mathbf{v} \in K^n$. A matriz H é chamada de matriz* teste de paridade *de C e o vetor $H\mathbf{v}^t$ de* síndrome *de* $\mathbf{v}$.

Os resultados a seguir são de grande importância para os códigos corretores de erros lineares, pois a partir da matriz teste de paridade conseguimos estabelecer uma maneira de calcular a sua distância mínima sem precisar calcular o peso de todas as suas palavras não nulas.

Proposição 40. *Seja H uma matriz de ordem $(n-k)\times n$, a matriz teste de paridade de um código linear C. Temos que $d = \omega(C) \geq s$ se, e somente se, quaisquer $s-1$ colunas de H são linearmente independentes.*

Demonstração. Suponha que $\omega(C) \geq s$, ou seja, $\omega(\mathbf{c}) \geq s$ para todo $\mathbf{c} = (c_1, \ldots, c_n) \in C$. Agora suponha por absurdo que existam $s-1$ colunas de H linearmente dependentes, a saber $h^{i_1}, \ldots, h^{i_{s-1}}$. Logo existem $c_{i_1}, \ldots, c_{i_{s-1}} \in K$, nem todos nulos, tais que

$$c_{i_1}h^{i_1} + \cdots + c_{i_{s-1}}h^{i_{s-1}} = \left[0\right]_{(n-k)\times 1}. \tag{4.5}$$

Note que existe um vetor $\mathbf{c} = (0, \ldots, c_{i_1}, \ldots, c_{i_{s-1}}, \ldots, 0) \in C$, pois $H\mathbf{c}^t = c_{i_1}h^{i_1} + \cdots + c_{i_{s-1}}h^{i_{s-1}} = \left[0\right]_{(n-k)\times 1}$, com no máximo $s-1$ coordenadas não nulas, isto é, $\omega(\mathbf{c}) \leq s-1$. Contradição, pois temos por hipótese que $\omega(C) \geq s$.

Reciprocamente, supondo que quaisquer $s-1$ colunas de H são linearmente independentes e sendo $\mathbf{c} \in C$, temos que $H\mathbf{c}^t = \left[0\right]_{(n-k)\times 1}$, logo

$$\begin{aligned}
\left[0\right]_{(n-k)\times 1} &= H\mathbf{c}^t = \begin{bmatrix} h_{11} & h_{12} & \ldots & h_{1n} \\ h_{21} & h_{22} & \ldots & h_{2n} \\ \vdots & \vdots & & \vdots \\ h_{(n-k)1} & h_{(n-k)2} & \ldots & h_{(n-k)n} \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \ldots & c_n \end{bmatrix}^t = \\
&= \begin{bmatrix} h_{11}c_1 & + & h_{12}c_2 & + & \ldots & + & h_{1n}c_n \\ h_{21}c_1 & + & h_{22}c_2 & + & \ldots & + & h_{2n}c_n \\ \vdots & & \vdots & & & & \vdots \\ h_{(n-k)1}c_1 & + & h_{(n-k)2}c_2 & + & \ldots & + & h_{(n-k)n}c_n \end{bmatrix} = \\
&= c_1 \begin{bmatrix} h_{11} \\ h_{21} \\ \vdots \\ h_{(n-k)1} \end{bmatrix} + c_2 \begin{bmatrix} h_{12} \\ h_{22} \\ \vdots \\ h_{(n-k)2} \end{bmatrix} + \ldots + c_n \begin{bmatrix} h_{1n} \\ h_{2n} \\ \vdots \\ h_{(n-k)n} \end{bmatrix}.
\end{aligned}$$

Denotando as colunas de H por $h^1, h^2, \ldots, h^n$, a combinação acima pode ser escrita como

$$\left[0\right]_{(n-k)\times 1} = c_1h^1 + c_2h^2 + \ldots + c_nh^n. \tag{4.6}$$

Como o peso de uma palavra é a quantidade de suas coordenadas não nulas, segue que quando $\omega(\mathbf{c}) \leq s-1$, temos no máximo $s-1$ coordenadas de $\mathbf{c}$ não nulas, deste modo, pela equação (4.6),

$$\left[0\right]_{(n-k)\times 1} = c_{i_1}h^{i_1} + c_{i_2}h^{i_2} + \ldots + c_{i_{s-1}}h^{i_{s-1}},$$

temos uma combinação linear nula de $s-1$ colunas de H, em que nem todos os coeficientes $c_{i_1}, \ldots, c_{i_{s-1}}$ são nulos, o que é um absurdo, pois a priori supomos que quaisquer $s-1$ colunas de H são linearmente independentes. Portanto $\omega(\mathbf{c}) \geq s$. □

Teorema 18. *Seja H a matriz teste de paridade de um código C. Temos que $\omega(C) = s$ se, e somente se, quaisquer $s-1$ colunas de H são linearmente independentes e existem s colunas de H linearmente dependentes.*

Demonstração. Suponha $\omega(C) = s$. Pela Proposição 40, temos que quaisquer $s-1$ colunas de H são linearmente independentes. Além disso, existem s colunas de H linearmente dependentes, pois caso contrário poderíamos ter que quaisquer s colunas da matriz H seriam linearmente

independentes e, dessa forma, teríamos pela Proposição 40 que $\omega(C) \geq s + 1$, contradizendo o que supomos a priori.

Reciprocamente, suponha que quaisquer $s - 1$ colunas de H sejam linearmente independentes e s colunas linearmente dependentes. Pela Proposição 40, $\omega(C) \geq s$, no entanto, não podemos ter $\omega(C) > s$, pois teríamos que quaisquer s colunas de H seriam linearmente independentes, o que contradiz a nossa hipótese. □

Corolário 11 (Cota de *Singleton*)**.** *Os parâmetros* (n, k, d) *de um código linear* C *satisfazem a desigualdade*

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

Demonstração. Seja H a matriz teste de paridade de C. Temos que H tem posto $n - k$, pois possui $n - k$ colunas linearmente independentes. Além disso, pelo Teorema 18, $d = \omega(C)$ e, consequentemente, quaisquer $d - 1$ colunas de H são linearmente independentes. Logo $d - 1 \leq n - k$ e, portanto, $d \leq n - k - 1$. □

4.3 Decodificação

Nesta seção, abordaremos um dos assuntos mais importantes desse trabalho, a *decodificação*. Nossos esforços em apresentar alguns conceitos e resultados matemáticos fundamentais na teoria que envolve os códigos corretores de erros, serão, de certo modo, recompensados nessa seção ao visualizarmos como ocorre o processo de detecção e correção de erros em códigos lineares.

Durante o processo de transmissão de dados, nos interessa saber se a informação que desejamos transmitir chegará integralmente ao destinatário. Todavia, é sabido que tanto por motivos internos, quanto por externos, podem ocorrer interferências nos canais de transmissão de modo a comprometer a integridade da informação. Nesse sentido, é de nosso interesse o desenvolvimento de um método, o qual chamaremos de decodificação, que seja capaz de detectar e, se possível, corrigir esses erros. Entretanto, antes de estruturar o *algoritmo da decodificação*, apresentaremos alguns resultados imprescindíveis em sua construção.

Definição 56. *Sejam* $C \subset K^n$ *um código linear e* $\mathbf{c} \in C$*. Se após uma transmissão da palavra* $\mathbf{c}$*, a palavra recebida for* $\mathbf{r}$*, o vetor erro é definido como* $\mathbf{e} = \mathbf{r} - \mathbf{c}$*.*

Proposição 41. *Se* H *é a matriz teste de paridade do código* C *e* $\mathbf{c} \in C$*, então* $H\mathbf{e}^t = H\mathbf{r}^t$*.*

Demonstração. Como $\mathbf{c} \in C$, então pela Proposição 39, segue que

$$H\mathbf{c}^t = \left[0\right]_{(n-k)\times 1}.$$

Logo,

$$H\mathbf{e}^t = H(\mathbf{r}-\mathbf{c})^t = H\mathbf{r}^t - H\mathbf{c}^t = H\mathbf{r}^t$$

e, portanto, a palavra recebida e o vetor erro tem a mesma síndrome. □

Lema 14. *Seja C um código linear em K^n com capacidade de correção κ. Se $\mathbf{r} \in K^n$ e $\mathbf{c} \in C$ são tais que $d(\mathbf{c},\mathbf{r}) \leq \kappa$, então existe um único vetor $\mathbf{e}$, com $\omega(\mathbf{e}) \leq \kappa$, cuja síndrome é igual à síndrome de $\mathbf{r}$ e tal que $\mathbf{c} = \mathbf{r} - \mathbf{e}$.*

Demonstração. Provemos a existência do vetor erro $\mathbf{e}$. Temos que, se $d(\mathbf{c},\mathbf{r}) \leq \kappa$, então $\omega(\mathbf{r}-\mathbf{c}) \leq \kappa$, logo existe $\mathbf{e} = \mathbf{r} - \mathbf{c}$, tal que $\omega(\mathbf{e}) \leq \kappa$. Provemos agora a unicidade do vetor erro. Para tanto, suponha que existem dois vetores erros $\mathbf{e}_1$ e $\mathbf{e}_2$, com a mesma síndrome de $\mathbf{r}$, tais que $\omega(\mathbf{e}_1) \leq \kappa$ e $\omega(\mathbf{e}_2) \leq \kappa$. Desse modo, temos que

$$H\mathbf{e}_1^t = H\mathbf{e}_2^t \Rightarrow H(\mathbf{e}_1^t - \mathbf{e}_2^t) = \left[0\right]_{(n-k)\times 1} \Rightarrow H(\mathbf{e}_1 - \mathbf{e}_2)^t = \left[0\right]_{(n-k)\times 1} \Rightarrow \mathbf{e}_1 - \mathbf{e}_2 \in C.$$

Agora, pelo item (iii) da Proposição 36 e pela Definição 46, temos

$$d(\mathbf{e}_1,\mathbf{e}_2) \leq d(\mathbf{e}_1,0) + d(0,\mathbf{e}_2) = \omega(\mathbf{e}_1) + \omega(\mathbf{e}_2) \leq \kappa + \kappa = 2\kappa \leq d - 1 < d.$$

Portanto a distância entre $\mathbf{e}_1$ e $\mathbf{e}_2$ é menor que a distância mínima do código C, o que é uma contradição, a menos que $\mathbf{e}_1 = \mathbf{e}_2$. □

Definição 57. *Seja $\mathbf{v} \in K^n$. Definimos por* classe lateral *de $\mathbf{v}$ segundo C o conjunto*

$$\mathbf{v} + C = \{\mathbf{v} + \mathbf{c}; \mathbf{c} \in C\}.$$

Lema 15. *Sejam $\mathbf{u}, \mathbf{v} \in K^n$ e $\mathbf{c} \in C$. Então os vetores $\mathbf{u}$ e $\mathbf{v}$ têm a mesma síndrome se, e somente se, $\mathbf{u} \in \mathbf{v} + C$.*

Demonstração. Com efeito,

$$H\mathbf{u}^t = H\mathbf{v}^t \Leftrightarrow H(\mathbf{u}-\mathbf{v})^t = \left[0\right]_{(n-k)\times 1} \Leftrightarrow \mathbf{u}-\mathbf{v} \in C \Leftrightarrow \mathbf{u}-\mathbf{v} = \mathbf{c},\ \mathbf{c} \in C \Leftrightarrow \mathbf{u} = \mathbf{v}+\mathbf{c},\ \mathbf{c} \in C \Leftrightarrow$$
$$\Leftrightarrow \mathbf{u} \in \mathbf{v} + C.$$

□

Proposição 42. *Seja C um código linear de dimensão k, cujas palavras têm comprimento n. Temos que*

(i) $\mathbf{v} + C = \mathbf{v}' + C$ *se, e somente se,* $\mathbf{v} - \mathbf{v}' \in C$;

(ii) se $(\mathbf{v} + C) \cap (\mathbf{v}' + C) \neq \emptyset$, *então* $\mathbf{v} + C = \mathbf{v}' + C$;

(iii) $\bigcup_{\mathbf{v}\in K^n} (\mathbf{v}+C) = K^n$;

(iv) $|(\mathbf{v}+C)| = |C| = q^k$.

Demonstração. (i) Suponha que $\mathbf{v}+C = \mathbf{v}'+C$. Temos que $\mathbf{v} = \mathbf{v}+\mathbf{0} \in \mathbf{v}+C$, pois $\mathbf{0} \in C$. Como $\mathbf{v}+C = \mathbf{v}'+C$, então $\mathbf{v} \in \mathbf{v}'+C$. Portanto, $\mathbf{v} = \mathbf{v}'+\mathbf{c}'$ para algum $\mathbf{c}' \in C$. Assim, $\mathbf{v}-\mathbf{v}' = \mathbf{c}' \in C$.

Reciprocamente, sejam $\mathbf{v}, \mathbf{v}' \in K^n$ tais que $\mathbf{v}-\mathbf{v}' \in C$. Considere agora $\mathbf{v}+\mathbf{c}$ com $\mathbf{c} \in C$ um elemento qualquer de $\mathbf{v}+C$. Como $\mathbf{v}-\mathbf{v}' \in C$, existe $\mathbf{c}' \in C$ tal que $\mathbf{v}-\mathbf{v}' = \mathbf{c}'$. Daí,

$$\mathbf{v}+\mathbf{c} = \mathbf{v}-\mathbf{v}'+\mathbf{v}'+\mathbf{c} = \mathbf{c}'+\mathbf{v}'+\mathbf{c} = \mathbf{v}'+\mathbf{c}'' \in \mathbf{v}'+C,$$

ou seja, $\mathbf{v}+C \subset \mathbf{v}'+C$. De modo análogo, prova-se que $\mathbf{v}+C \subset \mathbf{v}'+C$. Consequentemente, $\mathbf{v}+C = \mathbf{v}'+C$.

(ii) Suponha que $(\mathbf{v}+C) \cap (\mathbf{v}'+C) \neq \emptyset$, então existe $\mathbf{u} \in (\mathbf{v}+C) \cap (\mathbf{v}'+C)$ tal que

$$\mathbf{u} = \mathbf{v}+\mathbf{c}_1 \in \mathbf{v}+C, \text{ para algum } \mathbf{c}_1 \in C \tag{4.7}$$

e

$$\mathbf{u} = \mathbf{v}'+\mathbf{c}_2 \in \mathbf{v}'+C, \text{ para algum } \mathbf{c}_2 \in C. \tag{4.8}$$

Daí, substituindo (4.7) em (4.8), temos que $\mathbf{v}+\mathbf{c}_1 = \mathbf{v}'+\mathbf{c}_2$, com $\mathbf{c}_1, \mathbf{c}_2 \in C$. Logo, $\mathbf{v}-\mathbf{v}' = \mathbf{c}_1+\mathbf{c}_2 \in C$. Portanto, pelo item (i), segue que $\mathbf{v}+C = \mathbf{v}'+C$.

(iii) Seja $\mathbf{u} \in K^n$ qualquer. Temos que $\mathbf{u} \in \mathbf{u}+C \subset \bigcup_{\mathbf{v}\in K^n} (\mathbf{v}+C)$ e, portanto, $\bigcup_{\mathbf{v}\in K^n} (\mathbf{v}+C) = K^n$.

(iv) De fato, temos que

$$|(\mathbf{v}+C)| = |\{\mathbf{v}+\mathbf{c};\ \mathbf{c} \in C\}| = |\{\mathbf{v}+\mathbf{c}_1, \mathbf{v}+\mathbf{c}_2, \ldots, \mathbf{v}+\mathbf{c}_{q^k}\}| \leq q^k = |C|.$$

Suponha por absurdo que vale a desigualdade estrita acima. Então, existem $\mathbf{c}_i \neq \mathbf{c}_j$ tais que $\mathbf{v}+\mathbf{c}_i = \mathbf{v}+\mathbf{c}_j$, logo

$$\mathbf{v}+\mathbf{c}_i = \mathbf{v}+\mathbf{c}_j \Rightarrow \mathbf{c}_i = \mathbf{c}_j,$$

o que é um absurdo. Portanto,

$$|(\mathbf{v}+C)| = q^k = |C|.$$

□

Corolário 12. *Seja C um código linear de dimensão k, onde as palavras tem comprimento n. A quantidade de classes laterais de $\mathbf{v}$ segundo C é $\alpha = q^{n-k}$.*

Demonstração. Considere S um conjunto formado por um representante de cada classe lateral de C. Pelo item (iii) da Proposição 42, temos que $\bigcup_{\mathbf{v}\in K^n}(\mathbf{v}+C) = K^n$. Agora, suponha que existem $\mathbf{v}, \mathbf{v}' \in K^n$ tais que

$$\mathbf{v}+C \cap \mathbf{v}'+C \neq \emptyset.$$

Segue do item (ii) da Proposição 42 que $\mathbf{v}+C = \mathbf{v}'+C$. Logo,

$$K^n = \bigcup_{\mathbf{v}\in S}(\mathbf{v}+C)$$

é uma união disjunta das classes laterais de C. Como $|K^n| = q^n$, então pelos itens (iii) e (iv) da Proposição 42,

$$q^n = |K^n| = \left|\bigcup_{\mathbf{v}\in S}(\mathbf{v}+C)\right| = \sum_{\mathbf{v}\in S}|(\mathbf{v}+C)| = \sum_{\mathbf{v}\in S}|C| = \alpha q^k,$$

em que $\alpha = |S|$. Logo $\alpha q^k = q^n$ e, portanto, $\alpha = q^{n-k}$. □

Proposição 43. *Seja $C \in K^n$ um código linear e $\mathbf{v}+C$ uma classe lateral de $\mathbf{v}$ segundo C. Então*

$$\mathbf{v}+C = C \Leftrightarrow \mathbf{v} \in C.$$

Demonstração. Suponha que $\mathbf{v}+C = C$, então $\mathbf{v} = \mathbf{v}+\mathbf{0} \in \mathbf{v}+C = C$. Por outro lado, se $\mathbf{v} \in C$, então $\mathbf{v}+\mathbf{c} \in C$ para todo $\mathbf{c} \in C$ e, como pela Definição 57 temos $\mathbf{v}+\mathbf{c} \in \mathbf{v}+C$, segue que $\mathbf{v} \in (\mathbf{v}+C) \cap C$. Logo, pelo item (ii) da Proposição 42, temos $\mathbf{v}+C = C$. □

Definição 58. *Um vetor de peso mínimo numa classe lateral é chamado de* elemento líder *dessa classe.*

Proposição 44. *Seja C um código linear em K^n com distância mínima d. Se $\mathbf{u} \in K^n$ é tal que*

$$\omega(\mathbf{u}) \leq \left\lfloor \frac{d-1}{2} \right\rfloor = \kappa,$$

então $\mathbf{u}$ é o único elemento líder de sua classe.

Demonstração. Suponhamos por absurdo que existem dois elementos distintos, líderes da mesma classe, digamos $\mathbf{u}, \mathbf{v} \in K^n$. Daí teríamos que

$$\omega(\mathbf{u}-\mathbf{v}) \leq \omega(\mathbf{u}) + \omega(\mathbf{v}) \leq \left\lfloor \frac{d-1}{2} \right\rfloor + \left\lfloor \frac{d-1}{2} \right\rfloor \leq 2\kappa \leq d-1.$$

Portanto $\mathbf{u} = \mathbf{v}$, pois a distância mínima do código linear C é d. □

A Proposição 44 assegura a unicidade do elemento líder de uma classe quando este atende a desigualdade $\omega(\mathbf{u}) \leq \kappa$ e, mais ainda, que ele é líder de somente uma classe. Com isso, para determinarmos líderes de classes, selecionamos todos os elementos $\mathbf{u}$ tais que $\omega(\mathbf{u}) \leq \kappa$. Os elementos líderes que não atendem a essa condição serão desconsiderados, visto que seu peso supera a capacidade de correção.

Antes de iniciarmos o processo de detecção e correção de erros, determinamos todos os possíveis elementos líderes de peso menor ou igual a κ e calculamos suas respectivas síndromes. Em seguida, organizamos esses dados em uma tabela.

O processo de detecção e correção de erros é o que chamamos de decodificação, o qual baseia-se primeiramente calcular a síndrome da palavra recebida $\mathbf{r}$ e verificar se tal palavra atende a Proposição 39. Caso isso não aconteça, significa que ocorreram erros na transmissão dessa palavra e sua integridade foi comprometida. Contudo, podemos verificar a possibilidade de corrigir esses erros.

Observe que, pela Proposição 41, a síndrome do vetor erro $\mathbf{e}$ coincide com a síndrome da palavra recebida $\mathbf{r}$. Além disso, pelo Lema 15, o vetor erro $\mathbf{e}$ está na classe lateral $\mathbf{r} + C$, determinada por $\mathbf{r}$. Então, se $\omega(\mathbf{e}) \leq \kappa$, o vetor $\mathbf{e}$ é o único elemento líder de sua classe, como vimos na Proposição 44. Logo ele está na tabela de elementos líderes e, portanto, temos pelo Lema 14 que a palavra transmitida é determinada por $\mathbf{c} = \mathbf{r} - \mathbf{e}$.

Em síntese, o processo de decodificação pode ser descrito como um algoritmo, o qual apresentamos como o seguinte Teorema:

Teorema 19. (Algoritmo da Decodificação) *Seja* $\mathbf{r}$ *uma palavra recebida, o algoritmo da decodificação é dado por:*

1. *Calcule a síndrome* $\mathbf{s}^t = H\mathbf{r}^t$*;*
2. *Se* $\mathbf{s}$ *está na tabela das síndromes dos vetores erros (descrita acima), digamos o erro* ℓ*, troque* $\mathbf{r}$ *por* $\mathbf{c} = \mathbf{r} - \ell$*;*
3. *Se* $\mathbf{s}$ *não está na tabela, então na mensagem foram cometidos mais do que* κ *erros.*

Exemplo 12. Considere C, um código linear sobre $\mathbb{F}_2$. Suponha que C tenha dimensão 4 e que suas palavras tenham comprimento 7. Seja H a matriz teste de paridade de C, dada por

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

Note que quaisquer duas colunas da matriz H são linearmente independentes, uma vez que dois vetores são linearmente dependentes se e, somente se, um é múltiplo escalar do outro, mais ainda, como neste caso estamos sobre $\mathbb{F}_2$, dois vetores são linearmente dependentes se, e somente se são iguais, o que não acontece com as colunas de H. Além disso, temos que existem três colunas de H linearmente dependentes, pois a quarta coluna de H é a soma das duas primeiras colunas de H. Logo, pelo Teorema 18, $d = 3$ e, consequentemente, $\kappa = 1$. Daí, temos que os vetores de peso menor ou igual a $\kappa = 1$ em K^n e suas respectivas síndromes, são dados na tabela abaixo.

vetores erros	síndrome
0000000	000
0000001	101
0000010	111
0000100	011
0001000	110
0010000	001
0100000	010
1000000	100

.

Seja $\mathbf{r} = 1100111$, temos que

$$H\mathbf{r}^t = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^t = 111^t = \mathbf{s}^t.$$

Como $\mathbf{s}^t$ está na tabela, segue que o erro detectado é $\ell = 0000010$. Portanto, a palavra enviada é $\mathbf{c} = \mathbf{r} - \ell = 1100111 - 0000010 = 1100101$.

Convém observar que, se durante uma transmissão ocorrerem mais erros do que a capacidade de detecção do código C, pode acontecer uma correção equivocada, vejamos a seguinte situação: suponhamos que a palavra enviada seja $\mathbf{c} = 1110100$ e a recebida seja $\mathbf{r} = 1111011$, isto é, ocorreram 4 erros durante a transmissão. Temos que

$$H\mathbf{r}^t = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}^t = 011^t = \mathbf{s}^t.$$

Como $\mathbf{s}^t$ está na tabela, segue que o erro detectado é $\ell = 000100$ e, consequentemente, a correção é feita para a palavra $\mathbf{r} - \ell = 1111011 - 0000100 = 1111111$, do código C, que é diferente da palavra enviada $\mathbf{c} = 1110100$.

Exemplo 13. Considere o código binário C com matriz geradora

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

cujas palavras do código de canal são

$$\begin{array}{llll} A = 10000 & B = 01000 & C = 00100 & D = 00010 \\ E = 00001 & F = 11000 & G = 10100 & H = 10010 \\ I = 10001 & J = 01100 & L = 01010 & M = 01001 \\ N = 00110 & O = 00101 & P = 00011 & Q = 11100 \\ R = 10110 & S = 10101 & T = 11010 & U = 11001 \\ V = 01110 & X = 00111 & Z = 11110 & \text{Ç} = 01011 \\ ! = 10111 & \cdot = 01111 & , = 01101 & ; = 10011 \\ ? = 11011 & : 11101 & \ldots = 11111 & \text{espaço} = 00000. \end{array}$$

A transformação linear injetora $T : C \subset \mathbb{F}_2^5 \rightarrow \mathbb{F}_2^9$, em que $T(\mathbf{x}) = \mathbf{x}G$ para todo $\mathbf{x} \in C$, definida na seção 4.1, nos permite codificar essas palavras para o código de canal, assim obtemos a tabela abaixo

Símbolo	Código Fonte	Código de Canal	Símbolo	Código Fonte	Código de Canal
A	10000	110010000	R	10110	011111001
B	01000	100100010	S	10101	000011010
C	00100	111000001	T	11010	000011010
D	00010	010101000	U	11001	011010110
E	00001	001100100	V	01110	001001011
F	11000	010110010	X	00111	100001101
G	10100	001010001	Z	11110	111011011
H	10010	100111000	Ç	01011	111101110
I	10001	111110100	·	01111	000101111
J	01100	011100011	,	01101	010000111
L	01010	110001010	;	10011	101011100
M	01001	101000110	!	10111	010011101
N	00110	101101001	?	11011	001111110
O	00101	110100101	:	11101	100010111
P	00011	011001100	...	11111	110111111
Q	11100	101110011	espaço	00000	000000000

Note que o comprimento do código é $n = 9$ e sua dimensão é $k = 5$ sobre $\mathbb{F}_2$. Além disso, realizando operações elementares nas linhas da matriz G, obtemos uma $\tilde{G} = (Id_5 \mid A^t)$ na forma

padrão

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

A partir da matriz $\tilde{G}$ e pela Proposição 38, determinamos a matriz teste de paridade

$$H = (-A^t \mid Id_4) = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Daí, como quaisquer duas colunas de H são linearmente independentes, pois como estamos sobre $\mathbb{F}_2$, para duas colunas serem linearmente dependentes elas devem ser iguais, o que não acontece com a matriz H. Além disso, existem três colunas linearmente dependentes, pois a quinta coluna de H é a soma da sexta e oitava coluna de H. Logo, pelo Teorema, 18 $d = 3$, ou seja, $\kappa = 1$. Deste modo, os elementos líderes cujo $\omega(\ell) \leq 1$ e suas respectivas síndromes são dadas de acordo com a tabela abaixo

Líderes	Síndromes
000000000	0000
000000001	0001
000000010	0010
000000100	0100
000001000	1000
000010000	1010
000100000	1111
001000000	1011
010000000	0111
100000000	1101

.

Suponhamos que após uma transmissão foram recebidas as palavras $\mathbf{r}_1 = 011010111$, $\mathbf{r}_2 = 010110010$, $\mathbf{r}_3 = 110100101$, $\mathbf{r}_4 = 010001100$ e $\mathbf{r}_5 = 010011101$. Aplicando o Algoritmo da decodificação, temos que

$$H\mathbf{r}_1^t = 0001;\ H\mathbf{r}_2^t = 0000;\ H\mathbf{r}_3^t = 0000;\ H\mathbf{r}_4^t = 1011;\ H\mathbf{r}_5^t = 0000.$$

Daí, ao compararmos as síndromes encontradas com as síndromes da tabela, conseguimos corrigir as palavras que sofreram erros durante a transmissão para palavras do código C, como segue

$$\mathbf{c}_1 = \mathbf{r}_1 - \ell = 011010111 - 000000001 = 011010110;$$

$$\mathbf{c}_4 = \mathbf{r}_4 - \ell = 010001100 - 001000000 = 011001100.$$

Logo a mensagem corrigida é 011010110 010110010 110100101 011001100 010011101. Decodificando essa mensagem para o código de canal, verificamos que a mensagem transmitida é 11001 11000 00101 00011, cujos símbolos representam a mensagem *"UFOP!"*.

5 Códigos Cíclicos

Neste capítulo serão apresentados os códigos cíclicos, que são uma classe dos códigos lineares, que acabamos de ver no Capítulo 4. Contudo, a estrutura na qual conseguimos operar seus elementos é o grande diferencial desse capítulo, pois possibilita a criação de algoritmos de codificação e decodificação mais rápidos e eficientes, (HEFEZ; VILLELA, 2008). Nosso objetivo aqui é estruturar os códigos cíclicos a fim de mostrarmos, por meio de exemplos, o funcionamento dos seus algoritmos de codificação e decodificação.

Proposição 45. *Sejam K um corpo finito e $C \subset K^n$ um código linear. São equivalentes:*

(i) para todo $\mathbf{c} = (c_0, \ldots, c_{n-1}) \in C$ *temos* $(c_{n-1}, c_0, \ldots, c_{n-2}) \in C$*;*

(ii) $T_\pi(C) \subset C$*, em que* π *é a aplicação*

$$\pi(i) = \begin{cases} i-1, & se \quad i \geq 1 \\ n-1, & se \quad i = 0, \end{cases} \tag{5.1}$$

e T é a permutação de coordenadas definida em C, dada por

$$T_\pi(c_0, \ldots, c_{n-1}) = (c_{n-1}, c_0, \ldots, c_{n-2}). \tag{5.2}$$

Demonstração. $(i) \Rightarrow (ii)$: De fato, se para todo $\mathbf{c} = (c_0, \ldots, c_{n-1}) \in C$ temos $T_\pi(c_0, \ldots, c_{n-1}) = (c_{n-1}, c_0, \ldots, c_{n-2}) \in C$, segue imediatamente que $T_\pi(C) \subset C$.
$(ii) \Rightarrow (i)$: Se $T_\pi(C) \subset C$, então $T_\pi(\mathbf{c}) = T_\pi(c_0, \ldots, c_{n-1}) = (c_{n-1}, c_0, \ldots, c_{n-2}) \in C$ para todo $\mathbf{c} \in C$, como queríamos. □

Definição 59. *Seja $C \subset K^n$ um código linear. Dizemos que C é cíclico quando satisfaz algum dos itens da Proposição 45.*

Exemplo 14. O código $C = \{000, 110, 101, 011\} \subset \mathbb{F}_2^3$ é cíclico. De fato,

$$\begin{aligned} T_\pi(000) &= 000 \in C \\ T_\pi(110) &= 011 \in C \\ T_\pi(101) &= 110 \in C \\ T_\pi(011) &= 101 \in C. \end{aligned}$$

Logo $T_\pi(C) \subset C$ e, portanto, C é cíclico.

Exemplo 15. Seja $\mathbf{v} \in K^n$. O espaço vetorial V, dado por

$$V = \langle \mathbf{v} \rangle = K\mathbf{v} + KT_\pi(\mathbf{v}) + \ldots + KT_\pi^{n-1}(\mathbf{v}),$$

é um código cíclico. De fato, seja $\mathbf{v} = (a_0, a_1, \ldots, a_{n-1})$. Temos que

$$\begin{aligned}
\langle \mathbf{v} \rangle &= KId(a_0, a_1, \ldots, a_{n-1}) + \ldots + KT_\pi^{n-1}(a_0, a_1, \ldots, a_{n-1}) \\
&= K(a_0, a_1, \ldots, a_{n-1}) + \ldots + K(a_1, a_2, \ldots, a_0) \\
&= \{\lambda_0(a_0, a_1, \ldots, a_{n-1}); \lambda_0 \in K\} + \ldots + \{\lambda_{n-1}(a_1, a_2, \ldots, a_0);\ \lambda_{n-1} \in K\} \\
&= \{\lambda_0(a_0, a_1, \ldots, a_{n-1}) + \ldots + \lambda_{n-1}(a_1, a_2, \ldots, a_0);\ \lambda_0, \ldots, \lambda_{n-1} \in K\} \\
&= \{(\lambda_0 a_0, \lambda_0 a_1, \ldots, \lambda_0 a_{n-1}) + \ldots + (\lambda_{n-1} a_1, \lambda_{n-1} a_2, \ldots, \lambda_{n-1} a_0);\ \lambda_0, \ldots, \lambda_{n-1} \in K\} \\
&= \{(\lambda_0 a_0 + \lambda_1 a_{n-1} + \ldots + \lambda_{n-1} a_1, \ldots, \lambda_0 a_{n-1} + \ldots + \lambda_{n-1} a_0);\ \lambda_0, \ldots, \lambda_{n-1} \in K\}.
\end{aligned}$$

Daí, se $\mathbf{u} \in \langle \mathbf{v} \rangle$, então $\mathbf{u} = (\lambda_0 a_0 + \lambda_1 a_{n-1} + \ldots + \lambda_{n-1} a_1, \ldots, \lambda_0 a_{n-1} + \lambda_1 a_{n-2} + \ldots + \lambda_{n-1} a_0)$, com $\lambda_0, \ldots, \lambda_{n-1} \in K$. Deste modo,

$$\begin{aligned}
T_\pi(\mathbf{u}) &= T_\pi(\lambda_0 a_0 + \lambda_1 a_{n-1} + \ldots + \lambda_{n-1} a_1, \ldots, \lambda_0 a_{n-1} + \lambda_1 a_{n-2} + \ldots + \lambda_{n-1} a_0) \\
&= (\lambda_0 a_{n-1} + \lambda_1 a_{n-2} + \ldots + \lambda_{n-1} a_0, \ldots, \lambda_0 a_{n-2} + \lambda_1 a_{n-3} + \ldots + \lambda_{n-1} a_{n-1}) \\
&= \lambda_0(a_{n-1}, a_0, \ldots, a_{n-2}) + \ldots + \lambda_{n-1}(a_0, a_1, \ldots, a_{n-1}) \\
&= \lambda_0 T_\pi(\mathbf{v}) + \lambda_1 T_\pi^2(\mathbf{v}) + \ldots + \lambda_{n-1} Id(\mathbf{v}).
\end{aligned}$$

Logo $T_\pi(\mathbf{u}) \in \langle \mathbf{v} \rangle$ para todo $\mathbf{u} \in \langle \mathbf{v} \rangle$. Portanto, V é um código cíclico.

Veremos a partir de agora que a principal estrutura que utilizaremos para trabalharmos com os códigos cíclicos é o anel quociente de polinômios. Dito isso, chamaremos de R_n o anel quociente de $K[X]$, em que K é um corpo finito com q elementos, pelo ideal gerado pelo polinômio $p(X) = X^n - 1$, isto é,

$$R_n = \frac{K[X]}{I(X^n - 1)} = \{[r(X)];\ r(X) \in K[X],\ \text{com } r(X) = 0,\ \text{ou } gr(r(X)) < n\}.$$

Note que, de acordo com a Observação 3, R_n é um espaço vetorial de dimensão n sobre K. Em vista disso, a função

$$\begin{array}{rccc}
\Psi: & K^n & \to & R_n \\
& (a_0, \ldots, a_{n-1}) & \mapsto & [a_0 + a_1 X + \ldots + a_{n-1} X^{n-1}]
\end{array} \tag{5.3}$$

é um isomorfismo linear. Com efeito, dados $\mathbf{u}, \mathbf{v} \in K^n$ e $\lambda \in K$,

$$\begin{aligned}
\Psi(\mathbf{u}+\lambda\mathbf{v}) &= \Psi((a_0,\ldots,a_{n-1})+\lambda(b_0,\ldots,b_{n-1})) \\
&= \Psi(a_0+\lambda b_0,\ldots,a_{n-1}+\lambda b_{n-1}) \\
&= [(a_0+\lambda b_0)+(a_1+\lambda b_1)X+\ldots+(a_{n-1}+\lambda b_{n-1})X^{n-1}] \\
&= [a_0+a_1X+\ldots+a_{n-1}X^{n-1}+\lambda b_0+\lambda b_1X+\ldots+\lambda b_{n-1}X^{n-1}] \\
&= [a_0+a_1X+\ldots+a_{n-1}X^{n-1}]+\lambda[b_0+b_1X+\ldots+b_{n-1}X^{n-1}] \\
&= \Psi(a_0,\ldots,a_{n-1})+\lambda\Psi(b_0,\ldots,b_{n-1}) \\
&= \Psi(\mathbf{u})+\lambda\Psi(\mathbf{v}).
\end{aligned}$$

Logo Ψ é linear, além disso, como $dim_K K^n = n = dim_K R_n$, para provar que essa função é bijetiva, basta mostrar que Ψ é sobrejetiva. De fato, seja $[a_0+a_1X+\ldots+a_{n-1}X^{n-1}] \in R_n$, temos que $(a_0,\ldots,a_{n-1})$ é a pré imagem de $[a_0+a_1X+\ldots+a_{n-1}X^{n-1}]$, logo Ψ é sobrejetiva. Portanto Ψ é um isomorfismo linear. Como consequência desse isomorfismo, a função Ψ nos permite transportar o código $C \subset K^n$ para R_n, o que é de grande utilidade, pois assim conseguimos propriedades adicionais do código cíclico C quando o enxergamos como ideal de um anel quociente, visto que agora além de subespaço vetorial ele é também um ideal, sendo este o próximo assunto que iremos tratar.

Antes de tudo, cabe observar que aplicar a permutação de coordenadas T_π em K^n, corresponde a multiplicar por $[X]$ em R_n. De fato, tomando $\mathbf{c} = (c_0,\ldots,c_{n-1})$, temos

$$T_\pi(\mathbf{c}) = (c_{n-1}, c_0, \ldots, c_{n-2})$$

e

$$\begin{aligned}
\Psi(T_\pi(\mathbf{c})) &= [c_{n-1}+c_0X+\ldots+c_{n-2}X^{n-1}] \\
&= [X][c_0+c_1X+\ldots+c_{n-1}X^{n-1}] \\
&= [X]\Psi(\mathrm{c}).
\end{aligned}$$

Lema 16. *Seja V um subespaço vetorial de R_n. Então, V é um ideal de R_n se, e somente se, V é fechado pela multiplicação por $[X]$.*

Demonstração. Suponha que V seja um ideal de R_n. Daí, pela definição de ideal, segue que $[X][f(X)] \in V$ para todo $[f(X)] \in R_n$, logo V é fechado pela multiplicação por $[X]$.

Reciprocamente, mostremos por indução sobre o grau de X^n, que V é um ideal de R_n. Para isso, suponha que V seja fechado pela multiplicação por $[X]$. Como V é um subespaço vetorial de R_n, então o elemento nulo de R_n pertence a V e quaisquer que sejam $[f(X)], [g(X)] \in V$,

$[f(X)] + [g(X)] = [f(X) + g(X)] \in V$. Além disso, $a[f(X)] \in V$ para todo $[f(X)] \in V$ e $a \in K$. Como por hipótese

$$[Xf(X)] = [X][f(X)] \in V,$$

então

$$[X^2 f(X)] = [X][Xf(X)] \in V,$$

visto que $[Xf(X)] \in V$. Suponha que $[X^{n-1} f(X)] \in V$, com $n \in \mathbb{N}$. Mostremos que $[X^n f(X)] \in V$. De fato,

$$[X^n f(X)] = [X][X^{n-1} f(X)] \in V,$$

pois estamos supondo que $[X^{n-1} f(X)] \in V$ e V é fechado pela multiplicação por $[X]$. Observe que, para todo $[g(X)] = [a_0 + a_1 X + \ldots + a_{n-1} X^{n-1}] \in R_n$ e para todo $[f(X)] \in V$, temos

$$\begin{aligned}
[g(X)][f(X)] &= [a_0 + a_1 X + \ldots + a_{n-1} X^{n-1}][f(X)] \\
&= (a_0 + a_1[X] + \ldots + a_{n-1}[X^{n-1}])[f(X)] \\
&= a_0[f(X)] + a_1[X][f(X)] + \ldots + a_{n-1}[X^{n-1}][f(X)].
\end{aligned}$$

Logo $[g(X)][f(X)] \in V$, uma vez que cada parcela da soma acima pertence a V pelo que provamos por indução. Portanto, V é um ideal de R_n. □

Teorema 20. *Um subespaço $C \subset K^n$ é um código cíclico se, e somente se, $\Psi(C)$ é um ideal de R_n.*

Demonstração. Primeiramente vamos mostrar que $\Psi(C)$ é um ideal de R_n. Para tanto, suponha que o código C é cíclico. Logo $T_\pi(\mathbf{c}) \in C$ para todo $\mathbf{c} \in C$ e, como

$$\Psi(T_\pi(\mathbf{c})) = [X]\Psi(c), \quad \forall\, \mathbf{c} \in C,$$

segue que $[X]\Psi(c) \in \Psi(C)$ para todo $\mathbf{c} \in C$. Deste modo, como $\Psi(C)$ é um subespaço vetorial de R_n e como acabamos de ver, é também fechado pela multiplicação por $[X]$, segue do Lema 16 que $\Psi(C)$ é um ideal de R_n.

Reciprocamente, suponha $\Psi(C)$ um ideal de R_n. Temos que $\Psi(C)$ é fechado pela multiplicação por $[X]$, ou seja, $[X]\Psi(\mathbf{c}) \in \Psi(C)$ para todo $\mathbf{c} = (c_0, \ldots, c_n) \in C$. Além disso, como Ψ é um isomorfismo e $\Psi(T_\pi(\mathbf{c})) = [X]\Psi(\mathbf{c})$ para todo $\mathbf{c} \in C$, então

$$\begin{aligned}
T_\pi(\mathbf{c}) &= \Psi^{-1}([X]\Psi(\mathbf{c}) \\
&= \Psi^{-1}([X][c_0 + c_1 X + \ldots + c_{n-1} X^{n-1}) \\
&= \Psi^{-1}(c_{n-1} + c_0[X] + \ldots + c_{n-2}[X^{n-1}]) \\
&= (c_{n-1}, c_0, \ldots, c_{n-2}) \in C.
\end{aligned}$$

Como vale para todo $\mathbf{c} \in C$, segue que $T_\pi(C) \subset C$ e, portanto, C é um código cíclico. □

Segue dos Teoremas 9 e 20, que um código $C \in K^n$ é cíclico se, e somente se, $\Psi(C) = I([g(X)])$, em que $g(X) \in K[X]$ é um divisor de $X^n - 1$.

A caracterização de códigos cíclicos feita acima é essencial para o desenvolvimento desses códigos, visto que agora os trataremos como um ideal que, por ser uma estrutura algébrica diferente de espaços vetoriais, adiciona algumas propriedades que serão bastante úteis na construção dos códigos cíclicos.

Seja $p = car(K)$ e $n = mp^s$, m e p primos entre si e s um número natural, então por (2.35),

$$X^n - 1 = X^{mp^s} - 1 = (X^m - 1)^{p^s}.$$

Além disso,

$$(X^m - 1)' = mX^{m-1} \neq 0 \Rightarrow mdc(X^m - 1, mX^{m-1}) = 1,$$

com $m \neq 0$ em $\mathbb{F}_p$, pois m e p são primos entre si. Daí, pela Proposição 16, o polinômio $X^m - 1$ não possui fatores múltiplos não constantes e, assim,

$$X^m - 1 = f_1(X) \dots f_r(X),$$

com $f_i(X)$ irredutíveis e mônicos para todo $1 \leq i \leq r$, dois a dois distintos. Logo,

$$X^n - 1 = (X^m - 1)^{p^s} = (f_1(X) \dots f_r(X))^{p^s} = f_1(X)^{p^s} \dots f_r(X)^{p^s}.$$

Observação 10. Observe que, se $g(X) \in D(X^n - 1)$, em que $D(X^n - 1)$ é o conjunto dos divisores mônicos de $X^n - 1$, então

$$g(X) = f_1(X)^{\alpha_1} \dots f_r(X)^{\alpha_r},$$

com $0 \leq \alpha_i \leq p^s$ para todo $1 \leq i \leq r$. Logo

$$|D(X^n - 1)| = (p^s + 1)^r.$$

Proposição 46. *A função* Ω *definida como*

$$\begin{array}{rcl} \Omega : \mathcal{I}(R_n) & \to & D(X^n - 1) \\ I([g(X)]) & \mapsto & g(X) \end{array},$$

com $g(X)$ *mônico, é uma bijeção entre o conjunto* $\mathcal{I}(R_n)$ *dos ideais de* R_n *e o conjunto* $D(X^n - 1)$.

Demonstração. Primeiramente, vamos mostrar que a função Ω está bem definida. Temos pelo Teorema 9 que todo ideal de R_n é do tipo $I([g(X)])$, sendo $g(X)$ um divisor de $X^n - 1$. Daí,

$$I([g_1(X)]) = I([g_2(X)]) \Rightarrow \underbrace{g_1(X) = a(X)g_2(X)}_{(I)} \text{ e } \underbrace{g_2(X) = b(X)g_1(X)}_{(II)}$$

Substituindo (II) em (I), temos que

$$g_1(X) = a(X)b(X)g_1(X) \Leftrightarrow a(X)b(X) = 1.$$

Assim, segue que $a(X)$ e $b(X)$ são invertíveis, ou seja, $a(X) = a \in K \setminus \{0\}$ e $b(X) = b \in K \setminus \{0\}$. Logo, $g_1(X) = ag_2(X)$ e $g_2(X) = bg_1(X)$. Como $g_1(X)$ e $g_2(X)$ são mônicos, segue que $a = b = 1$. Portanto $g_1(X) = g_2(X)$, consequentemente, Ω está bem definida.

Vamos agora mostrar que Ω é bijetiva. De fato, sendo $g_1(X), g_2(X) \in D(X^n - 1)$ tais que $g_1(X) = g_2(X)$, temos que $I([g_1(X)]) = I([g_2(X)])$, logo Ω é injetiva. Por outro lado, se $g(X) \in D(X^n - 1)$, segue do Teorema 9 que $I([g(X)]) \in \mathcal{I}(R_n)$. Logo, $\Omega(I([g(X)])) = g(X)$, consequentemente, Ω é sobrejetiva. Portanto Ω é bijetiva. □

Note que, pela Observação 10, concluímos que o conjunto dos divisores mônicos de $X^n - 1$ possui $(p^s + 1)^r$ elementos. Além disso, pela Proposição 46, vimos que existe uma bijeção entre esse conjunto e o conjunto dos ideais de R_n. Portanto, R_n possui $(p^s + 1)^r$ ideais.

Sendo o polinômio $g(X) \in K[X]$ de grau s um divisor de $X^n - 1$, definimos o polinômio $h(X)$ por

$$h(X) = \frac{X^n - 1}{g(X)}. \tag{5.4}$$

Teorema 21. *Seja $I = I([g(X)])$, em que $g(X)$ é um divisor de $X^n - 1$ de grau s. Temos que $\mathcal{B} = \{[g(X)], [Xg(X)], [X^2g(X)], \ldots, [X^{n-s-1}g(X)]\}$ é uma base de I como espaço vetorial sobre K.*

Demonstração. Primeiro verificaremos que $\mathcal{B}$ é linearmente independente. Para isso, dados $a_0, \ldots, a_{n-s-1} \in K$, suponha que

$$a_0[g(X)] + a_1[Xg(X)] + \ldots + a_{n-s-1}[X^{n-s-1}g(X)] = [0].$$

Daí,

$$[a_0g(X) + \ldots + a_{n-s-1}X^{n-s-1}g(X)] = [0] \Leftrightarrow [g(X)(a_0 + \ldots + a_{n-s-1}X^{n-s-1})] = [0].$$

Segue que existe $d(X) \in K[X]$, tal que

$$\begin{aligned} & g(X)(a_0 + a_1X + \ldots + a_{n-s-1}X^{n-s-1}) = d(X)(X^n - 1) \\ \Leftrightarrow \quad & (a_0 + a_1X + \ldots + a_{n-s-1}X^{n-s-1}) = d(X)h(X). \end{aligned}$$

Observe que, como o $gr(h(X)) = n - s$, a última igualdade acima é verdadeira somente se $d(X) = 0$, então

$$a_0 + a_1X + \ldots + a_{n-s-1}X^{n-s-1} = 0,$$

o que implica em $a_i = 0$ para todo $0 \leq i \leq n - s - 1$. Portanto, $\mathcal{B}$ é linearmente independente.

Mostremos agora que $\mathcal{B}$ gera I sobre K. Tomando $[f(X)] \in I$, temos que existe $r(X) \in K[X]$ tal que

$$[f(X)] = [t(X)g(X)],$$

em I. Assim,

$$\begin{aligned}
f(X) \equiv t(X)g(X)\ (mod\ I(X^n - 1)) &\Leftrightarrow f(X) - t(X)g(X) \in I(X^n - 1) \\
&\Leftrightarrow f(X) - t(X)g(X) = c(X)(X^n - 1);\ c(X) \in K[X] \\
&\Leftrightarrow f(X) - t(X)g(X) = c(X)h(X)g(X) \\
&\Leftrightarrow f(X) = c(X)h(X)g(X) + t(X)g(X) \\
&\Leftrightarrow f(X) = g(X)(\underbrace{c(X)h(X) + t(X)}_{=d(X) \in K[X]}).
\end{aligned}$$

Daí, pelo Teorema 4, temos que existem únicos $q(X), r(X) \in K[X]$ tais que

$$d(X) = q(X)h(X) + r(X), \quad \text{com } r(X) = 0 \text{ ou } gr(r(X)) < gr(h(X)) = n - s,$$

isto é, o polinômio $r(X)$ é do tipo

$$r(X) = b_0 + b_1X + \ldots + b_{n-s-1}X^{n-s-1},$$

em que $b_0, b_1, \ldots, b_{n-s-1} \in K$. Deste modo,

$$\begin{aligned}
f(X) \equiv d(X)g(X)\ (mod\ I(X^n - 1)) &\equiv q(X)h(X)g(X) + r(X)g(X)\ (mod\ I(X^n - 1)) \\
&\equiv q(X)(x^n - 1) + r(X)g(X)\ (mod\ I(X^n - 1)) \\
&\equiv r(X)g(X)\ (mod\ I(X^n - 1)).
\end{aligned}$$

Assim,

$$\begin{aligned}
[f(X)] &= [r(X)g(X)] \\
&= [(b_0 + b_1X + \ldots + a_{n-s-1}X^{n-s-1})g(X)] \\
&= [b_0g(X) + b_1Xg(X) + \ldots + b_{n-s-1}X^{n-s-1}g(X)] \\
&= [b_0g(X)] + [b_1Xg(X)] + \ldots + [b_{n-s-1}X^{n-s-1}g(X)] \\
&= b_0[g(X)] + b_1[Xg(X)] + \ldots + b_{n-s-1}[X^{n-s-1}g(X)].
\end{aligned}$$

Logo $\mathcal{B}$ é um conjunto gerador de I sobre K. Portanto, como $\mathcal{B}$ é linearmente independente e gera I sobre K, segue que $\mathcal{B}$ é uma base de I sobre K. □

Corolário 13. *Dado um código cíclico C, existe $\mathbf{v} \in C$ tal que $C = \langle \mathbf{v} \rangle$.*

Demonstração. Seja $\Psi(C) = I([g(X)])$. Temos pelo Teorema 21 que $\mathcal{B}$ é uma base de $I = I([g(X)])$ e, como Ψ é um isomorfismo linear, existe um único $\mathbf{v} \in C$ tal que $\mathbf{v} = \Psi^{-1}([g(X)])$. Primeiro, vamos mostrar que $C = \Psi^{-1}(I[g(X)]) \subset \langle \mathbf{v} \rangle$. Para isso, seja $\Psi(\mathbf{c}) = [f(X)] \in I$, temos que

$$\begin{aligned} [f(X)] &= [a_0 g(X) + a_1 X g(X) + \ldots + a_{n-s-1} X^{n-s-1} g(X)] \\ &= a_0[g(X)] + a_1[Xg(X)] + \ldots + a_{n-s-1}[X^{n-s-1} g(X)], \end{aligned}$$

em que $a_0, \ldots, a_{n-s-1} \in K$. Daí,

$$\begin{aligned} \mathbf{c} &= \Psi^{-1}([f(X)]) = \Psi^{-1}(a_0[g(X)] + a_1[Xg(X)] + \ldots + a_{n-s-1}[X^{n-s-1} g(X)]) \\ &= a_0 \Psi^{-1}([g(X)]) + a_1 \Psi^{-1}([Xg(X)]) + \ldots + a_{n-s-1} \Psi^{-1}([X^{n-s-1} g(X)]) \\ &= a_0 \mathbf{v} + a_1 T_\pi(\mathbf{v}) + a_2 T_\pi^2(\mathbf{v}) + \ldots + a_{n-s-1} T_\pi^{n-s-1}(\mathbf{v}). \end{aligned}$$

Logo $\mathbf{c} = \Psi^{-1}([f(X)]) \in \langle \mathbf{v} \rangle$. Como $\mathbf{c}$ é arbitrário, segue que $C = \Psi^{-1}(I([g(X)])) \subset \langle \mathbf{v} \rangle$.

Vamos agora mostrar que $\langle \mathbf{v} \rangle \subset \Psi^{-1}(I[g(X)]) = C$. De fato, se $\mathbf{w} \in \langle \mathbf{v} \rangle$, então

$$\mathbf{w} = b_0 \mathbf{v} + b_1 T_\pi(\mathbf{v}) + b_2 T_\pi^2(\mathbf{v}) + \ldots + b_{n-1} T_\pi^{n-1}(\mathbf{v}),$$

com $b_0, \ldots, b_{n-1} \in K$. Aplicando Ψ em ambos os lados da igualdade, temos

$$\begin{aligned} \Psi(\mathbf{w}) &= \Psi(b_0 \mathbf{v} + b_1 T_\pi(\mathbf{v}) + b_2 T_\pi^2(\mathbf{v}) + \ldots + b_{n-1} T_\pi^{n-1}(\mathbf{v})) \\ &= b_0 \Psi(\mathbf{v}) + b_1 \Psi(T_\pi(\mathbf{v})) + b_2 \Psi(T_\pi^2(\mathbf{v})) + \ldots + b_{n-1} \Psi(T_\pi^{n-1}(\mathbf{v})) \\ &= b_0[g(X)] + b_1[Xg(X)] + \ldots + b_{n-1}[X^{n-s-1} g(X)] \\ &= [g(X)] \underbrace{[b_0 + b_1 X + \ldots + b_{n-1} X^{n-s-1}]}_{= d(X) \in K[X]} \\ &= [g(X)][d(X)]. \end{aligned}$$

Deste modo, $\Psi(\mathbf{w}) \in I([g(X)])$ e assim $\Psi^{-1}(\Psi(\mathbf{w})) \in \Psi^{-1}(I([g(X)]))$, ou seja, $\mathbf{w} \in \Psi^{-1}(I([g(X)]))$ para todo $\mathbf{w} \in \langle \mathbf{v} \rangle$, logo $\langle \mathbf{v} \rangle \subset \Psi^{-1}(I([g(X)])) = C$. Portanto, $C = \langle \mathbf{v} \rangle$. □

Corolário 14. *Seja $g(X) = g_0 + g_1 X + \ldots + g_s X^s \in K[X]$, tal que $g(X)$ é um divisor de $X^n - 1$. Se $I = I([g(X)])$, então*

$$dim_K I = n - s, \tag{5.5}$$

e a matriz

$$G = \begin{bmatrix} \Psi^{-1}([g(X)]) \\ \Psi^{-1}([Xg(X)]) \\ \vdots \\ \Psi^{-1}([X^{n-s-1} g(X)]) \end{bmatrix} = \begin{bmatrix} g_0 & g_1 & \ldots & g_s & 0 & \ldots & 0 \\ 0 & g_0 & g_1 & \ldots & g_s & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & g_0 & \ldots & \ldots & g_s \end{bmatrix}$$

é a matriz geradora do código $C = \Psi^{-1}(I)$.

Demonstração. A $dim_K I = n - s$ segue imediatamente do Teorema 21, pois I possui uma base com $n - s$ elementos. Como $C = \Psi^{-1}(I)$, segue que

$$\{\Psi^{-1}([g(X)]), \Psi^{-1}([Xg(X)]), \ldots, \Psi^{-1}([X^{n-s-1}g(X)])\}$$

é uma base para C. Logo

$$G = \begin{bmatrix} \Psi^{-1}([g(X)]) \\ \Psi^{-1}([Xg(X)]) \\ \vdots \\ \Psi^{-1}([X^{n-s-1}g(X)]) \end{bmatrix}.$$

Assim o resultado segue pela definição de matriz geradora, dada por (4.2). □

A definição e a proposição a seguir serão necessárias para construirmos a matriz teste de paridade de um código cíclico.

Definição 60. *Sejam K um corpo e $F(X) \in K[X]$ de grau t. Definimos polinômio recíproco de $F(X)$ por*

$$F^*(X) = X^t F\left(\frac{1}{X}\right). \tag{5.6}$$

Proposição 47. *Dado $h(X) = h_0 + h_1X + \ldots + h_tX^t \in K[X]$ um polinômio de grau t, então o polinômio recíproco de $h(X)$ atende as seguintes propriedades.*

(i) $h^*(X) = X^t h\left(\frac{1}{X}\right) = h_t + h_{t-1}X + \ldots + h_0X^t$;

(ii) Se $h(X) \mid g(X)$, *então* $h^*(X) \mid g^*(X)$, *com* $gr(g(X)) = s$;

(iii) Se $h(X) \mid (X^n - 1)$, *então* $h^*(X) \mid (X^n - 1)$.

Demonstração. (i) De fato, temos que

$$h(X) = \sum_{i=0}^{t} h_iX^i = h_0 + h_1X + \ldots + h_tX^t.$$

Logo

$$\begin{aligned} h^*(X) &= X^t h\left(\frac{1}{X}\right) = X^t \sum_{i=0}^{t} h_i \left(\frac{1}{X}\right)^i = X^t \left(h_0 + h_1\frac{1}{X} + \ldots + h_t\left(\frac{1}{X}\right)^t\right) \\ &= h_0X^t + h_1X^{t-1} + \ldots + h_{t-1}X + h_t = h_t + h_{t-1}X + \ldots + h_1X^{t-1} + h_0X^t. \end{aligned}$$

(ii) Se $h(X) \mid g(X)$, então existe $q(X) = q_0 + q_1X + \ldots + q_mX^m \in K[X]$ tal que

$$g(X) = q(X)h(X). \tag{5.7}$$

Daí,

$$g^*(X) = (q(X)h(X))^*. \tag{5.8}$$

Da equação (5.7) temos que

$$q(X)h(X) = \sum_{i=0}^{s} c_iX^i = c_0 + c_1X + \ldots + c_sX^s, \tag{5.9}$$

em que $c_i = q_0h_i + q_1h_{i-1} + \ldots + q_ih_0$. Assim, segue do item (i) que

$$(q(X)h(X))^* = X^s \sum_{i=0}^{s} c_i \left(\frac{1}{X}\right)^i = \sum_{i=0}^{s} c_iX^{s-i}.$$

Por outro lado,

$$\begin{aligned} q^*(X) \cdot h^*(X) &= \left(X^m \sum_{i=0}^{m} q_i \left(\frac{1}{X}\right)^i\right) \cdot \left(X^{s-m} \sum_{i=0}^{s-m} h_i \left(\frac{1}{X}\right)^i\right) = X^s \sum_{i=0}^{s} c_i \left(\frac{1}{X}\right)^i = \\ &= \sum_{i=0}^{s} c_iX^{s-i} = (q(X)h(X))^*. \end{aligned}$$

Logo,

$$g^*(X) = (q(X)h(X))^* = q^*(X)h^*(X).$$

Portanto, $h^*(X) \mid g^*(X)$.

(iii) Seja $f(X) = X^n - 1$, então pelo item (i)

$$f^*(X) = -X^n + 1 = -f(X).$$

Temos por hipótese que $h(X) \mid f(X)$ e, pelo item (ii), segue que $h^*(X) \mid f^*(X) = -f(X)$, o que implica em $h^*(X) \mid f(X)$. Portanto, $h^*(X) \mid (X^n - 1)$.

□

Cabe observar que, pelo item (iii) da Proposição 47 e o Teorema 20, o ideal gerado por $[h^*(X)]$ é gerador de um código cíclico que veremos a seguir.

Teorema 22. *Seja $C = \Psi^{-1}(I)$ um código cíclico, em que $I = I([g(X)])$, com $gr(g(X)) = s$ e $g(X)$ um divisor de $X^n - 1$. Então $C^\perp$ é cíclico e $C^\perp = \Psi^{-1}(J)$, sendo $J = I([h^*(X)])$, em que $h(X) = \dfrac{X^n - 1}{g(X)}$.*

Demonstração. Suponha que $C = \Psi^{-1}(I)$ é um código cíclico, em que $I = I([g(X)])$ com

$$g(X) = g_0 + g_1X + \ldots + g_sX^s \in K[X],$$

cujo grau é igual a s, e seja

$$h(X) = h_0 + h_1X + \ldots + h_{n-s}X^{n-s} \in K[X],$$

como definido em (5.4). Assim, temos que $g_s \neq 0$ e $h_{n-s} \neq 0$. Sejam G a matriz geradora do código C, e H uma matriz cujas entradas são os coeficientes do polinômio $h(X)$, como se segue

$$G = \begin{bmatrix} \Psi^{-1}([g(X)]) \\ \Psi^{-1}([Xg(X)]) \\ \vdots \\ \Psi^{-1}([X^{n-s-1}g(X)]) \end{bmatrix} = \begin{bmatrix} g_0 & g_1 & \ldots & g_s & 0 & \ldots & 0 \\ 0 & g_0 & g_1 & \ldots & g_s & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & g_0 & \ldots & \ldots & g_s \end{bmatrix}$$

e

$$H = \begin{bmatrix} h_{n-s} & h_{n-s-1} & \ldots & h_0 & 0 & \ldots & 0 \\ 0 & h_{n-s} & h_{n-s-1} & \ldots & h_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & h_{n-s} & \ldots & \ldots & h_0 \end{bmatrix}.$$

Temos que as linhas de H são linearmente independentes. Além disso, as linhas de G podem ser escritas como combinação linear da base canônica de K^n. De fato, seja

$$G_i = \begin{bmatrix} 0 & \ldots & g_0 & g_1 & \ldots & g_s & 0 & \ldots & 0 \end{bmatrix}$$

a i-ésima linha da matriz G, em que g_0 e g_s são, respectivamente, a i-ésima e a $(i+s)$-ésima coordenadas da linha G_i, isto é, a linha G_i pode ser associada ao vetor

$$G_i \leftrightarrow (0, \ldots, g_0, g_1, \ldots, g_s, 0, \ldots, 0) = g_0\mathbf{e}_i + g_1\mathbf{e}_{i+1} + \ldots + g_s\mathbf{e}_{i+s}.$$

De modo análogo, a j-ésima linha de H (ou a j-ésima coluna de H^t) é

$$h_{n-s}\mathbf{e}_j + h_{n-s-1}\mathbf{e}_{j+1} + \cdots + h_0\mathbf{e}_{j+n-s},$$

com $1 \leq j \leq s$. Daí segue que, se $i \leq j$, temos

$$\langle G_i, H_j \rangle = g_{j-i}h_{n-s} + g_{j-i+1}h_{n-s-1} + \cdots + g_{n-s}h_{j-i},$$

com $0 \leq j - i \leq s - 1$. Já se $j \leq i$, temos

$$\langle G_i, H_j \rangle = g_0h_{n-s-(i-j)} + g_1h_{n-s-(i-j)-1} + \cdots + g_kh_{n-s-(i-j)-k},$$

com $0 \leq i - j \leq s - 1$ e $0 \leq k \leq s$.

Temos pela definição de produto entre dois polinômios que em ambos os casos acima o resultado do produto interno é, respectivamente, o coeficiente do monômio $X^{n-s+j-i}$ e $X^{n-s+i-j}$ no produto entre os polinômios $g(X) \cdot h(X) (= X^n - 1)$. Como

$$1 \leq n - s + j - i,\ n - s + i - j \leq n - 1,$$

tal coeficiente é igual a zero. Então o produto matricial $G \cdot H^t = \left[0\right]_{k\times(n-k)}$ (pois este produto matricial é obtido calculando o produto interno entre as linhas G pelas linhas de H (ou colunas de H^t)) e, pelo Lema 13, segue que H é a matriz geradora do código $C^\perp$, ou seja,

$$H = \begin{bmatrix} \Psi^{-1}([h^*(X)]) \\ \Psi^{-1}([Xh^*(X)]) \\ \vdots \\ \Psi^{-1}([X^{n-gr(h^*(X))-1}h^*(X)]) \end{bmatrix}.$$

Logo, pelo Teorema 21, temos que $C^\perp = \Psi^{-1}(J)$, em que $J = I([h^*(X)])$. Portanto, a matriz H é a matriz teste de paridade do código C. □

Exemplo 16. Seja $X^7 - 1$ sobre $\mathbb{F}_2$, temos que

$$X^7 - 1 = (1 + X)(1 + X + X^3)(1 + X^2 + X^3).$$

Dessa forma, obtemos que $R_7 = \dfrac{\mathbb{F}_2[X]}{I(X^7 - 1)}$. Considerando $g(X) = 1 + X + X^3$, segue que $I = I([1 + X + X^3])$ e o código cíclico será dado por $C = \Psi^{-1}(I)$, em que Ψ é o isomorfismo entre $\mathbb{F}_2^7$ e R_7. Pelo Corolário 14, temos que a matriz geradora do código cíclico é

$$G = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

e, consequentemente, o código cíclico é
$C = \Psi^{-1}(I([1+X+X^3])) = \{0000000, 1101000, 0110100, 0011010, 0001101, 1011100, 1110010,$
$1100101, 0101110, 0111001, 0010111, 1000110, 1010001, 1111111, 0100011, 1001011\}$.
Pelo Teorema 22, temos $h(X) = \dfrac{X^7 - 1}{1 + X + X^3} = 1 + X + X^2 + X^4$ em $\mathbb{F}_2$. Deste modo,

$$h^*(X) = X^4 + X^3 + X^2 + 1$$

e assim,

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

Logo,

$$GH^t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

em $\mathbb{F}_2$ e, portanto, H é uma matriz geradora de $C^\perp$.

5.1 Decodificação em Códigos Cíclicos

Apresentaremos nesta seção uma maneira de determinar a matriz geradora de um código cíclico C na forma padrão $(R \mid Id)$ e um algoritmo de codificação. Mostraremos também como determinar a síndrome em códigos cíclicos.

Seja

$$\begin{array}{rccl} \mu: & K^s & \to & K[X]_{s-1} \subset K[X] \\ & (a_0, \ldots, a_{s-1}) & \mapsto & \displaystyle\sum_{i=0}^{s-1} a_i X^i. \end{array}$$

Temos que μ é um isomorfismo linear entre os K-espaços vetoriais K^s e $K[X]_{s-1}$, sendo $K[X]_{s-1}$ o espaço vetorial dos polinômios de grau menor ou igual a $s-1$ e o polinômio nulo. Esse isomorfismo será bastante útil para o que faremos a partir de agora.

Teorema 23. *Seja $C \subset K^n$ um código cíclico, tal que $C = \Psi^{-1}(I)$, em que Ψ é a aplicação dada em (5.3) e $I = I([g(X)])$, com $g(X)$ de grau s, um divisor de $X^n - 1$. Considere R uma matriz $(n-s) \times s$ cuja i-ésima linha é dada por*

$$R_i = -\mu^{-1}(r_i(X)), \quad 1 \leq i \leq n-s,$$

em que $r_i(X)$ é o resto da divisão de X^{s-1+i} por $g(X)$. Então $(R \mid Id_{n-s})$ é uma matriz geradora de C, na forma padrão.

Demonstração. Suponha que $C = \Psi^{-1}(I)$ é um código cíclico. Temos pelo algoritmo da divisão que existem únicos $q_i(X), r_i(X) \in K[X]$ para cada i, tais que

$$X^{s-1+i} = g(X)q_i(X) + r_i(X), \text{ com } r_i(X) = 0 \text{ ou } gr(r_i(X)) \leq s-1.$$

Daí,

$$X^{s-1+i} - r_i(X) = g(X)q_i(X) \Leftrightarrow [X^{s-1+i} - r_i(X)] \in I([g(X)]).$$

Tomando a imagem inversa de $[X^{s-1+i} - r_i(X)]$ pelo isomorfismo linear Ψ, temos

$$\begin{aligned}\Psi^{-1}([X^{s-1+i} - r_i(X)]) &= \Psi^{-1}([X^{s-1+i}]) - \Psi^{-1}([r_i(X)]) \\ &= \mathbf{e}_{s-1+i} - \Psi^{-1}([r_i(X)]).\end{aligned}$$

Agora, definindo a função injetiva

$$\begin{array}{rccc}\phi: & K^s & \to & K^n \\ & (a_0, \ldots, a_{s-1}) & \mapsto & (a_0, \ldots, a_{s-1}, 0, \ldots, 0),\end{array}$$

temos que $[r_i(X)] = [a_0 + a_1X + \ldots + a_{s-1}X^{s-1}]$, o que implica em $\Psi^{-1}([r_i(X)]) = (a_0, \ldots, a_{s-1}, 0, \ldots, 0)$, assim,

$$\Psi^{-1}([r_i(X)]) \in Im(\phi).$$

Dessa forma, aplicando a função ϕ em ambos os lados da última igualdade, temos

$$\begin{aligned}\phi^{-1}(\Psi^{-1}([r_i(X)])) &= \phi^{-1}(a_0, \ldots, a_{s-1}, 0, \ldots, 0) \\ &= (a_0, \ldots, a_{s-1}) \\ &= \mu^{-1}([r_i(X)]).\end{aligned}$$

Daí, segue que $\Psi^{-1}([r_i(X)]) = \phi(\mu^{-1}([r_i(X)]))$ e, consequentemente,

$$\Psi^{-1}([X^{s-1+i} - r_i(X)]) = \mathbf{e}_{s-1+i} - \phi(\mu^{-1}([r_i(X)])), \quad \forall\, i = 1, \ldots, n-s.$$

Logo,

$$G' = \begin{bmatrix} \Psi^{-1}([X^{s-1+1} - r_1(X)]) \\ \Psi^{-1}([X^{s-1+2} - r_2(X)]) \\ \vdots \\ \Psi^{-1}([X^{s-1+n-s} - r_{n-s}(X)]) \end{bmatrix} = \begin{bmatrix} -\mu^{-1}(r_1(X)) & 1 & 0 & \ldots & 0 \\ -\mu^{-1}(r_2(X)) & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\mu^{-1}(r_{n-s}(X)) & 0 & 0 & \ldots & 1 \end{bmatrix},$$

cujas linhas são linearmente independentes por causa do bloco identidade e, pelo Corolário 14,

$$G = (R \mid Id_{n-s})$$

é a matriz geradora do código C, sendo $R_i = -\mu^{-1}(r_i(X))$ para $1 \leq i \leq n-s$, as linhas da matriz R. □

Observação 11. Dado o vetor $(a_1, \ldots, a_{n-s}) \in K^{n-s}$ do código fonte, podemos codificá-lo como elemento do código C da seguinte forma:

$$(a_1, \ldots, a_{n-s})G = (b_0, \ldots, b_{s-1}, a_1, \ldots, a_{n-s}),$$

em que

$$\begin{aligned}(b_0,\ldots,b_{s-1}) &= -a_1\mu^{-1}(r_1(x)) - \ldots - a_{n-s}\mu^{-1}(r_{n-s}(x)) \\ &= -\mu^{-1}(a_1(r_1(x)) - \ldots - a_{n-s}(r_{n-s}(x)) \\ &= -\mu^{-1}\left(\sum_{i=1}^{n-s} a_i r^i(X)\right).\end{aligned}$$

Tal observação é de suma importância na construção de códigos cíclicos, visto que caracteriza um dos pilares no processo de transmissão de dados: a codificação.

Teorema 24. *Seja $C = \Psi^{-1}([g(X)]) \subset K^n$, em que Ψ é a aplicação dada em (5.3), um código cíclico, com $gr(g(X)) = s$ e matriz geradora da forma padrão $\tilde{G} = (R \mid Id_{n-s})$, sendo R a matriz formada pelas linhas $R_i = -\mu^{-1}(r_i(X))$, para $1 \leq i \leq n-s$. A matriz teste de paridade de C é dada por $H = (Id_s \mid -R^t)$. Além disso, se $\mathbf{v} = (v_0,\ldots,v_{n-1}) \in K^n$, então a síndrome de $\mathbf{v}$, com relação a matriz H, é dada por*

$$\mu^{-1}(r(X)),$$

em que $r(X)$ é o resto da divisão de $v_0 + v_1X + \ldots + v_{n-1}X^{n-1}$ por $g(X)$.

Demonstração. Suponha que $C = \Psi^{-1}([g(X)])$ é um código cíclico. Temos que a síndrome do vetor $\mathbf{v}$ é dada por $H\mathbf{v}^t = (Id_s \mid -R^t)\mathbf{v}^t$. Além disso, pelo Teorema 23, a matriz R é a matriz

$$R = \begin{bmatrix} -\mu^{-1}(r_1(X)) \\ -\mu^{-1}(r_2(X)) \\ \vdots \\ -\mu^{-1}(r_{n-s}(X)) \end{bmatrix} = \begin{bmatrix} -r_{11} & -r_{12} & \ldots & -r_{1s} \\ -r_{21} & -r_{22} & \ldots & -r_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ -r_{(n-s)1} & -r_{(n-s)2} & \ldots & -r_{(n-s)s} \end{bmatrix}.$$

Assim,

$$H = (Id_s \mid -R^t) = \begin{bmatrix} 1 & 0 & \ldots & 0 & r_{11} & r_{21} & \ldots & r_{(n-s)1} \\ 0 & 1 & \ldots & 0 & r_{12} & r_{22} & \ldots & r_{(n-s)2} \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & r_{1s} & r_{2s} & \ldots & r_{(n-s)s} \end{bmatrix}.$$

Logo,

$$
\begin{aligned}
H\mathbf{v}^t &= \begin{bmatrix} 1 & 0 & \dots & 0 & r_{11} & r_{21} & \dots & r_{(n-s)1} \\ 0 & 1 & \dots & 0 & r_{12} & r_{22} & \dots & r_{(n-s)2} \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & r_{1s} & r_{2s} & \dots & r_{(n-s)s} \end{bmatrix} \cdot \begin{bmatrix} v_0 & v_1 & \dots & v_{s-1} & v_s & v_{s+1} & \dots & v_{n-1} \end{bmatrix}^t \\
&= \begin{bmatrix} v_0 + v_s r_{11} + v_{s+1} r_{21} + \cdots + v_{n-1} r_{(n-s)1} \\ v_1 + v_s r_{12} + v_{s+1} r_{22} + \cdots + v_{n-1} r_{(n-s)2} \\ \vdots \\ v_{s-1} + v_s r_{1s} + v_{s+1} r_{2s} + \cdots + v_{n-1} r_{(n-s)s} \end{bmatrix} \\
&= v_0 \mathbf{e}_1^t + v_1 \mathbf{e}_2^t + \cdots + v_{s-1} \mathbf{e}_s^t + v_s (\mu^{-1}(r_1(X)))^t + \cdots + v_{n-1} (\mu^{-1}(r_{n-s}(X)))^t \\
&= (v_0 \mathbf{e}_1 + v_1 \mathbf{e}_2 + \cdots + v_{s-1} \mathbf{e}_s + v_s \mu^{-1}(r_1(X)) + \cdots + v_{n-1} \mu^{-1}(r_{n-s}(X)))^t \\
&= (v_0 \mu^{-1}(1) + v_1 \mu^{-1}(X) + \cdots + v_{s-1} \mu^{-1}(X^{s-1}) + v_s \mu^{-1}(r_1(X)) + \cdots + v_{n-1} \mu^{-1}(r_{n-s}(X)))^t \\
&= \mu^{-1}(v_0 + v_1 X + \cdots + v_{s-1} X^{s-1} + v_s r_1(X) + \cdots + v_{n-1} r_{n-s}(X))^t.
\end{aligned}
$$

Como

$$
\begin{aligned}
1 &= 0 \cdot g(X) + 1 \\
&\vdots \\
X^{s-1} &= 0 \cdot g(X) + X^{s-1} \\
X^s &= q_s(X) g(X) + r_1(X) \\
&\vdots \\
X^{n-1} &= q_{n-1}(X) g(X) + r_{n-s}(X),
\end{aligned}
$$

segue que

$$
\begin{aligned}
v_0 + \ldots + v_{s-1} X^{s-1} + v_s X^s + \ldots + v_{n-1} X^{n-1} = g(X)[0 + \ldots + 0 + v_s q_s(X) + \ldots + v_{n-1} q_{n-1}(X)] + \\
+ v_0 + \ldots + v_{s-1} X^{s-1} + v_s r_1(X) + \ldots + v_{n-1} r_{n-s}(X),
\end{aligned}
$$

em que $r(X) = v_0 + \ldots + v_{s-1} X^{s-1} + v_s r_1(X) + \ldots + v_{n-1} r_{n-s}(X)$ é tal que $r(X) = 0$ ou $gr(r(X)) < gr(g(X))$. Portanto,

$$
r(X) = v_0 + v_1 X + \cdots + v_{s-1} X^{s-1} + v_s r_1(X) + \cdots + v_{n-1} r_{n-s}(X)
$$

é o resto da divisão de $v_0 + v_1 X + \cdots + v_{n-1} X^{n-1}$ por $g(X)$. □

Exemplo 17. *Seja* $X^7 - 1$ *sobre* $\mathbb{F}_2$. *Temos que*

$$
X^7 - 1 = (1 + X)(1 + X + X^3)(1 + X^2 + X^3).
$$

Considerando o código $C = \Psi^{-1}(I) \subset \mathbb{F}_2^7$, em que $I = I([1 + X + X^3])$, temos pelo Corolário 14 que

$$dim_{\mathbb{F}_2^7} C = 7 - 3 = 4.$$

Vamos então determinar a matriz geradora de C na forma padrão. Dividindo X^{2+i} (com $i = 1, \ldots, 4$) por $g(X) = 1 + X + X^3$, temos

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

Daí,

$$\begin{aligned} [X^3 - (X + 1)] &= [X^3 + X + 1] \in I([g(X)]); \\ [X^4 - (X^2 + X)] &= [(X^3 + X + 1)][X] \in I([g(X)]); \\ [X^5 - (X^2 + X + 1)] &= [(X^3 + X + 1)][(X^2 + 1)] \in I([g(X)]); \\ [X^6 - (X^2 + 1)] &= [(X^3 + X + 1)][(X^3 + X + 1)] \in I([g(X)]). \end{aligned}$$

Aplicando a função Ψ^{-1} em $[X^{2+i} - r_i(X)]$, $i = 1, \ldots, 4$, segue que

$$\begin{aligned} \Psi^{-1}([X^3 - (X + 1)]) &= (1, 1, 0, 1, 0, 0, 0) \in C; \\ \Psi^{-1}([X^4 - (X^2 + X)]) &= (0, 1, 1, 0, 1, 0, 0) \in C; \\ \Psi^{-1}([X^5 - (X^2 + X + 1)]) &= (1, 1, 1, 0, 0, 1, 0) \in C; \\ \Psi^{-1}([X^6 - (X^2 + 1)]) &= (1, 0, 1, 0, 0, 0, 1) \in C. \end{aligned}$$

Logo,

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

Agora considere o vetor $(a_1, a_2, a_3, a_4) \in \mathbb{F}_2^4$ do código da fonte. Aplicando o algoritmo da codificação, temos

$$\begin{aligned} (a_1, a_2, a_3, a_4) \cdot G' &= (a_1, a_2, a_3, a_4) \cdot \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \\ &= (b_0, b_1, b_2, a_1, a_2, a_3, a_4), \end{aligned}$$

sendo b_0, b_1 e b_2 os coeficientes do polinômio

$$a_1(X+1) + a_2(X^2+X) + a_3(X^2+X+1) + a_4(X^2+1) =$$
$$= a_1 + a_3 + a_4 + (a_1 + a_2 + a_3)X + (a_2 + a_3 + a_4)X^2.$$

Portanto,

$$(a_1 + a_3 + a_4, a_1 + a_2 + a_3, a_2 + a_3 + a_4, a_1, a_2, a_3, a_4)$$

é uma codificação de (a_1, a_2, a_3, a_4). Além disso, pelo Teorema 24, a matriz teste de paridade de C é

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

Com isso, dado o vetor $\mathbf{u} = (1, 0, 1, 1, 0, 0, 1) \in \mathbb{F}_2^7$, temos que $r(X) = 1 + X$ é o resto da divisão de $1 + X^2 + X^3 + X^6$ por $g(X) = 1 + X + X^3$. Segue do Teorema 24 que a síndrome de $\mathbf{u}$ em relação a H é $\mu^{-1}(r(X)) = 110$.

Note que a síndrome do vetor $\mathbf{e} = (0, 0, 0, 1, 0, 0, 0)$ é $\mu^{-1}(r(X)) = 110$, que coincide com a síndrome de $\mathbf{u}$. Portanto $\mathbf{c} = \mathbf{u} - \mathbf{e} = (1, 0, 1, 0, 0, 0, 1)$, ou seja, $\mathbf{c} = 1010001$ é a palavra transmitida.

6 Códigos BCH

Neste capítulo discutiremos a subclasse de códigos cíclicos chamada de *códigos BCH*. Focaremos aqui em estabelecer uma maneira de corrigir eventuais erros de transmissão de dados via códigos BCH, sobretudo, em encontrar sua distância mínima, embora na maioria dos casos conseguimos apenas estimá-la, o que já é um feito importante, em vista de que para os códigos cíclicos em geral não existe uma maneira mais eficiente (com uma quantidade reduzida de cálculos) de determinar a distância mínima. (HEFEZ; VILLELA, 2008).

6.1 Códigos Cíclicos por Anulamento

Proposição 48. *Sejam $K = \mathbb{F}_q$ um corpo finito com q elementos e $C \subset K^n$ um código cíclico, com n e q primos entre si. E sejam F o corpo de decomposição do polinômio $X^n - 1$ de $K[X]$ e $\alpha_1, \ldots, \alpha_r$ todas as raízes distintas de $g(X)$ em F, em que $g(X) \in K[X]$ é um divisor de $X^n - 1$. Então*

$$\Psi(C) = I([g(X)]) = \{[f(X)] \in R_n;\ f(\alpha_1) = \ldots = f(\alpha_r) = 0\},$$

em que Ψ é a função definida em (5.3).

Demonstração. Sendo $[f(X)] \in I([g(X)])$, então existe $q(X) \in K[X]$ tal que

$$\begin{aligned}[f(X)] = [q(X)][g(X)] = [q(X)][0] = [0] &\Leftrightarrow f(X) \equiv 0 \ (mod\, I([g(X)])) \\ &\Leftrightarrow f(X) = h(X)g(X),\end{aligned}$$

com $h(X) \in K[X]$. Como α_i é raiz de $g(X)$ para todo $i = 1, \ldots, r$, então

$$f(\alpha_i) = g(\alpha_i) \cdot h(\alpha_i) = 0 \Rightarrow f(\alpha_i) = 0, \quad \forall\, i = 1, \ldots, r.$$

Portanto,

$$I([g(X)]) \subset \{[f(X)] \in R_n;\ f(\alpha_1) = \ldots = f(\alpha_r) = 0\}.$$

Por outro lado, sendo $[m(X)] \in \{[f(X)] \in R_n;\ f(\alpha_1) = \ldots = f(\alpha_r) = 0\}$, temos que $m(\alpha_i) = 0$ para todo $i = 1 \ldots, r$, consequentemente, $X - \alpha_i$ divide o polinômio $m(X)$, para todo $i = 1, \ldots, r$. Daí,

$$m(X) = (X - \alpha_1)(X - \alpha_2) \ldots (X - \alpha_r) \cdot t(X), \tag{6.1}$$

com $t(X) \in K[X]$. Como por hipótese, $\alpha_1, \ldots, \alpha_r$ são as raízes de $g(X)$ em F, duas a duas distintas, então

$$(X - \alpha_1)(X - \alpha_2) \ldots (X - \alpha_r) = g(X), \tag{6.2}$$

substituindo (6.1) em (6.2) obtemos $m(X) = t(X)g(X)$. Assim,

$$[m(X)] = [t(X)g(x)] \Leftrightarrow [m(X)] \in I([g(X)]).$$

Logo,

$$\{[f(X)] \in R_n;\ f(\alpha_1) = \ldots = f(\alpha_r) = 0\} \subset I([g(X)]).$$

Portanto, $\Psi(C) = I([g(X)])$. □

A partir desse resultado, conseguimos determinar a matriz teste de paridade de um código cíclico, como veremos a seguir. Para isso, seja $f(X) = \sum_{j=0}^{n-1} a_j X^j \in K[X]$. Temos, pela Proposição 48, que

$$[f(X)] \in I([g(X)]) \Leftrightarrow f(\alpha_i) = \sum_{j=0}^{n-1} a_j \alpha_i^j = 0, \quad \forall\, i = 1, \ldots, r. \tag{6.3}$$

E assim, sendo $\Psi(a_0, \ldots, a_{n-1}) = [a_0 + \cdots + a_{n-1}X^{n-1}] = [f(X)]$, temos que

$$(a_0, \ldots, a_{n-1}) \in C \Leftrightarrow [f(X)] \in I([g(X)]).$$

Logo, o código cíclico C definido por $g(X)$, é determinado pelo conjunto dos elementos $(a_0, \ldots, a_{n-1}) \in K^n$ que satisfazem o sistema de equações (6.3), o qual pode ser visto como o produto matricial

$$\tilde{H}[a_0, \ldots, a_{n-1}]^t = \left[0\right]_{r\times 1}, \tag{6.4}$$

em que

$$\tilde{H} = \begin{bmatrix} 1 & \alpha_1^1 & \ldots & \alpha_1^{n-1} \\ 1 & \alpha_2^1 & \ldots & \alpha_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_r^1 & \ldots & \alpha_r^{n-1} \end{bmatrix} = \begin{bmatrix} \alpha_1^0 & \alpha_1^1 & \ldots & \alpha_1^{n-1} \\ \alpha_2^0 & \alpha_2^1 & \ldots & \alpha_2^{n-1} \\ \vdots & \vdots & & \vdots \\ \alpha_r^0 & \alpha_r^1 & \ldots & \alpha_r^{n-1} \end{bmatrix} \tag{6.5}$$

é uma matriz com entradas em F.

Note que, como $\mathbf{a} = (a_0, \ldots, a_{n-1}) \in C$ é um elemento qualquer do código, então vale para qualquer $\mathbf{c} \in C$, que

$$\mathbf{c} \in C \Leftrightarrow \tilde{H}\mathbf{c}^t = \left[0\right]_{r\times 1}.$$

Outra observação importante é a de que, embora os elementos do código sejam determinados pelo produto $\tilde{H}[a_0, \ldots, a_{n-1}]^t = \left[0\right]_{r\times 1}$, a matriz $\tilde{H}$ não é a matriz teste de paridade de C, pois as suas entradas estão no corpo F, que por sua vez, é uma extensão do corpo K, sobre o qual o código C está definido. No entanto, podemos enxergar F como um espaço vetorial de

dimensão finita sobre K, digamos que F tem dimensão d sobre K e seja $\mathcal{B} = \{\rho_0, \rho_1, \ldots, \rho_{d-1}\}$ uma base de F sobre K. Temos então que $\alpha_i^j \in F$ se, e somente se,

$$\alpha_i^j = \lambda_{i0}^j \rho_0 + \lambda_{i1}^j \rho_1 + \ldots + \lambda_{i(d-1)}^j \rho_{d-1},$$

tal que $\lambda_{i\ell}^j \in K$ com $0 \leq \ell \leq d-1$. Deste modo, a matriz dos coeficientes de α_i^j na base $\mathcal{B}$ é dada por

$$[\alpha_i^j]_{\mathcal{B}} = \begin{bmatrix} \lambda_{i0}^j \\ \lambda_{i1}^j \\ \vdots \\ \lambda_{i(d-1)}^j \end{bmatrix} \in K^d \cong (\mathbb{F}_q)^d.$$

Logo, de acordo com (6.3), temos que

$$\begin{aligned} 0 &= f(\alpha_i) = \sum_{j=0}^{n-1} a_j \alpha_i^j = a_0 \alpha_i^0 + \cdots + a_{n-1} \alpha_i^{n-1} \\ &= a_0(\lambda_{i0}^0 \rho_0 + \cdots + \lambda_{i(d-1)}^0 \rho_{d-1}) + \cdots + a_{n-1}(\lambda_{i0}^{n-1} \rho_0 + \cdots + \lambda_{i(d-1)}^{n-1} \rho_{d-1}) \\ &= a_0 \lambda_{i0}^0 \rho_0 + \cdots + a_0 \lambda_{i(d-1)}^0 \rho_{d-1} + \cdots + a_{n-1} \lambda_{i0}^{n-1} \rho_0 + \cdots + a_{n-1} \lambda_{i(d-1)}^{n-1} \rho_{d-1} \\ &= \rho_0(a_0 \lambda_{i0}^0 + \ldots + a_{n-1} \lambda_{i0}^{n-1}) + \ldots + \rho_{d-1}(\lambda_{i(d-1)}^0 + \ldots + a_{n-1} \lambda_{i(d-1)}^{n-1}). \end{aligned}$$

Escrevendo a matriz dos coeficientes do sistema de equações acima, temos

$$\begin{aligned} \left[0\right]_{(d-1)\times 1} &= \left[\sum_{j=0}^{n-1} \mathbf{a}_j \alpha_i^j\right]_{\mathcal{B}} = \begin{bmatrix} a_0 \lambda_{i0}^0 + \ldots + a_{n-1} \lambda_{i0}^{n-1} \\ \vdots \\ a_0 \lambda_{i(d-1)}^0 + \ldots + a_{n-1} \lambda_{i(d-1)}^{n-1} \end{bmatrix}_{\mathcal{B}} \\ &= a_0 \begin{bmatrix} \lambda_{i0}^0 \\ \vdots \\ \lambda_{i(d-1)}^0 \end{bmatrix}_{\mathcal{B}} + \ldots + a_{n-1} \begin{bmatrix} \lambda_{i0}^{n-1} \\ \vdots \\ \lambda_{i(d-1)}^{n-1} \end{bmatrix}_{\mathcal{B}} = \sum_{j=0}^{n-1} \mathbf{a}_j \left[\alpha_i^j\right]_{\mathcal{B}}. \end{aligned}$$

Agora defina a matriz H', de ordem $dr \times n$, como

$$H' = \begin{bmatrix} [\alpha_1^0]_{\mathcal{B}} & [\alpha_1^1]_{\mathcal{B}} & \ldots & [\alpha_1^{n-1}]_{\mathcal{B}} \\ [\alpha_2^0]_{\mathcal{B}} & [\alpha_2^1]_{\mathcal{B}} & \ldots & [\alpha_2^{n-1}]_{\mathcal{B}} \\ \vdots & \vdots & & \vdots \\ [\alpha_r^0]_{\mathcal{B}} & [\alpha_r^1]_{\mathcal{B}} & \ldots & [\alpha_r^{n-1}]_{\mathcal{B}} \end{bmatrix}.$$

Temos que as entradas de H' são elementos de K. Além disso,

$$\mathbf{c} \in C \Leftrightarrow H'\mathbf{c}^t = \left[0\right]_{dr\times 1}. \tag{6.6}$$

É importante observar que não conseguimos garantir que todas as linhas da matriz H' são linearmente independentes. Com isso, escolhemos um conjunto maximal de linhas linearmente

independentes de H' e obtemos uma matriz H formada pelas linhas linearmente independentes de H'. Seja s o número máximo de colunas linearmente independentes de $\tilde{H}$. Escrevendo essa matriz como

$$\begin{bmatrix} \alpha^0 & \alpha^1 & \dots & \alpha^{n-1} \end{bmatrix}_{r\times n},$$

em que

$$\alpha^j = \begin{bmatrix} \alpha_1^j & \alpha_2^j & \dots & \alpha_r^j \end{bmatrix}^t_{1\times r},$$

com $0 \leq j \leq n-1$. Temos que a matriz H' é dada por

$$H' = \begin{bmatrix} [\alpha^0]_{\mathcal{B}} & [\alpha^1]_{\mathcal{B}} & \dots & [\alpha^{n-1}]_{\mathcal{B}} \end{bmatrix}_{dr\times n}$$

e possui no máximo s colunas linearmente independentes, pois os vetores colunas de H' são os mesmos de $\tilde{H}$, considerando as devidas representações dos seus elementos sobre o corpo K. Logo, pelo Teorema 1, a matriz H' possui no máximo s linhas linearmente independentes. Além disso, pela construção da matriz H e novamente pelo Teorema 1, segue que H possui no máximo s colunas linearmente independentes. Portanto, o número máximo de colunas linearmente independentes de H, H' e $\tilde{H}$ coincidem. Dessa forma, segue do Teorema 18 que a distância mínima do código C pode ser determinada encontrando o maior número d tal que quaisquer $d-1$ colunas de $\tilde{H}$ sejam linearmente independentes.

Observação 12. É conveniente observarmos que podemos determinar um código cíclico através de um conjunto de elementos de F, como se segue

Sejam $\beta_1, \dots, \beta_r \in F$, raízes de $X^n - 1$, o conjunto

$$\Psi^{-1}(\{f(X) \in K[X];\ f(\beta_1) = \dots = f(\beta_r) = 0\}),$$

é um código cíclico definido por $g(X) = mmc(m_{\beta_1}(X), \dots, m_{\beta_r}(X))$, sendo $m_{\beta_i}(X)$ o polinômio minimal de β_i sobre K, para todo $i = 1, \dots, r$. Note que o polinômio $g(X)$ é determinado pelo mínimo múltiplo comum dos polinômios minimais e não pelo produto desses minimais, pois poderíamos ter, para dois elementos distintos de F, o mesmo polinômio minimal e, consequentemente, não teríamos $g(X)$ um divisor de $X^n - 1$.

6.2 Códigos BCH

Descoberto pelos matemáticos indianos Bose e Ray-Chaudhuri em 1960, e de maneira independente pelo matemático francês Alexis Hocquenghem em 1959, os códigos BCH formam uma importante classe de códigos cíclicos, conhecidos por sua capacidade de corrigir múltiplos erros de transmissão e possuir algoritmos de decodificação bastante eficientes, (REED; CHEN, 2001, página 189, capítulo 5). Apresentaremos aqui o principal resultado desse capítulo, que nos dá informações sobre a distância mínima e a dimensão de um código.

Teorema 25. (Bose-Chaudhuri-Hocquenghem) *Sejam $K = \mathbb{F}_q$, n um inteiro maior do que 1 e primo com q, e F um corpo onde $X^n - 1$ se decompõe em fatores lineares, com $\gamma \in F$ uma raiz n-ésima primitiva da unidade. Dado C um código cíclico com polinômio gerador*

$$g(X) = mmc(m_{\gamma^a}(X), \ldots, m_{\gamma^{a+\delta-2}}(X)),$$

com $a \geq 0$ e $0 \leq \delta \leq n$. Então, a distância mínima de C é pelo menos δ e a sua dimensão é pelo menos $n - m(\delta - 1)$, em que $m = dim_K F$.

Demonstração. Suponhamos que $\{\gamma^a, \gamma^{a+1}, \ldots, \gamma^{a+\delta-2}\} \cup \{\beta_1, \ldots, \beta_s\}$ sejam as raízes distintas de $g(X)$. Realizando uma construção análoga à feita para obter a matriz dada em (6.5), temos

$$\tilde{H} = \begin{bmatrix} \gamma^0 & \gamma^a & (\gamma^a)^2 & \ldots & (\gamma^a)^{n-1} \\ \gamma^0 & \gamma^{a+1} & \left(\gamma^{a+1}\right)^2 & \ldots & \left(\gamma^{a+1}\right)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma^0 & \gamma^{a+\delta-2} & \left(\gamma^{a+\delta-2}\right)^2 & \ldots & \left(\gamma^{a+\delta-2}\right)^{n-1} \\ \beta_1^0 & \beta_1^1 & \beta_1^2 & \ldots & \beta_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \beta_s^0 & \beta_s^1 & \beta_s^2 & \ldots & \beta_s^{n-1} \end{bmatrix} = \begin{bmatrix} 1 & \gamma^a & (\gamma^a)^2 & \ldots & (\gamma^a)^{n-1} \\ 1 & \gamma^{a+1} & \left(\gamma^{a+1}\right)^2 & \ldots & \left(\gamma^{a+1}\right)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \gamma^{a+\delta-2} & \left(\gamma^{a+\delta-2}\right)^2 & \ldots & \left(\gamma^{a+\delta-2}\right)^{n-1} \\ 1 & \beta_1^1 & \beta_1^2 & \ldots & \beta_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \beta_s^1 & \beta_s^2 & \ldots & \beta_s^{n-1} \end{bmatrix}.$$

Note que a matriz $\tilde{H}$ é uma matriz de Vandermonde. Como queremos uma cota para a distância mínima, devemos encontrar um δ, de modo que quaisquer $\delta - 1$ colunas de $\tilde{H}$ sejam linearmente independentes, para então aplicarmos a Proposição 40. Motivados por isso, escolhemos $\delta - 1$ colunas arbitrárias de $\tilde{H}$ e em cada uma delas selecionamos as primeiras $\delta - 1$ linhas, a fim de obtermos uma matriz de ordem $(\delta - 1) \times (\delta - 1)$, como se segue

$$B = \begin{bmatrix} (\gamma^a)^{i_1} & (\gamma^a)^{i_2} & \ldots & (\gamma^a)^{i_{\delta-1}} \\ \left(\gamma^{a+1}\right)^{i_1} & \left(\gamma^{a+1}\right)^{i_2} & \ldots & \left(\gamma^{a+1}\right)^{i_{\delta-1}} \\ \vdots & \vdots & & \vdots \\ \left(\gamma^{a+\delta-2}\right)^{i_1} & \left(\gamma^{a+\delta-2}\right)^{i_2} & \ldots & \left(\gamma^{a+\delta-2}\right)^{i_{\delta-1}} \end{bmatrix}. \tag{6.7}$$

Assim,

$$\begin{aligned} detB &= (\gamma^a)^{(i_1+\ldots+i_{\delta-1})} \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \gamma^{i_1} & \gamma^{i_2} & \ldots & \gamma^{i_{\delta-1}} \\ \vdots & \vdots & & \vdots \\ \left(\gamma^{\delta-2}\right)^{i_1} & \left(\gamma^{\delta-2}\right)^{i_2} & \ldots & \left(\gamma^{\delta-2}\right)^{i_{\delta-1}} \end{bmatrix} \\ &= (\gamma^a)^{(i_1+\ldots+i_{\delta-1})} \prod_{1 \leq k < \ell \leq \delta-1} (\gamma^{i_\ell} - \gamma^{i_k}) \neq 0, \end{aligned}$$

pois $\gamma^{i_\ell} \neq \gamma^{i_k}$ quando $\ell \neq k$, uma vez que γ é uma raiz n-ésima *primitiva* da unidade. Logo quaisquer $\delta - 1$ colunas de $\tilde{H}$ são linearmente independentes sobre K. Portanto, pela Proposição 40, temos $d = \omega(C) \geq \delta$. Como o código C é, pela Proposição 48, o conjunto $\{[f(X)] \in R_n;\ f(\alpha_1) = \ldots = f(\alpha_r) = 0\}$, então

$$f(\gamma^a) = f(\gamma^{a+1}) = \ldots = f(\gamma^{a+\delta-2}) = 0 \Leftrightarrow m_{\gamma^{a+i}}(X) \mid f(X), \quad \forall\, i = 1, \ldots, \delta - 2. \tag{6.8}$$

De fato, temos pelo Teorema 4 que existem únicos $h_i(X), r_i(X) \in K[X]$ para cada i, tais que

$$f(X) = m_{\gamma^{a+i}}(X)h_i(X) + r_i(X) \quad \text{com } gr(r_i(X)) < m_{\gamma^{a+i}}(X) \text{ ou } r_i(X) = 0, \quad \forall\, i = 1, \ldots, \delta - 2.$$

Daí, se $r_i(X) = 0$, então $m_{\gamma^{a+i}} \mid f(X)$. Se $r_i(X) \neq 0$, então como $f(\gamma^{a+i}) = 0$ implica que

$$0 = f(\gamma^{a+i}) = m_{\gamma^{a+i}}(\gamma^{a+i})h_i(\gamma^{a+i}) + r_i(\gamma^{a+i}) = r_i(\gamma^{a+i}),$$

segue que $r(\gamma^{a+i}) = 0$, o que contradiz a minimalidade do grau de $m_{\gamma^{a+i}}(X)$ para todo $i = 1, \ldots, \delta - 2$. Logo, $r_i(X) = 0$ e, portanto, $m_{\gamma^{a+i}} \mid f(X)$.
Reciprocamente, se $m_{\gamma^{a+i}}(X) \mid f(X)$, então existem $h_i(X) \in K[X]$, para cada i, tais que

$$f(X) = h_i(X)m_{\gamma^{a+i}}(X), \quad \forall\, i = 1, \ldots, \delta - 2,$$

logo

$$f(\gamma^{a+i}) = h_i(\gamma^{a+i})m_{\gamma^{a+i}}(\gamma^{a+i}) = 0, \quad \forall\, i = 1, \ldots, \delta - 2.$$

Como consequência da equivalência (6.8), temos que $f(\beta_i) = 0$ para todo $i = 1, \ldots, s$. Além disso, pela caracterização de códigos cíclicos dada pela Proposição 48, segue que as palavras $\mathbf{c}$ do código C são tais que $\hat{H}\mathbf{c}^t = [0]_{(\delta-1)\times 1}$, em que

$$\hat{H} = \begin{bmatrix} \gamma^0 & \gamma^a & (\gamma^a)^2 & \ldots & (\gamma^a)^{n-1} \\ \gamma^0 & \gamma^{a+1} & \left(\gamma^{a+1}\right)^2 & \ldots & \left(\gamma^{a+1}\right)^{(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma^0 & \gamma^{a+\delta-2} & \left(\gamma^{a+\delta-2}\right)^2 & \ldots & \left(\gamma^{a+\delta-2}\right)^{n-1} \end{bmatrix}.$$

Considerando $\mathcal{B}$ uma base do corpo F sobre K como espaço vetorial m-dimensional, podemos escrever $\hat{H}$ como em (6.6), daí obtemos

$$H' = \begin{bmatrix} [\gamma^0]_{\mathcal{B}} & [\gamma^a]_{\mathcal{B}} & [(\gamma^a)^2]_{\mathcal{B}} & \ldots & [(\gamma^a)^{n-1}]_{\mathcal{B}} \\ [\gamma^0]_{\mathcal{B}} & [\gamma^{a+1}]_{\mathcal{B}} & \left[\left(\gamma^{a+1}\right)^2\right]_{\mathcal{B}} & \ldots & \left[\left(\gamma^{a+1}\right)^{n-1}\right]_{\mathcal{B}} \\ \vdots & \vdots & \vdots & & \vdots \\ [\gamma^0]_{\mathcal{B}} & [\gamma^{a+\delta-2}]_{\mathcal{B}} & \left[\left(\gamma^{a+\delta-2}\right)^2\right]_{\mathcal{B}} & \ldots & \left[\left(\gamma^{a+\delta-2}\right)^{n-1}\right]_{\mathcal{B}} \end{bmatrix}, \tag{6.9}$$

uma matriz de ordem $m(\delta-1)\times n$, com coeficientes em K, tal que $H'\mathbf{c}^t = \left[0\right]_{m(\delta-1)\times 1}$. Escolhendo um conjunto maximal de linhas linearmente independentes de H', digamos l, obtemos a matriz H, teste de paridade de C, ou seja, a matriz geradora de $C^\perp$, logo $dim_K C^\perp = l$. Por outro lado, $l \leq m(\delta-1)$ e portanto $dim_K C^\perp \leq m(\delta-1)$.

Observe que

$$\begin{array}{rccc} T: & K^n = C \oplus C^\perp & \rightarrow & K^l \\ & \mathbf{v} \oplus \mathbf{u} & \mapsto & \mathbf{u} \end{array},$$

em que $C = Ker(T)$ e $C^\perp = Im(T)$. Logo

$$dim_K K^n = dim_K C + dim_K C^\perp \leq dim_K C + m(\delta-1).$$

Portanto,

$$dim_K C \geq dim_K K^n - m(\delta-1) = n - m(\delta-1).$$

□

O número δ do Teorema 25 é chamado de *peso estimado* do código BCH por ser uma cota para a distância mínima de C.

Os resultados apresentados acima nos permitem dar uma nova caracterização para os códigos BCH. Para tanto, fixados um corpo K e uma extensão F de K, com uma raiz n-ésima primitiva da unidade $\gamma \in F$ e considerando $a = 1$, definimos

$$C_K(n,\delta) = \left\{(b_0,\ldots,b_{n-1}) \in K^n;\ H'[b_0,\ldots,b_{n-1}]^t = \left[0\right]_{m(\delta-1)\times 1}\right\}, \tag{6.10}$$

em que a matriz H', obtida como em (6.9), é a matriz associada ao código BCH, gerada pelo polinômio gerador

$$g(X) = mmc(m_\gamma(X),\ldots,m_{\gamma^{\delta-1}}(X)).$$

Note que

$$\delta < \delta' \Rightarrow C_K(n,\delta') \subset C_K(n,\delta), \tag{6.11}$$

visto que, diminuir o número δ, implica em diminuir o número de equações a serem satisfeitas do sistema gerado em (6.10) e, consequentemente, aumentar o número de soluções possíveis para este sistema.

Teorema 26. *Seja $C_K(n,\delta)$ um código BCH, definido pelo polinômio $g(X) = mmc(m_\gamma(X),\ldots,m_{\gamma^{\delta-1}}(X))$. Então,*

$$(b_0,\ldots,b_{n-1}) \in C_K(n,\delta) \Leftrightarrow \sum_{j=0}^{n-1} b_j \gamma^{j(\delta-1)} \frac{X^{\delta-1} - \gamma^{-j(\delta-1)}}{X - \gamma^{-j}} = 0. \tag{6.12}$$

Demonstração. Primeiramente, mostremos que

$$(b_0, \ldots, b_{n-1}) \in C_K(n, \delta) \Leftrightarrow \sum_{i=1}^{\delta-1} \left(\sum_{j=0}^{n-1} b_j \gamma^{ij} \right) X^i = 0.$$

Suponha que $(b_0, \ldots, b_{n-1}) \in C_K(n, \delta)$, logo

$$\sum_{j=0}^{n-1} b_j \gamma^{ij} = 0 \Rightarrow \sum_{i=1}^{\delta-1} \left(\sum_{j=0}^{n-1} b_j \gamma^{ij} \right) X^i = \sum_{i=1}^{\delta-1} 0 \cdot X^i = 0.$$

Reciprocamente, suponha que $\sum_{i=1}^{\delta-1} \left(\sum_{j=0}^{n-1} b_j \gamma^{ij} \right) X^i = 0$. Como o primeiro somatório refere-se ao polinômio nulo na incógnita X, temos que cada um dos seus coeficientes é igual a zero, ou seja,

$$\sum_{j=0}^{n-1} b_j \gamma^{ij} = 0 \Leftrightarrow H'[b_0, \ldots, b_{n-1}]^t = \left[0\right]_{m(\delta-1)\times 1}, \quad \forall\, i = 1, \ldots, \delta - 1$$

$$\Leftrightarrow (b_0, \ldots, b_{n-1}) \in C_K(n, \delta).$$

Logo,

$$(b_0, \ldots, b_{n-1}) \in C_K(n, \delta) \Leftrightarrow \sum_{i=1}^{\delta-1} \left(\sum_{j=0}^{n-1} b_j \gamma^{ij} \right) X^i = 0. \tag{6.13}$$

Além disso, como

$$\sum_{i=1}^{\delta-1} \left(\sum_{j=0}^{n-1} b_j \gamma^{ij} \right) X^i = \sum_{j=0}^{n-1} b_j \gamma^{j(\delta-1)} \sum_{i=0}^{\delta-1} \gamma^{-j(\delta-2-i)} X^i = \sum_{j=0}^{n-1} b_j \gamma^{j(\delta-1)} \frac{X^{\delta-1} - \gamma^{-j(\delta-1)}}{X - \gamma^{-j}},$$

segue de (6.13) que

$$(b_0, \ldots, b_{n-1}) \in C_K(n, \delta) \Leftrightarrow \sum_{j=0}^{n-1} b_j \gamma^{j(\delta-1)} \frac{X^{\delta-1} - \gamma^{-j(\delta-1)}}{X - \gamma^{-j}} = 0.$$

□

Vimos anteriormente que conseguimos estimar o peso de um código BCH através do Teorema 25. Veremos agora um caso particular de códigos BCH, que sob certas condições, conseguimos dar uma estimativa melhor para a sua distância mínima.

Dado C, um código BCH de K^n sobre K, definido pelo polinômio $g(X) = mmc(m_\gamma(X), \ldots, m_{\gamma^{\delta-1}}(X))$. Temos que se $n = q^m - 1$, para algum $m \in \mathbb{N}$, então o código C será chamado de *código primitivo*.

Observe que, quando se trata de códigos BCH primitivos, desde que $p = \text{car}(K)$ e $n = q^m - 1$, sendo q uma potência da característica de K, digamos $q = p^r$, para algum natural r, temos que

$$mdc(q, q^m - 1) = 1 \Leftrightarrow mdc(p, n) = 1,$$

ou seja, n e p são primos entre si.

Definição 61. *Sejam n e $q \in \mathbb{N}$ primos entre si e seja F uma extensão de K com raíz n-ésima primitiva da unidade γ. Considerando $\mathbf{a} = \sum_{i=1}^{n} a_i \mathbf{e}_i \in K^n$ com $\omega(\mathbf{a}) = w$, ou seja, $\mathbf{a}$ possui w coordenadas não nulas, a saber $a_{i_1}, \ldots, a_{i_w}$. Definimos o polinômio localizador de $\mathbf{a}$ por*

$$l_{\mathbf{a}}(X) = \prod_{j=1}^{w} (1 - \gamma^{i_j} X) \in F[X].$$

Note que ao expandirmos o produto dos fatores de $l_{\mathbf{a}}(X)$, o termo de grau w é dado por $-\gamma^{(i_1+i_2+\ldots+i_w)} X^w$, cujo coeficiente é não nulo devido γ ser uma raiz n-ésima da unidade. Portanto, vemos que o grau de $l_{\mathbf{a}}(X)$ é igual a $w = \omega(\mathbf{a})$. Além disso, as raízes de $l_{\mathbf{a}}(X)$ são raízes n-ésimas da unidade. De fato,

$$1 - \gamma^{i_j} X = 0 \Leftrightarrow \gamma^{i_j} X = 1 \Leftrightarrow \gamma^{n-i_j} \gamma^{i_j} X = \gamma^{n-i_j} \cdot 1 \Leftrightarrow \gamma^n X = \gamma^{n-i_j} \Leftrightarrow X = \gamma^{-i_j}. \tag{6.14}$$

Logo, as raízes do polinômio $l_{\mathbf{a}}(X)$ são raízes n-ésimas da unidade.

Lema 17. *Seja $l(X) = \sum_{j=0}^{w} l_j X^j \in F[X]$, de grau w. Então, $l(X)$ é o polinômio localizador de uma palavra de coordenadas iguais a 0 e 1 de peso w de $C_K(n, \delta)$ se, e somente se, são verificadas ambas as condições abaixo*

(i) As raízes de $l(X)$ são raízes n-ésimas da unidade, duas a duas distintas;

(ii) $l_j = 0$ para todo $j \leq \delta - 1$ e j não divisível por p.

Demonstração. Seja $l(X)$ o polinômio localizador de $\mathbf{a} \in K^n$, temos que as raízes de $l(X)$ são raízes n-ésimas da unidade e portanto a condição (i) é satisfeita.

Mostremos agora que a condição (ii) também é válida. De fato, sendo $\mathbf{a} \in K^n$, cujas coordenadas $a_{i_1}, \ldots, a_{i_w}$ são todas iguais a 1 e as demais todas nulas, temos por (6.10) que

$$\mathbf{a} \in C_K(n,\delta) \Leftrightarrow \begin{cases} P_1(\gamma^{i_1}, \ldots, \gamma^{i_w}) = \sum_{k=0}^{w} \gamma^{i_k} = 0, \\ P_2(\gamma^{i_1}, \ldots, \gamma^{i_w}) = \sum_{k=0}^{w} \gamma^{2i_k} = 0, \\ \vdots \\ P_{\delta-1}(\gamma^{i_1}, \ldots, \gamma^{i_w}) = \sum_{k=0}^{w} \gamma^{(\delta-1)i_k} = 0. \end{cases} \tag{6.15}$$

em que $P_j(\gamma^{i_1}, \ldots, \gamma^{i_w})$ representa a soma das j-ésimas potências de $\gamma^{i_1}, \ldots, \gamma^{i_w}$. Além disso, pelas definições de polinômio localizador, polinômio simétrico elementar e pelo Lema 6, temos

$$\begin{aligned} l_{\mathbf{a}}(X) &= \prod_{k=1}^{w}(1-\gamma^{i_k}X) = \prod_{k=1}^{w} -\gamma^{i_k}(X-\gamma^{-i_k}) \\ &= (-\gamma^{i_1})(-\gamma^{i_2})\ldots(-\gamma^{i_w})(X-\gamma^{-i_1})(X-\gamma^{-i_2})\ldots(X-\gamma^{-i_w}) \\ &= (-1)^w \cdot \gamma^{i_1}\gamma^{i_2}\ldots\gamma^{i_w} \cdot \sum_{l=0}^{w}(-1)^l S_l(\gamma^{-i_1}, \gamma^{-i_2}\ldots, \gamma^{-i_w})X^{n-l}. \end{aligned}$$

Para entendermos o que acontece a seguir, vamos isolar o termo S_l.

$$S_l = S_l(\gamma^{-i_1}, \gamma^{-i_2}\ldots, \gamma^{-i_w}) = \sum_{1\le k_1<\ldots<k_l\le w} \gamma^{-i_{k_1}}\gamma^{-i_{k_2}}\ldots\gamma^{-i_{k_l}}.$$

Multiplicando $(-1)^l S_l$ por $(-1)^w \cdot \gamma^{i_1}\gamma^{i_2}\ldots\gamma^{i_w}$, temos

$$(-1)^w \cdot \gamma^{i_1}\gamma^{i_2}\ldots\gamma^{i_w} \cdot (-1)^l S_l = (-1)^{w-l} \cdot \gamma^{i_1}\gamma^{i_2}\ldots\gamma^{i_w} \cdot \sum_{1\le k_1<\ldots<k_l\le w} \gamma^{-i_{k_1}}\gamma^{-i_{k_2}}\ldots\gamma^{-i_{k_l}}.$$

Calculando o produto em cada termo do somatório, obtemos

$$(-1)^{w-l}\left[\gamma^{i_1}\ldots\gamma^{i_w} \cdot (\gamma^{-i_1}\ldots\gamma^{-i_l}) + \ldots + \gamma^{i_1}\ldots\gamma^{i_w} \cdot (\gamma^{-i_{w-l+1}}\ldots\gamma^{-i_w})\right].$$

Note que ao efetuarmos o produto de cada parcela da soma entre colchetes, observamos que produto de todos os termos cujos expoentes são opostos é igual a 1, ou seja, $\gamma^{i_{k_1}} \cdot \gamma^{-i_{k_1}} = \ldots = \gamma^{i_{k_l}} \cdot \gamma^{-i_{k_l}} = \gamma^0 = 1$, com $k_1, \ldots, k_l \in \{1, \ldots, w\}$ dois a dois distintos. Com isso, temos

$$\begin{aligned} & (-1)^{w-l}\left[(\gamma^{i_{l+1}}\ldots\gamma^{i_w}) + \ldots + (\gamma^{i_1}\ldots\gamma^{i_{w-l}})\right] = \\ = \; & (-1)^{w-l} \cdot \sum_{1\le\alpha_1<\ldots<\alpha_{w-l}\le w} \gamma^{i_{\alpha_1}}\gamma^{i_{\alpha_2}}\ldots\gamma^{i_{\alpha_{w-l}}} = (-1)^{w-l} \cdot S_{w-l}(\gamma^{i_1}, \gamma^{i_2}\ldots, \gamma^{i_w}), \end{aligned}$$

com $\alpha_t \notin \{k_1, \ldots, k_l\}$, $t = 1, \ldots, w-l$. Logo

$$l_{\mathbf{a}}(X) = \sum_{w-l=0}^{w} (-1)^{w-l} S_{w-l}(\gamma^{i_1}, \gamma^{i_2} \ldots, \gamma^{i_w}) X^{w-l} \overset{(j=w-l)}{=} \sum_{j=0}^{w} (-1)^j S_j(\gamma^{i_1}, \gamma^{i_2} \ldots, \gamma^{i_w}) X^j$$

Agora, combinando os valores de P_j encontrados em (6.15) com as relações (2.29) do Corolário 19, temos

$$\begin{aligned}
&P_1 - S_1 = 0 \Rightarrow S_1 = 0;\\
&P_2 - S_1 P_1 + 2S_2 = 0 \Rightarrow 2S_2 = 0 \Rightarrow S_2 = 0;\\
&\vdots\\
&P_j - S_1 P_{j-1} + \cdots + (-1)^{j-1} S_{j-1} P_1 + (-1)^j j S_n = 0 \Rightarrow j S_j = 0.
\end{aligned}$$

Observe que, por (6.15) e pela Proposição 21, $j \leq \delta - 1$, para todo $j \leq \delta - 1$ e j não divisível por p, $jS_j = 0$ se, e somente se, $S_j = 0$, então $S_j = 0$ para todo $j \leq \delta - 1$ e j não divisível por p. Além disso, como por hipótese $l(X) = l_{\mathbf{a}}(X)$, segue que $S_j = 0$ se, e somente se, $l_j = 0$, para todo $j \leq \delta - 1$ e j não divisível por p. Portanto, satisfaz a condição (ii).

Reciprocamente, sendo $l(X)$ um polinômio que satisfaz as condições (i) e (ii), temos que a partir das raízes n-ésimas da unidade $l(X)$, é possível determinar $\gamma^{-i_1}, \gamma^{-i_2}, \ldots, \gamma^{-i_w}$, pois $\gamma^{i_1}, \gamma^{i_2}, \ldots, \gamma^{i_w}$ são as raízes n-ésimas de $l(X)$. Além disso, $\gamma^{-i_1}, \gamma^{-i_2}, \ldots, \gamma^{-i_w}$ determinam um vetor $\mathbf{a}$ com 1 nas coordenadas $i_1, \ldots, i_w$ e 0 nas demais. Temos ainda que, como o polinômio $l(X)$ satisfaz a condição (ii), então $jS_j = 0$ para todo $j \leq \delta - 1$, com j não divisível por p. Daí, novamente pelo Corolário 19, temos que

$$\begin{aligned}
&P_1 - S_1 = 0 \Rightarrow P_1 = 0;\\
&P_2 - S_1 P_1 + 2S_2 = 0 \Rightarrow P_2 = 0;\\
&\vdots\\
&P_j - S_1 P_{j-1} + \cdots + (-1)^{j-1} S_{j-1} P_1 + (-1)^j j S_n = 0 \Rightarrow P_j = 0.
\end{aligned}$$

Com isso, $P_j = 0$ para todo $j \leq \delta - 1$ e j não divisível por p. Logo, por (6.15), segue que $\mathbf{a} \in C_K(n, \delta)$ e $l(X)$ é o seu polinômio localizador. □

Vale observar que, em corpos com mais do que 2 elementos, torna-se difícil localizar com exatidão as palavras que queremos através do polinômio localizador, uma vez que ele nos permite determinar quais as coordenadas do vetor (palavra) são não nulas. Porém, o polinômio localizador não nos fornece a informação de qual é o valor dessa coordenada e, por essa razão, o submetemos às restrições do Lema 17, com o intuito de que ele seja, de fato, o localizador de uma palavra de um código BCH.

Veremos no resultado a seguir, que se fizermos uma exigência sobre o peso estimado, conseguiremos determinar de maneira exata qual é a distância mínima de certos códigos BCH.

Proposição 49. *Seja $C = C_K(n, \delta)$ um código BCH, em que K é um corpo com q elementos. Se $n = q^m - 1$ para algum inteiro m, e $\delta = q^h - 1$ para algum inteiro h com $h < m$, então C tem peso $d = \delta$.*

Demonstração. Temos por definição que, se γ é um elemento primitivo de $F = \mathbb{F}_{q^m}$, então

$$F^* = \{1, \gamma, \gamma^2, \ldots, \gamma^{n-1}\},$$

ou seja, $\text{ord}(\gamma) = n$, e γ é uma raiz n-ésima primitiva da unidade em F. Para provar o que se pede, é necessário exibir um elemento $\mathbf{a}$ de C com peso igual a δ, já que pelo Teorema 25 sabemos que a distância mínima é maior ou igual a δ. Assim, seja $V \subset F$ um subespaço de F, de dimensão h sobre o corpo K. Considere o polinômio

$$L(X) = \prod_{\beta \in V} (X - \beta).$$

Temos pelo Lema 9 que $L(X)$ é um polinômio q-linear de grau q^h, isto é,

$$L(X) = X^{q^h} + c_{q^{h-1}} X^{q^{h-1}} + c_{q^{h-2}} X^{q^{h-2}} + \cdots + c_0 X.$$

Além disso, $c_0 \neq 0$, pois como V é um espaço vetorial, então $\mathbf{0} \in V$ é raiz do polinômio $L(X)$ com multiplicidade 1. Considere o polinômio recíproco de $L(X)$,

$$L^*(X) = X^{q^h} L\left(\frac{1}{X}\right) = 1 + c_{q^{h-1}} X^{q^h - q^{h-1}} + c_{q^{h-2}} X^{q^h - q^{h-2}} + \cdots + c_0 X^{q^h - 1}.$$

O grau do polinômio $L^*(X)$ é igual a $q^h - 1$, que por hipótese é igual a δ. Agora, verifiquemos que o polinômio $L^*(X)$ satisfaz as condições (i) e (ii) do Lema 17. Temos pelo Corolário 6 que os elementos de F são raízes de $X^{q^m} - X$. Como V é um K-subespaço vetorial de F, então $V \subset F$, consequentemente, os q^h elementos de V também são raízes de $X^{q^m} - X$. Além disso, segue do item (iii) da Proposição 47 que

$$L^*(X) \mid (X^{q^m} - X) \Rightarrow L^*(X) \mid [X(X^{q^m - 1} - 1)] \Rightarrow L^*(X) \mid X \text{ ou } L^*(X) \mid (X^{q^m - 1} - 1).$$

Note que 0 não é uma raiz de $L^*(X)$, por consequência disso, temos que $mdc(X, L^*(X)) = 1$. Logo $L^*(X) \mid (X^{q^m - 1} - 1)$, isto é, as raízes de $L^*(X)$ são raízes de $X^{q^m - 1} - 1 = X^n - 1$. Portanto são raízes n-ésimas da unidade, satisfazendo assim a condição (i) do Lema 17. Além disso, temos que entre os termos $c_{q^{h-i}} X^{q^h - q^{h-i}}$ e $c_{q^{h-i-1}} X^{q^h - q^{h-i-1}}$, com $i = 1, \ldots, h-1$, existem $q^{h-i} - q^{h-i-1}$ termos cujos coeficientes são todos nulos. Como os índices dos coeficientes $c_{q^{h-i}}$ e $c_{q^{h-i-1}}$ são potências consecutivas de p, então os índices dos coeficientes dos termos que estão entre $c_{q^{h-i}} X^{q^h - q^{h-i}}$ e $c_{q^{h-i-1}} X^{q^h - q^{h-i-1}}$ não são múltiplos de p e, por serem todos nulos, segue que $c_j = 0$ para todo j não divisível por p, com $q^{h-i-1} < j < q^{h-i}$. Logo a condição (ii) é satisfeita. Portanto, segue do Lema 17 que existe $\mathbf{a} \in C_K(n, \delta)$ tal que $\omega(\mathbf{a}) = \delta$, com polinômio localizador $L^*(X)$ e, consequentemente, a distância mínima é $d = \delta$. □

Teorema 27. *Seja $C = C_K(n, \delta)$ um código BCH primitivo, em que K é um corpo com q elementos. Então C tem peso d no máximo igual a $q\delta - 1$.*

Demonstração. Considere h um inteiro positivo tal que $q^{h-1} \leq \delta < q^h - 1$. Temos por (6.11) que

$$\delta < q^h - 1 \Rightarrow C'_K(q^m - 1, q^h - 1) \subset C_K(n, \delta) = C_K(q^m - 1, \delta).$$

Observe que C' é um código BCH em que $n = q^m - 1$ e $\delta' = q^h - 1$, e pela Proposição 49, tem peso igual a $q^h - 1$. Disso, decorre que

$$q^{h-1} \leq \delta \Rightarrow q^h \leq q\delta \Rightarrow q^h - 1 \leq q\delta - 1.$$

Agora note que não podemos ter $q^h - 1 < d$, pois como $\delta \leq d$ pode valer a igualdade e, nesse caso, temos $q^h - 1 < \delta$, o que é uma contradição com $\delta < q^h - 1$. Logo $d \leq q^h - 1$. Portanto,

$$d \leq q^h - 1 \leq q\delta - 1.$$

□

6.3 Polinômio Gerador de um Código BCH

Na seção anterior vimos como obter um código BCH $C_K(n, \delta)$ através da construção feita no Teorema 25, em que δ é uma estimativa para o seu peso. Vimos ainda o caso específico de códigos BCH primitivos, em que conseguimos determinar com exatidão o seu peso. No entanto, vale destacar que todas as construções feitas partiram de polinômios, nos quais manipulamos suas raízes (ou raízes n-ésimas da unidade quando conveniente) a fim de obtermos informações sobre os códigos construídos relacionadas ao seu peso.

Nesta seção iremos estudar um código BCH, tendo como ponto de partida os parâmetros n e as raízes n-ésimas da unidade que o define, buscando determinar o seu polinômio gerador $g(X)$. Para tal, considere um código BCH contido em K^n e definido pelas raízes n-ésimas da unidade $\gamma^a, \ldots, \gamma^{a+\delta-2} \in F$. Para determinar $g(X)$, basta encontrar os polinômios minimais $m_{\gamma^j}(X)$ para cada j, com $j = a, a+1, \ldots, a+\delta-2$. Temos pelo item (ii) da Proposição 35 que, o polinômio minimal $m_{\gamma^j}(X)$ é definido por

$$m_{\gamma^a} = (X - \gamma^j)(X - (\gamma^j)^q) \ldots (X - (\gamma^j)^{q^{d_j - 1}}),$$

em que d_j é o menor inteiro positivo que satisfaz $(\gamma^j)^{q^{d_j}} = \gamma^j$, ou seja, d_j é o menor inteiro positivo tal que

$$[jq^{d_j}] = [j] \quad \text{em } \mathbb{Z}_n,$$

isto é,

$$jq^{d_j} \equiv j \ (mod\ n).$$

Como para determinarmos o polinômio minimal $m_{\gamma^j}(X)$ é necessário e suficiente que encontremos todas as suas raízes, então basta encontrarmos os expoentes que tornam o elemento $\gamma \in F$ uma raiz de $m_{\gamma^j}(X)$, ou seja, determinarmos o conjunto

$$C_j = \{[jq^t] \in \mathbb{Z}_n; t \in \mathbb{Z}, t > 0\}. \tag{6.16}$$

Proposição 50. *Os conjuntos C_i, definidos em (6.16), possuem as seguintes propriedades:*

(i) Se $C_i \cap C_j \neq \emptyset$, então $C_i = C_j$;

(ii) A união de todos os C_j é igual a $\mathbb{Z}_n$.

Demonstração. (i) Suponha que $C_i \cap C_j \neq \emptyset$. Então existem inteiros não negativos t e s, tais que

$$[iq^t] = [jq^s]. \tag{6.17}$$

Da igualdade (6.17),

$$iq^t \in [jq^s] \Rightarrow \exists\ k \in \mathbb{Z};\ iq^t = jq^k.$$

Sem perda de generalidade, podemos supor $t \geq k$, dessa forma, temos

$$iq^{t-k} = j \Rightarrow [j] = [iq^{t-k}] \in C_i.$$

Logo, $C_j \subset C_i$. Por outro lado, multiplicando a última igualdade acima por $[q^r]$, para algum inteiro r tal que $[q^{t-k+r}] = 1$, de modo que $[iq^{t-k+r}] = [i]$, obtemos

$$[i] = [iq^{t-k+r}] = [jq^r] \in C_j.$$

Logo, $C_i \subset C_j$. Portanto vale a igualdade.

(ii) É imediato que a união de todos os C_j está contido em $\mathbb{Z}_n$, pois cada C_j está contido em $\mathbb{Z}_n$. Por outro lado, sendo $[m] \in \mathbb{Z}_n$ arbitrário, temos que $[m] = [mq^0] \in C_m$. Logo $\mathbb{Z}_n$ está contido na união de todos os C_j. Portanto, vale a igualdade.

□

O conjunto C_j definido em (6.16) é chamado de *classe de ciclotomia* de j módulo n.

Desde que um código BCH também é um código cíclico, sua dimensão é dada da mesma maneira que a dos códigos cíclicos, sendo

$$dim_K C = n - gr(g(X)),$$

como vimos no Corolário 14. Vamos agora dar um exemplo de um código BCH sobre $\mathbb{F}_2$, cujas palavras tenham comprimento $n = 15$ e peso estimado $\delta = 5$, ou seja, é o código $C_K(15, 5)$.

Exemplo 18. Considere $\alpha = [X]$ um elemento primitivo de $F = \mathbb{F}_{16}$, como visto no Exemplo 4. Temos que $\gamma = \alpha^{\frac{16-1}{15}} = \alpha$ é uma raiz 15-ésima primitiva da unidade em F.

Como o menor valor de m tal que $2^m \equiv 1 \ (mod\ 15)$ é 4, passamos a determinar as classes de ciclotomia i módulo 15. Temos por (6.16) que,

- $i = 0 : C_0 = \{[0]\}$.
- $i = 1 : C_1 = \{[1], [2], [4], [8]\}$, pois
$$\begin{cases} [1 \cdot 2^0] = [1]; \\ [1 \cdot 2^1] = [2]; \\ [1 \cdot 2^2] = [4]; \\ [1 \cdot 2^3] = [8]; \\ [1 \cdot 2^4] = [1]. \end{cases}$$
- $i = 3 : C_3 = \{[3], [6], [9], [12]\}$, pois
$$\begin{cases} [3 \cdot 2^0] = [3]; \\ [3 \cdot 2^1] = [6]; \\ [3 \cdot 2^2] = [12]; \\ [3 \cdot 2^3] = [9]; \\ [3 \cdot 2^4] = [3]. \end{cases}$$

De modo análogo, obtemos $C_5 = \{[5], [10]\}$ e $C_7 = \{[7], [11], [13], [14]\}$. Agora vamos determinar os polinômios minimais de todas as potências de γ. Para tanto, façamos

$$\begin{aligned} m_\gamma(X) &= m_{\gamma^2}(X) = m_{\gamma^4}(X) = m_{\gamma^8}(X) = (X-\gamma)(X-\gamma^2)(X-\gamma^4)(X-\gamma^8) = \\ &= X^4 + X^3 \underbrace{(\gamma^8 + \gamma^4 + \gamma^2 + \gamma)}_{(I)} + X^2 \underbrace{(\gamma^{12} + \gamma^{10} + \gamma^9 + \gamma^6 + \gamma^5 + \gamma^3)}_{(II)} + \\ &\quad + X \underbrace{(\gamma^{14} + \gamma^{13} + \gamma^{11} + \gamma^7)}_{(III)} + \underbrace{\gamma^{15}}_{(IV)}. \end{aligned}$$

Daí, determinamos cada coeficiente usando a tabela de Zech do Exemplo 4, como se segue.
(I) :

$$\begin{aligned} &\gamma^4 + \gamma^8 = \gamma^4 \cdot \gamma^{z(4)} = \gamma^4 \cdot \gamma = \gamma^5 \\ &\gamma^2 + \gamma^5 = \gamma^2 \cdot \gamma^{z(3)} = \gamma^2 \cdot \gamma^{14} = \gamma^{16} = \gamma \\ &\gamma + \gamma = 2\gamma = 0. \end{aligned}$$

Assim, temos que $\gamma^8 + \gamma^4 + \gamma^2 + \gamma$ resulta em 0, consequentemente, temos $0 \cdot X^3$. Analogamente, concluímos que em (II), $\gamma^{12} + \gamma^{10} + \gamma^9 + \gamma^6 + \gamma^5 + \gamma^3$ resulta em 0 e, consequentemente, temos $0 \cdot X^2$; em (III), $\gamma^{14} + \gamma^{13} + \gamma^{11} + \gamma^7$ resulta em 1 e, portanto, temos $1 \cdot X$; em (IV), $\gamma^{15} = 1$. Portanto,

$$m_\gamma(X) = m_{\gamma^2}(X) = m_{\gamma^4}(X) = m_{\gamma^8}(X) = X^4 + X + 1.$$

Procedendo de maneira análoga para os outros minimais, obtemos

$$\begin{aligned} m_{\gamma^3}(X) &= m_{\gamma^6}(X) = m_{\gamma^9}(X) = m_{\gamma^{12}}(X) = X^4 + X^3 + X^2 + X + 1; \\ m_{\gamma^5}(X) &= m_{\gamma^{10}}(X) = X^2 + X + 1; \\ m_{\gamma^7}(X) &= m_{\gamma^{11}}(X) = m_{\gamma^{13}}(X) = m_{\gamma^{14}}(X) = X^4 + X^3 + 1. \end{aligned}$$

Pelo Teorema 25, segue que o código BCH é gerado pelo polinômio

$$\begin{aligned} g(X) &= mmc(m_\gamma(X), m_{\gamma^2}(X), m_{\gamma^3}(X), m_{\gamma^4}(X)) = mmc(m_\gamma(X), m_{\gamma^3}(X)) = \\ &= (X^4 + X + 1)(X^4 + X^3 + X^2 + X + 1) \\ &= X^8 + X^7 + X^6 + X^4 + 1. \end{aligned}$$

Logo, pelo Corolário 14, a dimensão de $C = n - gr(g(X)) = 7$ e, portanto, $|C| = 2^7 = 128$. Agora, para encontrarmos a matriz geradora de C na forma padrão, fazemos

$$\begin{aligned} X^8 &= (X^8 + X^7 + X^6 + X^4 + 1) + (X^7 + X^6 + X^4 + 1); \\ X^9 &= (X^8 + X^7 + X^6 + X^4 + 1)(X + 1) + (X^6 + X^5 + X^4 + X + 1); \\ X^{10} &= (X^8 + X^7 + X^6 + X^4 + 1)(X^2 + X) + (X^7 + X^6 + X^5 + X^2 + X); \\ X^{11} &= (X^8 + X^7 + X^6 + X^4 + 1)(X^3 + X^2 + 1) + (X^4 + X^3 + X^2 + 1); \\ X^{12} &= (X^8 + X^7 + X^6 + X^4 + 1)(X^4 + X^3 + X^2) + (X^5 + X^4 + X^3 + X); \\ X^{13} &= (X^8 + X^7 + X^6 + X^4 + 1)(X^5 + X^4 + X^3) + (X^6 + X^5 + X^4 + X^2); \\ X^{14} &= (X^8 + X^7 + X^6 + X^4 + 1)(X^6 + X^5 + X^4) + (X^7 + X^6 + X^5 + X^3). \end{aligned}$$

E assim, pelo Teorema 23, a matriz G' é dada por

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

Além disso, temos que

$$\begin{aligned} &\begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \end{bmatrix} G' = \\ &= \begin{bmatrix} b_0 & b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \end{bmatrix}, \end{aligned}$$

sendo $b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7$ coeficientes do polinômio

$$\begin{aligned} & a_1(X^7+X^6+X^4+1)+a_2(X^6+X^5+X^4+X+1)+a_3(X^7+X^6+X^5+X^2+X)+ \\ & +a_4(X^4+X^3+X^2+1)+a_5(X^5+X^4+X^3+X)+a_6(X^6+X^5+X^4+X^2)+ \\ & +a_7(X^7+X^6+X^5+X^3)= \\ = \; & a_1+a_2+a_4+(a_2+a_3+a_5)X+(a_3+a_4+a_6)X^2+ \\ & +(a_4+a_5+a_7)X^3+(a_1+a_2+a_4+a_5+a_6)X^4+(a_2+a_3+a_5+a_6+a_7)X^5 \\ & +(a_1+a_2+a_3+a_6+a_7)X^6+(a_1+a_3+a_7)X^7. \end{aligned}$$

Logo, pelo algoritmo de codificação dos códigos cíclicos, o elemento $(a_1, a_2, a_3, a_4, a_5, a_6, a_7)$ de $\mathbb{F}_2^7$ é codificado como

$$(b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7, a_1, a_2, a_3, a_4, a_5, a_6, a_7).$$

Segue do Teorema 24 que a matriz teste de paridade é

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

Suponhamos que foi recebida por um canal de comunicação a seguinte mensagem:

0 1 0 1 0 1 0 0 1 0 1 1 0 1 0	0 1 1 0 1 1 0 1 1 0 1 1 0 1 1	1 0 1 0 1 0 1 0 0 1 1 0 1 0 1
0 0 1 0 1 0 0 0 0 1 1 1 0 1 1	1 0 0 0 0 1 1 1 0 1 1 0 0 1 0	1 0 0 0 1 0 1 1 1 0 1 1 1 1 1
0 0 1 1 0 0 0 1 1 0 1 1 1 1 1	1 1 0 0 0 0 1 0 1 0 1 0 0 1 0	0 1 1 1 0 0 0 1 1 0 0 1 1 1 1
0 0 0 0 0 1 0 0 0 1 1 1 0 1 1 0	0 1 0 1 1 0 0 0 0 1 0 1 0 1 0	0 1 0 1 1 1 1 0 0 0 1 0 0 1 1

Verificando se as palavras estão no código, obtemos que

$$H\begin{bmatrix}0&1&0&1&0&1&0&0&1&0&1&1&0&1&0\end{bmatrix}^t = \begin{bmatrix}0&0&1&0&1&1&1&0\end{bmatrix};$$

$$H\begin{bmatrix}0&1&1&0&1&1&0&1&1&0&1&1&0&1&1\end{bmatrix}^t = \begin{bmatrix}0&0&0&0&0&0&0&0\end{bmatrix};$$

$$H\begin{bmatrix}1&0&1&0&1&0&1&0&0&1&1&0&1&0&1\end{bmatrix}^t = \begin{bmatrix}0&1&0&0&1&0&0&0\end{bmatrix};$$

$$H\begin{bmatrix}0&0&1&0&1&0&0&0&0&1&1&1&0&1&1\end{bmatrix}^t = \begin{bmatrix}0&0&0&0&0&0&0&0\end{bmatrix};$$

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

Segue que ocorreram erros na transmissão das palavras

$$0\,1\,0\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,1\,0 \quad 1\,0\,1\,0\,1\,0\,1\,0\,0\,1\,1\,0\,1\,0\,1$$
$$1\,0\,0\,0\,1\,0\,1\,1\,1\,0\,1\,1\,1\,1\,1 \quad 0\,1\,1\,1\,0\,0\,0\,1\,1\,0\,0\,1\,1\,1\,1.$$

Aplicando o Teorema 19 para detectar e, se possível, corrigir esses erros, vemos que ocorreram no máximo dois erros para cada uma dessas palavras e, com isso, conseguimos corrigi-las da seguinte maneira

$$0\,1\,0\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,\mathbf{1}\,0 \mapsto 0\,1\,0\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,\mathbf{0}\,0$$
$$1\,\mathbf{0}\,1\,0\,\mathbf{1}\,0\,1\,0\,0\,1\,1\,0\,1\,0\,1 \mapsto 1\,\mathbf{1}\,1\,0\,\mathbf{0}\,0\,1\,0\,0\,1\,1\,0\,1\,0\,1$$
$$1\,0\,0\,0\,1\,0\,\mathbf{1}\,1\,1\,0\,1\,\mathbf{1}\,1\,1\,1 \mapsto 1\,0\,0\,0\,1\,0\,\mathbf{0}\,1\,1\,0\,1\,\mathbf{0}\,1\,1\,1$$
$$0\,\mathbf{1}\,1\,1\,0\,0\,0\,1\,1\,0\,\mathbf{0}\,1\,1\,1\,1 \mapsto 0\,\mathbf{0}\,1\,1\,0\,0\,0\,1\,1\,0\,\mathbf{1}\,1\,1\,1\,1$$

Logo, a mensagem transmitida foi

$$0\,1\,0\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,\mathbf{0}\,0 \quad 0\,1\,1\,0\,1\,1\,0\,1\,1\,0\,1\,1\,0\,1\,1 \quad 1\,\mathbf{1}\,1\,0\,\mathbf{0}\,0\,1\,0\,0\,1\,1\,0\,1\,0\,1$$
$$0\,0\,1\,0\,1\,0\,0\,0\,0\,1\,1\,1\,0\,1\,1 \quad 1\,0\,0\,0\,0\,1\,1\,1\,0\,1\,1\,0\,0\,1\,0 \quad 1\,0\,0\,0\,1\,0\,\mathbf{0}\,1\,1\,0\,1\,\mathbf{0}\,1\,1\,1$$
$$0\,0\,1\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,1\,1 \quad 1\,1\,0\,0\,0\,0\,1\,0\,1\,0\,1\,0\,0\,1\,0 \quad 0\,\mathbf{0}\,1\,1\,0\,0\,0\,1\,1\,0\,\mathbf{1}\,1\,1\,1\,1$$
$$0\,0\,0\,0\,0\,1\,0\,0\,0\,1\,0\,1\,1\,1\,0 \quad 0\,1\,0\,1\,1\,0\,0\,0\,0\,1\,0\,1\,0\,1\,0 \quad 0\,1\,0\,1\,1\,1\,1\,0\,0\,0\,1\,0\,0\,1\,1$$

Seguem nas próximas duas páginas as tabelas de codificação e de síndromes dos vetores erros, respectivamente. As contas feitas para gerar as palavras codificadas e as síndromes dos vetores erros foram todas executadas por uma implementação em Python e estão disponíveis em (SOARES, 2020).

Tabela de Codificação					
Código Fonte	Código de Canal	Código Fonte	Código de Canal	Código Fonte	Código de Canal
0000000	0000000000000000	0101011	0100111101010011	1010110	1001111010010110
0000001	0001011100000001	0101100	0010101001011100	1010111	1000100110101111
0000010	0010111000000010	0101101	0011110101011101	1011000	0101010010110000
0000011	0011100100000011	0101110	0000010001011110	1011001	0100001110110001
0000100	0101110000000100	0101111	0001001101011111	1011010	0111101010110010
0000101	0100101100000101	0110000	1010100101100000	1011011	0110110110110011
0000110	0111001000000110	0110001	1011111001100001	1011100	0000100010111100
0000111	0110010100000111	0110010	1000011101100010	1011101	0001111111011101
0001000	1011110000001000	0110011	1001000001100011	1011110	0010011010111110
0001001	1010111110001001	0110100	1111010101101000	1011111	0011000110111111
0001010	1001011000001010	0110101	1110001001101010	1100000	0100010111000000
0001011	1000000100011011	0110110	1101101101101110	1100001	0101001011000001
0001100	1110010000011100	0110111	1100110001101111	1100010	0110101111000010
0001101	1111001100011101	0111000	0001000101110000	1100011	0111110011000011
0001110	1100101000011110	0111001	0000011001110011	1100100	0001100111001000
0001111	1101110100011111	0111010	0011111101110100	1100101	0000111011001001
0010000	0110011100100000	0111011	0010100001110110	1100110	0011011111001100
0010001	0111000000100001	0111100	0100110101111000	1100111	0010000011001111
0010010	0100100100100100	0111101	0101101001111010	1101000	1111110111010000
0010011	0101111000100111	0111110	0110001101111100	1101001	1110101011010011
0010100	0011101100101000	0111111	0111010001111111	1101010	1101001111010100
0010101	0010110000101010	1000000	1000101110000000	1101011	1100010011010110
0010110	0001010100101100	1000001	1001110010000001	1101100	1010000111011000
0010111	0000001000101111	1000010	1010010110000010	1101101	1011011011011010
0011000	1101111100110000	1000011	1011001010000011	1101110	1000111111011100
0011001	1100100000110011	1000100	1101011110001000	1101111	1001100011011111
0011010	1111000100110100	1000101	1100000010001010	1110000	0010001011100000
0011011	1110011000110111	1000110	1111100110001100	1110001	0011010111100011
0011100	1000001100111000	1000111	1110111010001111	1110010	0000110011100100
0011101	1001010000111011	1001000	0011001110001000	1110011	0001101111100111
0011110	1010110100111110	1001001	0010010010010011	1110100	0111111011101000
0011111	1011101000111111	1001010	0001110110010100	1110101	0110100111101011
0100000	1100111001000000	1001011	0000101010010011	1110110	0101000011101100
0100001	1101100101000001	1001100	0110111111001100	1110111	0100011111101111
0100010	1110000001000010	1001101	0111100010011011	1111000	1001101011111000
0100011	1111011101000011	1001110	0100000110011100	1111001	1000110111111001
0100100	1001001001001000	1001111	0101011010011111	1111010	1011010011111010
0100101	1000010101001011	1010000	1110110010100000	1111011	1010001111111011
0100110	1011111000100110	1010001	1111101110100011	1111100	1100011011111100
0100111	1010101101001111	1010010	1100001010100010	1111101	1101000111111101
0101000	0111011001010000	1010011	1101010110100011	1111110	1110100011111110
0101001	0110000101010011	1010100	1011000010101000	1111111	1111111111111111
0101010	0101100001010100	1010101	1010011110101011		

Tabela de síndromes			
Erro	Síndrome	Erro	Síndrome
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0	0 0 0 0 0 0 1 0 0 0 0 1 0 0 0	1 0 1 1 1 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1	0 0 0 1 0 1 1 1	0 0 0 0 0 0 0 1 0 0 0 1 0 0 0	1 0 1 1 1 0 0 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0	0 0 1 0 1 1 1 0	0 0 0 0 0 0 0 0 1 0 0 1 0 0 0	0 0 1 1 0 0 1 1
0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0	0 1 0 1 1 1 0 0	0 0 0 0 0 0 0 0 0 1 0 1 0 0 0	0 1 1 1 0 1 1 0
0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0	1 0 1 1 1 0 0 0	0 0 0 0 0 0 0 0 0 0 1 1 0 0 0	1 1 0 1 1 1 1 1
0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0	0 1 1 0 0 1 1 1	1 0 0 0 0 0 0 0 0 0 1 0 0 0 0	1 1 1 0 0 1 1 1
0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0	1 1 0 0 1 1 1 0	0 1 0 0 0 0 0 0 0 0 1 0 0 0 0	0 0 1 0 0 1 1 1
0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0	1 0 0 0 1 0 1 1	0 0 1 0 0 0 0 0 0 0 1 0 0 0 0	0 1 0 0 0 1 1 1
0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0	0 0 0 0 0 0 0 1	0 0 0 1 0 0 0 0 0 0 1 0 0 0 0	0 1 1 1 0 1 1 1
0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0	0 0 0 0 0 0 1 0	0 0 0 0 1 0 0 0 0 0 1 0 0 0 0	0 1 1 0 1 1 1 1
0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0	0 0 0 0 0 1 0 0	0 0 0 0 0 1 0 0 0 0 1 0 0 0 0	0 1 1 0 0 0 1 1
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 1 0 0 0	0 0 0 0 0 0 1 0 0 0 1 0 0 0 0	0 1 1 0 0 1 0 1
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 1 0 0 0 0	0 0 0 0 0 0 0 1 0 0 1 0 0 0 0	0 1 1 0 0 1 1 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 1 0 0 0 0 0	0 0 0 0 0 0 0 0 1 0 1 0 0 0 0	1 1 1 0 1 1 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 1 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 1 1 0 0 0 0	1 0 1 0 1 0 0 1
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1 0 0 0 0 0 0 0	1 0 0 0 0 0 0 0 0 0 1 0 0 0 0	0 1 0 0 1 1 1 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1	1 0 0 1 0 1 1 1	0 1 0 0 0 0 0 0 0 0 1 0 0 0 0	1 0 0 0 1 1 1 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1	0 1 0 1 0 1 1 1	0 0 1 0 0 0 0 0 0 0 1 0 0 0 0	1 1 1 0 1 1 1 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1	0 0 1 1 0 1 1 1	0 0 0 1 0 0 0 0 0 0 1 0 0 0 0	1 1 0 1 1 1 1 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1	0 0 0 0 0 1 1 1	0 0 0 0 1 0 0 0 0 0 1 0 0 0 0	1 1 0 0 0 1 1 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1	0 0 0 1 1 1 1 1	0 0 0 0 0 1 0 0 0 1 0 0 0 0 0	1 1 0 0 1 0 1 0
0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1	0 0 0 1 0 0 1 1	0 0 0 0 0 0 1 0 0 1 0 0 0 0 0	1 1 0 0 1 1 0 0
0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1	0 0 0 1 0 1 0 1	0 0 0 0 0 0 0 1 0 1 0 0 0 0 0	1 1 0 0 1 1 1 1
0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1	0 0 0 1 0 1 1 0	0 0 0 0 0 0 0 0 1 1 0 0 0 0 0	0 1 0 0 0 1 0 1
0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1	1 0 0 1 1 1 0 0	1 0 0 0 0 0 0 0 1 0 0 0 0 0 0	0 0 0 0 1 0 1 1
0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1	1 1 0 1 1 0 0 1	0 1 0 0 0 0 0 0 1 0 0 0 0 0 0	1 1 0 0 1 0 1 1
0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1	0 1 1 1 0 0 0 0	0 0 1 0 0 0 0 0 1 0 0 0 0 0 0	1 0 1 0 1 0 1 1
0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1	1 0 1 0 1 1 1 1	0 0 0 1 0 0 0 0 1 0 0 0 0 0 0	1 0 0 1 1 0 1 1
0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1	0 1 0 0 1 0 1 1	0 0 0 0 1 0 0 0 1 0 0 0 0 0 0	1 0 0 0 0 0 1 1
0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1	0 0 1 1 1 0 0 1	0 0 0 0 0 1 0 0 1 0 0 0 0 0 0	1 0 0 0 1 1 1 1
1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0	1 0 1 0 1 1 1 0	0 0 0 0 0 0 1 0 1 0 0 0 0 0 0	1 0 0 0 1 0 0 1
0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0	0 1 1 0 1 1 1 0	0 0 0 0 0 0 0 1 1 0 0 0 0 0 0	1 0 0 0 1 0 1 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0	0 0 0 0 1 1 1 0	1 0 0 0 0 0 0 1 0 0 0 0 0 0 0	1 0 0 0 0 0 0 1
0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0	0 0 1 1 1 1 1 0	0 1 0 0 0 0 0 1 0 0 0 0 0 0 0	0 1 0 0 0 0 0 1
0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0	0 0 1 0 0 1 1 0	0 0 1 0 0 0 0 1 0 0 0 0 0 0 0	0 0 1 0 0 0 0 1
0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0	0 0 1 0 1 0 1 0	0 0 0 1 0 0 0 1 0 0 0 0 0 0 0	0 0 0 1 0 0 0 1
0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0	0 0 1 0 1 1 0 0	0 0 0 0 1 0 0 1 0 0 0 0 0 0 0	0 0 0 0 1 0 0 1
0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0	0 0 1 0 1 1 1 1	0 0 0 0 0 1 0 1 0 0 0 0 0 0 0	0 0 0 0 0 1 0 1
0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0	0 1 0 0 1 0 1 0	0 0 0 0 0 0 1 1 0 0 0 0 0 0 0	0 0 0 0 0 0 1 1
0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0	1 0 1 0 0 1 0 0	1 0 0 0 0 0 1 0 0 0 0 0 0 0 0	1 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0	1 1 1 0 0 0 0 0	1 0 0 0 0 0 1 0 0 0 0 0 0 0 0	1 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0	0 1 0 0 1 0 0 1	0 1 0 0 0 0 1 0 0 0 0 0 0 0 0	0 1 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0	1 0 0 1 0 1 1 0	0 0 1 0 0 0 1 0 0 0 0 0 0 0 0	0 0 1 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0	0 1 1 1 0 0 1 0	0 0 0 1 0 0 1 0 0 0 0 0 0 0 0	0 0 0 1 0 0 1 0
1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0	1 1 0 1 1 1 0 0	0 0 0 0 1 0 1 0 0 0 0 0 0 0 0	0 0 0 0 1 0 1 0
0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0	0 0 0 1 1 1 0 0	0 0 0 0 0 1 1 0 0 0 0 0 0 0 0	0 0 0 0 0 1 1 0
0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0	0 1 1 1 1 1 0 0	1 0 0 0 0 1 0 0 0 0 0 0 0 0 0	1 0 0 0 0 1 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0	0 1 0 0 1 1 0 0	0 1 0 0 0 1 0 0 0 0 0 0 0 0 0	0 1 0 0 0 1 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0	0 1 0 1 0 1 0 0	0 0 1 0 0 1 0 0 0 0 0 0 0 0 0	0 0 1 0 0 1 0 0
0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0	0 1 0 1 1 0 0 0	0 0 0 1 0 1 0 0 0 0 0 0 0 0 0	0 0 0 1 0 1 0 0
0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0	0 1 0 1 1 1 1 0	0 0 0 0 1 1 0 0 0 0 0 0 0 0 0	0 0 0 0 1 1 0 0
0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0	0 1 0 1 1 1 0 1	1 0 0 0 1 0 0 0 0 0 0 0 0 0 0	1 0 0 0 1 0 0 0
0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0	1 1 0 1 0 1 1 1	0 1 0 0 1 0 0 0 0 0 0 0 0 0 0	0 1 0 0 1 0 0 0
0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0	1 0 0 1 0 0 1 0	0 0 1 0 1 0 0 0 0 0 0 0 0 0 0	0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0	0 0 1 1 1 0 1 1	0 0 0 1 1 0 0 0 0 0 0 0 0 0 0	0 0 0 1 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0	1 1 1 0 0 1 0 0	1 0 0 1 0 0 0 0 0 0 0 0 0 0 0	1 0 0 1 0 0 0 0
1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0	0 0 1 1 1 0 0 0	0 1 0 1 0 0 0 0 0 0 0 0 0 0 0	0 1 0 1 0 0 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0	1 1 1 1 1 0 0 0	0 0 1 1 0 0 0 0 0 0 0 0 0 0 0	0 0 1 1 0 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0	1 0 0 1 1 0 0 0	1 0 1 0 0 0 0 0 0 0 0 0 0 0 0	1 0 1 0 0 0 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0	1 0 1 0 1 0 0 0	0 1 1 0 0 0 0 0 0 0 0 0 0 0 0	0 1 1 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0	1 0 1 1 0 0 0 0	1 1 0 0 0 0 0 0 0 0 0 0 0 0 0	1 1 0 0 0 0 0 0
0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0	1 0 1 1 1 1 0 0		

7 Conclusão

Nesse trabalho estudamos os códigos lineares, e algumas de suas classes, sendo eles os códigos cíclicos e os códigos BCH, sendo este último uma classe de códigos cíclicos, tendo como principal base o material citado em (HEFEZ; VILLELA, 2008). Em um primeiro momento nos dedicamos em introduzir, com o devido rigor matemático, alguns conceitos da álgebra linear e da álgebra abstrata que foram essenciais para a construção desses códigos, em virtude do nosso foco principal ter sido estudá-los sob o ponto de vista matemático e não computacional.

Pudemos perceber no capítulo 3 que uma das dificuldades encontradas é a de determinar a distância mínima de um código, devido à grande quantidade de cálculos necessários para isso, o que, por consequência, dificulta a determinação de seus parâmetros. Já no capítulo 4, vimos que quando tratamos de códigos lineares é possível descrevê-los como núcleo ou imagem de uma transformação linear, o que diminui consideravelmente a quantidade de cálculos necessários para determinação da distância mínima, entretanto, ainda é necessário uma grande quantidade de cálculos quando se trata de códigos de muitas palavras. Por esse motivo estudamos as matrizes geradoras de um código e, através da matriz teste de paridade (obtida a partir da matriz geradora), foi possível estabelecermos um método que exige menos cálculos para determinar a distância mínima. Com os parâmetros do código encontrados, apresentamos o Teorema 19 (Algoritmo da Decodificação), a fim de estabelecer um método para a detecção e correção de erros para códigos lineares, o qual se mostrou bastante eficiente para o que propomos.

No capítulo 5, observamos que um código cíclico, por ser uma subclasse de um código linear, possui a estrutura de espaço vetorial, entretanto, vimos no Teorema 20 que conseguimos definir um código cíclico como um ideal de R_n, o que enriquece a estrutura sobre a qual um código cíclico está definido, que por sua vez, possibilita darmos um algoritmo de codificação mais eficiente do que o de códigos lineares. Apesar disso, determinar a distância mínima de um código cíclico é um problema parcialmente em aberto. Por consequência, não conseguimos aplicar o Teorema 19 de modo eficiente para os códigos cíclicos.

Tendo em vista a impossibilidade de encontrarmos a distância mínima de um código cíclico, estudamos os códigos BCH no capítulo 6, que por intermédio do Teorema 25, possibilita estimarmos a distância mínima de um código cíclico. Além disso, vimos na Proposição 49, que quando tratamos de códigos BCH primitivos, conseguimos determinar de maneira exata qual a sua distância mínima. Ainda no capítulo 6, foi possível percebermos, através do Exemplo 18, que os cálculos necessários para construirmos um código BCH de muitas palavras são significativamente mais simples e rápidos do que os códigos lineares vistos no capítulo 4.

Este trabalho, sob uma perspectiva pessoal, possibilitou-me desenvolver algumas habilidades essenciais, principalmente para a carreira acadêmica, como o raciocínio lógico, um senso crítico em constante desenvolvimento, uma maior capacidade de pesquisa cada vez mais independente e, como autocrítica, a importância do contato constante com tudo o que foi estudado desde o primeiro dia de trabalho, juntamente com revisões completas do texto redigido.

Como perspectiva para continuidade dos estudos, é relevante a generalização dos códigos BCH, dada pelo estudo dos códigos de Goppa, com o objetivo de obtermos outros tipos de codificação e decodificação, para fins comparativos. Por um outro lado, poderíamos também seguir um estudo computacional para a implementação dos códigos estudados, a fim de verificarmos a atuação desses na prática.

Referências

BERLEKAMP, E. *Algebraic coding theory, revised edition*. Singapore: World Scientific, 2015. Citado na página 11.

BHATTACHARYA, P. B.; JAIN, S. K.; NAGPAUL, S. *Basic abstract algebra*. New York, NY: Cambridge University Press, 1994. Citado 2 vezes nas páginas 12 e 24.

COELHO, F. U.; LOURENCO, M. L. *Um Curso de Álgebra Linear, Vol. 34*. 2. ed. São Paulo/SP: Edusp, 2013. Citado 4 vezes nas páginas 12, 16, 20 e 21.

GEISEL, W. A. Tutorial on reed-solomon error correction coding. *NASA Thecnical Memorandum*, n. 102162, p. 1–8, 1990. Citado na página 11.

GONÇALVES, A. *Introdução à álgebra*. 6. ed. Rio de Janeiro/RJ: IMPA, 2017. Citado na página 12.

HEFEZ, A. *Curso de álgebra, volume 1*. 3. ed. Rio de Janeiro/RJ: Impa, 2002. (Coleção Matemática Universitária). Citado 4 vezes nas páginas 12, 15, 22 e 27.

HEFEZ, A.; VILLELA, M. L. T. *Códigos corretores de erros*. 2. ed. Rio de Janeiro/RJ: IMPA, 2008. Citado 14 vezes nas páginas 12, 13, 26, 40, 46, 62, 75, 77, 79, 83, 90, 105, 123 e 143.

IEEE. *Claude E. Shannon*. 2020. IEEE - Computer Society. Disponível em: <https://www.itsoc.org/about/shannon>. Acesso em: 07 fev 2020. Citado na página 11.

LABS, N. B. *History*. 2019. Nokia Bell labs. Disponível em: <https://www.bell-labs.com/about/history-bell-labs/>. Acesso em: 08 abr 2020. Citado na página 11.

LEE, J. A. N. *Computer Pioneers*. 2019. IEEE - Computer Society. Disponível em: <https://history.computer.org/pioneers/hamming.html>. Acesso em: 07 fev 2020. Citado na página 11.

OLIVEIRA, O. *Teorema do Posto (para matrizes) e Teorema do Núcleo e da Imagem*. 2018. 1–4 p. Citado na página 19.

RAAPHORST, S. Reed-muller codes. *Carleton University, May*, Citeseer, v. 9, 2003. Citado na página 11.

REED, I. S.; CHEN, X. *Error-control coding for data networks*. New York, NY: Springer Science & Business Media, 2001. v. 508. Citado 2 vezes nas páginas 11 e 126.

SOARES, H. *Correcting Codes*. 2020. GitHub. Disponível em: <https://github.com/hugoazevedosoares/correcting_codes>. Citado na página 140.