UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI
Pós-Graduação em Matemática em Rede Nacional

Pedro Henrique Diolindo de Paiva

Códigos Corretores de Erros e Criptografia Baseada em Códigos.

Teófilo Otoni - MG
2018

Pedro Henrique Diolindo de Paiva

Códigos Corretores de Erros e Criptografia Baseada em Códigos.

Dissertação apresentada ao programa de Pós-Graduação em Matemática em Rede Nacional da Universidade Federal dos Vales do Jequitinhonha e Mucuri, como requisito para obtenção do título de Mestre.

Orientador: Prof. Dr. Nolmar Melo

Teófilo Otoni

2018

Elaborado com os dados fornecidos pelo(a) autor(a).

P149c Paiva, Pedro Henrique Diolindo de
Códigos corretores de erros e criptografia baseada em códigos / Pedro Henrique Diolindo de Paiva, 2018.
77 p. : il.

Orientador: Nolmar Melo

Dissertação (Mestrado Profissional – Programa de Pós-Graduação em Matemática em Rede Nacional) - Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, 2018.

1. Códigos corretores de erros. 2. Criptografia. 3. Álgebra. 4. Educação básica. I. Melo, Nolmar. II. Título. III. Universidade Federal dos Vales do Jequitinhonha e Mucuri.

CDD 510

Ficha Catalográfica – Serviço de Bibliotecas/UFVJM
Bibliotecária Nádia Santos Barbosa, CRB-6/3468

.

PEDRO HENRIQUE DIOLINDO DE PAIVA

Códigos Corretores de Erros e Criptografia Baseada em Códigos.

Dissertação apresentada ao MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL, nível de MESTRADO como parte dos requisitos para obtenção do título de MAGISTER SCIENTIAE EM MATEMÁTICA

Orientador (a): Prof. Dr. Nolmar Melo De Souza

Data da aprovação : 04/10/2018

Prof.Dr. NOLMAR MELO DE SOUZA - UFVJM

Prof.Dr. GERALDO MOREIRA DA ROCHA FILHO - UFVJM

Prof.Drª KARINA KFOURI SARTORI - UESC

Prof. WEVERSSON DALMASO SELLIN - UFVJM

Dedico esse trabalho à minha mãe Ana Maria Luíza de Paiva.

AGRADECIMENTOS

À minha família pelo apoio, companheirismo e incentivo ao qual sem eles não conseguiria prosseguir.

Aos meus colegas e amigos do PROFMAT pela parceria e companheirismo, pelos momentos de descontração e aprendizagem.

À CAPES e FAPEMIG pela ajuda financeira.

Ao Prof. Dr. Nolmar Melo por me introduzir a Teoria dos Códigos Corretores de Erros e das Estruturas Algébricas, pela orientação, paciência e sugestões durante o processo.

Aos professores do PROFMAT que contribuíram de alguma forma para meu avanço no conhecimento.

À todos que torceram e torcem pelo meu sucesso e pela minha felicidade incondicionalmente.

Somos o que há de melhor
Somos o que dá pra fazer
O que não dá pra evitar
E não se pode escolher.
(Humberto Gessinger)

RESUMO

O principal objetivo desta dissertação é apresentar o primeiro criptosistema baseado em códigos corretores de erros, sugerido por McEliece e baseado em códigos de Goppa binários irredutíveis. Durante o texto é abordado conceitos básicos de Álgebra e suas aplicações nos Códigos Corretores de Erros e na Criptografia. Algumas estruturas algébricas são apresentadas no primeiro capítulo e utilizadas posteriormente para enriquecer o estudo dos códigos lineares e de criptosistemas, simplificando a codificação e decodificação. Apontamos também a importância de trabalhar estes temas na educação básica, mostrando aos alunos que sua grade curricular tem aplicação prática e relacionada ao seu dia-a-dia.

Palavras chave: Códigos Corretores de Erros. Criptografia. Álgebra. Educação Básica.

ABSTRACT

The main objective of this dissertation is to present the first cryptosystem based on error-correcting codes, suggested by McEliece and based on irreducible binary Goppa codes. During the text the basic concepts of Algebra and its applications in Error Correcting Codes and Cryptography are discussed. Some algebraic structures are presented in the first chapter and later used to enrich the study of linear codes and cryptosystems, simplifying coding and decoding. We also point out the importance of working on these themes in basic education, showing students that their curriculum has practical application and related to their daily life.

Keywords: Error Correcting Codes. Cryptography. Algebra. Basic Education.

Sumário

Introdução

Pense em sua vida moderna onde satélites e computadores estão comunicando entre si 24 horas por dia, o telefone sinaliza a quase todo minuto que uma mensagem chegou no Whatsapp e sempre tem alguma coisa legal que foi compartilhada no Facebook! Sem o surgimento da Teoria dos Códigos Corretores de Erros nada disto funcionaria bem e sem a Álgebra não existiria a Teoria dos Códigos Corretores de Erros. Mesmo assim, pouquíssimas pessoas sabem o que é um Código Corretor de Erros e quando comparando com outros campos de estudo que envolvem informação o mesmo é pouco difundido.

Imagine agora um cenário onde qualquer mensagem que você envia é lida por qualquer um que consiga interceptar a transmissão, além disto, este interceptor poderia modificar o conteúdo de sua mensagem ou até mesmo enviar outras mensagens como se fosse você. Seu e-mail não seria mais só seu, fazer compras ou olhar o saldo bancário precisaria ser presencialmente, seria mais fácil clonar um cartão de crédito, seus vizinhos saberiam a senha da sua internet, entre outras complicações. Os objetivos principais da Criptografia é garantir confidencialidade, integridade, autenticação e irretratabilidade, evitando que um cenário assim ocorra. Os criptosistemas que mais utilizamos atualmente dependem da dificuldade de computar logaritmos discretos e fatorar grandes números, porém quando computadores quânticos deixarem de fazer parte de um modelo teórico e com o auxílio da física quântica simplificar estes problemas haverá uma ameaça aos criptosistemas atuais. Fazendo assim com que criptólogos pesquisem avanços em métodos não-quânticos que sejam resistentes a ataques de algoritmos quânticos.

Relacionando ambos os campos de pesquisa(Teoria dos Códigos Corretores de Erros e Criptografia) o matemático e engenheiro McEliece em 1978 publicou o artigo [8] apresentando à primeira proposta de um criptosistema baseado em códigos corretores de erros sugerindo a utilização de um código de Goppa binário aleatório. Este criptosistema é considerado um método não-quântico que resiste a ataques de algoritmos quânticos, pois até o presente momento não existe algoritmo que leva pouco tempo para descifrá-lo.

Podemos ver que a Teoria dos Códigos Corretores de Erros e a Criptografia afetam de forma relevante a humanidade e possivelmente continuaram afetando no futuro, porém é fácil constatar que estas áreas de pesquisa não são divulgadas durante o ensino básico, mesmo estando relacionada a conteúdos que fazem parte da Proposta Curricular.

Por outro lado os Parâmetros Curriculares Nacionais do Ensino Médio nos informa que a aprendizagem das Ciências da Natureza, Matemática e suas Tecnologias no Ensino Médio ganha níveis de generalidade e deixando de ser conceitos espontâneos passando a serem entendidos e formulados de modo abstrato, levando os alunos a se perguntarem "por que estudar?", "pra que serve?" e "onde se encontra?", em busca de uma transição do abstrato para o concreto e particular.

O primeiro capítulo desse texto, traz teorias sobre estruturas algébricas básicas que são utilizadas no estudo dos códigos corretores de erros e em alguns sistemas criptográficos.

No segundo capítulo apresentamos os códigos corretores de erros, este capítulo é dividido em três seções onde a primeira é reservada aos conceitos básicos da teoria dos códigos corretores de erros. Na segunda seção apresentamos a classe mais utilizada na prática dos códigos corretores de erros, os códigos lineares. Já na terceira seção apresentamos uma subclasse dos códigos lineares chamada de códigos cíclicos que possuem famílias de códigos com bons algoritmos de codificação e decodificação, uma família especifica de códigos cíclicos chamada de Códigos de Goppa é apresentada no final da seção junto com o algoritmo de Patterson eficiente em decodificação.

No terceiro capítulo apresentamos sobre criptografia, este capítulo é dividido em duas seções onde a primeira manifesta o objetivo principal da criptografia, define criptosistema e esclarece as diferenças entre chaves públicas e simétricas. Na segunda seção apresentamos o primeiro criptosistema baseado em códigos lineares, o criptosistema de McEliece.

No quarto capítulo é oferecido propostas que apresentam o conceito de Criptografia e de Códigos Corretores de Erros à alunos na educação básica.

Capítulo 1

Estruturas Algébricas

Algumas teorias matemáticas sobre estruturas algébricas básicas são essências para o estudo dos códigos corretores de erros. Estas estruturas são modelos abstratos que tratam de várias situações matemáticas, cada estrutura possui diversas propriedades que enriquecem os códigos os tornando muito mais eficientes computacionalmente.

Definição 1.0.1. *(Anel). Um anel ou anel comutativo $(A, +, \cdot)$ é um conjunto A com pelo menos dois elementos, munido de duas operações chamadas de adição $+$*

$$\begin{array}{rccc} + : & A \times A & \to & A \\ & (x, y) & \mapsto & x + y \end{array}$$

e multiplicação $\cdot$

$$\begin{array}{rccc} \cdot : & A \times A & \to & A \\ & (x, y) & \mapsto & x \cdot y \end{array}$$

que possuem as seguintes propriedades:

A1) *Associatividade da adição:*

$$\forall x, y, z \in A, \qquad (x + y) + z = x + (y + z);$$

A2) *Existência de elemento neutro para adição: Existe um elemento no anel A chamado zero e denotado por 0, tal que*

$$\forall x \in A, \qquad x + 0 = 0 + x = x;$$

A3) *Existência de elemento inverso para adição: Para todo elemento $x \in A$, existe um elemento também pertencente A chamado simétrico de x e denotado por $-x$, tal que*

$$x + (-x) = -x + x = 0;$$

A4) *Comutatividade da adição:*

$$\forall x, y \in A, \qquad x + y = y + x;$$

M1) *Associatividade da multiplicação:*

$$\forall x, y, z \in A, \qquad (x \cdot y) \cdot z = x \cdot (y \cdot z);$$

M2) *Existência de elemento neutro para a multiplicação: Existe um elemento no anel A chamado unidade e denotado por 1, tal que*

$$\forall x \in A, \qquad x \cdot 1 = 1 \cdot x = x;$$

M3) *Comutatividade da multiplicação:*

$$\forall x, y \in A, \qquad x \cdot y = y \cdot x;$$

AM) *Distributividade da multiplicação com relação à adição:*

$$\forall x, y, z \in A, \qquad x \cdot (y + z) = x \cdot y + x \cdot z.$$

Exemplo 1.0.2. $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ *e* $(\mathbb{C}, +, \cdot)$ *são anéis.*
$(\mathbb{N}, +, \cdot)$ *não forma um anel, pois não satisfaz as propriedades A2) e A3).*

Definição 1.0.3. *(Domínio). Um anel $(D, +, \cdot)$ é chamado de domínio ou domínio de integridade se possuir a seguinte propriedade:*

M4). *O produto de quaisquer dois elementos não-nulos de D é um elemento não-nulo de D, ou seja:*

$$\forall x, y \in D \setminus \{0\}, \qquad x \cdot y \neq 0.$$

Definição 1.0.4. *(Corpo). Um anel $(K, +, \cdot)$ é chamado corpo, se possuir a seguinte propriedade:*

M4'). *Qualquer elemento diferente de 0 em K possui um inverso com respeito à multiplicação, ou seja:*

$$\forall x \in K \setminus \{0\}, \qquad \exists y \in K, \text{tal que } x \cdot y = 1.$$

Exemplo 1.0.5. $(\mathbb{Z}, +, \cdot)$ *é um domínio mas não é um corpo, pois com exceção dos*

números 1 *e* -1*, que possuem inversos*

$$1 \cdot 1 = 1 \ e \ -1 \cdot (-1) = 1,$$

seus demais elementos não possuem. Por outro lado $(\mathbb{Q}, +, \cdot)$*,* $(\mathbb{R}, +, \cdot)$ *e* $(\mathbb{C}, +, \cdot)$ *são corpos.*

Definição 1.0.6. *(Anéis dos inteiros módulo n). Definimos a relação* $x \equiv y \mod n$ *onde se diz* x *é côngruo a* y *módulo* n*, para* $x, y \in \mathbb{Z}$ *e* $n \in \mathbb{Z}^+$*, da seguinte maneira:*

$$x \equiv y \mod n \Leftrightarrow x - y \text{ é um multiplo de } n.$$

Esta é uma relação de equivalência, ou seja:

$$\begin{cases} x \equiv x \mod n \\ x \equiv y \mod n \Leftrightarrow y \equiv x \mod n \\ x \equiv y \mod n, y \equiv z \mod n \Rightarrow x \equiv z \mod n \end{cases}$$

Portanto, se $x \in \mathbb{Z}$*, sua classe de equivalência módulo* n *é o conjunto*

$$\{y \in \mathbb{Z} : y \equiv x \mod n\} = \{x + \lambda n : \lambda \in \mathbb{Z}\}$$

este subconjunto é denotado por $\overline{x}$*. Chamamos o conjunto das classes de equivalência módulo* n *de* $\mathbb{Z}_n$*, isto é* $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$*, e ao definirmos as operações:*

$$\begin{array}{rccc} \oplus : & \mathbb{Z}_n \times \mathbb{Z}_n & \to & \mathbb{Z}_n \\ & (\overline{x}, \overline{y}) & \mapsto & \overline{x + y} \end{array}$$

e

$$\begin{array}{rccc} \odot : & \mathbb{Z}_n \times \mathbb{Z}_n & \to & \mathbb{Z}_n \\ & (\overline{x}, \overline{y}) & \mapsto & \overline{x \cdot y} \end{array}$$

obtemos o conjunto finito $(\mathbb{Z}_n, \oplus, \odot)$ *chamado anel dos inteiros módulo* n*.*

Exemplo 1.0.7. *O conjunto* $GF(2) = \{\overline{0}, \overline{1}\}$ *munido das operações:*

$$\begin{array}{c|cc} \oplus & \overline{0} & \overline{1} \\ \hline \overline{0} & \overline{0} & \overline{1} \\ \overline{1} & \overline{1} & \overline{0} \end{array} \quad e \quad \begin{array}{c|cc} \odot & \overline{0} & \overline{1} \\ \hline \overline{0} & \overline{0} & \overline{0} \\ \overline{1} & \overline{0} & \overline{1} \end{array}$$

é um corpo finito, pois possui apenas dois elementos, é um anel ao qual seu único elemento não nulo $\overline{1}$ *é seu próprio inverso multiplicativo. Este é chamado de Corpo de Galois em homenagem ao matemático Evariste Galois.*

É usual omitirmos a barra "$\overline{\ }$" que denota a classe de equivalência, como também utilizar o simbolo $(+)$ para denotar a operação $(\oplus)$ e o simbolo $(\cdot)$ para denotar a

operação ($\odot$), quando isto não causa desordem no texto.

Proposição 1.0.8. $(\mathbb{Z}_n, +, \cdot)$ *anel dos inteiros módulo n é um corpo se, e somente se, n é primo.*

Demonstração. ($\Rightarrow$) Se $\mathbb{Z}_n$ é um corpo, então $\mathbb{Z}_n \setminus \{0\}$ é um grupo multiplicativo, ou seja, ele possui as propriedades $M1$ e $M2$, além disto todos os seus elementos possuem inverso multiplicativo. Portanto, o MDC$(a, n) = 1$, para todo $1 \leq a < n$, segue que não existe $a < n$ diferente de 1 que divida n. Logo, n é primo.

($\Leftarrow$) Se n é primo, então o MDC$(a, n) = 1$, para todo $1 \leq a < n$. Portanto, o conjunto dos elementos invertíveis de $\mathbb{Z}_n$ é $\mathbb{Z}_n \setminus \{0\}$. Logo, $\mathbb{Z}_n$ é um corpo finito. □

Se n é primo, então $\mathbb{Z}_n$ é um corpo. É usual utilizar a letra p no lugar do n em algumas ocasiões e denota-lo por $\mathbb{Z}_p$, $GF(p)$ ou $\mathbb{F}_p$.

Definição 1.0.9. *(Ideal). Sejam $(A, +, \cdot)$ um anel e I um subconjunto não-vazio de A. Chamamos o conjunto I de ideal de A se*

i) $\forall x, y \in I, \quad x + y \in I;$

ii) $\forall x \in I \ e \ \lambda \in A, \ x\lambda \in I.$

Se $x \in A$, chamamos o conjunto

$$I(x) = \{x\lambda : \lambda \in A\}$$

de ideal principal gerado por x. E de forma mais geral, se $\alpha_1, \ldots, \alpha_n \in A$, então o conjunto

$$I(\alpha_1, \ldots, \alpha_n) = \{a_1\alpha_1 + \ldots + a_n\alpha_n : a_1, \ldots, a_n \in A\}$$

é um ideal de A, onde os elementos $\alpha_1, \ldots, \alpha_n$ são chamados de geradores.

Exemplo 1.0.10. *Seja o anel $(\mathbb{Z}, +, \cdot)$, utilizando o número $2 \in \mathbb{Z}$ como gerador, obtemos o ideal principal*

$$I(2) = \{2\lambda : \lambda \in \mathbb{Z}\} = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$$

Um anel é chamado de principal, se todos os seus ideais são principais.

Definição 1.0.11. *(Polinômios). Dados um anel comutativo A e uma indeterminada X, o conjunto $A[X]$ é formado por todas as expressões*

$$F(X) = \sum_{i=0}^{n} a_i X^i = a_0 + a_1 X + \cdots + a_n X^n$$

com $n \in \mathbb{N}$ e $a_i \in A$, a_i é denominado coeficiente de X^i em $F(X)$. As partes $a_i X^i$ são chamadas de monômio e as expressões $F(X) \in A[X]$ são chamadas polinômios.

Definição 1.0.12. *(Igualdade de polinômios). Sejam* $F(X) = a_0 + a_1X + \cdots + a_nX^n$ *e* $G(X) = b_1 + b_2X + \cdots + b_mX^m$*, dois polinômios, ou seja,* $F(X), G(X) \in A[X]$*. Dizemos que* $F(X) = G(X)$ *se, somente se,* $a_i = b_i$*, para todo* i*.*

Definição 1.0.13. *(Grau de um polinômio). Seja* $F(X) \in A[X]$*, tal que* $F(X) = \sum_{i=0}^{n} a_iX^i$*, com* $n \in \mathbb{N}$ *e* $a_n \neq 0$*. Dizemos que o grau do polinômio* $F(X)$ *é* n*, ou simplesmente*

$$gr(F(X)) = n.$$

Se $a_n = 1$*, diremos que* $F(X)$ *é um polinômio mônico.*

Definição 1.0.14. *(Anéis de Polinômios). No conjunto* $A[X]$*, dos polinômios numa variável sobre* A*, definimos as operações de soma e produto*

$$\begin{array}{rccc} \oplus: & A[X] \times A[X] & \to & A[X] \\ & (a_0 + a_1X + \cdots + a_nX^n, b_0 + b_1X + \cdots + b_mX^m) & \mapsto & \sum_{i=0}^{max\{n,m\}}(a_i + b_i)X^i \end{array}$$

e

$$\begin{array}{rccc} \odot: & A[X] \times A[X] & \to & A[X] \\ & (a_0 + a_1X + \cdots + a_nX^n, b_0 + b_1X + \cdots + b_mX^m) & \mapsto & \sum_{i=0}^{n+m} c_iX^i \end{array}$$

onde

$$\begin{array}{lcl} c_0 & = & a_0 \cdot b_0 \\ c_1 & = & a_0 \cdot b_1 + a_1 \cdot b_0 \\ \cdots & & \\ c_i & = & \sum_{j=0}^{i} a_j \cdot b_{i-j} = a_0 \cdot b_i + a_1 \cdot b_{i-1} + \cdots + a_i \cdot b_0 \\ \cdots & & \\ c_{n+m} & = & a_n \cdot b_m. \end{array}$$

O algoritmo da divisão entre polinômios com coeficientes no corpo $(\mathbb{R}, +, \cdot)$ é normalmente lecionado no ensino básico, podemos fazer exatamente igual em polinômios com coeficientes num corpo K arbitrário.

Definição 1.0.15. *(Polinômio Irredutível). Seja* $F(X) \in K[X]$*, um polinômio com coeficientes no corpo* K*. Chamamos* $F(X)$ *de irredutível sobre* K*, se para todo* $Q_1(X), Q_2(X) \in K[X]$ *tais que* $F(X) = Q_1(X)Q_2(X)$*, tenhamos que* Q_1 *ou* Q_2 *seja constante. Ou seja,* $F(X)$ *não pode ser fatorado em polinômios de grau menores.*

Exemplo 1.0.16. *O polinômio* $X^2 + X + 1$ *pode ser fatorado sobre* $\mathbb{F}_3$*, pois*

$$X^2 + X + 1 = (X + 2) \cdot (X + 2)$$

no corpo $\mathbb{F}_3$*. Dizemos então que* $X^2 + X + 1 \in \mathbb{F}_3[X]$ *não é irredutível. Por outro lado* $X^2 + X + 1$ *é irredutível em* $\mathbb{R}$*, pois não existem* $a, b \in \mathbb{R}$*, tais que*

$$(X + a) \cdot (X + b) = X^2 + X + 1.$$

Ou seja, o sistema $\begin{cases} 2ab = 1 \\ ab = 1 \end{cases}$ *, não possui solução em* $\mathbb{R}$*. Dizemos então que* $X^2+X+1 \in \mathbb{R}$ *é irredutível.*

Definição 1.0.17. *Chamamos de* $K[X]/F(X)$ *a classe residual de* $K[X]$ *módulo* $F(X)$*, onde* $F(X)$ *é um polinômio não constante.*

Exemplo 1.0.18. *Seja* $K = \mathbb{F}_2$ *e* $F(X) = X^3 + X + 1$*, então o conjunto*

$$\mathbb{F}_2[X]/X^3 + X + 1 = \{\overline{0}, \overline{1}, \overline{X}, \overline{X+1}, \overline{X^2}.\overline{X^2+1}, \overline{X^2+X}, \overline{X^2+X+1}\}$$

é o conjunto dos polinômios $\mathbb{F}_2[X]$ *munido das operações de adição e multiplicação módulo* X^3+X+1*, onde estão definidas as operações de soma e multiplicação da seguinte maneira:*

$\oplus$	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
0	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
1	1	0	$x+1$	x	x^2+1	x^2	x^2+x+1	x^2+x
x	x	$x+1$	0	1	x^2+x	x^2+x+1	x^2	x^2+1
$x+1$	$x+1$	x	1	0	x^2+x+1	x^2+x	x^2+1	x^2
x^2	x^2	x^2+1	x^2+x	x^2+x+1	0	1	x	$x+1$
x^2+1	x^2+1	x^2	x^2+x+1	x^2+x	1	0	$x+1$	1
x^2+x	x^2+x	x^2+x+1	x^2	x^2+1	x	$x+1$	0	1
x^2+x+1	x^2+x+1	x^2+x	x^2+1	x^2	$x+1$	x	1	0

e

$\odot$	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
0	0	0	0	0	0	0	0	0
1	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
x	0	x	x^2	x^2+x	$x+1$	1	x^2+x+1	x^2+1
$x+1$	0	$x+1$	x^2+x	x^2+1	x^2+x+1	x^2	1	x
x^2	0	x^2	$x+1$	x^2+x+1	x^2+x	x	x^2+1	1
x^2+1	0	x^2+1	1	x^2	x	x^2+x+1	$x+1$	x^2+x
x^2+x	0	x^2+x	x^2+x+1	1	x^2+1	$x+1$	x	x^2
x^2+x+1	0	x^2+x+1	x^2+1	x	1	x^2+x	x^2	$x+1$

Obs: *Foi omitido a barra "−" que representaria a classe de equivalência módulo* X^3+X+1*.*

A seguinte proposição é consequência do algoritmo de divisão em $K[X]$, pois dado $F(X), G(X) \in K[X]$ com $G(X) \neq 0$, existem únicos polinômios $Q(X), R(X) \in K[X]$ tais que $F(X) = Q(X)G(X) + R(X)$, com $R(X) = 0$ ou $gr(R(X)) < gr(G(X))$.

Proposição 1.0.19. *Seja* $F(X)$ *é um polinômio de grau* n*. O quociente* $K[X]/F(X)$ *pode ser identificado como* $\{Q(X) \in K[X] \mid gr(Q(X)) < n\}$ *com adição e multiplicação módulo* $F(X)$*.*

Proposição 1.0.20. *$K[X]/F(X)$ é um corpo se, somente se, $F(X)$ é irredutível sobre o corpo K.*

Demonstração. ($\Rightarrow$) Seja $K[X]/F(X)$ um corpo e suponhamos que $F(X)$ não é irredutível sobre o corpo K, então existe $Q_1(X), Q_2(X) \in K[X]$ ambos não constantes, tais que $F(X) = Q_1(X)Q_2(X)$. O produto de $\overline{Q_1(X)} = \{Q_1(X) + \lambda(X)F(X) : \lambda(X) \in K[X]\}$ e $\overline{Q_2(X)} = \{Q_2(X) + \lambda(X)F(X) : \lambda(X) \in K[X]\}$ é $\overline{Q_1(X)Q_2(X)} = \{Q_1(X)Q_2(X) + \lambda(X)F(X) : \lambda(X) \in K[X]\} = \{F(X)+\lambda(X)F(X) : \lambda(X) \in K[X]\} = \{F(X)(1+\lambda(X)) : \lambda(X) \in K[X]\} = \overline{0}$, temos portanto uma contradição, pois $\overline{Q_1(X)}$ e $\overline{Q_2(X)}$ são divisores de 0 no corpo $K[X]/F(X)$. Logo $F(X)$ é irredutível sobre o corpo K.

($\Leftarrow$) Seja $F(X)$ é um polinômio irredutível de grau n, então para todo $Q(X) \in K[X]\setminus\{0\}$ com $gr(Q(X))$ menor do que n, MDC$(Q(X), F(X)) = 1$. Portanto, o conjunto dos elementos invertíveis de $K[X]/F(X)$ é $(K[X]/F(X)) \setminus \{0\}$. Logo, $K[X]/F(X)$ é um corpo. □

Definição 1.0.21. *(Raiz Primitiva). Um elemento α de um corpo finito $\mathbb{F}_q$, cujas potências enumeram todos os elementos de $\mathbb{F}_q \setminus \{0\}$ é chamado de elemento primitivo ou uma raiz primitiva.*

$$\mathbb{F}_q \setminus \{0\} = \{\alpha^0, \alpha^1, \ldots, \alpha^{q-2}\}.$$

Definição 1.0.22. *(Derivada formal) Dado um corpo K e um polinômio $F(X) \in K[X]$, onde $F(X) = \sum_{i=0}^{n} a_i X^i$, chamamos de derivada formal e denotamos $F'(X)$, o polinômio*

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

Definição 1.0.23. *(Produto Interno). Seja V um espaço vetorial sobre um corpo K. Um produto interno sobre V é uma função que a cada par de vetores $u = (u_1, \ldots, u_n)$ e $v = (v_1, \ldots, v_n)$ em V associa um escalar em K, definido como*

$$\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i.$$

Observação 1.0.24. $\langle v, v\rangle = 0 \Rightarrow v = 0$, *vale apenas para corpos de característica zero onde qualquer soma do elemento neutro multiplicativo com ele mesmo não tem como resultado o elemento neutro aditivo.*

Capítulo 2

Códigos Corretores de Erros

Na nossa sociedade atual o uso de dispositivos eletrônicos está presente em momentos de descansos e encargos da maior parte da população. Estes dispositivos requerem sistemas onde a informação é transmitida e processada de forma digital, nestes sistemas existem possibilidades de ocorrer interferências eletromagnéticas ou erros humanos, provocando problemas na transmissão ou leitura dos dados que representam a informação. Para evitar que estes erros, que normalmente são chamados de ruídos, prejudiquem o uso dos dispositivos eletrônicos, estudos são feitos para desenvolver métodos que permitam detectá-los e corrigi-los. Este capítulo é destinado a apresentar a Teoria dos Códigos Corretores de Erros, o campo de estudo que visa desenvolver estes métodos de detecção e correção de erros.

2.1 Conceitos Básicos

Esta seção é dedicada à apresentação dos conceitos básicos da teoria dos códigos corretores de erros, servindo assim como uma introdução ao assunto. Nela é definido o que são códigos corretores de erros, sistemas de comunicação, a métrica de Hamming, os parâmetros de um código, a equivalência de códigos e os principais problemas e objetivos da teoria dos códigos.

Definição 2.1.1. *(Códigos). É um conjunto de palavras utilizadas para emissão e recepção de mensagens.*

Exemplo 2.1.2. *Os mais comuns dos códigos utilizados pelos seres humanos são os idiomas, além dos idiomas os números de contas bancárias e de CPF são exemplos de códigos importantes em nossa vida moderna.*

Definição 2.1.3. *(Alfabeto). É um conjunto que possui uma quantidade finita de elementos.*

Exemplo 2.1.4. *O conjunto que representa o alfabeto utilizado para escrever um código na língua portuguesa é formado pelas 26 letras do alfabeto da língua portuguesa o ç e as vogais acentuadas, todos sendo considerados duas vezes, pois podem ser escritos de forma maiúscula e minusculas, além do espaço em branco. Os corpos finitos $\mathbb{F}_q$, onde $q = p^k$ com p um número primo e k um número natural serão os alfabetos mais utilizados neste trabalho.*

Observação 2.1.5. *Quando escrevemos um código utilizando um corpo finito $\mathbb{F}_q$ como alfabeto, dizemos que o código é $q-$ário.*

Exemplo 2.1.6. *Códigos escrito utilizando o corpo finito $\mathbb{F}_2 = \{\overline{0}, \overline{1}\}$ são chamado de códigos binários.*

Definição 2.1.7. *Chamamos de palavras sequências finitas de símbolos de um alfabeto.*

Exemplo 2.1.8. *"Matemática" e "Educação" fazem parte da língua portuguesa e do código que é nosso idioma. Utilizando o corpo finito $\mathbb{F}_2 = \{\overline{0}, \overline{1}\}$ como alfabeto, podemos escrever palavras como* 00001*,* 111110*, etc.*

Observação 2.1.9. *A quantidade de elementos ou coordenadas que uma palavra possui é chamada de comprimento.*

Exemplo 2.1.10. *"Matemática" possui dez letras e portanto seu comprimento é dez, já a palavra "Educação" possui oito letras e seu comprimento é oito.*

Definição 2.1.11. *(Código Trivial). Chamamos um código $\mathcal{C}$ de código trivial se $\mathcal{C}$ possui apenas uma palavra, ou seja $|\mathcal{C}| = 1$.*

Observação 2.1.12. *Quando um código $\mathcal{C}$ é escrito utilizando um certo alfabeto A e suas palavras possuem sempre um comprimento especifico n, este código $\mathcal{C}$ é um subconjunto de A^n:*

$$\mathcal{C} \subset A^n = A \times A \times \cdots \times A.$$

Definição 2.1.13. *(Sistema de Comunicação). Entidade que processa sinais na entrada e produz outros sinais na saída, transmitindo assim informação. Possui os seguintes elementos; Fonte, Emissor, Canal, Receptor e Destino.*

Fonte: Origina a mensagem que será enviada.

Emissor: Dispositivo que emite a mensagem para o receptor através de um sinal adequado ao canal utilizado.

Canal: Meio usado para transportar uma mensagem do emissor ao receptor.

Receptor: Dispositivo que recebe a mensagem enviada pelo emissor.

Destino: Local ou pessoa ao qual a mensagem foi destinada.

Exemplo 2.1.14. *Comunicações via satélite, ou internas de um computador.*

Definição 2.1.15. *(Código Fonte). Código obtido por uma conversão de todas as palavras possíveis que possam se originar da fonte.*

Exemplo 2.1.16. *Imagine um robô que se move em quatro direções (Norte, Sul, Leste ou Oeste). Podemos converter estes movimentos para elementos de* $\mathbb{F}_2^2$*,*

$$\begin{aligned} \textit{Norte} &\longmapsto 00 \\ \textit{Sul} &\longmapsto 01 \\ \textit{Leste} &\longmapsto 10 \\ \textit{Oeste} &\longmapsto 11 \end{aligned}$$

Obtendo o código fonte e binário $\mathcal{C} = \{00, 01, 10, 11\}$*, viável para ser utilizado no dispositivo eletrônico emissor dos comandos ao robô. Observe que este código não possui capacidade de corrigir erros, pois substituindo em qualquer palavra um elemento do alfabeto por outro a palavra toma outro significado.*

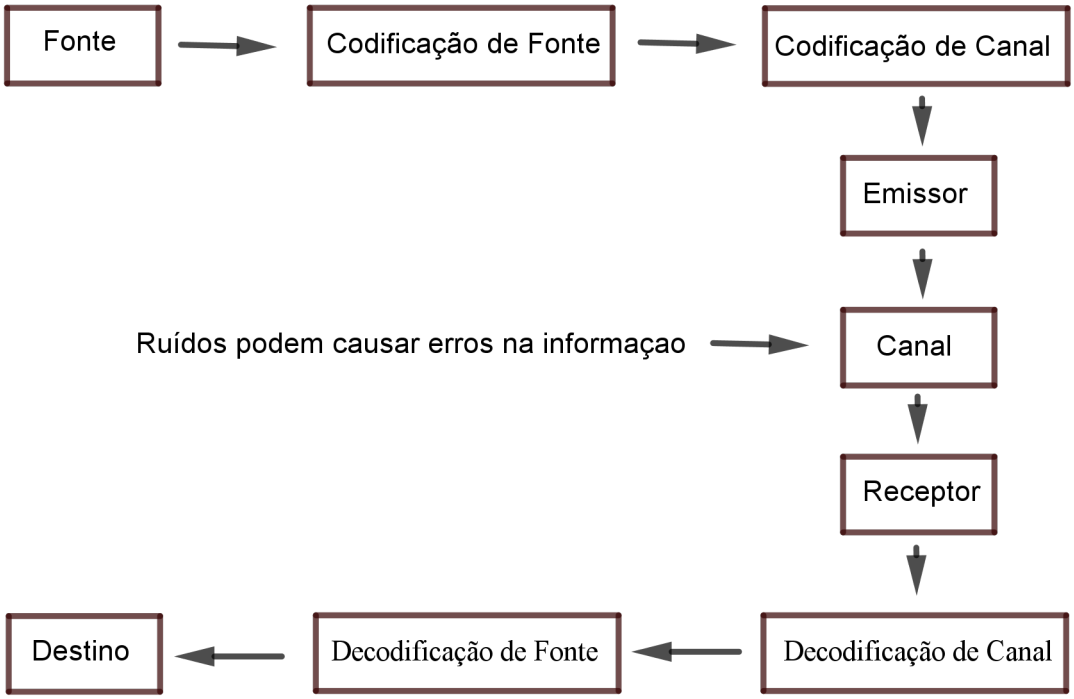
Definição 2.1.17. *(Código Canal). Código corretor de erros obtido por uma conversão das palavras do código fonte, as palavras do código canal tem redundâncias que tornam seu comprimento maior do que as palavras do código fonte e as fornecem capacidade de corrigir erros que são introduzidos quando a mensagem passa pelo canal.*

Exemplo 2.1.18. *Dado o código fonte* $\mathcal{C} = \{00, 01, 10, 11\}$*, podemos através de uma tripla repetição obter o código canal* $\mathcal{C}' = \{000000, 010101, 101010, 111111\}$*. Ao qual possui 3 vezes a informação, logo quando um erro ocorre a informação poderá ser corrigida separando as 3 informações recebidas e considerando como verdadeira a que se repete duas vezes, como se segue:*

Palavra recebida	*Verificação*	*Informação Repetida*	*Palavra corrigida*
000001	00 00 01	00	000000

A teoria dos códigos corretores de erros consiste em estudar meios para transformar o código fonte em código canal, de forma que dado o canal de transporte, a informação recebida com erro possa ser corrigida e o código canal ser decodificado em código da fonte.

Figura 2.1: Diagrama de um sistema de comunicação



Definição 2.1.19. *(Métrica de Hamming) Dado um alfabeto A e os elementos*

$$x = (x_1, x_2, \cdots, x_n) \text{ e } y = (y_1, y_2, \cdots, y_n)$$

pertencentes ao espaço A^n, chama-se distância de Hamming de x a y o número de coordenadas em que estes elementos se diferem, ou seja:

$$\begin{array}{rccl} d_h: & A^n \times A^n & \to & \mathbb{R}^+ \\ & (x, y) & \mapsto & |\{i : x_i \neq y_i, 1 \leq i \leq n\}|. \end{array}$$

Exemplo 2.1.20. *No espaço A^4 das palavras em português que possuem quatro letras, temos que: $d_h(gato, bota) = 3$, $d_h(gato, rato) = 1$, $d_h(bota, rato) = 3$, $d_h(dedo, gato) = 3, d_h(dedo, bota) = 4, d_h(dedo, rato) = 3$. No espaço $\mathbb{F}_2^3$: $d_h(001, 000) = 1$, $d_h(111, 000) = 3$, $d_h(101, 110) = 2$, $d_h(011, 011) = 0$.*

Proposição 2.1.21. *A métrica de Hamming é uma função distância.*

Demonstração. **(i) Positividade:** A cardinalidade de um conjunto é sempre um número maior do que ou igual a zero. Além disto:

$$d_h(x, y) = 0 \Leftrightarrow x_1 = y_1, x_2 = y_2, \ldots, x_n = y_n \Leftrightarrow x = y.$$

Logo $d_h(x, y) \geq 0$ e $d_h(x, y) = 0 \Leftrightarrow x = y$.

(ii) Simetria: Seque de imediato, pois para todo $i \in \{1, 2, \cdots, n\}$:

$$x_i = y_i \Leftrightarrow y_i = x_i$$
$$x_i \neq y_i \Leftrightarrow y_i \neq x_i$$

Logo $d_h(x, y) = d_h(y, x)$.

(iii) Desigualdade Triangular: Analisemos as contribuições de uma coordenada i qualquer, nas distâncias $d_h(x, y)$ e $d_h(x, z) + d_h(z, y)$:
Se $x_i = y_i$, $x_i = z_i$ e $z_i = y_i$, temos que $0 \leq 0 + 0$;
Se $x_i = y_i$, $x_i \neq z_i$ e $z_i \neq y_i$, temos que $0 \leq 1 + 1$;
Se $x_i \neq y_i$, $x_i = z_i$ e $z_i \neq y_i$, temos que $1 \leq 0 + 1$;
Se $x_i \neq y_i$, $x_i \neq z_i$ e $z_i = y_i$, temos que $1 \leq 1 + 0$;
Se $x_i \neq y_i$, $x_i \neq z_i$ e $z_i \neq y_i$, temos que $1 \leq 1 + 1$.
Como a desigualdade é mantida na contribuição de qualquer coordenada, ela também será mantida na soma das contribuições de todas as coordenadas. Logo

$$d_h(x, y) \leq d_h(x, z) + d_h(z, y).$$

□

Definição 2.1.22. *(Distância mínima). Dado um código $\mathcal{C}$ qualquer, sua distância mínima $d(\mathcal{C})$ é o menor valor da métrica de Hamming entre duas palavras diferentes deste código, ou seja:*

$$d(\mathcal{C}) = min\{d_h(x, y) | x, y \in \mathcal{C}, x \neq y\}.$$

Exemplo 2.1.23. *No código binário $\mathcal{C}_1 = \{000, 011, 110\}$:*

$$d_h(000, 011) = 2,\ d_h(000, 110) = 2 \text{ e } d_h(011, 110) = 2.$$

Portanto $d(\mathcal{C}_1) = 2$.
No código binário $\mathcal{C}_2 = \{0000, 1111, 1010, 0101, 0001\}$:

$$\begin{aligned} &d_h(0000, 1111) = 4, \quad d_h(0000, 1010) = 2, \\ &d_h(0000, 0101) = 2, \quad d_h(0000, 0001) = 1, \\ &d_h(1111, 1010) = 2, \quad d_h(1111, 0101) = 2, \\ &d_h(1111, 0001) = 1, \quad d_h(1010, 0101) = 4, \\ &d_h(1010, 0001) = 3, \quad d_h(0101, 0001) = 1. \end{aligned}$$

Existem quatro distâncias de Hamming diferentes entre as palavras deste código porém a menor delas é 1, portanto $d(\mathcal{C}_2) = 1$.

Definição 2.1.24. *(Esfera e Bola fechada em A^n). Dados um elemento $x = (x_1, x_2, \cdots, x_n)$ pertencente ao espaço A^n e um natural r, chama-se esfera de centro x e raio r em A^n, o*

conjunto:

$$S(x,r) = \{y \in A^n : d_h(y,x) = r\}$$

e a bola fechada de centro x e raio r em A^n, o conjunto:

$$B[x,r] = \{y \in A^n : d_h(y,x) \leq r\}.$$

Exemplo 2.1.25. *Em $\mathbb{F}_2^2 = \{00, 01, 10, 11\}$:*

$$\begin{array}{ll} S(00,1) = \{01,10\}, & B[01,1] = \{00,01,11\}, \\ S(11,2) = \{00\}, & B[10,2] = \{00,01,10,11\}. \end{array}$$

Lema 2.1.26. *Sejam $a, b, q \in \mathbb{N}$, tal que $a = bq + r$, com $0 \leq r < b$. Então*

$$b\left\lfloor \frac{a}{b} \right\rfloor = a - r$$

onde $\lfloor x \rfloor$, representa uma função piso definida como:

$$\begin{array}{rl} f: & \mathbb{R} \to \mathbb{Z} \\ & x \mapsto max\{y \in \mathbb{Z} | y \leq x\}. \end{array}$$

Demonstração. Se a é múltiplo de b então $r = 0$ e $a = bq$, portanto

$$b\left\lfloor \frac{a}{b} \right\rfloor = b\lfloor q \rfloor = bq = a.$$

Se a não é múltiplo de b então $0 < r < b \Rightarrow 0 < \frac{r}{b} < 1$ e $bq = a - r$, portanto

$$b\left\lfloor \frac{a}{b} \right\rfloor = b\left\lfloor q + \frac{r}{b} \right\rfloor = bq = a - r.$$

□

Lema 2.1.27. *Dados um código $\mathcal{C}$ e as duas palavras x e y distintas e pertencentes ao código, então:*

$$B\left[x, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right] \bigcap B\left[y, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right] = \emptyset.$$

Demonstração. Suponhamos por absurdo que existe

$$z \in B\left[x, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right] \bigcap B\left[y, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right]$$

então pela definição de bola fechada $d_h(z,x) \leq \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$ e $d_h(z,y) \leq \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$, enquanto que pela definição de distância mínima $d(\mathcal{C}) \leqslant d_h(x,y)$, assim pelo fato da métrica de Hamming ser uma função distância e pelo lema anterior, temos que

$$d(\mathcal{C}) \leqslant d_h(x,y) \leq d_h(x,z) + d_(z,y) = d_h(z,x) + d_h(z,y) \leq 2\left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor \leq d(\mathcal{C}) - 1.$$

Que é impossível. Logo $B\left[x, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right] \bigcap B\left[y, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right] = \emptyset$. □

Teorema 2.1.28. *Um código $\mathcal{C}$ pode corrigir até $\left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$ erros e detectar $d(\mathcal{C})-1$ erros.*

Demonstração. Suponhamos que uma palavra a do código $\mathcal{C}$ foi transmitida e a palavra b com uma quantidade de erros $c \leqslant \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$ foi recebida. Segue que

$$d_h(a,b) = c \Leftrightarrow d_h(a,b) \leqslant \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor.$$

Enquanto que, o lema anterior nos informa que não existe outra palavra do código contida na bola fechada de centro a e raio $\left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$, sendo assim a única bola fechada com centro em uma palavra do código $\mathcal{C}$ e raio $\left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$ que contem b é

$$B\left[a, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right].$$

Logo determinamos de forma univoca que a palavra transmitida foi a.
Por outro lado, se a palavra b difere no mínimo uma e no máximo $d(\mathcal{C})-1$ posições da palavra enviada, então $1 \leqslant d_h(a,b) \leqslant d(\mathcal{C})-1$, enquanto que, pela definição de distância mínima se p é também uma palavra do código $d_h(a,p) \geqslant d(\mathcal{C})$, pelo fato da métrica de Hamming ser uma função distância com desigualdade triangular:

$$d_h(a,p) \leqslant d_h(a,b) + d_h(b,p) \Leftrightarrow d_h(b,p) \geqslant d_h(a,p) - d_h(a,b) \geqslant d(\mathcal{C}) - (d(\mathcal{C})-1) = 1.$$

Logo a palavra recebida será diferente de todas as outras palavras do código $\mathcal{C}$, possibilitando a detecção de erros. □

Definição 2.1.29. *(Código Perfeito). Um código $\mathcal{C} \subset A^n$ é dito $\left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$ - perfeito se*

$$\bigcup_{x \in \mathcal{C}} B\left[x, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right] = A^n$$

Exemplo 2.1.30. *O código binário $\mathcal{C}_1 = \{000, 110, 011\} \subset \mathbb{F}_2^3$, possui, como já visto anteriormente, distância mínima $d(\mathcal{C}_1) = 2$, ou seja $\left\lfloor \frac{d(\mathcal{C}_1)-1}{2} \right\rfloor = \left\lfloor \frac{2-1}{2} \right\rfloor = \left\lfloor \frac{1}{2} \right\rfloor = 0$. E portanto*

$$\begin{aligned} \bigcup_{x \in \mathcal{C}_1} B\left[x, \left\lfloor \frac{d(\mathcal{C}_1)-1}{2} \right\rfloor\right] &= \bigcup_{x \in \mathcal{C}_1} B[x,0] \\ &= B[000,0] \bigcup B[110,0] \bigcup B[011,0] \\ &= \{000, 110, 011\} \neq \mathbb{F}_2^3. \end{aligned}$$

Logo este código não é perfeito.

Já o código binário $\mathcal{C}_2 = \{000, 111\} \subset \mathbb{F}_2^3$*, possui distância mínima* $d(\mathcal{C}_2) = d_h(000, 111) = 3$*, ou seja* $\left\lfloor \frac{d(\mathcal{C}_2)-1}{2} \right\rfloor = \left\lfloor \frac{3-1}{2} \right\rfloor = \lfloor 1 \rfloor = 1$*. E portanto*

$$\begin{aligned} \bigcup_{x \in \mathcal{C}_2} B\left[x, \left\lfloor \tfrac{d(\mathcal{C}_2)-1}{2} \right\rfloor\right] &= \bigcup_{x \in \mathcal{C}_2} B[x, 1] \\ &= B[000, 1] \bigcup B[111, 1] \\ &= \{000, 001, 010, 100\} \bigcup \{111, 110, 101, 011\} \\ &= \{000, 001, 010, 100, 111, 110, 101, 011\} \\ &= \mathbb{F}_2^3. \end{aligned}$$

Logo este código é 1-perfeito.

Definição 2.1.31. *(Isometria de Hamming). Sejam o alfabeto* A *e* $n \in \mathbb{N}$*. Uma função* $\varphi : A^n \to A^n$ *é chamada de uma isometria de Hamming ou isometria de* A^n *se ela preservar distâncias de Hamming em* A^n

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

Exemplo 2.1.32. *Sejam* $\mathbb{F}_2^2 = \{00, 01, 10, 11\}$ *e a função*

$$\begin{array}{rccc} \varphi: & \mathbb{F}_2^2 & \to & \mathbb{F}_2^2 \\ & x & \mapsto & x + 01 \end{array}$$

segue que

$$\begin{aligned} d_h(\varphi(00), \varphi(01)) &= d_h(01, 00) = 1 = d_h(00, 01); \\ d_h(\varphi(00), \varphi(01)) &= d_h(01, 00) = 1 = d_h(00, 01); \\ d_h(\varphi(00), \varphi(10)) &= d_h(01, 11) = 1 = d_h(00, 10); \\ d_h(\varphi(00), \varphi(11)) &= d_h(01, 10) = 2 = d_h(00, 11); \\ d_h(\varphi(01), \varphi(10)) &= d_h(00, 11) = 2 = d_h(01, 10); \\ d_h(\varphi(01), \varphi(11)) &= d_h(00, 10) = 1 = d_h(01, 11); \\ d_h(\varphi(10), \varphi(11)) &= d_h(11, 10) = 1 = d_h(10, 11). \end{aligned}$$

Logo a função φ *é uma isometria de Hamming.*

Proposição 2.1.33. *Toda isometria de Hamming em* A^n *é uma bijeção.*

Demonstração. Seja $|A^n| = q$, como o domínio e contradomínio desta função é o mesmo espaço que possui q elementos, tem então mesma cardinalidade. Por outro lado, seja $x, y \in A^n$, tal que $\varphi(x) = \varphi(y)$, segue da positividade de uma função distância que $\varphi(x) = \varphi(y) \Leftrightarrow d_h(\varphi(x), \varphi(y)) = 0$, enquanto que, pelo fato de φ ser uma isometria de Hamming $d_h(\varphi(x), \varphi(y)) = d_h(x, y)$, e portanto $d_h(x, y) = 0$, novamente pela positividade de uma função distância $d_h(x, y) = 0 \Leftrightarrow x = y$. Assim como $\varphi(x) = \varphi(y) \Rightarrow x = y$ a

função φ é injetiva, além disto a cardinalidade do seu domínio e contradomínio são iguais, portanto ela também é sobrejetiva. Logo φ é uma bijeção. □

Definição 2.1.34. *(Código Geometricamente Uniforme). Um código $\mathcal{C} \in A^n$ é chamado geometricamente uniforme se, e somente se, dadas duas palavras a e b do código, existe uma isometria de Hamming $\varphi : A^n \to A^n$ tal que:*

(i). $\varphi(\mathcal{C}) = \mathcal{C}$;

(ii). $\varphi(a) = b$.

Exemplo 2.1.35. *Dados o código $\mathcal{C} = \{00, 01, 10, 11\} \in \mathbb{F}_2^2$ e as duas palavras $a = (00)$ e $b = (11)$. Da isometria de Hamming $\varphi : \mathbb{F}_2^2 \to \mathbb{F}_2^2$, definida por $\varphi(x) = x + 11$ com $x \in \mathbb{F}_2^2$, temos que*

$$\varphi(00) = 11, \varphi(01) = 10, \varphi(10) = 01 \text{ e } \varphi(11) = 00.$$

Logo o código $\mathcal{C}$ é um código geometricamente uniforme.

Proposição 2.1.36. *A função identidade $Id : A^n \to A^n$ é uma isometria de Hamming em A^n.*

Demonstração. Sejam $x, y \in A^n$ e uma função identidade $Id : A^n \to A^n$. Sabemos que $Id(x) = x$ e $Id(y) = y$, segue que $d_h(Id(x), Id(y)) = d_h(x, y)$. □

Proposição 2.1.37. *Se φ é uma isometria de Hamming em A^n, então φ^{-1} é uma isometria de Hamming em A^n.*

Demonstração. Sabemos que φ é uma bijeção, portanto existe a função inversa φ^{-1} de φ evidente que φ é inversa de φ^{-1}, segue pela definição de isometria de Hamming

$$d_h(\varphi^{-1}(a), \varphi^{-1}(b)) = d_h(\varphi(\varphi^{-1}(a)), \varphi(\varphi^{-1}(b))); \quad \forall\ a, b \in A^n$$

e pela definição de uma função inversa $\varphi(\varphi^{-1}(x)) = x$ para qualquer $x \in A^n$, portanto

$$d_h(\varphi(\varphi^{-1}(a)), \varphi(\varphi^{-1}(b))) = d_h(a, b); \quad \forall\ a, b \in A^n$$

juntando as duas igualdades:

$$d_h(\varphi^{-1}(a), \varphi^{-1}(b)) = d_h(a, b); \quad \forall\ a, b \in A^n.$$

Logo a inversa de φ é uma isometria de Hamming em A^n. □

Proposição 2.1.38. *Se φ_1 e φ_2 são isometrias de Hamming em A^n, então $\varphi_1 \circ \varphi_2$ é uma isometria Hamming em A^n.*

Demonstração. Pela definição de isometria de Hamming, temos que

$$d_h(\varphi_1(\varphi_2(a)), \varphi_1(\varphi_2(b))) = d_h(\varphi_2(a), \varphi_2(b)) = d_h(a, b); \quad \forall\ a, b \in A^n$$

portanto $\varphi_1 \circ \varphi_2$ é uma isometria de Hamming em A^n. □

Definição 2.1.39. *(Equivalência de Códigos). Dados dois códigos $\mathcal{C}_1$ e $\mathcal{C}_2$ em A^n, diremos que $\mathcal{C}_2$ é equivalente a $\mathcal{C}_1$ se existir uma isometria de Hamming φ tal que $\varphi(\mathcal{C}_1) = \mathcal{C}_2$.*

Observação 2.1.40. *Segue das três proposições demonstradas anteriormente que a equivalência de códigos é uma relação de equivalência.*

Exemplo 2.1.41. *Sejam os códigos binários*

$$\mathcal{C}_1 = \{000, 011, 110, 101\} \text{ e } \mathcal{C}_2 = \{111, 100, 001, 010\},$$

da isometria de Hamming $\varphi : \mathcal{C}_1 \to \mathcal{C}_2$, definida por $\varphi(x) = x + 111$, temos que

$$\varphi(000) = 111, \varphi(011) = 100, \varphi(110) = 001 \text{ e } \varphi(101) = 010.$$

Portanto estes códigos são equivalentes.

Definição 2.1.42. *(Taxa de Informação). Seja $\mathcal{C} \subsetneq A^n$ um código corretor de erros. Chamamos de taxa de informação do código $\mathcal{C}$ sua proporção de dados não redundante, definida como o número real*

$$R(\mathcal{C}) = \frac{\log_{|A|} |\mathcal{C}|}{n}.$$

Exemplo 2.1.43. *O código de tripla repetição $\mathcal{C}_1 = \{000000, 010101, 101010, 111111\} \subsetneq \mathbb{F}_2^6$, possui taxa de informação*

$$R(\mathcal{C}_1) = \frac{\log_2 4}{6} = \frac{2}{6} = \frac{1}{3}.$$

O código $\mathcal{C}_2 = \{00000, 01011, 10110, 11101\} \subsetneq \mathbb{F}_2^5$, possui taxa de informação

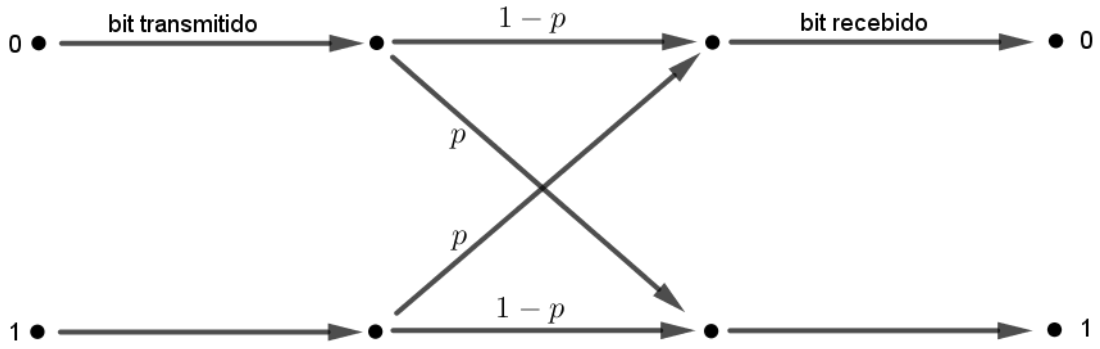
$$R(\mathcal{C}_2) = \frac{\log_2 4}{5} = \frac{2}{5}.$$

Observação 2.1.44. *Os códigos corretores de erros possuem taxa de informação sempre menor do que um, porém sabendo o canal de transmissão que será utilizado e a probabilidade de ocorrer erros neste canal é possível utilizar um código corretor de erros que seja adequado ao canal. Proporcionando assim uma maior probabilidade de não ocorrer erros.*

Definição 2.1.45. *(Canal Binário Simétrico). Um canal binário simétrico transmite apenas códigos binários, sendo assim quando ocorre um erro o bit recebido é invertido,*

transformando 0 em 1 ou 1 em 0. A probabilidade de acontecer essa inversão é igual e independente em todos os bits enviados, além disto esta probabilidade p está compreendida no intervalo $\left(0, \frac{1}{2}\right)$.

Figura 2.2: BSC (binary symmetric channel)



Exemplo 2.1.46. *Sejam os códigos binários*

$$\mathcal{C}_1 = \{00, 01, 10, 11\} \text{ e } \mathcal{C}_2 = \{000000, 010101, 101010, 111111\},$$

sendo o segundo uma tripla repetição do primeiro. A probabilidade de recebermos a informação de forma correta utilizando o código $\mathcal{C}_1$ *é igual a probabilidade de enviarmos dois bits e recebermos a informação sem erro, ou seja*

$$(1-p)^2.$$

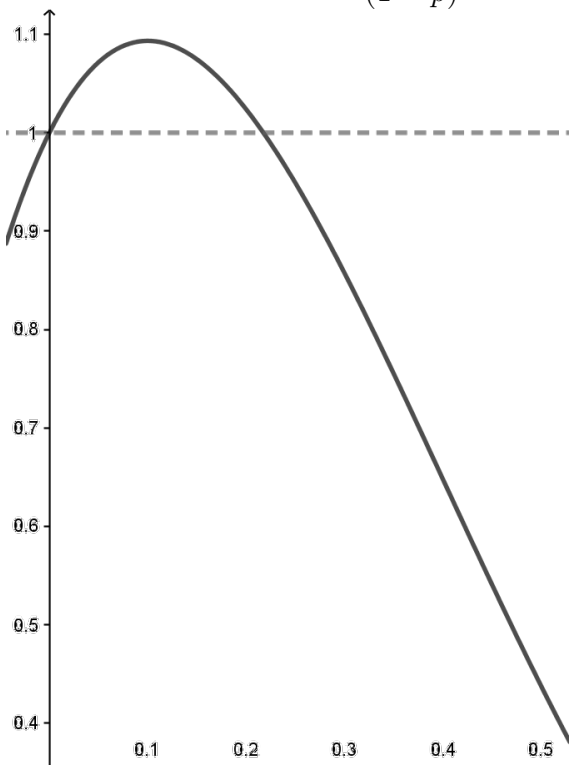
Enquanto que o código de tripla repetição $\mathcal{C}_2$ *tem capacidade de corrigir um erro, sendo assim a probabilidade de recebermos a informação sem erros utilizando o código* $\mathcal{C}_2$ *é a soma da probabilidade de recebermos os seis bits corretos com a probabilidades de termos cinco bits corretos e um incorreto, ou seja*

$$(1-p)^6 + \binom{6}{5}(1-p)^5 p.$$

Então o polinômio $f(p) = \dfrac{(1-p)^6 + \binom{6}{5}(1-p)^5 p}{(1-p)^2}$ *nos dá uma medida do ganho ou perca na probabilidade de não haver erros ao utilizarmos o código* $\mathcal{C}_2$ *ao invés do código* $\mathcal{C}_1$. *Em um canal de transmissão viável a ser utilizado, a probabilidade de ocorrer erros é menor do que* $\frac{1}{2}$, *portanto devemos observar os valores de p no intervalo* $\left(0, \frac{1}{2}\right)$, *ocorrendo ganho na probabilidade de não ocorrer erros quando o valor obtido para* $f(p)$ *for maior do que 1 e perda quando menor do que 1. Observe figura* 2.3.

Observação 2.1.47. *O principal objetivo da teoria dos códigos corretores de erros é desenvolver códigos eficientes com taxa de informação alta e que quando utilizados como*

Figura 2.3: $f(p) = \dfrac{(1-p)^6 + \binom{6}{5}(1-p)^5 p}{(1-p)^2}$



código canal em um sistema de comunicação, forneça ganhos consideráveis na probabilidade de não ocorrer erros quando comparados com o código fonte.

2.2 Códigos Lineares

Esta seção é destinada a falar sobre a classe de códigos mais utilizada na prática, chamada de classe dos códigos lineares, ao qual o alfabeto utilizado é sempre um corpo finito $\mathbb{F}_q$, com q elementos. Nela são apresentados os parâmetros dos códigos lineares, um algoritmo geral de correção de erros e a família dos código de Hamming como um exemplo de código binário linear simples.

Definição 2.2.1. *(Código Linear). Sejam o corpo finito $\mathbb{F}_q$ e o numero natural n, chamamos o código $\mathcal{C} \subset \mathbb{F}_q^n$ de $(n,k)-$código de bloco linear, se $\mathcal{C}$ for um subespaço vetorial de $\mathbb{F}_q^n$ com dimensão k.*

Observação 2.2.2. *Seja*

$$\mathcal{B} = \{b_1 = (b_{11}, b_{12}, \cdots, b_{1n}), b_2 = (b_{21},\ b_{22}, \cdots, b_{2n}), \cdots,\ b_k = (b_{k1}, b_{k2}, \cdots, b_{kn})\}$$

uma base de um código linear $\mathcal{C}$ subespaço vetorial de $\mathbb{F}_q^n$, então todos as palavras de $\mathcal{C}$ se

escrevem de modo único na forma

$$\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_k b_k,$$

onde λ_i, $i = 1, \ldots, k$, *pertencem ao corpo finito* $\mathbb{F}_q$ *que possui* q *elementos distintos, portanto a quantidade de palavras deste código é* $|\mathcal{C}| = q^k$ *e como o comprimento de suas palavras é* n*, segue que sua taxa de informação é*

$$R(\mathcal{C}) = \frac{\log_q(q)^k}{n} = \frac{k}{n}.$$

Definição 2.2.3. *(Peso de uma Palavra). Seja a palavra* $x = (x_1, x_2, \cdots, x_n)$*, onde* x_i, $i = 1, \ldots, n$*, são elementos de um corpo finito* $\mathbb{F}_q$*, chamamos de peso de* x *o número natural*

$$\omega(x) = |\{i : x_i \neq 0\}| = d_h(x, 0).$$

Exemplo 2.2.4. *Dadas as palavras* $a = (0, 0, 1)$, $b = (1, 0, 0, 1)$, $c = (1, 0, 1)$ *e* $d = (1, 1, 1, 0)$ *todas escritas sobre o alfabeto* $\mathbb{F}_2$*, temos que*

$$\omega(a) = 1, \quad \omega(b) = 2, \quad \omega(c) = 2 \text{ e } \quad \omega(d) = 3.$$

Definição 2.2.5. *(Peso de um Código). Dado um código linear* $\mathcal{C}$*, o número natural* $\omega(\mathcal{C})$ *que representa o mínimo peso das palavras não nulas de* $\mathcal{C}$ *é chamado de peso do código* $\mathcal{C}$*, ou seja*

$$\omega(\mathcal{C}) = min\{\omega(x) : x \in \mathcal{C} \setminus \{0\}\}$$

Exemplo 2.2.6. *Seja* $\mathcal{C} = \{(0, 0, 0, 0, 0, 0), (0, 1, 0, 1, 0, 1), (1, 0, 1, 0, 1, 0), (1, 1, 1, 1, 1, 1)\}$*, segue que* $\omega((0, 1, 0, 1, 0, 1)) = 3$*,* $\omega((1, 0, 1, 0, 1, 0)) = 3$ *e* $\omega((1, 1, 1, 1, 1, 1)) = 6$*, portanto* $\omega(\mathcal{C}) = min\{3, 6\} = 3$*.*

Lema 2.2.7. *Dado um código de bloco linear* $\mathcal{C}$ *subespaço vetorial de* $\mathbb{F}_q^n$*, para todo* $a \in \mathbb{F}_q^n$ *as funções* $f_a : \mathcal{C} \to \mathbb{F}_q^n$*, definidas por* $f_a(x) = x + a$*,* $x \in \mathcal{C}$ *são isometrias de Hamming.*

Demonstração. Sejam $x, y \in \mathcal{C}$, temos que

$$d_h(f_a(x), f_a(y)) = d_h(x + a, y + a) = d_h(x + a - (y + a), 0) = d_h(x - y, 0) = d_h(x, y).$$

Logo, para todo $a \in \mathbb{F}_q^n$ as funções $f_a : \mathcal{C} \to \mathbb{F}_q^n$, definidas por $f_a(x) = x + a$ são isometrias de Hamming. □

Proposição 2.2.8. *Todo código de bloco linear é geometricamente uniforme.*

Demonstração. Seja o código linear $\mathcal{C}$ subespaço vetorial de $\mathbb{F}_q^n$, dadas as palavras $a, b \in \mathcal{C}$. Se $a = b$ então temos a função identidade $Id : \mathbb{F}_q^n \to \mathbb{F}_q^n$ que é uma isometria e satisfaz

(i). $Id(\mathcal{C}) = \mathcal{C}$;

(ii). $Id(a) = b$.

Se $a \neq b$, então existe $c \in \mathbb{F}_q^n \setminus \{0\}$ tal que $a + c = b$ sobre $\mathbb{F}_q^n$ e como $\mathcal{C}$ é um subespaço vetorial $b - a = a + c - a = c \in \mathcal{C}$ e $a - b = a - (a + c) = -c \in \mathcal{C}$,temos então pelo lema anterior a isometria de Hamming $\varphi : \mathcal{C} \to \mathcal{C}$, definida por $\varphi(x) = x + c,\ x \in \mathcal{C}$ ao qual

(i). $\varphi(\mathcal{C}) = \mathcal{C}$;

(ii). $\varphi(a) = a + c = b$.

Logo, todo código de bloco linear é geometricamente uniforme. □

Proposição 2.2.9. *A distância mínima de um código linear $\mathcal{C}$ é o mínimo peso das palavras não nulas de $\mathcal{C}$, ou seja*

$$d(\mathcal{C}) = \omega(\mathcal{C}).$$

Demonstração. Seja um código linear $\mathcal{C}$, dados $a, b \in \mathcal{C}$ tais que $d(\mathcal{C}) = d_h(a, b) = d_h(a - b, 0) = \omega(a - b)$, por definição de código linear $\mathcal{C}$ é um subespaço vetorial e portanto $a - b \in \mathcal{C}$, temos então que $\omega(\mathcal{C}) = \omega(a - b)$. Logo juntando as duas igualdades $d(\mathcal{C}) = \omega(\mathcal{C})$. □

Definição 2.2.10. *(Parâmetros de um Código Linear). Sejam $\mathcal{C}$ um $(n, k)-$código de bloco linear e d um número natural tal que $d(\mathcal{C}) = \omega(\mathcal{C}) = d$. Nessas condições é dito que o código $\mathcal{C}$ tem parâmetros (n, k, d).*

Definição 2.2.11. *(Matriz Geradora). Seja $\mathcal{C}$ um $(n, k)-$código de bloco linear. A matriz*

$$G_{k\times n} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix}$$

é uma matriz geradora do código $\mathcal{C}$ se $\mathcal{B} = \{b_1, \cdots, b_k\}$ for uma base para $\mathcal{C}$.

Observação 2.2.12. *Se $\mathcal{C}$ é um $(n, k)-$código de bloco linear e $G_{k\times n}$ é uma matriz geradora do código $\mathcal{C}$, então $\mathcal{C}$ é a imagem da transformação linear injetiva:*

$$\begin{array}{rccc} \Phi: & \mathbb{F}_q^k & \to & \mathbb{F}_q^n \\ & x & \mapsto & x \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix}. \end{array}$$

Exemplo 2.2.13. *Seja* $\mathcal{B} = \{(1,0,1,0,1,1,0),(0,1,0,1,0,1)\}$ *uma base do código linear* $\mathcal{C}$ *subespaço vetorial de* $\mathbb{F}_2^6$*, segue que* $\mathcal{C}$ *é a imagem da transformação linear:*

$$\begin{array}{rccl} \Phi: & \mathbb{F}_2^2 & \to & \mathbb{F}_2^6 \\ & x & \mapsto & x\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}, \end{array}$$

observe que

$$\begin{bmatrix} 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

$$\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix};$$

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

Portanto $\mathcal{C} = \{(0,0,0,0,0,0),(0,1,0,1,0,1),(1,0,1,0,1,0),(1,1,1,1,1,1)\}$.

Definição 2.2.14. *(Forma padrão de uma Matriz Geradora). Dizemos que uma matriz geradora* G *de um* $(n,k)-$*código de bloco linear, esta escrita na forma padrão ou sistemática se as primeiras* k *colunas de* G *formam uma matriz identidade de ordem* k*, ou seja*

$$G = \left[\begin{array}{c|c} I_{k\times k} & B_{k\times(n-k)} \end{array}\right].$$

Observação 2.2.15. *Matrizes geradoras de um código linear* $\mathcal{C}$ *podem ser obtidas de outra matriz geradora do mesmo código através de sequências de operações do tipo; permutação de duas linhas, multiplicação de uma linha por um escalar não nulo e substituição de uma linha por ela mesmo somada com o resultado da multiplicação de outra linha por um escalar não nulo.*

Exemplo 2.2.16. *Dados o alfabeto* $\mathbb{F}_2$ *e a matriz geradora*

$$G_1 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

Substituindo a primeira linha por ela mesma somada com a segunda e terceira linha obtemos a matriz

$$G_2 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

Substituindo a terceira linha por ela mesma somada com a segunda linha obtemos a matriz

$$G_3 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Por fim, substituindo a segunda linha por ela mesma somada com a primeira linha obtemos a matriz na forma padrão

$$G_p = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} = \left[\begin{array}{c|c} I_{3\times 3} & B_{3\times 2} \end{array} \right]$$

onde $B_{3\times 2} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$.

Observação 2.2.17. *Nem todo código linear possui uma matriz geradora G na forma padrão.*

Exemplo 2.2.18. *Dado o código binário $\mathcal{C}$ com matriz geradora*

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

Através das operações permitidas não é possível obter uma matriz geradora na forma padrão para este código, pois ao tentarmos obter zero na terceira coordenada da primeira linha multiplicando ela por um escalar par ou substituindo ela pela soma dela mesma com o resultado da multiplicação da terceira linha por um escalar impar a linha obtida possuirá zero como primeira coordenada, além disto ao tentarmos obter zero na primeira coordenada da terceira linha também teremos um problema semelhante.

Definição 2.2.19. *(Códigos Linearmente Equivalentes). Sejam $\mathcal{C}_1$ e $\mathcal{C}_2$ códigos lineares subespaços vetoriais de $\mathbb{F}_q^n$, chamamos $\mathcal{C}_1$ e $\mathcal{C}_2$ de linearmente equivalentes se existir uma isometria linear $\varphi : \mathcal{C}_1 \to \mathcal{C}_2$ tal que $\varphi(\mathcal{C}_1) = \mathcal{C}_2$, ou seja o código $\mathcal{C}_2$ é obtido de $\mathcal{C}_1$ através de aplicações sucessivas das operações:*

(i). *Permutar a ordem das coordenadas de todas as palavras do código por meio de uma permutação π dos índices $\{1, \ldots, n\}$.*

(ii). *Multiplicar cada coordenada i fixa de todas as palavras do código por um escalar $\lambda_i \in \mathbb{F}_q \backslash \{0\}$.*

Observação 2.2.20. *Seja k a dimensão do código linear $\mathcal{C}_1 \subset \mathbb{F}_q^n$ e G_1 sua matriz geradora, toda palavra $(x_1, \ldots, x_n) \in \mathcal{C}_1$ pode ser obtida pela multiplicação de um determinado vetor $(y_1, \ldots, y_k) \in \mathbb{F}_q^k$ pela matriz geradora G_1, sendo assim se $\mathcal{C}_2$ é linearmente equivalente a $\mathcal{C}_1$, ele será obtido através das operações descritas na definição de Códigos Linearmente Equivalentes, a saber:*

(i). *Permutar a ordem das coordenadas de todas as palavras do código $\mathcal{C}_1$ equivale a permutar a ordem das colunas da matriz geradora G_1.*

(ii). *Multiplicar cada coordenada i fixa de todas as palavras do código por um escalar $\lambda_i \in \mathbb{F}_q \backslash \{0\}$, equivale a multiplicar todos os elementos de um coluna i fixa da matriz geradora G_1 por um escalar $\lambda_i \in \mathbb{F}_q \backslash \{0\}$.*

Tornando assim possível obter uma matriz geradora G_2 do código $\mathcal{C}_2$ através da matriz geradora G_1 do código $\mathcal{C}_1$.

Exemplo 2.2.21. *Sejam os códigos $\mathcal{C}_1 = \{000, 121, 212\}$ e $\mathcal{C}_2 = \{000, 111, 222\}$ subespaços vetoriais de $\mathbb{F}_3^3$. Note as igualdades dos produtos internos*

$$\langle 000, 121\rangle = 0 = \langle 000, 111\rangle, \quad \langle 000, 212\rangle = 0 = \langle 000, 222\rangle \quad e \quad \langle 121, 212\rangle = 6 = \langle 111, 222\rangle.$$

Além disto o código $\mathcal{C}_2$ pode ser obtido do código $\mathcal{C}_1$ ao multiplicamos a segunda coordenada de cada palavra do código $\mathcal{C}_1$ pelo escalar $2 \in \mathbb{F}_3 \backslash \{0\}$, ou seja

$$\begin{array}{rccc} \varphi: & \mathcal{C}_1 & \to & \mathcal{C}_2 \\ & x & \mapsto & x \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

é a isometria linear tal que $\varphi(\mathcal{C}_1) = \mathcal{C}_2$. Portanto os códigos $\mathcal{C}_1$ e $\mathcal{C}_2$ são linearmente equivalentes.

Observação 2.2.22. *Uma isometria de Hamming ou isometria em A^n é definida utilizando a métrica de Hamming, enquanto que uma isometria linear é definida utilizando o produto interno canônico. Portanto, os códigos equivalentes a um código linear podem não ser subespaços vetoriais.*

Exemplo 2.2.23. *Sejam os códigos*

$$\mathcal{C}_1 = \{000, 111, 222, 333, 444\} \text{ e } \mathcal{C}_2 = \{010, 121, 232, 343, 404\}$$

subconjuntos do espaço $\mathbb{F}_5^3$, a isometria de Hamming

$$\begin{array}{rccc} \varphi: & \mathcal{C}_1 & \to & \mathcal{C}_2 \\ & x & \mapsto & x + \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \end{array}$$

nos prova que $\mathcal{C}_1$ e $\mathcal{C}_2$ são equivalentes, porém é fácil ver que $\mathcal{C}_2$ não é um subespaço linear, um dos motivos é que $0 \notin \mathcal{C}_2$.

Teorema 2.2.24. *Seja $\mathcal{C}_1$ um $(n,k)-$código de bloco linear não trivial, existe um código linearmente equivalente $\mathcal{C}_2$ com matriz geradora na forma padrão.*

Demonstração. Seja G uma matriz geradora de $\mathcal{C}_1$ que não se encontra na forma padrão

$$G = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{12} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{bmatrix}.$$

Como G é uma matriz geradora $\mathcal{B} = \{b_1, b_2, \cdots, b_k\}$ é uma base do subespaço vetorial $\mathcal{C}$, portanto $b_i \neq 0, i = 1, 2, \ldots, k$, pois são linearmente independentes. Então existe uma coordenada $b_{1j} \neq 0$, $j \in \{1, 2, \cdots, n\}$ na primeira linha, como a matriz G foi escrita utilizando um corpo finito existe um inverso multiplicativo para qualquer elemento diferente de 0 em G, sendo assim multiplicando cada linha ao qual a coordenada $b_{ij} \neq 0$ pelo respectivo inverso de b_{ij} e substituindo as linhas em que foi feito a multiplicação pela soma delas mesmas com o inverso da primeira linha, obtemos a matriz

$$G' = \begin{bmatrix} b'_{11} & \cdots & b'_{1j} & \cdots & b'_{1n} \\ b'_{21} & \cdots & b'_{2j} & \cdots & b'_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b'_{k1} & \cdots & b'_{kj} & \cdots & b'_{kn} \end{bmatrix} = \begin{bmatrix} b'_{11} & \cdots & 1 & \cdots & b'_{1n} \\ b'_{21} & \cdots & 0 & \cdots & b'_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b'_{k1} & \cdots & 0 & \cdots & b'_{kn} \end{bmatrix}.$$

Repetindo o mesmo processo sucessivamente na segunda, terceira, ..., k-ésima linha, obtemos a matriz G" ao qual existe k colunas que quando colocadas na forma escalonada formam uma matriz identidade $(k \times k)$. Logo permutando as colunas de G" de modo a colocar a matriz identidade $I_{k\times k}$ nas primeiras k colunas, obtemos a matriz geradora $\overline{G}$ na forma padrão, e mais o código $\mathcal{C}_2$ gerado pela matriz $\overline{G}$ é linearmente equivalente ao código $\mathcal{C}_1$. □

Exemplo 2.2.25. *Seja o código binário $\mathcal{C}_1$, gerado pela matriz*

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

Não é possível colocar esta matriz na forma padrão, um dos motivos é que a única linha que possui um elemento diferente de zero na terceira coordenada é a primeira

linha, sendo assim ela deve ser permutada para ocupar o lugar da terceira linha, porém o único meio de zerarmos a segunda coordenada desta linha sem zerar a terceira coordenada é a substituindo pelo resultado da soma dela mesma com a quarta linha multiplicada por um escalar impar, o que resultaria em sua quarta coordenada ser diferente de zero. Portanto, não é possível criar uma terceira linha de forma que suas primeiras quatro coordenadas sejam $(0, 0, 1, 0)$*. Logo este código não tem uma matriz geradora na forma padrão. Porém ao permutamos suas colunas podemos obter a matriz geradora na forma padrão de um código* $\mathcal{C}_2$ *linearmente equivalente a* $\mathcal{C}_1$*, façamos como descrito na demonstração anterior. Substituindo a segunda e terceira linha pela soma delas mesmas com o inverso da primeira linha, obtemos a matriz*

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

Substituindo a primeira, terceira e quarta linha pela soma delas mesmas com o inverso da segunda linha, obtemos a matriz

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

Substituindo a primeira e segunda linha pelo soma delas mesmas com o inverso da terceira linha, obtemos a matriz

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

Substituindo a terceira linha pela soma dela mesma com o inverso da quarta linha, obtemos a matriz

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

Observe agora que a primeira, segunda, quarta e quinta colunas desta matriz formam uma matriz identidade (4×4)*, permutando então de forma sucessivas a terceira*

coluna com a quarta e a quarta com a quinta, obtemos a matriz

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

Que se encontra na forma padrão e gera o código linear $\mathcal{C}_2$ linearmente equivalente ao código $\mathcal{C}_1$.

Definição 2.2.26. *(Código Dual). Dado um código linear $\mathcal{C} \subset \mathbb{F}_q^n$, chamamos o conjunto*

$$\mathcal{C}^{\perp} = \{a \in \mathbb{F}_q^n : \langle a, b\rangle = 0, \ \forall \ b \in \mathcal{C}\}$$

de código dual de $\mathcal{C}$.

Lema 2.2.27. *Dado um código linear $\mathcal{C} \subset \mathbb{F}_q^n$, o seu código dual $\mathcal{C}^{\perp}$ também é um subespaço vetorial de $\mathbb{F}_q^n$*

Demonstração. Pela definição de subespaço vetorial $\mathcal{C}^{\perp}$ é um subespaço vetorial do espaço vetorial $\mathbb{F}_q^n$, se

(i) $0 \in \mathcal{C}^{\perp}$. De fato $\langle 0, x\rangle = 0$, para todo $x \in \mathcal{C}$. Portanto $0 \in \mathcal{C}^{\perp}$;

(ii) $a, b \in \mathcal{C}^{\perp} \Rightarrow a + b \in \mathcal{C}^{\perp}$. Se $a, b \in \mathcal{C}^{\perp}$, então para todo $x \in \mathcal{C}$

$$\begin{cases} \langle a, x\rangle = 0 \\ \langle b, x\rangle = 0 \end{cases} \Rightarrow \langle a, x\rangle + \langle b, x\rangle = \langle a + b, x\rangle = 0 \Rightarrow a + b \in \mathcal{C}^{\perp};$$

(iii) $a \in \mathcal{C}^{\perp} \Rightarrow \lambda a \in \mathcal{C}^{\perp}, \forall \lambda \in \mathbb{F}_q$. Se $a \in \mathcal{C}^{\perp}$, para todo $x \in \mathcal{C}$ e $\lambda \in \mathbb{F}_q$, temos

$$\langle a, x\rangle = 0 \Rightarrow \lambda \langle a, x\rangle = 0 \Rightarrow \langle \lambda a, x\rangle = 0 \Rightarrow \lambda a \in \mathcal{C}^{\perp}.$$

Logo $\mathcal{C}^{\perp}$ é um subespaço vetorial de $\mathbb{F}_q^n$.

□

Definição 2.2.28. *(Matriz de teste de paridade). Seja o código linear $\mathcal{C}$ e seu código dual $\mathcal{C}^{\perp}$, a matriz H geradora do código dual $\mathcal{C}^{\perp}$ é chamada matriz de teste de paridade de $\mathcal{C}$.*

Observação 2.2.29. *Dado um código linear $\mathcal{C}$ com matriz geradora G e matriz teste de paridade H, então $\mathcal{C}$ é o espaço das linhas de G e o núcleo de H. Analogamente, o código dual $\mathcal{C}^{\perp}$ é o espaço das linhas de H e o núcleo de G.*

Exemplo 2.2.30. *Seja o código linear* $\mathcal{C} \subset \mathbb{F}_3^4$ *com matriz geradora*

$$G = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix}.$$

Então se $\mathcal{B}'$ *é uma base para o código dual* $\mathcal{C}^\perp$, $|\mathcal{B}'| = (4-2) = 2$ *e* $x = (x_1, x_2, x_3, x_4) \in \mathcal{B}'$, *se*

$$\begin{cases} x_1 + 2x_3 = 0 \\ 2x_2 + x_4 = 0 \end{cases}$$

Analisando a expressão $x_1 + 2x_3 = 0$ *sobre* $\mathbb{F}_3$, *temos que* $x_1 = -2x_3 = x_3$. *De forma análoga a expressão* $2x_2 + x_4 = 0$ *sobre* $\mathbb{F}_3$ *nos fornece que* $x_2 = x_4$. *Logo* $\{(0,1,0,1),(1,0,1,0)\}$ *é uma base do código dual* $\mathcal{C}^\perp$.

Proposição 2.2.31. *Seja* $\mathcal{C} \subset \mathbb{F}_n^q$ *um* $(n,k)-$ *código linear com matriz geradora na forma padrão* $G = \left[\begin{array}{c|c} I_{k\times k} & B_{k\times(n-k)} \end{array}\right]$, *então* $H = \left[\begin{array}{c|c} -B^t_{k\times(n-k)} & I_{(n-k)\times(n-k)} \end{array}\right]$ *é uma matriz geradora de seu código dual* $\mathcal{C}^\perp$.

Demonstração. Note que

$$GH^t = \left[\begin{array}{c|c} I_{k\times k} & B_{k\times(n-k)} \end{array}\right] \begin{bmatrix} -B_{k\times(n-k)} \\ I_{(n-k)\times(n-k)} \end{bmatrix} = -B_{k\times(n-k)} + B_{k\times(n-k)} = 0.$$

Sendo assim o espaço gerado pelas linhas de H pertence ao código dual $\mathcal{C}^\perp$, além disto é fácil ver que todas as $n-k$ linhas de H são linearmente independentes. Logo H é uma matriz geradora do código dual $\mathcal{C}^\perp$. □

Corolário 2.2.32. *Todo código linear* $\mathcal{C}$ *é ortogonal ao seu código dual* $\mathcal{C}^\perp$, *ou seja* $(C^\perp)^\perp = \mathcal{C}$.

Demonstração. Seja $\mathcal{C}$ um código linear com matriz geradoras G e teste de paridade H, tem-se que H é a matriz geradora de $\mathcal{C}^\perp$ e $GH^t = 0$, aplicando o transposta nesta igualdade temos que $HG^t = 0$, logo G é matriz geradora de $(C^\perp)^\perp$. □

Definição 2.2.33. *(Síndrome). Dados um código linear* $\mathcal{C} \subset \mathbb{F}_q^n$ *com matriz de teste de paridade* H, *para todo* $x \in \mathbb{F}_q^n$ *chamamos* Hx^t *de síndrome de* x.

Proposição 2.2.34. *Seja* $\mathcal{C} \subset \mathbb{F}_q^n$ *um* $(n,k)-$*código linear com matriz de teste de paridade* H, *então a síndrome* $x \in \mathbb{F}_q^n$ *é igual a zero se, e somente se,* $x \in \mathcal{C}$, *ou seja* $Hx^t = 0 \Leftrightarrow x \in \mathcal{C}$.

Demonstração. Como $\mathcal{C} = (C^\perp)^\perp$, temos que

$$x \in \mathcal{C} \Leftrightarrow x \in (C^\perp)^\perp \Leftrightarrow \langle x, Hy\rangle = 0,\ \forall y \in \mathbb{F}_q^{n-k} \Leftrightarrow Hx^t = 0.$$

□

Definição 2.2.35. *(Vetor Erro). Ao transmitir um vetor x e receber um vetor y por um canal onde existe possibilidade de se introduzir erros, chamamos de vetor erro a diferença entre o vetor recebido e o vetor transmitido, ou seja*

$$e = y - x.$$

Observação 2.2.36. *O peso do vetor erro equivale a quantidade de erros introduzidos na palavra enviada durante a transmissão pelo canal.*

Proposição 2.2.37. *O vetor erro possui a mesma síndrome do vetor recebido.*

Demonstração. Seja $\mathcal{C}$ um código linear com matriz de teste de paridade H, utilizado como código canal em um canal onde existe possibilidade de ocorrer erros. Dado um vetor erro e referente ao vetor transmitido x e ao vetor recebido y, sabemos que

$$e = y - x \Leftrightarrow e^t = (y^t - x^t)$$

além disto a proposição anterior nos afirma que $Hx^t = 0$, portanto

$$He^t = H(y^t - x^t) = Hy^t - Hx^t = Hy^t - 0 = Hy^t.$$

□

Proposição 2.2.38. *Sejam o código linear $\mathcal{C} \subset \mathbb{F}_q^n$ com matrizes teste de paridade H e $x, y \in \mathbb{F}_q^n$, então x e y possuem a mesma síndrome, se e somente se, $x - y \in \mathcal{C}$, ou seja $Hx^t = Hy^t \Leftrightarrow x - y \in \mathcal{C}$.*

Demonstração. $Hx^t = Hy^t \Leftrightarrow Hx^t - Hy^t = 0 \Leftrightarrow H(x-y)^t = 0 \Leftrightarrow x - y \in \mathcal{C}$. □

Proposição 2.2.39. *Seja o código linear $\mathcal{C} \subset \mathbb{F}_q^n$ com dimensão k e matriz de teste de paridade H então $\omega(\mathcal{C}) \geq d$ se, e somente se, quaisquer $d-1$ colunas de H são linearmente independentes.*

Demonstração. Seja $H = \left[\begin{array}{c|c|c} h_1 & \cdots & h_n \end{array} \right]$.

($\Rightarrow$) Suponhamos que existe um conjunto de $d-1$ colunas distintas de H linearmente dependentes, sem perda de generalidade podemos supor que são as $d-1$ primeiras colunas. Desta forma, pela definição de conjunto linearmente dependente existe $\lambda_1, \lambda_2, \cdots, \lambda_{d-1} \in \mathbb{F}_q$, não todos nulos, tais que $\sum_{i=1}^{d-1} h_i \lambda_i = 0$. Sendo assim o vetor $x = (\lambda_1, \lambda_2, \cdots, \lambda_{d-1}, 0, \cdots, 0) \in \mathcal{C}$, pois $Hx^t = 0$. Como $\omega(x) \leq d-1 < d \leq \omega(\mathcal{C})$, temos uma contradição, pois $x \in \mathcal{C} \setminus \{0\}$ e $\omega(x) < \omega(\mathcal{C})$. Logo qualquer conjunto de $d-1$ colunas de H são linearmente independentes.

($\Leftarrow$) Seja $x = (x_1, \cdots, x_n) \in \mathcal{C} \setminus \{0\}$ e suponhamos que qualquer conjunto de $d-1$ colunas de H são linearmente independentes. Como $Hx^t = \sum_{i=1}^{n} h_i x_i = 0$, tem-se

que $\omega(x) \geq d$, pois se $\omega(x) \leq d-1$ existira uma combinação nula de $d-1$ vetores linearmente independentes.

□

Corolário 2.2.40. *(Cota de Singleton). Seja $\mathcal{C}$ um código linear com parâmetros (n, k, d), então*

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

Demonstração. Seja $H_{(n-k)\times n}$ a matriz de teste de paridade do código $\mathcal{C}$. Pela proposição anterior temos que qualquer conjunto de $\omega(\mathcal{C}) - 1 = d(\mathcal{C}) - 1 = d - 1$ colunas de $H_{(n-k)\times n}$ são linearmente independentes, por outro lado o posto de H é $n-k$. Logo $d-1 \leq n-k \Leftrightarrow d \leq n - k + 1$. □

Definição 2.2.41. *(Classe Lateral). Dados um código linear $\mathcal{C} \subset \mathbb{F}_q^n$ e um elemento $x \in \mathbb{F}_q^n$, chamamos de classe lateral de x segundo $\mathcal{C}$ o conjunto*

$$x + \mathcal{C} = \{x + y : y \in \mathcal{C}\}.$$

Exemplo 2.2.42. *Seja o código binário linear $\mathcal{C} = \{000, 011, 110, 101\} \subset \mathbb{F}_2^3$, as classes laterais segundo $\mathcal{C}$ são os conjuntos*

$$000 + \mathcal{C} = 011 + \mathcal{C} = 110 + \mathcal{C} = 101 + \mathcal{C} = \{000, 011, 110, 101\}$$

e

$$001 + \mathcal{C} = 010 + \mathcal{C} = 111 + \mathcal{C} = 100 + \mathcal{C} = \{001, 010, 111, 100\}.$$

Teorema 2.2.43. *Sejam $\mathcal{C} \subset \mathbb{F}_q^n$ um código linear de dimensão k e os vetores arbitrários $x, y \in \mathbb{F}_q^n$, então:*

(i) $y \in x + \mathcal{C} \Leftrightarrow x + \mathcal{C} = y + \mathcal{C}$;

(ii) *Qualquer vetor $x \in \mathbb{F}_q^n$ pertence a uma classe lateral segundo $\mathcal{C}$;*

(iii) *Todas as classes laterais segundo $\mathcal{C}$ possuem q^k elementos;*

(iv) $x + \mathcal{C} = y + \mathcal{C} \Leftrightarrow x - y \in \mathcal{C}$;

(v) $(x + \mathcal{C}) \cap (y + \mathcal{C}) \neq \emptyset \Leftrightarrow x + \mathcal{C} = y + \mathcal{C}$;

(vi) *Existem q^{n-k} classes laterais distintas segundo $\mathcal{C}$.*

Demonstração. **(i)** Se $x = y$ nada se tem a demonstrar, suponhamos então que $x \neq y$. Se $y \in x + \mathcal{C}$, então existe $z \in \mathcal{C} \backslash \{0\}$ tal que $x + z = y \Leftrightarrow x = y - z$, como $\mathcal{C}$ é um subespaço vetorial $z \in \mathcal{C} \Leftrightarrow -z \in \mathcal{C}$. Logo $x + \mathcal{C} = y + \mathcal{C}$.

(ii) Como $x \in \mathbb{F}_q^n$ e $0 \in \mathcal{C}$ então $x \in x + \mathcal{C}$.

(iii) Já vimos anteriormente que $|\mathcal{C}| = q^k$ além disto para todo $a \in \mathbb{F}_q^n$ as funções $f_a : \mathcal{C} \to a + \mathcal{C}$, definida por $f_a(x) = x + a, x \in \mathcal{C}$ são isometrias de Hamming e portanto bijeções, sendo assim a quantidade de elementos do domínio e da imagem são iguais, logo toda classe lateral possui q^k elementos.

(iv) Seque imediatamente de (i), pois

$$x + \mathcal{C} = y + \mathcal{C} \Leftrightarrow x \in y + \mathcal{C} \Leftrightarrow x - y \in \mathcal{C}$$

(v) Suponhamos que existe $z \in (x + \mathcal{C}) \cap (y + \mathcal{C})$ então existem $z', z" \in \mathcal{C}$ tais que $x + z' = y + z" = z$, além disto $z' + \mathcal{C} = z" + \mathcal{C} = \mathcal{C}$ portanto $x + \mathcal{C} = y + \mathcal{C}$.

(vi) Como $|\mathbb{F}_q^n| = q^n$ e para qualquer vetor $x \in \mathbb{F}_q^n$ temos $|x + \mathcal{C}| = q^k$, a quantidade de classes laterais distintas é o resultado da divisão

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

□

Definição 2.2.44. *(Elemento Líder). Dados um código linear $\mathcal{C} \subset \mathbb{F}_q^n$ e o vetor $x \in \mathbb{F}_q^n$, chamamos $y \in x + \mathcal{C}$ de um elemento líder desta classe lateral se*

$$\omega(y) = min\{\omega(z) : z \in x + \mathcal{C}\}$$

Exemplo 2.2.45. *Dado o código binário linear $\mathcal{C} = \{000, 011, 110, 101\} \subset \mathbb{F}_2^3$, sabemos que suas classes laterais são $\{000, 011, 110, 101\}$ onde o vetor 000 possui o menor peso e portanto é seu líder e $\{001, 010, 111, 100\}$ ao qual as três palavras 001, 010 e 100 possuem o peso mínimo da classe e são seus elementos líderes.*

Proposição 2.2.46. *Seja o código linear $\mathcal{C} \subset \mathbb{F}_q^n$. Se $x \in \mathbb{F}_q^n$ é tal que $\omega(x) \leq \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$, então x é o único elemento líder de sua classe.*

Demonstração. Suponhamos que existe $x, y \in \mathbb{F}_q^n$, onde $x \neq y$ e ambos pertencentes a mesma classe lateral e possuem peso menor do que ou igual a $\left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$. Pelo teorema anterior sabemos que x e y pertencem a mesma classe lateral se, e somente se, $x - y \in \mathcal{C}$, portanto $d_h(x - y, 0) = d_h(x, y) \geq d(\mathcal{C})$, por outro lado pela definição de peso $\omega(x) = d_h(x, 0) \leq \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$ e $\omega(y) = d_h(y, 0) = d_h(0, y) \leq \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$, segue que

$$d(\mathcal{C}) \leq d_h(x, y) \leq d_h(x, 0) + d_h(0, y) \leq 2 \left\lfloor \frac{d(\mathcal{C}) - 1}{2} \right\rfloor \leq d(\mathcal{C}) - 1.$$

Isto é impossível, logo $x = y$ é o único elemento líder de sua classe. □

Definição 2.2.47. *(Decodificação). Ao utilizarmos um método qualquer para detecção e correção de erros num determinado código estamos realizando um procedimento chamado de decodificação.*

Observação 2.2.48. *Existem diversos métodos que podem ser utilizados para decodificar um código linear, dentre eles esta por exemplo a Tabela de D. Slepian inventado na década de 60, o algoritmo que será apresentado a seguir possuíra algumas diferenças do mesmo, pois foi aperfeiçoado para se tornar mais eficiente(possuir um custo computacional menor) e seguirá critérios de verossimilhança, como a probabilidade de ocorrer um erro no canal será sempre menor do que $\frac{1}{2}$ a hipótese de que a palavra corrigida seja a mais próxima em relação a distância de Hamming da palavra recebida possui maior probabilidade e quanto menor for a probabilidade de erro do canal maior será a força de evidência em favor deste acontecimento.*

Algoritmo de Decodificação

Sejam $\mathcal{C} \subset \mathbb{F}_q^n$ um código linear com parâmetros (n, k, d), matriz de teste de paridade H e os vetores $x \in \mathcal{C}$ e $y \in \mathbb{F}_q^n$. Se $d < 3$ então não é possível detectar e corrigir erros segundo critérios de verossimilhança em $\mathcal{C}$ em alguns casos de apenas um único erro, suponhamos então que $d \geq 3$ para que $\mathcal{C}$ seja um código viável a ser utilizado como código canal.

Suponhamos que a palavra x foi enviada por um canal onde existe possibilidade de ocorrer erros na transmissão e a palavra y foi recebida. Se $Hy^t = 0$ é aceito que $x = y$, ou seja y é aceito como a palavra enviada, no caso contrário em que $Hy^t = z^t \neq 0$ fica claro que ouve um erro durante a transmissão, portanto existe o vetor erro $e \in \mathbb{F}_q^n$, com $\omega(e) > 0$, que possui a mesma síndrome e pertence a mesma classe lateral de y, como estamos utilizando um critério de verossimilhança aceitaremos que e seja o líder da classe, possuindo assim o menor peso possível. Visto que nem toda classe lateral possui vetor líder único, a correção por verossimilhança só poderá ser feita quando e for o único líder de sua classe, neste caso é aceito a igualdade $x = y - e$, caso contrário uma quantidade de erros maior do que $\left\lfloor \frac{d-1}{2} \right\rfloor$ foram inseridos durante a transmissão tornando impossível realizar a decodificação.

Exemplo 2.2.49. *Seja $\mathcal{B} = \{111\}$ uma base de um código linear binário $\mathcal{C}$, então a matriz geradora de $\mathcal{C}$ é $G = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, segue que $\begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ e $\begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ portanto $\mathcal{C} = \{000, 111\}$, como $\omega(\mathcal{C}) = d(\mathcal{C}) = d_h(111, 000) = 3$ este código possui capacidade de corrigir $\left\lfloor \frac{3-1}{2} \right\rfloor = 1$ erro, se $\mathcal{B}'$ é uma base do código dual $\mathcal{C}^\perp$ então $|\mathcal{B}'| = 3 - 1 = 2$ e $b = (b_1, b_2, b_3) \in \mathcal{B}'$, somente se $b_1 + b_2 + b_3 = 0$ em $\mathbb{F}_2$, logo*

$\{(1,1,0),(1,0,1)\}$ *é uma base do código dual* $\mathcal{C}^{\perp}$ *e* $H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ *é uma das matrizes de teste de paridade de* $\mathcal{C}$*. Suponhamos que uma palavra* x *deste código foi transmitida por um canal binário simétrico e a palavra* $y = 001$ *foi recebida. Utilizemos o algoritmo descrito anteriormente para decodificar a palavra* y *e descobrir assim a palavra* x*. Note que as classes laterais segundo* $\mathcal{C}$*, seus lideres e as síndromes referentes aos lideres de cada classe segundo a matriz* H *são respectivamente*

Classe laterais	*Líder*	*Síndrome do Líder da Classe*
$000 + \mathcal{C} = \{000, 111\}$	000	00
$001 + \mathcal{C} = \{001, 110\}$	001	01
$010 + \mathcal{C} = \{010, 101\}$	010	10
$100 + \mathcal{C} = \{100, 011\}$	100	11

Por outro lado

$$Hy^t = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq 0$$

e portanto existe um vetor erro $e = \{e_1, e_2, e_3\} \in \mathbb{F}_2^3$ *que possui a mesma síndrome que* y *e este vetor é o líder de uma das classes laterais de* $\mathcal{C}$*, consultando a tabela podemos ver que* $e = 001$*. Logo a palavra enviada* $x = y - e = 001 - 001 = 000$*.*

Definição 2.2.50. *(Código de Hamming). Um código binário linear* $\mathcal{C} \subset \mathbb{F}_2^n$ *é chamado de código de Hamming se as colunas da sua matriz de teste de paridade* $H_{k\times n}$ *são todos os elementos de* $\mathbb{F}_2^k \backslash \{0\}$ *repetidos uma única vez em uma ordem qualquer.*

Observação 2.2.51. *Todo código de Hamming é binário escrito assim sobre* $\mathbb{F}_2$*, portanto o comprimento das palavras de um código de Hamming* $\mathcal{C}$ *com matriz de teste de paridade* $H_{k\times n}$ *é* $n = 2^k - 1$ *e pela proposição referente a base de um código dual, temos que sua dimensão é* $n - k = 2^k - k - 1$*.*

Proposição 2.2.52. *Todo código de Hamming é 1-perfeito.*

Demonstração. Dado um código de Hamming $\mathcal{C}(2^k - 1, 2^k - k - 1)$ com matriz de teste de paridade $H_{k\times 2^k-1}$, pela definição de código de Hamming as colunas de H são formadas pelos elementos de $\mathbb{F}_2^k \backslash \{0\}$, onde é fácil observar que para todo elemento $a \in \mathbb{F}_2^k$ com peso i, $i \in \{2, \ldots, k\}$, existe outros dois elementos $b, c \in \mathbb{F}_2^k$, tais que $b + c = a$. Portanto, existem três colunas linearmente dependentes e a distância miníma $d(\mathcal{C}) = 3$ para todo código de Hamming.
Por outro lado dado $x \in \mathbb{F}_2^{2^k-1}$, temos que

$$\left| B\left[x, \left\lfloor \frac{d(\mathcal{C}) - 1}{2} \right\rfloor\right] \right| = \left| B\left[x, \left\lfloor \frac{3 - 1}{2} \right\rfloor\right] \right| = |B[x, 1]| = 1 + |S(x, 1)| = 1 + 2^k - 1 = 2^k$$

Além disto, a quantidade de palavras existentes em $\mathbb{F}_2^{2^k-1}$ são 2^{2^k-1} enquanto que a quantidade de palavras do código é $|\mathcal{C}| = 2^{2^k-k-1}$, portanto

$$\left|\bigcup_{x\in\mathcal{C}} B\left[x, \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor\right]\right| = (2^{2^k-k-1})2^k = 2^{2^k-1} = \left|\mathbb{F}_2^{2^k-1}\right|.$$

Logo todo código de Hamming é 1-perfeito. □

Exemplo 2.2.53. *Seja o código binário $\mathcal{C}$ com matriz de teste de paridade*

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

Como as colunas de H são todos os elementos de

$$\mathbb{F}_2^3 \backslash \{0\} = \{(1,0,0),(0,1,0),(0,0,1),(1,1,0),(1,0,1),(1,1,1)\}$$

$\mathcal{C}$ e um código de Hamming com $k = 3$, ou seja um código de Hamming $(7,4)$.
Por outro lado, o código $\mathcal{C}'$ com matriz de teste de paridade

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

não é um código de Hamming pois entre as colunas de H' não existe o elemento $(1,1,1,1) \in \mathbb{F}_2^4$.

2.3 Códigos Cíclicos

Nesta seção é apresentado a subclasse dos códigos lineares chamada de códigos cíclicos, estes códigos são muito utilizados em diversas aplicações por possuírem bons algoritmos de codificação e decodificação, além de requererem uma quantidade pequena de informação para se descrever todas as palavras. Existem famílias importantes de códigos como o BCH, Reed-Solomon e Goppa que são códigos cíclicos. Falaremos da família dos códigos de Goppa binários irredutíveis que possuem alta capacidade de correção de erros e do algoritmo de Patterson, que se mostra um dos mais eficientes para decodificação desta família de códigos.

Definição 2.3.1. *(Códigos Cíclicos). Um código linear $\mathcal{C} \subset \mathbb{F}_q^n$ é chamado de cíclico se*

$$\forall\ c = (c_0, c_1, \ldots, c_{n-1}) \in \mathcal{C} \Rightarrow (c_{n-1}, c_0, c_1, \ldots, c_{n-2}) \in \mathcal{C}.$$

Para que possamos lidar com os códigos cíclicos devemos enriquecer a estrutura do espaço vetorial $\mathbb{F}^n$, sendo assim será preciso fazer algumas definições

Definição 2.3.2. *(R_n).Chamamos de R_n o anel das classes residuais em $\mathbb{F}_q[X]$ módulo (X^n-1), ou seja*

$$R_n = \mathbb{F}_q[X]/X^n - 1$$

Observe que R_n é isomorfo a $\mathbb{F}_q^n$ através da transformação linear

$$\begin{array}{rccc} \nu: & \mathbb{F}_q^n & \to & R_n \\ & (c_0, c_1, \ldots, c_{n-2}, c_{n-1}) & \mapsto & c_0 + c_1X + \cdots + c_{n-2}X^{n-2} + c_{n-1}X^{n-1} \end{array}$$

Portanto qualquer código linear $\mathcal{C} \subset \mathbb{F}_q^n$ pode ser transportado para R_n através do isomorfismo ν. Ao estudarmos $\nu(\mathcal{C}) \subset R_n$ obtemos a vantagem de estarmos lidando com uma estrutura de anel, a qual permite caracterizar os códigos cíclicos.

Teorema 2.3.3. *$\mathbb{F}_q[X]$ é um anel principal. Ou seja, se $I \neq \{0\}$ é um ideal de $\mathbb{F}_q[X]$, então existe um único polinômio mônico $F(X)$ de grau mínimo em I, tal que $I = I(F(X))$.*

Demonstração. Sejam $I \neq \{0\}$ um ideal de $\mathbb{F}_q[X]$, $F(X)$ um polinômio não nulo de grau mínimo em I e $G(X)$ um elemento qualquer de I, pelo algoritmo da divisão em $\mathbb{F}_q[X]$, existem polinômios $Q(X)$ e $R(X)$, tais que $G(X) = F(X)Q(X) + R(X)$, com $gr(R(X)) < gr(F(X))$. Então como $F(X) \in I$ e $Q(X) \in \mathbb{F}_q[X]$, pela condição *ii* da definição de ideal temos $-F(X)Q(X) \in I$, além disto $G(X) \in I$ portanto pela condição *i* da definição de ideal $R(X) = G(X) - F(X)Q(X) \in I$. Temos então que $R(X) = 0$, pois caso contrário teríamos um elemento de grau menor do que o grau de $F(X)$ no ideal I, o que não é possível pois $F(X)$ é de grau mínimo. Como resultado o polinômio arbitrário $G(X) = F(X)Q(X) \in I(F(X))$, portanto $I \subset I(F(X))$. Como $F(X) \in I$, temos também a inclusão inversa $I \supset I(F(X))$. Logo $I = I(F(X))$.

Provaremos agora que o polinômio mônico gerador do ideal I é único. Para toda constante não nula $\lambda \in \mathbb{F}_q$, temos que $gr(\lambda F(X)) = gr(F(X))$ e como $\lambda \in \mathbb{F}_q[X]$, pela condição *ii* da definição de ideal $\lambda F(X) \in I$. Portanto, todo $\lambda F(X)$ é um polinômio de grau mínimo em I, sendo assim um gerador de I. Por outro lado, como $\mathbb{F}_q$ é um corpo todos os seus elementos não nulos possuem um único inverso com respeito à multiplicação, além disto o inverso do elemento chamado unidade é a própria unidade, que possui propriedade de ser neutra para a multiplicação. Seja $gr(F(X)) = g$ então $F(X)$ é da forma $c_0 + c_1X + \cdots + c_gX^g$ onde $c_i \in \mathbb{F}_q, \forall i \in \{0, \ldots, g\}$, existe então um único elemento inverso multiplicativo de $c_g \neq 0$ e quando $F(X)$ é multiplicado por este inverso encontramos o único polinômio mônico gerador de I. □

Proposição 2.3.4. *Todo ideal de $\mathbb{F}_q[X]/P(X)$ é da forma $I(\overline{F(X)})$, onde $\overline{F(X)} = \{F(X) + \lambda(X)P(X) : \lambda(X) \in \mathbb{F}_q[X]\}$ é um elemento de $\mathbb{F}_q[X]/P(X)$ e $F(X)$ é um divisor de $P(X)$.*

Demonstração. Seja I um ideal de $\mathbb{F}_q[X]/P(X)$. Provemos primeiramente que o conjunto $J = \{G(X) \in \mathbb{F}_q[X] : \overline{G(X)} \in I\}$ é um ideal de $\mathbb{F}_q[X]$.

i. Se $G_1(X), G_2(X) \in J$, então pela definição do conjunto J temos que $\overline{G_1(X)}, \overline{G_2(X)} \in I$, e pela condição i da definição de ideal $\overline{G_1(X) + G_2(X)} = \overline{G_1(X)} + \overline{G_2(X)} \in I$, consequentemente pela definição do conjunto J, $G_1(X) + G_2(X) \in J$.

ii. Se $G(X) \in J$ e $\lambda(X) \in \mathbb{F}_q[X]$, pela definição do conjunto J temos que $\overline{G(X)} \in I$, e pela condição ii da definição de ideal $\overline{G(X)\lambda(X)} = \overline{G(X)} \cdot \overline{\lambda(X)} \in I$, consequentemente pela definição do conjunto J, $G(X)\lambda(X) \in J$.

Portanto, J é um ideal de $\mathbb{F}_q[X]$ e como $P(X) \in J$, $J \neq \{0\}$, o teorema anterior nos informa que existe $F(X) \in \mathbb{F}_q[X]\backslash\{0\}$, tal que $J = I(F(X))$. Dada a definição do conjunto $J = I(F(X))$, temos

$$I = \{\overline{G(X)} : G(X) \in J\} = \{\overline{\lambda(X)} \cdot \overline{F(X)} : \overline{\lambda(X)} \in \mathbb{F}_q[X]/P(X)\} = I(\overline{F(X)}).$$

Além disto $P(X) = \lambda(X)F(X)$, para algum $\lambda(X) \in \mathbb{F}_q[X]$, logo $F(X)$ é um divisor de $P(X)$. □

Lema 2.3.5. *Um subespaço vetorial de R_n, é um ideal de R_n se, e somente se, este subespaço for fechado pela multiplicação por $\overline{X} = \{X + \lambda(X)(X^n - 1) : \lambda(X) \in \mathbb{F}_q[X]\}$.*

Demonstração. ($\Rightarrow$) Seja I um ideal de R_n, então I é fechado para a soma e para o produto por escalares de R_n, como os polinômios constantes de R_n formam os escalares de $\mathbb{F}_q$, I é um subconjunto vetorial de R_n. Além disto $\overline{X}$ é um escalar de R_n, portanto I é fechado pela multiplicação por $\overline{X}$.

($\Leftarrow$) Seja A um subespaço vetorial de R_n, fechado para a multiplicação em $\overline{X}$. Seja $\overline{F(X)} \in A$, então $\overline{XF(X)} = \overline{X} \cdot \overline{F(X)} \in A$ e consequentemente $\overline{X^2F(X)} = \overline{X} \cdot \overline{XF(X)} \in A$. Portanto, indutivamente obtemos que

$$\overline{X^kF(X)} = \overline{X^k} \cdot \overline{F(X)} \in A, \forall\ k \in \mathbb{N}$$

Observe que para todo escalar $\overline{\lambda(X)} = \overline{\alpha_0 + \alpha_1X + \cdots + \alpha_{n-1}X^{n-1}} \in R_n$, temos que

$$\overline{\lambda(X)} \cdot \overline{F(X)} = \overline{\lambda(X)F(X)} = \alpha_0\overline{F(X)} + \alpha_1\overline{X} \cdot \overline{F(X)} + \cdots + \alpha_{n-1}\overline{X^{n-1}} \cdot \overline{F(X)} \in A.$$

Pois A é subespaço vetorial e $\alpha_i, i = 0, \ldots, n-1$, são escalares de $\mathbb{F}_q$. Logo A é um ideal de R_n. □

Teorema 2.3.6. *Um subespaço $\mathcal{C} \subset \mathbb{F}_q^n$ é um código cíclico se, e somente se, $\nu(\mathcal{C})$ é um ideal de R_n.*

Demonstração. Como $\mathcal{C}$ é um subespaço vetorial de $\mathbb{F}_q^n$ e $\mathbb{F}_q^n$ é isomorfo a R_n através da transformação linear ν, $\nu(\mathcal{C})$ é um subespaço vetorial de R_n, valendo então o lema anterior. Resta então provar que $\nu(\mathcal{C})$ é fechado pela multiplicação por $[X]$, se e somente se, $\mathcal{C}$ for um código cíclico, que é evidente pela definição de código cíclico. □

Teorema 2.3.7. *Seja $\mathcal{C} \subset \mathbb{F}_q^n$ um código cíclico, onde $\nu(\mathcal{C}) = I(\overline{G(X)})$ e $G(X)$ é um divisor de X^{n-1} com $gr(G(X)) = \delta$. Então $\overline{G(X)}, \overline{X \cdot G(X)}, \ldots, \overline{X^{n-\delta-1} \cdot G(X)}$ é uma base de $\nu(\mathcal{C})$.*

Demonstração. Como todo código cíclico é um código linear, $\mathcal{C}$ é um subespaço vetorial de $\mathbb{F}_q^n$ e portanto $\nu(\mathcal{C})$ é subespaço vetorial de R_n, devemos então provar que $\overline{G(X)}, \overline{X \cdot G(X)}, \ldots, \overline{X^{n-\delta-1} \cdot G(X)}$ são linearmente independentes para que seja uma base de $\nu(C)$.

Dados $\alpha_0, \alpha_1, \ldots, \alpha_{n-\delta-1} \in \mathbb{F}_q$ tais que

$$\alpha_0 \cdot \overline{G(X)} + \alpha_1 \cdot \overline{X \cdot G(X)} + \cdots + \alpha_{n-\delta-1} \cdot \overline{X^{n-\delta-1} \cdot G(X)} = \overline{0}$$

Colocando $\overline{G(X)}$ em evidencia, temos

$$\overline{G(X)} \cdot \overline{\alpha_0 + \alpha_1 \cdot X + \cdots + \alpha_{n-\delta-1} \cdot X^{n-\delta-1}} = \overline{0}$$

Como

$$\begin{cases} gr(\alpha_0 + \alpha_1 \cdot X + \cdots + \alpha_{n-\delta-1} \cdot X^{n-\delta-1}) \leq n - \delta - 1 \\ gr(G(X)) = \delta \end{cases}$$

segue que

$$gr(G(X) \cdot (\alpha_0 + \alpha_1 \cdot X + \cdots + \alpha_{n-\delta-1} \cdot X^{n-\delta-1})) \leq n - \delta - 1 + \delta = n - 1 < n = gr(X^n - 1).$$

Logo, $\alpha_0 = \alpha_1 = \cdots = \alpha_{n-\delta-1} = 0$ e $\overline{G(X)}, \overline{X \cdot G(X)}, \ldots, \overline{X^{n-\delta-1} \cdot G(X)}$ são linearmente independentes, pois caso contrário existiria um polinômio de grau menor do que n em R_n igual a $\overline{0}$ e diferente de 0. □

Este teorema é de grande importância para representação de uma matriz geradora para códigos cíclicos. Pois como $\overline{G(X)}, \overline{X \cdot G(X)}, \ldots, \overline{X^{n-\delta-1} \cdot G(X)}$ é uma base de $\nu(\mathcal{C})$, segue que $\nu^{-1}(\overline{G(X)}), \nu^{-1}(\overline{X \cdot G(X)}), \ldots, \nu^{-1}(\overline{X^{n-\delta-1} \cdot G(X)})$ é uma base de $\nu^{-1}(\nu(\mathcal{C})) = \mathcal{C}$. Sendo assim, chamamos $G(X) = g_0 + g_1 X + \cdots + g_\delta X^\delta$ de polinômio

gerador do código $\mathcal{C}$ e a matriz geradora é

$$G_{(n-\delta)\times n} = \begin{bmatrix} \nu^{-1}(\overline{G(X)}) \\ \nu^{-1}(\overline{X \cdot G(X)}) \\ \vdots \\ \nu^{-1}(\overline{X^{n-\delta-2} \cdot G(X)}) \\ \nu^{-1}(\overline{X^{n-\delta-1} \cdot G(X)}) \end{bmatrix} = \begin{bmatrix} g_0 & g_1 & \cdots & g_\delta & 0 & \cdots & 0 & 0 \\ 0 & g_0 & \cdots & g_{\delta-1} & g_\delta & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & g_0 & \dots & g_{\delta-1} & g_\delta & 0 \\ 0 & \dots & 0 & 0 & g_0 & \dots & g_{\delta-1} & g_\delta \end{bmatrix}$$

Para que possamos descrever a matriz teste de paridade de um código cíclico é necessário conhecer a definição de polinômio recíproco e uma propriedade que estes polinômios carregam.

Definição 2.3.8. *(Polinômio Recíproco). Chamamos de polinômio recíproco de $F(X) \in K[X]$, onde $gr(F(X)) = n$, o polinômio*

$$F^*(X) = X^n \cdot F\left(\frac{1}{X}\right)$$

Proposição 2.3.9. *Seja $F(X), G(X), H(X) \in K[X]$ tais que $F(X) = G(X)H(X)$, então $F^*(X) = G^*(X)H^*(X)$.*

Demonstração. Suponha que $gr(G(X)) = m$ e $gr(H(X)) = n$, então

$$gr(F(X)) = gr(G(X)H(X)) = m + n.$$

Segue da definição de polinômio recíproco que

$$\begin{aligned} F^*(X) &= X^{m+n} \cdot F\left(\tfrac{1}{X}\right) \\ &= X^{m+n} \cdot G\left(\tfrac{1}{X}\right) H\left(\tfrac{1}{X}\right) \\ &= \left[X^m \cdot G\left(\tfrac{1}{X}\right)\right]\left[X^n \cdot H\left(\tfrac{1}{X}\right)\right] \\ &= G^*(X) \cdot H^*(X). \end{aligned}$$

□

Seja $G(X)H(X) = X^n - 1 \Leftrightarrow G^*(X)H^*(X) = X^n \cdot \left(\frac{1}{X}^n - 1\right) = 1 - X^n$. Multiplicando a segunda igualde por (-1), obtemos

$$H^*(X) \cdot (-1)G^*(X) = (1 - X^n) \cdot (-1) = X^n - 1.$$

Portanto se $H(X)$ é um divisor de $X^n - 1$, então $H^*(X)$ é também um divisor de $X^n - 1$, pois $(-1)G^*(X) \in K[X]$.

Para o próximo lema e teorema vamos considerar $G(X)H(X) = X^n - 1$, onde $G(X)$ é o polinômio gerador de um código cíclico e n o comprimento das palavras deste

código. Nestas condições o polinômio $H(X)$ é chamado polinômio de paridade de $\mathcal{C}$. Enquanto que o polinômio $H^*(X)$ que também é um divisor de $X^n - 1$ será o gerador de outro código cíclico ao qual identificaremos adiante.

Lema 2.3.10. *$F(X) \in \nu(\mathcal{C})$ se, somente se, $F(X) \cdot H(X) = \overline{0}$ em R_n.*

Demonstração. Seja $F(X) \in \mathbb{F}_q[X]$, então

$$\begin{aligned} F(X) \cdot H(X) = \overline{0} &\Leftrightarrow F(X)H(X) = Q(X)(X^n - 1) \text{ para algum } Q(X) \in \mathbb{F}_q[X] \\ &\Leftrightarrow F(X) = Q(X)G(X) \\ &\Leftrightarrow F(X) \in I(\overline{G(X)}) = \nu(\mathcal{C}) \end{aligned}$$

□

Teorema 2.3.11. *Seja $\mathcal{C} = \nu^{-1}(I(\overline{G(X)}))$ um código cíclico. Então o código dual $\mathcal{C}^\perp$, também é cíclico e $\mathcal{C}^\perp = \nu^{-1}(I(\overline{H^*(X)}))$.*

Demonstração. Seja $gr(G(X)) = \delta$ então $gr(H(X)) = gr(X^n - 1) - gr(G(X)) = n - \delta$. Escrevendo $H(X) = h_0 + h_1X + \cdots + h_{n-\delta}X^{n-\delta}$, segue do lema anterior que $F(X) = f_0 + f_1X + \cdots + f_{n-1}X^{n-1} \in \nu(\mathcal{C})$ se, somente se, $F(X) \cdot H(X) = \overline{0}$ em R_n, ou seja

$$\sum_{i=0}^{2n-\delta-1} \left(\sum_{j=0}^{i} f_i \cdot h_{i-j} \right) X^i = \overline{0}$$

Portanto os coeficientes de $F(X)$ são solução do sistema de equações lineares

$$\begin{cases} f_0h_{n-\delta} + f_1h_{n-\delta-1} + \cdots + f_{n-\delta}h_0 = 0 \\ f_1h_{n-\delta} + f_2h_{n-\delta-1} + \cdots + f_{n-\delta+1}h_0 = 0 \\ \ddots \quad \ddots \quad \ddots \quad \ddots = 0 \\ f_{n-\delta-1}h_{n-\delta} + f_{n-\delta}h_{n-\delta-1} + \cdots + f_{n-1}h_0 = 0 \end{cases}$$

Sendo assim

$$H_{\delta \times n} = \begin{bmatrix} \nu^{-1}(\overline{H^*(X)}) \\ \nu^{-1}(\overline{X \cdot H^*(X)}) \\ \vdots \\ \nu^{-1}(\overline{X^{\delta-2} \cdot H^*(X)}) \\ \nu^{-1}(\overline{X^{\delta-1} \cdot H^*(X)}) \end{bmatrix} = \begin{bmatrix} h_{n-\delta} & h_{n-\delta-1} & \cdots & h_0 & 0 & \cdots & 0 & 0 \\ 0 & h_{n-\delta} & \cdots & h_1 & h_0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_{n-\delta} & \cdots & h_1 & h_0 & 0 \\ 0 & \cdots & 0 & 0 & h_{n-\delta} & \cdots & h_1 & h_0 \end{bmatrix}$$

é uma matriz teste de paridade de $\mathcal{C}$. Logo, uma matriz geradora de seu código dual $\mathcal{C}^\perp$. Além disto como $\overline{H^*(X)}, \overline{X \cdot H^*(X)}, \ldots, \overline{X^{\delta-1} \cdot H^*(X)}$ é uma base para o ideal $I(\overline{H^*(X)})$, segue que $\mathcal{C}^\perp = \nu^{-1}(I(\overline{H^*(X)}))$. □

Exemplo 2.3.12. *Seja* $R_n = \mathbb{F}_2[X]/X^{15} - 1$. *Em* $\mathbb{F}_2[X]$, *temos que*

$$X^{15} - 1 = (X^8 + X^7 + X^5 + X^4 + X^3 + X + 1)(X^7 + X^6 + X^5 + X^2 + X + 1),$$

podemos então criar um código binário cíclico de comprimento 15, *com polinômio gerador e paridade sendo respectivamente*

$$G(X) = X^8 + X^7 + X^5 + X^4 + X^3 + X + 1 \text{ e } H(X) = X^7 + X^6 + X^5 + X^2 + X + 1$$

segue que a matriz

$$G_{7\times 15} = \begin{bmatrix} \nu^{-1}(\overline{G(X)}) \\ \nu^{-1}(\overline{X \cdot G(X)}) \\ \nu^{-1}(\overline{X^2 \cdot G(X)}) \\ \nu^{-1}(\overline{X^3 \cdot G(X)}) \\ \nu^{-1}(\overline{X^4 \cdot G(X)}) \\ \nu^{-1}(\overline{X^5 \cdot G(X)}) \\ \nu^{-1}(\overline{X^6 \cdot G(X)}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}$$

é sua matriz geradora, enquanto que

$$H_{8\times 15} \begin{bmatrix} \nu^{-1}(\overline{H^*(X)}) \\ \nu^{-1}(\overline{X \cdot H^*(X)}) \\ \nu^{-1}(\overline{X^2 \cdot H^*(X)}) \\ \nu^{-1}(\overline{X^3 \cdot H^*(X)}) \\ \nu^{-1}(\overline{X^4 \cdot H^*(X)}) \\ \nu^{-1}(\overline{X^5 \cdot H^*(X)}) \\ \nu^{-1}(\overline{X^6 \cdot H^*(X)}) \\ \nu^{-1}(\overline{X^7 \cdot H^*(X)}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

é sua matriz teste de paridade.

Código de Goppa

Em 1970, V. D. Goppa, publicou o artigo [4], descrevendo uma subclasse de códigos alternantes, ao qual posteriormente em 1978 no [8] foi sugerida por R.J. McEliece para ser utilizada em sistemas de criptografia, devido a existência de algoritmos eficientes de decodificações. Estes códigos são descritos em termos de um polinômio gerador, chamado polinômio de Goppa e de um conjunto de elementos distintos L.

Definição 2.3.13. *(Códigos de Goppa). Seja* $F = \mathbb{F}_{q^m}$ *um corpo finito, extensão de um corpo* $K = \mathbb{F}_q$. *Sejam* $\varphi(X) \in F[X]$ *e* $L = \{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\} \subset F$, *onde* α_i *são dois a*

dois distintos e tais que $\varphi(\alpha_i) \neq 0$ *para* $i = 0, \ldots, n-1$. *Defini-se como Código de Goppa*

$$\Gamma_K(L, \varphi) = \left\{ (c_0, \ldots, c_{n-1}) \in K^n; \sum_{i=0}^{n-1} c_i \varphi(\alpha_i)^{-1} \cdot \frac{\varphi(X) - \varphi(\alpha_i)}{X - \alpha_i} = 0 \right\}.$$

Nosso estudo estará restrito a códigos de Goppa binário irredutíveis, nestes $q = 2$ e $\varphi(X)$ é irredutível. Esta classe foi especificamente a sugerida por McEliece para ser utilizada em sistemas de criptografia.

A síndrome de uma palavra que pertence a um código de Goppa é um polinômio, o qual é utilizado no processo de decodificação desta palavra e na correção de erros, quando necessário.

Definição 2.3.14. *(Função Síndrome).Dado um vetor* $c \in K^n$, *chamamos de função síndrome de c:*

$$S_c(X) = \sum_{i=0}^{n-1} \frac{c_i}{X - \alpha_i}$$

Segue que c é uma palavra do código se, somente se, sua síndrome for congruente a 0 módulo o polinômio de Goppa $\varphi(X)$. Para que possamos construir uma matriz teste de paridade adequada a um código de Goppa genérico, devemos então analisar os coeficientes de φ, continuaremos utilizando o simbolo δ para representar o grau do polinômio gerador:

$$\varphi(X) = \sum_{i=0}^{\delta} \varphi_i \cdot X^i,$$

segue que

$$\begin{array}{rcl} \dfrac{\varphi(X) - \varphi(\alpha)}{X - \alpha} & = & \sum_{i=0}^{\delta} \varphi_i \cdot \dfrac{X^i - \alpha^i}{X - \alpha} \\ & = & \sum_{i=0}^{\delta-1} \left(\sum_{j=i+1}^{\delta} \varphi_j \alpha^{j-i-1} \right) X^i, \quad 0 \leq i \leq \delta - 1. \end{array}$$

Pela definição de código de Goppa, $c \in K^n$ é uma palavra do código se, somente se

$$\sum_{i=0}^{n-1} c_i \varphi(\alpha_i)^{-1} \cdot \frac{\varphi(X) - \varphi(\alpha_i)}{X - \alpha_i} = 0.$$

Efetuando as devidas substituições

$$\begin{array}{rcl} \sum_{k=0}^{n-1} c_k \varphi(\alpha_k)^{-1} \cdot \dfrac{\varphi(X) - \varphi(\alpha_i)}{X - \alpha_i} & = & \sum_{k=0}^{n-1} c_k \varphi(\alpha_k)^{-1} \cdot \sum_{i=0}^{\delta-1} \left(\sum_{j=i+1}^{\delta} \varphi_j \alpha^{j-i-1} \right) X^i \\ & = & \sum_{i=0}^{\delta-1} \left(\sum_{k=0}^{n-1} \left(\varphi(\alpha_k)^{-1} \sum_{j=i+1}^{\delta} \varphi_j \alpha^{j-i-1} \right) c_i \right) X^i \\ & = & 0. \end{array}$$

Como este é um polinômio nulo, com todos os coeficientes iguais a 0. O segundo somatório é portanto igual a 0, ou seja

$$\sum_{k=0}^{n-1}\left(\varphi(\alpha_k)^{-1}\sum_{j=i+1}^{\delta}\varphi_j\alpha^{j-i-1}\right)c_i = 0$$

Escrevendo esta igualdade em forma matricial onde $k+1$ representa a coluna j e linha $i+1$, obtemos

$$\begin{bmatrix} \varphi(\alpha_0)^{-1}\varphi_\delta & \cdots & \varphi(\alpha_{n-1})^{-1}\varphi_\delta \\ \varphi(\alpha_0)^{-1}(\varphi_{\delta-1}+(\varphi_\delta\alpha_0) & \cdots & \varphi(\alpha_{n-1})^{-1}\varphi_{\delta-1}+\varphi_\delta\alpha_{n-1}) \\ \vdots & & \vdots \\ \varphi(\alpha_0)^{-1}\sum_{j=1}^{\delta}\varphi_j\alpha_0^{j-1} & \cdots & \varphi(\alpha_{n-1})^{-1}\sum_{j=1}^{\delta}\varphi_j\alpha_{n-1}^{j-1} \end{bmatrix}\cdot c^t = 0.$$

Como $gr(\varphi(X)) = \delta$, $\varphi_\delta \neq 0$. Além disto, $\varphi_\delta \in F$ um corpo finito, portanto $(\varphi_\delta)^{-1} \neq 0$. Podemos então efetuar operações elementares como multiplicar uma linha por $(\varphi_\delta)^{-1}$ ou substituir uma linha por ela mesma somada com o resultado da multiplicação de outra linha por $(\varphi_\delta)^{-1}$ na matriz ao qual estamos multiplicando c^t. Através destas operações podemos obter a igualdade

$$\begin{bmatrix} \varphi(\alpha_0)^{-1} & \cdots & \varphi(\alpha_{n-1})^{-1} \\ \varphi(\alpha_0)^{-1}\alpha_0 & \cdots & \varphi(\alpha_{n-1})^{-1}\alpha_{n-1} \\ \vdots & & \vdots \\ \varphi(\alpha_0)^{-1}\alpha_0^{\delta-1} & \cdots & \varphi(\alpha_{n-1})^{-1}\alpha_{n-1}^{\delta-1} \end{bmatrix}\cdot c^t = 0.$$

Por definição os coeficientes do polinômio de Goppa $\varphi(X)$ e os elementos do suporte L, pertencem ao corpo F, uma extensão de K. Portanto a matriz formada não é uma matriz teste de paridade, contudo escrevendo cada entrada desta matriz como vetor coluna com coeficiente em K, obtemos uma matriz H', tal que $c \in \Gamma_K \Leftrightarrow H'c^t = 0$, onde H' não é necessariamente uma matriz teste de paridade do código, pois suas linhas não são inevitavelmente linearmente independentes. A matriz H teste de paridade é obtida através da matriz H', retirando linhas linearmente dependentes.

Vejamos agora o formato de uma matriz geradora de um código de Goppa $\Gamma_F(L,\varphi)$. Por definição a função síndrome de uma palavra $c \in K^n$ é

$$\begin{aligned} S_c(X) &= \sum_{i=0}^{n-1}\frac{c_i}{X-\alpha_i} \\ &= \frac{\sum_{j=0}^{n-1}\left(c_j\cdot\frac{\prod_{k=0}^{n-1}(X-\alpha_k)}{X-\alpha_j}\right)}{\prod_{i=0}^{n-1}(X-\alpha_i)}. \end{aligned}$$

Seja $H(X) = \prod_{i=0}^{n-1}(X - \alpha_i)$, então

$$\begin{aligned} S_c(X) &= \sum_{i=0}^{n-1} \frac{c_i}{X - \alpha_i} \\ &= \frac{\sum_{j=0}^{n-1} \left(c_j \cdot \frac{H(X)}{X - \alpha_j} \right)}{H(X)} \end{aligned}$$

sabemos que

$$\begin{aligned} c \in \Gamma_F(L, \varphi) &\Leftrightarrow \sum_{i=0}^{n-1} \frac{c_i}{X - \alpha_i} \equiv 0 \mod \varphi(X) \\ &\Leftrightarrow \sum_{j=0}^{n-1} \left(c_j \cdot \frac{H(X)}{X - \alpha_j} \right) = Q(X)\varphi(X) \\ &\Leftrightarrow \sum_{j=0}^{n-1} c_j H'(X) = Q(X)\varphi(X) \end{aligned}$$

para algum $Q(X) \in F[X]$. Da ultima igualdade, temos para todo elemento $\alpha_i \in L$,

$$\begin{aligned} Q(\alpha_i)\varphi(\alpha_i) &= \sum_{j=0}^{n-1} c_j H'(\alpha_i) \\ &= c_i H'(\alpha_i). \end{aligned}$$

Dividindo ambos os lados da igualdade por $\varphi(\alpha_i)$ que é diferente de 0 por definição de código de Goppa, obtemos

$$Q(\alpha_i) = \frac{c_i H'(\alpha_i)}{\varphi(\alpha_i)}.$$

Além disto, $gr(\varphi(X)) = \delta$ e $gr(H(X)) = n - 1$, portanto $gr(Q(X)) \leq gr(H(X)) - gr(\varphi(X)) = n - 1 - \delta$, podemos então escrever $Q(X) = \sum_{k=0}^{n-\delta-1} q_k X^k$, resultando em

$$\sum_{k=0}^{n-\delta-1} q_k (\alpha_i)^k = \frac{c_i H'(\alpha_i)}{\varphi(\alpha_i)} \quad \Leftrightarrow \quad c_i = \frac{\varphi(\alpha_i)}{H'(\alpha_i)} \sum_{k=0}^{n-\delta-1} q_k (\alpha_i)^k = \sum_{k=0}^{n-\delta-1} q_k \cdot \frac{\varphi(\alpha_i)(\alpha_i)^k}{H'(\alpha_i)}.$$

Como este somatório representa o elemento de coordenada i de um vetor do código $\Gamma_F(L, \varphi)$, podemos interpretar o índice $k+1$ como as linhas e o índice $i+1$ como as colunas de uma matriz, desta forma $c \in \Gamma_F(L, \varphi)$ se, somente se, existe $\begin{bmatrix} q_0 & \cdots & q_{n-1-\delta} \end{bmatrix} \in F^{n-\delta}$ tal que

$$c = \begin{bmatrix} q_0 & \cdots & q_{n-1-\delta} \end{bmatrix} \cdot \begin{bmatrix} \frac{\varphi(\alpha_0)}{H'(\alpha_0)} & \cdots & \frac{\varphi(\alpha_{n-1})}{H'(\alpha_{n-1})} \\ \frac{\varphi(\alpha_0)\alpha_o}{H'(\alpha_0)} & \cdots & \frac{\varphi(\alpha_{n-1})\alpha_{n-1}}{H'(\alpha_{n-1})} \\ \vdots & & \vdots \\ \frac{\varphi(\alpha_0)\alpha_0^{n-\delta-1}}{H'(\alpha_0)} & \cdots & \frac{\varphi(\alpha_{n-1})\alpha_{n-1}^{n-\delta-1}}{H'(\alpha_{n-1})} \end{bmatrix}.$$

Concluímos então que

$$\begin{bmatrix} \frac{\varphi(\alpha_0)}{H'(\alpha_0)} & \cdots & \frac{\varphi(\alpha_{n-1})}{H'(\alpha_{n-1})} \\ \frac{\varphi(\alpha_0)\alpha_0}{H'(\alpha_0)} & \cdots & \frac{\varphi(\alpha_{n-1})\alpha_{n-1}}{H'(\alpha_{n-1})} \\ \vdots & & \vdots \\ \frac{\varphi(\alpha_0)\alpha_0^{n-\delta-1}}{H'(\alpha_0)} & \cdots & \frac{\varphi(\alpha_{n-1})\alpha_{n-1}^{n-\delta-1}}{H'(\alpha_{n-1})} \end{bmatrix} \text{ e } \begin{bmatrix} \varphi(\alpha_0)^{-1} & \cdots & \varphi(\alpha_{n-1})^{-1} \\ \varphi(\alpha_0)^{-1}\alpha_0 & \cdots & \varphi(\alpha_{n-1})^{-1}\alpha_{n-1} \\ \vdots & & \vdots \\ \varphi(\alpha_0)^{-1}\alpha_0^{\delta-1} & \cdots & \varphi(\alpha_{n-1})^{-1}\alpha_{n-1}^{\delta-1} \end{bmatrix}$$

são, respectivamente, a matriz geradora e a matriz teste de paridade de $\Gamma_F(L, \varphi)$.

Códigos de Goppa Binários Irredutíveis

Os códigos de Goppa são uma família de códigos cujas distâncias mínimas possuem cotas inferiores obtidas a priori, determinaremos agora a cota inferior para códigos de Goppa binários irredutíveis.

Definição 2.3.15. *(Polinômio Localizador).Sejam $c \in \mathbb{F}_2^n$ e $T_c = \{i : c_i \neq 0\}$. Chamamos de polinômio localizador de c:*

$$\sigma_c(X) = \prod_{i \in T_c} (X - \alpha_i).$$

O polinômio localizador de um vetor nulo é por definição 1.

Lema 2.3.16. *Seja $\sigma_c'(X)$ a derivada do polinômio localizador de um vetor $c \in \mathbb{F}_2^n$, vale a igualdade*

$$S_c(X) \cdot \sigma_c(X) = \sigma_c'(X).$$

Demonstração. Da multiplicação entre $S_c(X)$ e $\sigma_c(X)$, temos

$$\begin{aligned} S_c(X) \cdot \sigma_c(X) &= \sum_{i=0}^{n-1} \frac{c_i}{X - \alpha_i} \cdot \prod_{i \in T_c} (X - \alpha_i) \\ &= \sum_{i \in T_c} \frac{c_i}{X - \alpha_i} \cdot \prod_{i \in T_c} (X - \alpha_i) \\ &= \sum_{i \in T_c} \left(\frac{\prod_{j \in T_c} (X - \alpha_j)}{X - \alpha_i} \right) \\ &= \sigma_c'(X). \end{aligned}$$

□

Lema 2.3.17. *Seja $F(X) \in \mathbb{F}_{2^m}[X]$, onde $F(X) = \sum_{i=0}^{n} f_i X^i$. Vale a igualdade*

$$(F(X))^2 = \sum_{1=0}^{2n} (f_i)^2 X^{2i}.$$

Demonstração. Aplicando indução no número de termos de $F(X)$: Se $F(X)$ tem apenas um termo, $F(X) = X^a, a \in [0, n]$. Então $(F(X))^2 = X^{2a}, 2a \in [0, 2n]$. Se $F(X)$ tem

dois termos, $F(X) = X^a + X^b, a \neq b$ e $a, b \in [0, n]$. Então $(F(X))^2 = X^{2a} + 2X^{ab} + X^{2b} = X^{2a} + X^{2b}$, $2a \neq 2b$ e $2a, 2b \in [0, 2n]$. Vemos então que a igualdade é validade para $F(X)$ tendo um e dois termos, suponhamos que seja valida para $F(X)$ tal que $|T_{F(X)}| = |\{i \in [0, n] : f_i \neq 0\}| = k \in [1, n-1]$, temos que provar que a igualdade é valida para $F_{k+1}(X)$ com $k+1$ termos. Podemos escrever $F_{k+1}(X) = X^j + \sum_{i \in T_{F(X)}} f_i X^i, j \neq i$ e $j \in [0, n]$. Logo, seu quadrado é

$$\begin{aligned}
(F_{k+1}(X))^2 &= \left(X^j + \textstyle\sum_{i \in T_{F(X)}} f_i X^i\right)^2 \\
&= X^{2j} + 2X^j \textstyle\sum_{i \in T_{F(X)}} f_i X^i + \left(\sum_{i \in T_{F(X)}} f_i X^i\right)^2 \\
&= X^{2j} + \left(\textstyle\sum_{i \in T_{F(X)}} f_i X^i\right)^2 \\
&= X^{2j} + \textstyle\sum_{i \in T_{F(X)}} f_i^2 X^{2i} \\
&= \textstyle\sum_{i=0}^{2n} f_i^2 X^{2i}.
\end{aligned}$$

□

Teorema 2.3.18. *Os códigos de Goppa binários irredutíveis são uma família de código cujas distâncias mínimas possuem cotas inferiores iguais a $2\delta + 1$, onde δ é o grau do polinômio de Goppa.*

Demonstração. Seja $\Gamma_{\mathbb{F}_2}(L, \varphi)$ um código de Goppa com $gr(\varphi) = \delta$ e φ irredutível, então $\Gamma_{\mathbb{F}_2}(L, \varphi)$ é um código de Goppa binário irredutível. Pela definição de síndrome e pelo lema 2.3.16:

$$\begin{cases} c \in \Gamma_{\mathbb{F}_2}(L, \varphi) \Leftrightarrow S_c(X) \equiv 0 \mod \varphi(X) \\ c \in \mathbb{F}_2^n \Leftrightarrow S_c(X) \cdot \sigma_c(X) = \sigma_c'(X) \end{cases}$$

portanto

$$c \in \Gamma_{\mathbb{F}_2}(L, \varphi) \Leftrightarrow \sigma_c'(X) \equiv 0 \mod \varphi(X).$$

Deve ser observado que no corpo $\mathbb{F}_2$, $\overline{0} = \{2k : k \in \mathbb{Z}\}$, e pela definição de derivada de um polinômio os coeficientes dos termos de potências ímpar $2k-1, k \in \mathbb{N}$ são obtidos através de uma multiplicação por $2k$, sendo assim nulos. Portanto, podemos escrever

$$\sigma_c'(X) = \sum_{i=0}^{2i} o_i X^{2i}$$

onde $i < \frac{n}{2}$. Pelo lema 2.3.17, existe um polinômio $F(X)$ tal que $(F(X))^2 = \sigma_c'(X)$, onde $gr(F(X)) = i$. Além disto como $\varphi(X)$ é irredutível $\sigma_c'(X) \equiv 0 \mod \varphi(X) \Rightarrow F(X) \equiv 0 \mod \varphi(X)$. Suponhamos então que c seja a palavra de menor peso não nula do código $\Gamma_{\mathbb{F}_2}(L, \varphi)$, então $gr(F(X)) \geq gr(\varphi(X))$ e portanto

$$d(\Gamma_{\mathbb{F}_2}(L, \varphi)) = \omega(c) = gr(\sigma_c(X)) \geq gr(\sigma_c'(X)) + 1 = 2i + 1 \geq 2\delta + 1$$

pelo teorema 2.1.28 podemos concluir que os códigos de Goppa binários irredutíveis podem corrigir até $\left\lfloor \frac{d(\Gamma_{\mathbb{F}_2}(L,\varphi))-1}{2} \right\rfloor \geq \left\lfloor \frac{2\delta+1-1}{2} \right\rfloor = \left\lfloor \frac{2\delta}{2} \right\rfloor = \delta$ erros. □

Proposição 2.3.19. *A derivada de um polinômio localizador $\sigma_c(X)$ de um vetor $c \in \mathbb{F}_2^n$, pode ser escrito como*

$$\sigma_c'(X) = (\sigma_{c-impar}(X))^2$$

onde $\sigma_{c-impar}(X)$ é a raiz quadrada do resultado da soma dos termos de $\sigma_c(X)$ cujo expoente de X é impar divididos por X.

Demonstração. Primeiramente note que o valor do coeficiente de $\sigma_c(X)$ que multiplica X^0 não interfere no resultado de sua derivada, podemos então ignorar o valor deste coeficiente e escrever

$$\sigma_c(X) = X(\sigma_{c-impar}(X))^2 + (\sigma_{c-par}(X))^2$$

onde $(\sigma_{c-par}(X)$ é a raiz quadrada do resultado da soma dos termos de $\sigma_c(X)$ cujo expoente de X é par. Efetuando o calculo da derivada obtemos

$$\begin{aligned} \sigma_c'(X) &= 2(\sigma_{c-par}(X))(\sigma_{c-par}'(X)) + (\sigma_{c-impar}(X))^2 + 2X(\sigma_{c-impar}(X))(\sigma_{c-impar}'(X)) \\ &= (\sigma_{c-impar}(X))^2 \end{aligned}$$

□

Algoritmo de Patterson

Descreveremos um algorítimo específico chamado Algoritmo de Patterson, que decodifica eficientemente um código de Goppa binário irredutível, com capacidade de corrigir uma quantidade δ de erros igual ao grau do polinômio de Goppa utilizado na criação do código e com operações de simplicidade razoável. Este algoritmo foi proposto em 1975 por N.J. Patterson em [10]. Posteriormente foi sugerido por McEliece para ser utilizado no primeiro criptosistema baseado em códigos. É possível ver resultados quantitativos da eficiência de uma versão moderna do algoritmo de Patterson em [13] escrito por J. G. Vasquez, onde o mesmo apresenta a melhor performance em um dos teste.

Descreveremos agora o algoritmo de Patterson em sua primeira versão. Sejam $x \in \Gamma_{\mathbb{F}_2}(L,\varphi)$ uma palavra de um código de Goppa binário irredutível que foi enviada por um sistema de comunicação, $y \in \mathbb{F}_2^n$ a palavra recebida e $e \in \mathbb{F}_2^n$ o vetor erro, ou seja $y = x + e$. Pela definição de função síndrome, encontramos a síndrome de e através de y:

$$S_y(X) \equiv S_{x+e}(X) \equiv S_x(X) + S_e(X) \equiv S_e(X) \mod \varphi(X).$$

Se $S_e(X) \equiv 0 \mod \varphi(X)$ aceitamos $x = y$, caso contrário pela lema 2.3.16 temos:

$$\sigma_e(X) \cdot S_e(X) \equiv \sigma_e'(X) \mod \varphi(X)$$

onde utilizando a proposição anterior obtemos:

$$(X(\sigma_{e-impar}(X))^2 + (\sigma_{e-par}(X))^2) \cdot S_e(X) \equiv (\sigma_{e-impar}(X))^2 \mod \varphi(X)$$

aplicando a distributividade da multiplicação com relação a adição:

$$X(\sigma_{e-impar}(X))^2 S_e(X) + (\sigma_{e-par}(X))^2) S_e(X) \equiv (\sigma_{e-impar}(X))^2 \mod \varphi(X)$$

subtraindo $X(\sigma_{e-impar}(X))^2 S_e(X)$ de ambos os membros da congruência :

$$(\sigma_{e-par}(X))^2) S_e(X) \equiv (\sigma_{e-impar}(X))^2 - X(\sigma_{e-impar}(X))^2 S_e(X) \mod \varphi(X)$$

como $-1 = 1$ em $\mathbb{F}_2$:

$$(\sigma_{e-par}(X))^2) S_e(X) \equiv (\sigma_{e-impar}(X))^2 + X(\sigma_{e-impar}(X))^2 S_e(X) \mod \varphi(X)$$

colocando $(\sigma_{e-impar}(X))^2$ em evidencia no segundo membro da congruência:

$$(\sigma_{e-par}(X))^2) S_e(X) \equiv (\sigma_{e-impar}(X))^2 (1 + X S_e(X)) \mod \varphi(X)$$

multiplicando ambos os membros da congruência por $F(X)$, onde $F(X)S_e(X) \equiv 1 \mod \varphi(X)$:

$$\sigma_{e-par}(X))^2 \equiv (F(X) + X)(\sigma_{e-impar}(X))^2 \mod \varphi(X)$$

como $F(X) \neq 0$, $(F(X) + 1)$ possui raiz quadrada em $\mathbb{F}_2$, seja $\tau(X)$ esta raiz, obtemos

$$\sigma_{e-par}(X))^2 \equiv (\tau(X))^2 (\sigma_{e-impar}(X))^2 \mod \varphi(X)$$

calculando a raiz quadrada de ambos os membros da congruência:

$$\sigma_{e-par}(X) \equiv (\tau(X))(\sigma_{e-impar}(X)) \mod \varphi(X).$$

Sendo assim, através da síndrome $S_e(X)$ encontramos $F(X)$ e calculamos

$$\sqrt{F(X) + X} = \tau(X).$$

De posse de $\tau(X)$ e $\varphi(X)$ podemos identificar $\sigma_{e-par}(X)$ e $\sigma_{e-impar}(X)$ e obter através deles o polinômio localizador $\sigma_e(X)$.

Pela definição de polinômio localizador $\sigma_e(\alpha_i) = 0$ se, somente se, existe um erro na posição i do vetor y. Podemos então corrigir os erros de y somando 1 nas coordenadas i tais que $\sigma_e(\alpha_i) = 0$ e obter assim o vetor x.

Capítulo 3

Criptografia

Olhando historicamente para o ser humano podemos notar diversas tentativas de invenção de códigos secretos, visando cifrar as mensagens que serão transmitida para que não sejam inteligíveis por um interceptor. A criação de tais códigos é um problema muito difícil, pois o interceptor pode possuir capacidades de análise que o possibilita decifrar a mensagem.

O estudo dos princípios e técnicas para a criação de tais códigos secretos é chamado de Criptografia, enquanto a arte de tentar decifrar uma mensagem é a Criptoanálise. Podemos classificar a criptografia de acordo com os métodos utilizados, chamando assim de criptografia clássica aquela onde os recursos utilizados eram apenas a caneta e papel, ou dispositivos mecânicos simples, já após o surgimento de máquinas especializadas e dos computadores uma evolução significativa ocorreu nas técnicas de criptografia iniciando assim a era da criptografia moderna ou contemporânea.

3.1 Conceitos Básicos

O objetivo principal da criptografia é o de modificar mensagens para que possam ser usadas em uma comunicação segura na presença de terceiros. Em outras palavras é permitir que duas pessoas, normalmente denotados por Alice e Bob, possam se comunicar através de um canal inseguro de forma que um adversário, Eve, não consiga entender suas conversas. Para alcançar esse objetivo Alice utiliza uma chave pré definida para encriptar suas mensagens, assim caso ela seja interceptada por Eve, que não possui conhecimento da chave utilizada por Alice, não conseguira decifrá-la. Por outro lado, Bob, tendo conhecimento prévio da chave utilizada por Alice, decifra a mensagem e obtém o conteúdo original. São exemplos de canais inseguros muito utilizados durante o dia-a-dia das pessoas as redes de computadores e as linhas telefônicas. Estas ideias são representadas por D.R. Stinson em [12] de uma forma mais formal por meios de notações matemáticas como veremos a seguir.

Definição 3.1.1. *(Criptosistema). Um criptosistema é uma quíntupla* $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$*, onde as seguintes condições são satisfeitas:*

1. $\mathcal{P}$ *é um conjunto finito de possíveis textos simples de mensagens;*
2. $\mathcal{C}$ *é um conjunto finito de possíveis textos cifrados de mensagens;*
3. $\mathcal{K}$ *é um conjunto finito de possíveis chaves;*
4. *Para cada chave* $K \in \mathcal{K}$*, existe uma regra de encriptação* $e_K \in \mathcal{E}$ *e uma regra de decriptação correspondente* $d_K \in \mathcal{D}$*. Onde* $e_K : \mathcal{P} \to \mathcal{C}$ *e* $d_K : \mathcal{C} \to \mathcal{P}$ *são funções tais que* $d_K(e_k(x)) = x$ *para cada texto simples* $x \in \mathcal{P}$*.*

Na criptografia contemporânea supomos que o interceptor faça uso de um computador e um algoritmo para ajudá-lo na análise de uma forma de decifrar a mensagem levando assim, a necessidade de se usar técnicas de criptografia onde algoritmos de complexidade simples não possam decifrar o código. Ou seja, o algorítimo utilizado para criar as regras de encriptação d_K e de decriptação e_K deve ser baseado em um problema de complexidade suficientemente alta. Sendo o logaritmo discreto um dos problemas mais utilizados na criptografia moderna.

Definição 3.1.2. *(Logaritmo Discreto). Seja* $(G, \cdot)$ *um grupo cíclico finito com* n *elementos. Seja* α *um gerador do grupo* G*, ou seja* $G = \{e_G, \alpha^1, \ldots, \alpha^{n-1}\}$*. Dado um elemento* $y \in G$*, chamamos* $x_1 \in \mathbb{Z}$ *de logaritmo discreto de* y *na base* α *se* $\alpha^{x_1} = y$*.*

Sendo assim se x_2 também for um representante do logaritmo discreto de y na base α, então $x_1 \equiv x_2 \mod n$.

Exemplo 3.1.3. *Dado o corpo finito* K*. Definimos a ordem de* α *em* K^* *e denotamos* $ord_K\alpha$*, como sendo o menor* $x \in \mathbb{Z}_+$*, tal que* $\alpha^x = 1 \in K$ *e dizemos que* α *é um gerador em* K *se* $ord_K\alpha = |K| - 1$*. Sendo assim dado o corpo finito* $\mathbb{Z}_7$*, temos que* $ord_{\mathbb{Z}_7}3 = 6 = 7 - 1 = |\mathbb{Z}_7| - 1$ *e portanto* 3 *é um gerador do grupo multiplicativo* $\mathbb{Z}_7^*$*, ou seja* $\mathbb{Z}_7^* = \{1 = 3^6, 3^1, 3^2, 3^3, 3^4, 3^5\}$*, além disto tomando* $y \in \mathbb{Z}_7 \setminus \{0\}$*, podemos encontrar o logaritmo discreto de* y *na base* 3 *por meio de tentativa e erro. Tomando* $y = 4$ *podemos encontrar um logaritmo discreto* $x_1 \in \mathbb{Z}^*$*, tal que* $3^x = 4$ *por meio de tentativa e erro, uma destas soluções é* $x_1 = 4$*, e podemos generalizar que todo logaritmo discreto de* 4 *na base* 3 *em* $\mathbb{Z}_7$ *é da forma* $x = 4 + 7k, k \in \mathbb{Z}$*.*

Os protocolos Diffie-Hellman para troca de chaves e RSA (que utiliza a primeira letra do nome de seus inventores Rivest, Shamir e Adleman) são exemplos de algoritmos baseados no problema do logaritmo discreto. Este problema é considerado de complexidade suficientemente alta devido a dificuldade computacional envolvida no processo de fatoração de um número grande.

Um criptossistema pode ser chamada de chave simétrica ou assimétrica(privada ou pública), de acordo com sua regra de encriptação e_K e decriptação d_K. Para compreender essa diferença é necessário entender quando um algoritmo possui complexidade de tempo polinomial ou exponencial.

Definição 3.1.4. *(Complexidade de Tempo). Quantificação do número de operações básicas necessárias para rodar um determinado algoritmo em relação ao tamanho da entrada.*

A complexidade de tempo é representada quando descrita assintoticamente por O-grande, que apresenta o máximo de operações do algoritmo em relação ao tamanho da entrada . É chamado de tempo polinomial e considerado eficiente quando para uma entrada de tamanho n, temos $O(n^k), k \leq 10$. Por outro lado $O(k^n), k > 1$ é chamado de tempo exponencial e considerado intratável.

Criptosistemas com chave simétrica

Quando d_K é igual ou determinada facilmente a partir de e_K em tempo polinomial por algum algoritmo, dizemos que o criptosistema que utiliza d_K e e_K possui chaves simétricas. Nestes criptosistemas a descoberta de d_K ou e_K por um interceptor torna o sistema inseguro, sendo assim a troca da chave K deve ser realizada previamente por um canal seguro o que pode ser muito difícil de se conseguir, principalmente se Alice e Bob não podem se encontrar fisicamente. O protocolo Diffie-Hellman de troca de chaves serviu como primeiro algorítimo a conseguir introduzir na prática o conceito de criptografia assimétrica, desenvolvido por Whifiel Diffie e Martin Hellman e publicado em 1976, este algoritmo é utilizado exclusivamente para troca de chaves.

Algoritmo Diffie-Hellman:

- Alice e Bob devem escolher um número p primo e um gerador α de $\mathbb{Z}_p^*$.
- Alice escolhe um número a confidencial e aleatório para enviar o resultado $\alpha^a \mod p$ por uma rede insegura para Bob.
- Bob escolhe um número b confidencial e aleatório para enviar o resultado $\alpha^b \mod p$ por uma rede insegura para Alice.
- Alice e Bob elevam o número que receberam pelo seus respectivos números confidenciais e calculam o resultado módulo p, obtendo assim uma chave K compartilhada.

Exemplo 3.1.5. • *Alice e Bob escolhem o número* 17 *e o gerador* 3 *de* $\mathbb{Z}_{17}^*$.

- *Alice escolhe de forma aleatória o número* 7 *para calcular* $3^7 \mod 17 = 11$ *e envia o número* 11 *para Bob.*
- *Bob escolhe de forma aleatória o número* 12 *para calcular* $3^{12} \mod 17 = 4$ *e envia o número* 4 *para Alice.*
- *Alice e Bob obtém a chave* $K = 11^{12} \mod 17 = 4^7 \mod 17 = 13$ *compartilhada ao elevar o número que receberam pelo seus respectivos números confidenciais.*

Observe que a chave compartilhada K sempre será igual para Alice e Bob, pois

$$(\alpha^a \mod p)^b \mod p = \alpha^{ab} \mod p = (\alpha^b \mod p)^a \mod p.$$

Os criptosistemas com chave simétrica não são considerados os mais simples e eficientes, mesmo quando utilizando um protocolo de troca de chaves, pois quando Alice desejar se comunicar com muitas pessoas, será necessário gerir diversas chaves e enviar diversas mensagens para estabelecer chaves diferentes para cada um de seus contatos.

Criptosistemas de chaves assimétricas

Nos criptosistemas de chaves assimétricas, também conhecidos por criptosistemas assimétricos, as regras de encriptação e decriptação e_K e d_K são diferentes e determinar d_K a partir de e_K é inviável, entretanto existe uma informação especial K chamada de armadilha ou segredo, que quando utilizado possibilita inverter a função e_K e obter a função d_K de forma fácil.

Neste modelo de criptografia cada parte envolvida na comunicação utiliza duas chaves assimétricas e complementares e_K e d_K, sendo e_K pública e d_K privada, a regra e_K deve ser fácil de ser computada e difícil de ser invertida, as funções com estas características são chamadas de funções de via única.

Devemos observar que um criptosistema de chave pública não garante segurança incondicional, pois um adversário pode utilizar a chave pública e_K para encriptar todos os textos simples $x \in \mathcal{P}$ até que seja encontrado o texto cifrado $y \in \mathcal{C}$ correspondente ao que foi interceptado. Por outro lado, o sistema de criptografia ser público é um benefício, pois a troca de detalhes de um sistema simétrico possui o maior perigo de interceptação, perigo este que não existe em criptosistemas de chaves públicas, os tornando assim a melhor abordagem para situações com diversos usuários. Outra vantagem dos criptosistemas de chaves públicas é permitir uma "assinatura", de forma que o receptor possa estar certo da integridade do emissor.

Assinando uma mensagem.

Sejam e_A e d_A, respectivamente, as chaves pública e privada de Alice e sejam e_B e d_B, respectivamente, as chaves pública e privada de Bob. Suponhamos então que Alice deseja enviar a mensagem $x \in \mathcal{P}$ assinada para Bob, então:

- Alice assina sua mensagem com sua chave privada: $d_A(x)$.
- Alice codifica sua assinatura com a chave pública de Bob: $e_B(d_A(x))$. E envia para Bob.
- Bob realiza a desencriptação utilizando sua chave privada: $d_B(e_B(d_A(x))) = d_A(x)$.
- Com a chave pública de Alice, Bob verifica a mensagem assinada: $e_A(d_A(x)) = x$.

A segurança dos criptosistemas são baseadas na dificuldade de resolver algum determinado problema matemático, dentre estes problemas a fatoração de inteiros e o logaritmo discreto são extensivamente usados atualmente como base de sistemas criptográficos para transmitir dados na internet.

3.2 Criptografia Baseada em Códigos Lineares

Esta seção tem como objetivo apresentar os principais conceitos envolvidos na criptografia baseada em códigos corretos de erros e apresentar o criptosistema de chave pública de McEliece, o primeiro criptosistema baseado em códigos corretores de erros.

Criptografia Pós-Quântica.

Com o desenvolvimento da computação quântica, fornecendo ferramental teóricos, em 1994 foi formulado o algoritmo de Shor para fatoração de inteiros, que utiliza computador quântico e possui complexidade de tempo polinomial. Como muitos criptosistemas atuais dependem da dificuldade em resolver estes problemas, criptólogos tentam conseguir avanços em métodos criptográficos não-quânticos que sejam resistentes a ataques de algoritmos quânticos como o de Shor, esta área é chamada de Criptografia Pós-Quântica.

Em 1978, Berlekamp, McEliece e Tilborg mostram em [1] que os problemas de decodificação geral e de encontrar o peso de um código linear são NP-difícil. Isto sugere fortemente, mas não necessariamente implica que não existe algoritmo com complexidade de tempo polinomial adequado para resolver estes problemas. Sendo assim, enquanto não descoberto um algoritmo clássico ou quântico de complexidade de tempo polinomial capaz de resolver estes problemas, toda criptografia baseada em código linear pode ser

considerada Pós-Quântica.

O Criptosistema de chave pública de McEliece

A primeira proposta de um criptosistema baseado em códigos corretores foi sugerida por Robert J. McEliece em 1978 no artigo [8], recomendando a utilização de código de Goppa binário aleatório para gerar um criptosistema onde a chave privada usa-se do código de Goppa devido a existência de algoritmos de codificações eficiente e a chave pública é obtida a partir da chave privada através de um processo que disfarça o código de Goppa em um código linear genérico, através da multiplicação de sua matriz geradora por duas matrizes invertíveis geradas por um processo aleatório.

É necessário destacar os seguintes valores, sendo o código de Goppa binário aleatório utilizado no sistema com parâmetros (n, k) e escrito sobre um corpo finito K, nossos textos simples de mensagens tem k bits, enquanto que os textos cifrados de mensagens tem n bits. O parâmetro δ representa o número máximo de erros que podem ser introduzidos no texto durante a codificação, sendo que o código de Goppa binário irredutível corrigi até δ erros quando o grau do polinômio irredutível utilizado para sua construção é δ.

Para criar a chave pública e privada do criptosistema os seguintes elementos são necessários:

G : Uma matriz $k \times n$ geradora do código $\mathcal{C}$ de Goppa binário aleatório.

S : Qualquer matriz $k \times k$ invertível.

P : Qualquer matriz $n \times n$ de permutação.

D : Um algoritmo de decodificação eficiente, com capacidade de corrigir até δ erros do código $\mathcal{C}$. Juntamente com as informações necessárias para seu devido funcionamento.

Com estes componentes criamos a matriz $G_{pub} = S \cdot G \cdot P$ que mascara a estrutura algébrica da matriz G, para que ela possa ser utilizada de forma segura, como resultado o código de Goppa fica disfarçado em um código linear genérico. Podemos agora descrever um criptosistema de McEliece:

- Chave Pública: (G_{pub}, δ)
- Chave Privada: (S, P, D)
- Encriptação: Dado o texto simples de mensagem $x \in K^k$, escolha aleatoriamente um vetor $e \in K^n$, com $\omega(e) = \delta$ e calcule o texto cifrado de mensagem y como:

$$y = xG_{pub} + e$$

- Decriptação: Dado o texto cifrado de mensagem $y \in K^n$, calcule

$$yP^{-1} = xSG + eP^{-1}.$$

  Como $\omega(e) = \omega(eP^{-1}) = \delta$ e $S \cdot G$ gera o mesmo subespaço que G, pois suas linhas são uma combinação linear das linhas de G. Utilize o algoritmo D para remover eP^{-1} e obtenha xS. Calcule $xSS^{-1} = x$ e finalmente obtenha x.

Este sistema pode ser facilmente implementado, pois se baseia em operações elementares de álgebra linear como multiplicação de vetores por matriz. Porém o tamanho da chave pública é uma desvantagem na utilização de um sistema de criptografia de McEliece, esta desvantagem pode ser mitigada utilizando diversas estrategias para redução de chaves, como a utilização de códigos quase-diádicos que foi sugerida em [9].

Capítulo 4

Aplicações na Educação Básica

Este capítulo tem como objetivo apresentar propostas de exemplos: Criptosistemas com chave simétrica a se apresentar para discentes do primeiro ano do ensino médio; Um código de Hamming $\mathcal{C}(7,4)$ e criptosistema baseado em código linear que utiliza o exemplo anterior de código de Hamming $\mathcal{C}(7,4)$ a se apresentar para discentes do segundo ano do ensino médio.

Criptosistemas com chave simétrica no ensino médio

Para uma melhor definição de criptografia os alunos necessitam compreender sobre funções, esta matéria faz parte do conteúdo de álgebra estudado no primeiro ano do ensino médio, sendo o conceito de função um dos mais importantes conceitos da matemática e das ciências em geral. Apresentaremos algumas definições que devem ser de conhecimento dos discentes:

Definição 4.0.1. *(Função de X em Y). Dados dois conjuntos não-vazios X e Y, uma regra f que associa cada elemento $x \in X$ a um único elemento $y \in Y$ é chamada uma função de X em Y. Denotamos*

$$\begin{array}{rccc} f: & X & \longrightarrow & Y \\ & x & \mapsto & y = f(x) \end{array}$$

O conjunto X é chamado de domínio da função e o conjunto Y, contradomínio da função. Cada $x \in X$ é associado a um $y \in Y$ chamado de imagem de x pela função f e representado por $f(x)$. Chamamos de conjunto imagem da função f o conjunto $Im(f) = \{f(x) : x \in X\}$.

Definição 4.0.2. *(Função Injetiva). Uma função $f : X \to Y$ é injetiva quando os elementos diferentes de X são transformados por f em elementos diferentes de Y. Ou*

seja, f é injetiva quando

$$\forall\ x, y \in X, x \neq y \Rightarrow f(x) \neq f(y).$$

Ou pela forma contrapositiva

$$\forall\ f(x), f(y) \in Y, f(x) = f(y) \Rightarrow x = y.$$

Definição 4.0.3. *(Função Sobrejetiva). Uma função $f : X \to Y$ é sobrejetiva quando para qualquer elemento $y \in Y$, pode-se encontrar ao menos um elemento $x \in X$ tal que $f(x) = y$.*

Definição 4.0.4. *(Função Bijetiva). Uma função $f : X \to Y$ é chamada de bijetiva se for, simultaneamente, injetiva e sobrejetiva.*

Se existe uma função $f : X \to Y$ bijetiva, dizemos que existe uma bijeção ou uma correspondência biunívoca entre X e Y.

Definição 4.0.5. *(Função Composta). Dadas as funções $f : X \to Y$ e $g : Y \to Z$, chamamos de composta de g e f a função $g \circ f : X \to Z$, onde $(g \circ f)(x) = g(f(x))$.*

Definição 4.0.6. *(Função Identidade). Uma função identidade Id tem como imagem o mesmo elemento do domínio, ou seja:*

$$\begin{array}{rccc} Id: & X & \to & X \\ & x & \mapsto & x \end{array}$$

Definição 4.0.7. *(Função Inversa). Uma função $g : Y \to X$ é a inversa da função bijetiva $f : X \to Y$ quando as composta $g \circ f$ e $f \circ g$ são funções identidades.*

É usual utilizamos f^{-1} para representar a função inversa de f. O processo apresentado nos livros didáticos do ensino médio para se encontrar uma função inversa de uma função bijetiva $f : X \to Y$, consiste em colocar x em evidência no primeiro membro da igualdade na fórmula da função f e substituir $f(x)$ por y e x por $f^{-1}(y)$.

Exemplo 4.0.8. *Dada a função bijetiva*

$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & -2x + 3 \end{array},$$

temos que

$$\begin{array}{rcl} f(x) = -2x + 3 & \Leftrightarrow & -2x = f(x) - 3 \\ & \Leftrightarrow & 2x = 3 - f(x) \\ & \Leftrightarrow & x = \dfrac{3 - f(x)}{2} \end{array}$$

substituindo $f(x)$ *por* y *e* x *por* $f^{-1}(y)$*, encontramos* $f^{-1}(y) = \dfrac{3-y}{2}$*, obtendo assim a função inversa*

$$\begin{array}{rccc} f^{-1}: & \mathbb{R} & \to & \mathbb{R} \\ & y & \mapsto & \dfrac{3-y}{2} \end{array}.$$

Observe que de fato f^{-1} *é a inversa de* f*, pois*

$$(f^{-1} \circ f)(x) = \frac{3-(-2x+3)}{2} = \frac{2x}{2} = x$$

e

$$(f \circ f^{-1})(y) = -2\frac{3-y}{2} + 3 = \frac{-6+2y}{2} + 3 = -3 + y + 3 = y$$

portanto $f^{-1} \circ f$ *e* $f \circ f^{-1}$ *são funções identidade.*

Tendo conhecimento destas definições podemos então explicar para os discentes o que é um criptosistema, utilizando uma função bijetiva $f : X \to Y$ qualquer como regra de encriptação. Chamamos de textos simples de mensagens, que podem ser lidos e entendidos os elementos do conjunto X domínio da função f, os textos cifrados de mensagens são elementos do conjunto $Im(f) = Y$. A função inversa f^{-1} é nossa regra de decriptação, pois aplicando ela em um texto cifrado de mensagem $y \in Y$ encontramos $x \in X$ que foi utilizado para obter y. Daremos a seguir um exemplo de dinâmica que pode ser trabalhada a partir do primeiro ano do ensino médio, após esta explicação sobre criptosistema.

Exemplo 4.0.9. *Separando a classe em uma quantidade par de grupos, suponha uma situação onde que se deseja enviar mensagens de forma secreta. Um exemplo de situação seria enviar mensagens a um carro-forte informando a direção em que ele deve ir, dado* $\{0,1\}$ *o conjunto dos textos simples de mensagens onde* 0 *representa que o carro deve virar a sua direita e* 1 *representa que o carro deve virar a sua esquerda. Informe uma função bijetiva de fácil inversão e uma sequência de textos simples diferente para metade dos grupos e os peça para que utilizando a função calcule a sequência de textos cifrados correspondente a sequência de textos simples, exemplo:*

$$\begin{array}{rccc} f_1: & \{0,1\} & \to & \{2,5\} \\ & x & \mapsto & 3x+2 \end{array}, \quad \begin{array}{rccc} f_2: & \{0,1\} & \to & \{2,5\} \\ & x & \mapsto & \begin{cases} x+5 \ se & x=0 \\ x+1 \ se & x=1 \end{cases} \end{array}, ...$$

onde suponhamos que um grupo recebe a função f_1 *e a sequência de textos simples* $0,0,1,1,0$*, os alunos deste grupo podem obter facilmente a sequência de textos cifrados* $2,2,5,5,2$*.*

Para a outra metade dos grupos entregue as mesmas funções e peça para que encontrem suas inversas, assim o grupo que receber f_1 *encontra* f_1^{-1} *e o grupo que receber*

f_2 encontra f_2^{-1}:

$$f_1^{-1}: \begin{array}{ccc} \{2,5\} & \to & \{0,1\} \\ x & \mapsto & \dfrac{x-2}{3} \end{array}, \qquad f_2^{-1}: \begin{array}{ccc} \{2,5\} & \to & \{0,1\} \\ x & \mapsto & \begin{cases} x-5 \ se \ \ x=5 \\ x-1 \ se \ \ x=2 \end{cases} \end{array}, ...$$

Chame um dos grupos que calculou uma sequência de textos cifrados por vez para ir ao quadro e mostra a sequência de textos que seria enviada, peça aos demais grupos que calcularam as funções inversas para tentar decodificar as mensagens cifradas e obter as mensagens simples. Nesta situação apenas o grupo que recebeu a mesma função do grupo que codificou consegue decodificar de forma correta as mensagens.

Espera-se com esta dinâmica fornecer a turma um conceito básico sobre criptosistema, outra coisa que pode ser abordada com a turma é o fato das funções f_i escolhida como regra de encriptação serem facilmente invertidas precisando assim ficar escondidas. Podemos informar de uma forma inteligível para os discentes que todo criptosistema que utiliza uma função de fácil inversão é considerado de chave privada, sua regra de encriptação f_i deve ser mantida em segredo e deve ser comunicado de forma segura a quem se deseja ter acesso a informação codificada. Além disto, caso se deseja comunicar com várias pessoas, para cada pessoa será necessário uma regra de encriptação diferente. Por estes motivos, os criptosistemas mais utilizados na prática são os de chave pública, onde a regra de encriptação não é invertida de forma fácil.

Código de Hamming $\mathcal{C}(7,4)$ no ensino médio

Os códigos de Hamming $\mathcal{C}(7,4)$ são modernos e $1-$perfeito, podendo então sempre corrigir mensagens ao qual foi introduzido um único erro. Porém para que se possa discutir sobre códigos de Hamming deve-se saber trabalhar com o corpo finito de dois elementos $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$. Corpo este não relacionado a grade curricular do ensino médio. Sugere-se então que este conteúdo seja apresentado em uma oficina extraclasse a alunos interessados. Como pré-requisito para participar da oficina o estudante deverá ter conhecimentos sobre matrizes e sistemas lineares presentes na álgebra estudada no segundo ano do ensino médio. Apresentaremos algumas definições que devem ser de conhecimento dos discentes:

Definição 4.0.10. *(Matriz). Dados dois números $m, n \in \mathbb{N}$. Chamamos de matriz $m \times n$ uma tabela retangular com $m \cdot n$ números, dispostos em m linhas e n colunas.*

Em uma matriz qualquer A do tipo $m \times n$, cada elemento é indicado por a_{ij}, onde o índice $i \in \{1, \ldots, m\}$ indica a linha e o índice $j \in \{1, \ldots, n\}$ a coluna na qual o elemento se encontra. As linhas são numeradas de cima para baixo e as colunas da esquerda para direita, sendo assim A é representada de forma genérica por:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \text{ ou } \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

e pode também ser chamada de matriz $A = (a_{ij})_{m\times n}$. Algumas matrizes recebem nomes especiais conforme suas características, como:

- Matriz - linha: é toda matriz que tem uma única linha, sendo assim do tipo $1 \times n$.
- Matriz - coluna: é toda matriz que tem uma única coluna, sendo assim do tipo $m \times 1$.
- Matriz - nula: é toda matriz $A = (a_{ij})_{m\times n}$, tal que $a_{ij} = 0$.
- Matriz - quadrada: é toda matriz que possui quantidade igual de linhas e colunas, sendo assim do tipo $n \times n$ e chamada de ordem n.

Uma matriz quadrada $A = (a_{ij})_{n\times n}$, possui conjuntos de elementos que são denominados:

- Diagonal principal - são os elementos que possuem dois índices iguais, ou seja $\{a_{11}, a_{22}, \ldots, a_{nn}\}$.
- Diagonal secundária - são os elementos que têm a soma dos índices igual a $n + 1$, ou seja $\{a_{1n}, a_{2,n-1}, \ldots, a_{n1}\}$.

Ainda falando sobre as matrizes quadradas, podemos chama-las de:

- Matriz - triangular: se todos os elementos acima ou abaixo da diagonal principal são nulos.
- Matriz - diagonal: se todos os elementos acima e abaixo da diagonal principal são nulos.
- Matriz - identidade: é toda matriz diagonal em que todos os elementos da diagonal principal são iguais a 1, uma matriz identidade de ordem n é representada por I_n.

Definição 4.0.11. *(Igualdade de Matrizes). Dadas as matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$, dizemos que A e B são iguais se, somente se, $a_{ij} = b_{ij}$ para todo $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, n\}$.*

Definição 4.0.12. *(Soma de Matrizes). Dadas as matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$, chamamos de adição ou soma da matriz A com a matriz B, que representamos por $A+B$, a matriz $C = (c_{ij})_{m\times n}$, tal que $c_{ij} = a_{ij} + b_{ij}$.*

Definição 4.0.13. *(Oposta de A). Dado uma matriz $A = (a_{ij})_{m\times n}$, chamamos de oposta de A e representamos por $-A$, a matriz tal que $A + (-A) = 0$, onde 0 é uma matriz nula de ordem $m \times n$.*

Definição 4.0.14. *(Subtração de Matrizes). Dadas as matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$, chamamos de diferença ou subtração da matriz A com a matriz B, que representamos por $A - B$, a soma da matriz A com a matriz oposta de B.*

Definição 4.0.15. *(Transposta de A). Dado uma matriz $A = (a_{ij})_{m\times n}$, chamamos de transposta de A e representamos por $A^t = (a'_{ji})_{n\times m}$, a matriz tal que $a'_{ji} = a_{ij}$.*

Definição 4.0.16. *(Produto de Matrizes). Dadas as matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{jk})_{n\times p}$, chamamos de produto ou multiplicação da matriz A com a matriz B, que representamos por AB, a matriz $C = (c_{ik})_{m\times p}$, tal que $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$.*

Definição 4.0.17. *(Equação Linear). Chamamos de equação linear, nas incógnitas $x_1, x_2, \ldots, x_n$, toda equação que pode ser escrita na forma*

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = a$$

onde os números $\lambda_1, \lambda_2, \ldots, \lambda_n$ são coeficientes das incógnitas e a é o termo independente.

Definição 4.0.18. *(Sistema Linear). Chamamos de sistema linear $m \times n$ um conjunto S de $m \geq 1$ equações, nas incógnitas $x_1, x_2, \ldots, x_n$, que pode ser escrito na forma*

$$S = \begin{cases} \lambda_{11}x_1 + \lambda_{12}x_2 + \ldots + \lambda_{1n}x_n = a_1 \\ \lambda_{21}x_1 + \lambda_{22}x_2 + \ldots + \lambda_{2n}x_n = a_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \lambda_{m1}x_1 + \lambda_{m2}x_2 + \ldots + \lambda_{mn}x_n = a_m \end{cases}$$

onde os números λ_{ij} são coeficientes das incógnitas e a_i os termos independentes, para todo $i \in \{1, \ldots, m\}$ e $j \in \{1 \ldots, n\}$.

Pela definição de produto de matrizes, podemos escrever este sistema linear de forma matricial como

$$\begin{bmatrix} \lambda_{11} & \lambda_{12} & \ldots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \ldots & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m1} & \lambda_{m2} & \ldots & \lambda_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

A definição de corpo não é apresentada a discentes da educação básica, porém as operações de adição e multiplicação são trabalhadas em $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$. Meramente não é utilizado a palavra corpo, com base neste fato que prova a capacidade cognitiva dos

estudantes de ensino médio, propomos fornecer as regras de adição e multiplicação sobre o conjunto dos dígitos binários $\{0,1\}$ que satisfazem as mesmas regras que são satisfeitas pelos corpos $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$ e $(\mathbb{C},+,\cdot)$, através das tabelas de soma e multiplicação já definidas no exemplo 1.0.7.

Conceitos básicos de teoria da informação devem ser transmitidos como o que é um código, alfabeto, palavras e sistema de comunicação, para que a definição de código corretor de erros fique bem compreendida e exemplos de aplicações possam ser apresentados junto a definição de sistema de comunicação. Estas definições e alguns exemplos podem ser encontrados no capítulo 2 na seção 1.

Podemos agora apresentar um exemplo de um código de Hamming $\mathcal{C}(7,4)$. Este código é obtido por uma conversão do código canal $\mathcal{C}_c = \mathbb{Z}_2^4$ com palavras de comprimento 4, escritas com o alfabeto $\mathbb{Z}_2$, ao adicionarmos mais 3 símbolos que são redundantes e que permitem o receptor a corrigir um erro quando único. Mas como adicionamos esses símbolos? As palavras do código canal $\mathcal{C}_c$ podem ser apresentadas para os discentes como as matrizes linha:

$$\begin{bmatrix}0&0&0&0\end{bmatrix},\begin{bmatrix}1&0&0&0\end{bmatrix},\begin{bmatrix}0&1&0&0\end{bmatrix},\begin{bmatrix}0&0&1&0\end{bmatrix},\begin{bmatrix}0&0&0&1\end{bmatrix},\begin{bmatrix}1&1&0&0\end{bmatrix},$$
$$\begin{bmatrix}1&0&1&0\end{bmatrix},\begin{bmatrix}1&0&0&1\end{bmatrix},\begin{bmatrix}0&1&1&0\end{bmatrix},\begin{bmatrix}0&1&0&1\end{bmatrix},\begin{bmatrix}0&0&1&1\end{bmatrix},\begin{bmatrix}1&1&1&0\end{bmatrix},$$
$$\begin{bmatrix}1&1&0&1\end{bmatrix},\begin{bmatrix}1&0&1&1\end{bmatrix},\begin{bmatrix}0&1&1&1\end{bmatrix}\text{ e }\begin{bmatrix}1&1&1&1\end{bmatrix}.$$

Limitando assim o vocabulário a apenas 16 palavras. Seja $\begin{bmatrix}x_1&x_2&x_3&x_4\end{bmatrix} \in \mathcal{C}_c$ uma palavra que se deseja transmitir, em seu lugar transmitiremos a palavra

$$\begin{bmatrix}y_1&y_2&y_3&y_4&y_5&y_6&y_7\end{bmatrix} \in \mathcal{C}(7,4)$$

onde

$$\begin{aligned}
y_1 &= x_1;\\
y_2 &= x_2;\\
y_3 &= x_3;\\
y_4 &= x_4;\\
y_5 &= x_1+x_2+x_3;\\
y_6 &= x_1+x_2+x_4;\\
y_7 &= x_1+x_3+x_4.
\end{aligned}$$

Pela definição de produto de matrizes, podemos transformar $\begin{bmatrix}x_1&x_2&x_3&x_4\end{bmatrix} \in \mathcal{C}_c$ em $\begin{bmatrix}y_1&y_2&y_3&y_4&y_5&y_6&y_7\end{bmatrix} \in \mathcal{C}(7,4)$, através do produto

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 \end{bmatrix}$$

A matriz

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

recebe o nome de matriz geradora. Resta explicar como o receptor consegue corrigir um erro quando único. Suponhamos então que o canal utilizado para transportar a mensagem do emissor ao receptor cause no máximo um erro a palavra enviada e que a matriz linha $Z = \begin{bmatrix} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 \end{bmatrix}$ foi recebida pelo receptor, há então oito possibilidades consideradas pelo receptor:

i. Todos os símbolos estão corretos;

ii. z_1 está incorreto;

iii. z_2 está incorreto;

iv. z_3 está incorreto;

v. z_4 está incorreto;

vi. z_5 está incorreto;

vii. z_6 está incorreto;

viii. z_7 está incorreto.

O receptor pode determinar qual possibilidade se verifica

i. Se $z_5 = z_1 + z_2 + z_3$, $z_6 = z_1 + z_2 + z_4$ e $z_7 = z_1 + z_3 + z_4$, neste caso podemos observar na tabela de soma em $\mathbb{Z}_2$ que $0 + 0 = 0$ e $1 + 1 = 0$, portanto dois elementos iguais somados resulta sempre em 0, temos então que $z_1 + z_2 + z_3 + z_5 = 0$, $z_1 + z_2 + z_4 + z_6 = 0$ e $z_1 + z_3 + z_4 + z_7 = 0$. Podemos escrever este caso como o seguinte produto entre

matrizes:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

ii. Se $z_5 \neq z_1+z_2+z_3$, $z_6 \neq z_1+z_2+z_4$ e $z_7 \neq z_1+z_3+z_4$, neste caso podemos observar na tabela de soma em $\mathbb{Z}_2$ que $0+1 = 1$ e $1+0 = 1$, portanto dois elementos diferentes somados resulta sempre em 1, temos então que $z_1+z_2+z_3+z_5 = 1$, $z_1+z_2+z_4+z_6 = 1$ e $z_1 + z_3 + z_4 + z_7 = 1$. Podemos escrever este caso como o seguinte produto entre matrizes:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

iii. Se $z_5 \neq z_1 + z_2 + z_3$, $z_6 \neq z_1 + z_2 + z_4$ e $z_7 = z_1 + z_3 + z_4$, consequentemente $z_1 + z_2 + z_3 + z_5 = 1$, $z_1 + z_2 + z_4 + z_6 = 1$ e $z_1 + z_3 + z_4 + z_7 = 0$, portanto:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

iv. Se $z_5 \neq z_1 + z_2 + z_3$, $z_6 = z_1 + z_2 + z_4$ e $z_7 \neq z_1 + z_3 + z_4$, consequentemente

$z_1 + z_2 + z_3 + z_5 = 1$, $z_1 + z_2 + z_4 + z_6 = 0$ e $z_1 + z_3 + z_4 + z_7 = 1$, portanto:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

v. Se $z_5 = z_1 + z_2 + z_3$, $z_6 \neq z_1 + z_2 + z_4$ e $z_7 \neq z_1 + z_3 + z_4$, consequentemente $z_1 + z_2 + z_3 + z_5 = 0$, $z_1 + z_2 + z_4 + z_6 = 1$ e $z_1 + z_3 + z_4 + z_7 = 1$, portanto:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

vi. Se $z_5 \neq z_1 + z_2 + z_3$, $z_6 = z_1 + z_2 + z_4$ e $z_7 = z_1 + z_3 + z_4$, consequentemente $z_1 + z_2 + z_3 + z_5 = 1$, $z_1 + z_2 + z_4 + z_6 = 0$ e $z_1 + z_3 + z_4 + z_7 = 0$, portanto:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

vii. Se $z_5 = z_1 + z_2 + z_3$, $z_6 \neq z_1 + z_2 + z_4$ e $z_7 = z_1 + z_3 + z_4$, consequentemente

$z_1 + z_2 + z_3 + z_5 = 0$, $z_1 + z_2 + z_4 + z_6 = 1$ e $z_1 + z_3 + z_4 + z_7 = 0$, portanto:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

viii. Se $z_5 = z_1 + z_2 + z_3$, $z_6 = z_1 + z_2 + z_4$ e $z_7 \neq z_1 + z_3 + z_4$, consequentemente $z_1 + z_2 + z_3 + z_5 = 0$, $z_1 + z_2 + z_4 + z_6 = 0$ e $z_1 + z_3 + z_4 + z_7 = 1$, portanto:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \end{bmatrix} = \begin{bmatrix} z_1 + z_2 + z_3 + z_5 \\ z_1 + z_2 + z_4 + z_6 \\ z_1 + z_3 + z_4 + z_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

A matriz

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

recebe o nome de matriz teste de paridade. A multiplicação da matriz teste de paridade pela transposta de uma matriz linha $Z = \begin{bmatrix} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 \end{bmatrix}$ recebida pelo receptor, HZ^t recebe o nome de síndrome de Z, quando a síndrome HZ^t é uma matriz nula Z é aceito como a palavra enviada, por outro lado quando a síndrome não é nula, considera-se que ouve um erro em Z, observe que HZ^t é igual a uma das colunas de H de acordo com o erro que acontece, portanto se HZ^t é igual a coluna $i \in \{1, \ldots, 7\}$ de H aceita-se que ouve um erro na coluna i de Z, devendo então ser substituída por $z_i + 1$.

Criptosistema baseado em códigos lineares no ensino médio.

Os problemas de decodificação geral para códigos lineares como também encontrar a quantidade de erros que um código linear pode corrigir são NP-difícil, este fato além da existência de algoritmos de decodificação eficientes para códigos de Goppa, são as justificativas apresentadas por McEliece para propor o primeiro criptosistema baseado em códigos corretores de erros. Este sistema que pode ser facilmente implementado, por se

basear em operações elementares como multiplicação entre matrizes, contudo a estrutura algébrica utilizada na definição de códigos de Goppa, não é elementar a ponto de poder ser trabalhada na educação básica. Sugere-se então apresentar um criptosistema semelhante ao de McEliece, baseado no código de Hamming $\mathcal{C}(7,4)$ apresentado anteriormente, como uma segunda oficina extraclasse a discentes do segundo ano do ensino médio. Durante a definição do criptosistema é utilizado uma matriz invertível e uma matriz de permutação, tornando então necessário que os discentes tenham conhecimento das seguintes definições:

Definição 4.0.19. *(Matriz invertível). Dada uma matriz quadrada $A_{n\times n}$, dizemos que $A_{n\times n}$ é uma matriz invertível se existe uma matriz $B_{n\times n}$ tal que*

$$A_{n\times n}B_{n\times n} = B_{n\times n}A_{n\times n} = I_n.$$

Definição 4.0.20. *(Inversa de A). Dada uma matriz invertível A de ordem n, chamamos de inversa de A a única matriz A^{-1} também de ordem n, tal que*

$$AA^{-1} = A^{-1}A = I_n.$$

Definição 4.0.21. *(Matriz de Permutação). Dada uma matriz quadrada A de ordem n, chamamos a matriz A de matriz de permutação se A é obtida a partir da permutação de linhas ou colunas da matriz identidade I_n.*

Alice e Bob desejam se comunicar de forma segura, utilizando um criptosistema de chave pública. Para isto cada um utiliza uma chave pública e uma privada, as chaves públicas são disponibilizadas utilizando o canal de transmissão e são as regras que devem ser utilizadas para encriptar os textos de mensagens simples antes de enviados, por outro lado a chave privada de cada um é mantida escondida, pois elas são as regras utilizadas para decriptação das mensagens. Para que Alice e Bob possam enviar sua regra de encriptação de forma pública esta regra não pode ser de fácil inversão. Criaremos agora uma chave pública e privada de um criptosistema baseado em um código linear para Alice, utilizando a matriz geradora e teste de paridade

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \text{ e } H = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

do código de Hamming $\mathcal{C}(7,4)$ dado como exemplo anteriormente, e as seguintes matrizes

S e P invertível e de permutação, respectivamente,

$$S = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \text{ e } P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Com estes componentes e realizando as operações de acordo com a tabela de soma e multiplicação em $\mathbb{Z}_2$, criamos a matriz

$$\begin{aligned} G_{pub} &= S \cdot G \cdot P \\ &= \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Como os códigos de Hamming são $1-$perfeito, Alice envia como chave pública $(G_{pub}, 1)$ para as pessoas com quem ela deseja comunicar-se. As matrizes (S, P, H) são chamadas de chave privada e mantidas em segredo. Suponhamos que alguém deseja enviar a mensagem

de texto simples $\begin{bmatrix}1 & 1 & 0 & 1\end{bmatrix}$ para Alice, então esta pessoa calcula

$$\begin{aligned}\begin{bmatrix}1 & 1 & 0 & 1\end{bmatrix} G_{pub} &= \begin{bmatrix}1 & 1 & 0 & 1\end{bmatrix} \begin{bmatrix}0 & 0 & 1 & 1 & 1 & 0 & 1\\ 0 & 1 & 0 & 1 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 1 & 1\\ 1 & 0 & 0 & 0 & 1 & 0 & 1\end{bmatrix}\\ &= \begin{bmatrix}1 & 1 & 1 & 0 & 1 & 0 & 0\end{bmatrix}\end{aligned}$$

e adiciona um erro aleatório a mensagem somando uma matriz

$$E = \begin{bmatrix}e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7\end{bmatrix}$$

onde existe apenas um elemento $e_i, i \in \{1, \ldots, 7\}$ diferente de 0. Para continuarmos o exemplo utilizaremos $E = \begin{bmatrix}0 & 1 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$ e a mensagem de texto cifrado que será enviada se torna

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

Após receber a mensagem cifrada $\begin{bmatrix}1 & 0 & 1 & 0 & 1 & 0 & 0\end{bmatrix}$, Alice realiza o processo de decriptação, começando por calcular

$$\begin{aligned}\begin{bmatrix}1 & 0 & 1 & 0 & 1 & 0 & 0\end{bmatrix} P^{-1} &= \begin{bmatrix}1 & 0 & 1 & 0 & 1 & 0 & 0\end{bmatrix} \begin{bmatrix}0 & 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0\end{bmatrix}\\ &= \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 1 & 1\end{bmatrix}\end{aligned}$$

utilizando a matriz de teste de paridade H descobre-se onde o erro E foi introduzido,

calculando

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

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

Como o resultado é igual a quinta coluna da matriz H aceitamos que o erro foi introduzido na quinta coluna de $\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 1 & 1\end{bmatrix}$ sendo assim a mensagem corrigida é $\begin{bmatrix}1 & 0 & 0 & 0 & 1 & 1 & 1\end{bmatrix}$ onde os quatro primeiros dígitos $\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}$ são informações e os demais são redundâncias, resta eliminar a matriz S que foi multiplicada pela matriz G, este problema é resolvido multiplicando

$$\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}S^{-1} = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & 1 & 0 & 1\\ 1 & 1 & 0 & 0\\ 0 & 1 & 1 & 1\\ 1 & 0 & 0 & 1\end{bmatrix}$$

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

Obtendo finalmente a mensagem de texto simples que foi enviada.

Considerações Finais

Neste texto são apresentados os conhecimentos básicos para compreensão do criptosistema de McEliece baseado em códigos de Goppa binários irredutíveis, este criptosistema apesar de ser considerado até o presente momento resistente a computação quântica, possui uma chave pública com tamanho bastante superior aos criptosistemas mais utilizados na atualidade, como o RSA. Espera-se que este texto contribua com a difusão de conhecimentos relacionados a Teoria dos Códigos Corretores de Erros e Criptografia, áreas estas relacionadas a Teoria da Informação e que possuem importância significativa na vida das pessoas. No decorrer do texto estruturas algébricas são utilizadas de diversas formas, mostrando assim a importância de se estudar modelos abstratos e como propriedades destes modelos enriquecem diversas áreas de pesquisa. O capítulo 4 é voltado especialmente a docentes do ensino médio, apresentando propostas para que o conceito de Criptografia e de Códigos Corretores de Erros possam ser apresentados aos discentes na educação básica, de forma que os mesmos possam ter um meio de relacionar conteúdos considerados abstratos com exemplos concretos e particulares.

Referências Bibliográficas

[1] E. Berlekamp, R. McEliece, and H. Van Tilborg. On the inherent intractability of certain coding problems (corresp.). *IEEE Transactions on Information Theory*, 24(3):384–386, 1978.

[2] L. R. Dante. *Matemática: volume único*. Ática, 2009.

[3] A. Garcia and Y. Lequain. *Elementos de álgebra*. Instituto de Matematica Pura e Aplicada, 2013.

[4] V. D. Goppa. A new class of linear error-correcting codes. *Probl. Inf. Transm.*, 6:300–304, 1970.

[5] A. Hefez and M. L. T. Villela. *Códigos corretores de erros*. Instituto de Matematica Pura e Aplicada, 2008.

[6] G. Iezzi and S. Hazzan. *Fundamentos de matemática elementar, 4: sequências, matrizes, determinantes, sistemas*. Atual, 2004.

[7] E. L. Lima. Algebra linear, 2a. ediçao. *IMPA, Rio de Janeiro*, 1996.

[8] R. J. Mceliece. A public-key cryptosystem based on algebraic. *Coding Thv*, 4244:114–116, 1978.

[9] R. Misoczki and P. S. Barreto. Compact mceliece keys from goppa codes. In *International Workshop on Selected Areas in Cryptography*, pages 376–392. Springer, 2009.

[10] N. Patterson. The algebraic decoding of goppa codes. *IEEE Transactions on Information Theory*, 21(2):203–207, 1975.

[11] J. C. Pellegrini. *Introdução à Criptografia e seus Fundamentos notas de aula*. 2016.

[12] D. R. Stinson. *Cryptography: theory and practice*. CRC press, 2005.

[13] J. G. Vasquez, N. Melo, and R. Portugal. Estudo comparativo de algoritmos de decodificaçao para códigos de goppa aplicados no mceliece.

[14] J. Ventura. *Notas de Combinatória e Teoria de Códigos*. 2013.