UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
DEPARTAMENTO DE COMUNICAÇÕES

UMA CONTRIBUIÇÃO À CONSTRUÇÃO E DECODIFICAÇÃO DE CÓDIGOS DE BLOCO LINEARES SOBRE ANÉIS FINITOS

Antonio Aparecido de Andrade

Orientador: **Prof. Dr. Reginaldo Palazzo Jr.**

Co-orientador: **Prof. Dr. Trajano Pires da Nóbrega Neto**

94/0431

Este exemplar corresponde à redação final da tese defendida por Antonio Aparecido de Andrade e aprovada pela comissão Julgadora em 26/12/96
Orientador

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação, FEEC - UNICAMP, como requisito parcial para obtenção do título de DOUTOR EM ENGENHARIA ELÉTRICA.

Dezembro - 1996
Campinas - SP

UNICAMP
BIBLIOTECA CENTRAL

UNIDADE BC
N.º CHAMADA:
T/UNICAMP
An24c
V. Ex.
TOMBO BC/ 29894
PROC. 281/97
C [ ] D [x]
PREÇO R$ 11,00
DATA 29/04/97
N.º CPD

CM-00098042-9

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

An24c

Andrade, Antonio Aparecido de
Uma contribuição à construção e decodificação de códigos de bloco lineares sobre anéis finitos / Antonio Aparecido de Andrade.--Campinas, SP: [s.n.], 1996.

Orientadores: Reginaldo Palazzo Jr.; Trajano Pires da Nobrega Neto.
Tese (Doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Teoria da codificação. 2. Códigos de controle de erros (Teoria da informação). 3. Teoria da informação. 4. Anéis de grupo. I. Palazzo Jr., Reginaldo. II. Nobrega Neto, Trajano Pires da. III. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. IV. Título.

Agradecimentos

Ao concluir o presente trabalho agradeço:

- Ao orientador Professor Doutor Reginaldo Palazzo Junior, pela paciência, amizade e pelo constante estímulo, que com seu enorme conhecimento nos trilhou em caminhos seguros e produtivos.
- Ao Professor Doutor Trajano Pires da Nóbrega Neto (Departamento de Matemática - Ibilce - Unesp - São José do Rio Preto - SP) pela co-orientação, apoio e incentivo.
- Ao Professor Doutor José Carmelo Interlando (Departamento de Matemática - Ibilce - Unesp - São José do Rio Preto - SP) pela colaboração e auxílio na compreensão de alguns resultados.
- À Banca Examinadora, constituída pelos Profs. Drs. Paulo Roberto Brumatti do Imecc - Unicamp, Dalton Soares Arantes e Max Henrique Machado Costa do Decom - Feec - Unicamp e, pelo Prof. Dr. José Carmelo Interlando do Departamento de Matemática do Ibilce - Unesp - São José do Rio Preto - SP.
- À minha esposa Maria Gorete pela compreensão e apoio inestimável nos momentos difíceis.
- Aos Professores do Departamento de Comunicações - Feec - Unicamp, pela acolhida, formação e estímulo, e aos colegas do curso de doutoramento.
- Aos Professores do Departamento de Matemática do Ibilce - Unesp - São José do Rio Preto - SP, que me concederam condições para a realização deste trabalho.
- À Coordenação do Aperfeiçoamento do Pessoal de Nível Superior (CAPES) pelo suporte financeiro.
- À todos que direta e indiretamente colaboraram para a realização deste trabalho.
- À Deus, por tudo.

O mundo seria mais digno se tivéssemos dado ouvidos aos nossos mestres. Esta é a verdadeira lição que devemos passar aos nossos filhos. Sendo assim, confio em Deus com todas as minhas forças, e peço a Ele que ilumine o meu caminho concedendo-me a graça que tanto desejo.

Dedicação

À minha esposa
Maria Gorete e
ao meu filho
Matheus.

Resumo

Neste trabalho apresentamos extensões de construções de códigos cíclicos (via anéis de grupo), códigos de Hamming, códigos Reed-Solomon, códigos BCH e códigos alternantes sobre anéis comutativos finitos locais com identidade sob a métrica de Hamming e códigos BCH sobre anéis de inteiros finitos locais sob a métrica de Lee. Os códigos de Hamming, Reed-Solomon, BCH e alternantes são construídos em termos de suas matrizes verificação de paridade e a derivação dos códigos BCH e dos códigos alternantes é baseada na fatoração de $x^s - 1$ sobre o grupo das unidades de uma extensão apropriada do anel finito. Também, apresentamos processos de decodificação eficientes para os códigos de Hamming, Reed-Solomon, BCH e alternantes sob a métrica de Hamming e um algoritmo de decodificação alternativo para os códigos BCH sob a métrica de Lee. Os algoritmos de decodificação para os códigos Reed-Solomon, BCH e alternantes são baseados no algoritmo de Berlekamp-Massey modificado.

Abstract

In this research we present extensions of constructions of cyclic codes (via group rings), Hamming codes, Reed-Solomon codes, BCH codes and alternant codes over arbitrary local finite rings for the Hamming metric, and BCH codes over local finite integer rings for the Lee metric. Hamming codes, Reed-Solomon codes, BCH codes and alternant codes are constructed in terms of their parity-check matrices and the derivation of BCH codes and alternant codes are based on the factorization of $x^s - 1$ over the unit ring of an appropriate extension of the finite ring. We present efficient decoding procedure for the Hamming codes, Reed-Solomon codes, BCH codes and alternant codes for the Hamming metric and an alternative decoding procedure for the BCH codes for the Lee metric. The algorithms for the Reed-Solomon codes, BCH codes and alternant codes are based on the modified Berlekamp-Massey algorithm.

Conteúdo

1 Introdução **1**

1.1 Histórico 1

1.2 Descrição do Trabalho 2

2 Conceitos Básicos **5**

2.1 Conceitos Básicos de Álgebra Abstrata 5

2.2 Inteiros Algébricos sobre um Anel 14

2.2.1 Domínios de Dedekind 17

2.3 Códigos de Bloco 22

2.3.1 Códigos sobre Corpos Finitos 23

2.4 Conclusões 30

3 Códigos sobre Anéis Comutativos Finitos com Identidade **31**

3.1 Códigos Cíclicos 32

3.2 Anéis de Grupos 36

3.2.1 Códigos via Anéis de Grupos 37

3.3 Códigos de Hamming 42

3.3.1 Construção da Matriz Verificação de Paridade 42

3.3.2 Algoritmo de Decodificação 46

3.4 Códigos Reed-Solomon 47

3.5 Proposta de Construção e Decodificação de Códigos sobre Anéis Finitos 52

3.6 Conclusões 53

4 Códigos BCH sobre Anéis Comutativos Finitos com Identidade **55**

4.1 Códigos BCH sobre Anéis Finitos Locais 56

4.1.1 Códigos BCH como Códigos Cíclicos 62

4.2 Códigos BCH sobre Anéis Finitos 65

4.3 Algoritmo de Decodificação 71

4.4 Aplicação do Algoritmo de Berlekamp-Massey Modificado 86

4.5 Conclusões 88

5 Códigos Alternantes e BCH sobre Anéis Comutativos Finitos Locais com Identidade **90**

5.1 Códigos Alternantes 91

5.1.1 Algoritmo de Decodificação 94

5.2 Códigos BCH sobre $\mathbb{Z}_q$ para a Métrica de Lee 98

5.2.1 Algoritmo de Decodificação 102

5.3 Conclusões 109

6 Conclusões e Sugestões **110**

6.1 Conclusões 110

6.2 Sugestões para Futuras Pesquisas 112

Capítulo 1

Introdução

Neste capítulo faremos um breve histórico da teoria de códigos e a descrição dos demais capítulos que compõem este trabalho.

1.1 Histórico

A teoria de códigos é um ramo da matemática em franca atividade possuindo várias ramificações que utilizam ferramentas bastante diversas, como por exemplo, teoria dos números, teoria dos grupos, combinatória, geometrias finitas e geometria algébrica, dentre outras. Códigos são utilizados sempre que se deseja transmitir ou armazenar dados. Por exemplo, nas comunicações via satélite, nas comunicações internas de um computador, no armazenamento de dados em fitas ou disquetes magnéticos, ou no armazenamento óptico de dados.

A teoria de códigos foi fundada pelo matemático Claude E. Shannon [47], na década de 40. Sendo que a teoria de códigos corretores de erros teve início nesta mesma década com os trabalhos de Golay [19], Hamming [21] e Shannon [47]. A grande descoberta da época surgiu, principalmente devido a Shannon, com os modelos de códigos capazes de detectar e corrigir erros num sistema de comunicações. Shannon mostrou em 1948 que, através de uma codificação apropriada da informação, os erros introduzidos por um canal ruidoso poderiam ser reduzidos a nível desejado sem sacrificar a taxa de transmissão (desde que a taxa de informação R (expressa em dígitos por segundo) for menor que um valor C. A partir de então pesquisadores vêm procurando encontrar famílias de bons códigos (previstos pela Teoria de Shannon) e também projetar decodificadores eficientes para os mesmos.

Em teoria de códigos o problema que se põe é: dada uma mensagem recebida que contém um certo

número de erros, como corrigir estes erros e recuperar a mensagem enviada? Se a mensagem enviada contém redundâncias e se a quantidade de erros é pequena, em muitos casos podemos recuperar a mensagem. Esta é a filosofia dos códigos corretores de erros. O que torna interessante a teoria de códigos corretores de erros é como um problema, a princípio tão "aplicado", se relaciona intimamente com vários tópicos da matemática motivando novos problemas e novos resultados em ambos os lados.

1.2 Descrição do Trabalho

A demanda por alta confiabilidade da informação a ser transmitida decorrente das necessidades dos novos serviços a serem prestados aos usuários dos sistemas de telecomunicações, requer que novas classes de códigos sejam encontradas. As classes de códigos corretores de erros estão fortemente fundamentadas na estrutura algébrica de corpos finitos e suas extensões. Contudo, os códigos decorrentes destas classes, por si só não são capazes de satisfazer as condições de alta taxa e alta confiabilidade. Por outro lado, as classes de códigos sobre anéis finitos e suas extensões são suficientemente capazes de satisfazer à condição exigida pelas novas demandas, mas se carecem de uma sistematização na determinação dos melhores códigos e também de algoritmos de decodificação.

Códigos sobre anéis finitos têm recentemente despertado um grande interesse em teoria de codificação pelas suas bem sucedidas aplicações em modulação codificada e pelo fato de serem capazes de satisfazerem às condições exigidas pelas novas demandas. Um trabalho notável de Hammons et al. [22], mostrou que certos códigos não lineares binários com boa capacidade de correção de erros, por exemplo os códigos Kerdock, Preparata, Goethals e outros, podem ser vistos, através do mapeamento de Gray, como códigos lineares sobre $\mathbb{Z}_4$. Trabalhos anteriores por Calderbank e Sloane [13], Blake [8] e [9], e Shankar [46] estenderam a noção de códigos cíclicos, códigos de Hamming, códigos Reed-Solomon e códigos BCH sobre corpos finitos para as classes de códigos sobre anéis finitos de inteiros residuais. Biglieri e Elia, [6], mostraram que códigos sobre grupos abelianos podem ser estudados via códigos lineares sobre anéis finitos. Interlando, Palazzo e Elia [33] apresentaram algoritmos de decodificação para códigos Reed-Solomon e códigos BCH quando definidos sobre anéis de inteiros residuais.

A maioria dos códigos não binários são determinados através da métrica de Hamming. Códigos com a métrica de Hamming são ideais para canais balanceados em que a probabilidade de erro para cada símbolo é a mesma. A métrica de Lee foi desenvolvida como uma alternativa para a métrica de

Hamming para transmissão de sinais não binários sobre certos canais ruidosos. Ulrich [51] considerou códigos que podem corrigir um erro de $+1$ ou -1 numa palavra-código. Códigos com a métrica de Lee sobre corpos finitos, foram descritos por Lee [34], Berlekamp [5], Chiang e Wolf [14], Satyanarayana [44], e recentemente por Roth e Siegel [42].

Códigos algébricos lineares podem ser definidos através de alfabetos escolhidos em um conjunto de tamanho arbitrário. Contudo, muitos dos resultados em Teoria de Codificação têm sido obtidos assumindo que os símbolos são elementos de um corpo finito. Deste modo, neste trabalho propomos novas técnicas de construções de códigos cíclicos, códigos de Hamming, códigos Reed-Solomon, códigos BCH e códigos alternantes sobre anéis comutativos finitos com identidade sob a métrica de Hamming e códigos BCH sobre anéis de inteiros residuais módulo q, onde q é uma potência de um primo, sob a métrica de Lee. Também analisamos o problema da decodificação destes códigos. A questão relativa à existência de um algoritmo de decodificação simples, sob a métrica de Lee, permanece aberta. Um processo sistemático para decodificação de códigos sob a métrica de Lee tem sido a motivação das pesquisas recentes.

A organização básica dos capítulos que compõem o escopo deste trabalho está dividida como se segue.

Inicialmente, no Capítulo 2, estabelecemos a terminologia e a notação adotadas neste trabalho, revisando os conceitos de álgebra abstrata, teoria dos números e de códigos de bloco sobre corpos finitos.

No Capítulo 3, apresentamos construções de códigos cíclicos (via anéis de grupo) sobre anéis comutativos finitos, códigos de Hamming e códigos Reed-Solomon sobre anéis comutativos finitos locais. Esta proposta estende as construções apresentadas por Blake em [8] e [9], quando estes códigos são definidos sobre anéis dos inteiros residuais módulo q, onde q é uma potência de um primo.

No capítulo 4, estendemos os códigos BCH sobre anéis comutativos finitos com identidade. Esta proposta é bastante similar à construção destes códigos quando definidos sobre corpos finitos e estende a construção proposta por Shankar [46]. Também, através do algoritmo de Berlekamp-Massey modificado proposto por Interlando, Palazzo e Elia [33] apresentamos um algoritmo de decodificação para os códigos Reed-Solomon e os códigos BCH definidos sobre anéis finitos.

No Capítulo 5, apresentamos construções de códigos alternantes sob a métrica de Hamming e códigos BCH sob a métrica de Lee. Os códigos alternantes foram estendidos sobre quaisquer anéis comutativos finitos locais com identidade enquanto os códigos BCH foram estendidos sobre anéis

locais de inteiros residuais módulo q, onde q é uma potência de um primo. Também apresentamos algoritmos de decodificação para estes códigos. O algoritmo de decodificação para os códigos BCH combina o algoritmo de Berlekamp-Massey modificado proposto por Interlando, Palazzo e Elia [33] com o método proposto por Roth e Siegel [42].

Finalmente, no Capítulo 6, faremos nossas conclusões e indicamos alguns tópicos para estudos futuros.

Capítulo 2

Conceitos Básicos

Este capítulo introduz os conceitos básicos sobre álgebra abstrata, teoria dos números algébricos e teoria de códigos sobre corpos finitos, necessários para o entendimento dos demais capítulos. Na Seção 2.1, faremos um breve resumo da teoria básica de álgebra abstrata necessária para o desenvolvimento deste e dos demais capítulos. Na Seção 2.2, faremos uma breve revisão dos principais fatos básicos sobre o anel dos inteiros algébricos de um corpo numérico. Veremos os corpos quadráticos e seus anéis de inteiros algébricos. Analisamos, também, que o anel dos inteiros algébricos é um domínio de Dedekind e focalizamos as decomposições de ideais nesses anéis. Os teoremas-definição das extensões e contrações de ideais primos serão analisados. Também veremos o fato que a partir dos domínios de Dedekind é possível encontrar um grande número de anéis finitos através do conceito de norma de um ideal. Na Seção 2.3, faremos uma breve revisão de códigos sobre corpos finitos enfocando seus principais resultados. Definindo seus parâmetros tais como comprimento, dimensão, taxa, distância mínima e conjuntos de informação, matrizes geradora e verificação de paridade e síndrome. O algoritmo de decodificação por máxima verossimilhança também será analisado. Em geral, este algoritmo é razoavelmente complexo, mas códigos com propriedades particulares (por exemplo, os códigos cíclicos) podem ter algoritmos de decodificação mais eficientes.

2.1 Conceitos Básicos de Álgebra Abstrata

Nesta seção, faremos uma breve revisão dos conceitos elementares sobre grupos, anéis, corpos, módulos e espaços vetoriais, estabelecendo a notação e a terminologia adotada neste trabalho. A referência adotada nesta seção é [27].

Definição 2.1.1 *Um* **grupo** $(G, *)$ *é um conjunto não vazio* G *junto com uma operação* $*$ *sobre* G *que satisfaz as seguintes propriedades:*

- $(a * b) * c = a * (b * c), \quad \forall a, b, c \in G \quad$ *(associativa)*
- *existe* $e \in G$ *tal que* $a * e = e * a = a, \quad \forall a \in G \quad$ *(elemento identidade)*
- $\forall a \in G$, *existe* $a' \in G$ *tal que* $a * a' = a' * a = e \quad$ *(elemento inverso)*

Se além disso, $a * b = b * a, \ \forall a, b \in G$ *(comutativa), dizemos que o grupo* G *é* **comutativo** *ou* **abeliano.** ■

Às vezes, por simplificação de linguagem, diz-se apenas "G é um grupo" ou o "grupo G".

Definição 2.1.2 *Seja* $(G, *)$ *um grupo. Dizemos que um subconjunto não vazio* $H \subseteq G$ *é um* **subgrupo** *se satisfaz as seguintes condições:*

- $\forall a, b \in H$ *tem-se que* $a * b \in H \quad$ *(fechamento)*
- $(H, *)$ *também é um grupo. (Aqui a operação é a mesma de* G *só que restrita a* H*).* ■

Os subconjuntos $\{e\}$ e G de G são subgrupos chamados **subgrupos triviais**, enquanto que os demais são ditos não triviais.

Definição 2.1.3 *Um grupo* G *é chamado um* **grupo cíclico** *se existir um elemento* $a \in G$ *tal que* $G = \{a^m \mid m \in \mathbb{Z}\}$. *O elemento* a *é chamado* **gerador** *de* G *e usamos a notação* $G =< a >$. ■

Usaremos a notação $a^m = a.\cdots.a$ (m vezes) se G é um grupo multiplicativo e a notação $a^m = a + \cdots + a = ma$ (m vezes) se G é um grupo aditivo.

Definição 2.1.4 *Dados dois grupos* $(G, *)$ *e* $(H, \triangle)$, *dizemos que uma aplicação* $\phi : G \to H$ *é um* **homomorfismo** *de grupos de* G *em* H *se para todo* $a, b \in G$ *tem-se que* $\phi(a * b) = \phi(a) \triangle \phi(b)$. *Se além disso,* ϕ *for bijetora, dizemos que* ϕ *é um* **isomorfismo** *de grupos e que os grupos* G *e* H *são* **isomorfos.** ■

Observe na Definição 2.1.4 que a operação $a * b$ é a do grupo G, enquanto que a operação $\phi(a) \triangle \phi(b)$ é a do grupo H.

Definição 2.1.5 *Um* **anel** $(A, +, \cdot)$ *é um conjunto não vazio* A *munido das operações* $+$ *(adição) e* $\cdot$ *(produto ou multiplicação) que satisfaz as seguintes propriedades:*

- $(A, +)$ *é um grupo abeliano.*
- *A operação multiplicação é associativa, isto é,* $(ab)c = a(bc)$, $\forall a, b, c \in A$.
- *A multiplicação é distributiva em relação à adição, isto é,* $a(b+c) = ab+ac$ *e* $(a+b)c = ac+bc$, *para todo* $a, b, c \in A$. ■

Às vezes dizemos apenas "A é um anel" ou falamos do "anel A", por simplificação de linguagem. Isto pressupõe, naturalmente, um par de operações em A (com as propriedades citadas) sobre as quais não há nenhuma dúvida.

Definição 2.1.6 *Um anel* A *em que a multiplicação é comutativa, isto é,* $ab = ba$ *para todo* $a, b \in A$, *é chamado um* **anel comutativo**. *Se além disso,* A *possuir elemento identidade em relação à multiplicação, que será denotado por 1, dizemos que* A *é um* **anel comutativo com identidade**. ■

Os anéis considerados neste trabalho, salvo menção contrária, serão sempre comutativos com identidade. Um **subanel** de A é um subconjunto não vazio de A que, considerado com as operações de A é um anel.

Definição 2.1.7 *Sejam* A *e* B *dois anéis. Uma aplicação* $\phi : A \to B$ *é chamada um* **homomorfismo** *de anéis se para todo* $a, b \in A$ *tem-se que* $\phi(a + b) = \phi(a) + \phi(b)$ *e* $\phi(ab) = \phi(a)\phi(b)$. *Se além disso,* ϕ *for bijetora, dizemos que* ϕ *é um* **isomorfismo** *de anéis e que os anéis* A *e* B *são* **isomorfos**. ■

Sejam $A_1, A_2, \cdots, A_r$ anéis e $A = A_1 \times A_2 \times \cdots \times A_r$ seu produto cartesiano. Sejam $\mathbf{a} = (a_1, a_2, \cdots, a_r)$ e $\mathbf{b} = (b_1, b_2, \cdots, b_r)$ dois elementos de A. Definimos a soma e o produto em A por

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \cdots, a_r + b_r) \quad e \quad \mathbf{ab} = (a_1 b_1, a_2 b_2, \cdots, a_r b_r). \tag{2.1}$$

Com estas duas operações, temos que A é um anel. O anel, assim obtido, é denominado **produto (soma) direto** de $A_1, A_2, \cdots, A_r$. Quando nos referimos ao anel $A = A_1 \times A_2 \times \cdots \times A_r$, sem nenhuma menção, iremos considerar o produto direto de $A_1, A_2, \cdots, A_r$.

Definição 2.1.8 *Seja A um anel. Dizemos que um subconjunto não vazio $I \subseteq A$ é um* **ideal** *em A se satisfaz as seguintes condições:*

- *Para todo $a, b \in I$ tem-se que $a + b \in I$;*
- *Para todo $a \in I$ e $r \in A$ tem-se que $ra \in I$.* ■

Os subconjuntos $\{0\}$ e A são ideais de A, chamados **ideais triviais**, enquanto que os demais são chamados não triviais. Se I é um ideal de A, o conjunto $A/I = \{r + I \mid r \in A\}$, munido com as operações $(a+I)+(b+I) = (a+b)+I$ e $(a+I)(b+I) = ab+I$ para todo $a, b \in A$, forma um anel, chamado **anel quociente** de A em relação ao ideal I. Entre os ideais I de A distinguem-se os ideais **principais**, isto é, quando $I =< a >= aA = Aa$, ou seja, quando I é gerado por um único elemento $a \in A$. Um anel A é chamado **principal** se todo ideal de A for principal.

Definição 2.1.9 *Seja P um ideal num anel A. Dizemos que P é um ideal* **primo** *se $P \neq A$ e para todo $a, b \in A$ com $ab \in P$ implicar que $a \in P$ ou $b \in P$. Um ideal M em A é chamado* **maximal** *se $M \neq A$ e se A é o único ideal em A que contém M e é diferente de M.* ■

Teorema 2.1.1 *Seja A um anel. Então todo ideal maximal de A é primo.* ■

Definição 2.1.10 *Seja A um anel finito.*

- *Um elemento $a \in A$ é dito uma* **unidade (ou invertível)** *se existe um elemento $b \in A$ tal que $ab = 1$.*
- *Um elemento $a \in A$ é dito uma* **não unidade (ou um divisor de zero)** *se existe um elemento $b \in A$, não nulo, tal que $ab = 0$.* ■

Se A não possuir divisores de zero, isto é, se $ab = 0$ implicar que $a = 0$ ou $b = 0$, então A é chamado um **domínio** (de integridade). As unidades do anel A formam um grupo multiplicativo comutativo, denotado por A^*. Se A for um domínio, um elemento não nulo e não invertível $a \in A$ é chamado **irredutível** se não for um produto de dois elementos de $A - A^*$, e é chamado **primo** se o ideal principal $< a >$ for um ideal primo. Todo elemento primo é irredutível. O domínio A é chamado **fatorial** se todo elemento $a \in A$, não nulo e não inversível, pode ser fatorado como $a = p_1 p_2 \cdots p_r$, onde $r \geq 1$ e os p_i irredutíveis, e se $a = q_1 q_2 \cdots q_s$, onde $s \geq 1$ e os q_j são também irredutíveis, então $r = s$ e cada fator p_i é associado de um fator q_j, ou seja, $p_i = uq_j$, para alguma unidade $u \in A$.

Definição 2.1.11 *Um anel* A *é dito* **local** *se possui apenas um ideal maximal* $\mathcal{M}$. ■

Teorema 2.1.2 *Seja* A *um anel finito. As seguintes condições são equivalentes.*

- A *é um anel local.*
- *Para todo* $a \in A$ *tem-se que* a *ou* $1+a$ *é uma unidade.*
- *As não unidades de* A *formam um grupo abeliano aditivo.* ■

Seja A um anel. Para quaisquer ideais $\mathcal{a}$, $\mathcal{b}$ de A, a **interseção** $\mathcal{a} \cap \mathcal{b}$ é o maior ideal de A que está contido em $\mathcal{a}$ e $\mathcal{b}$, e a **soma** $\mathcal{a} + \mathcal{b} = \{a+b : a \in \mathcal{a}, b \in \mathcal{b}\}$ é o menor ideal de A que contém $\mathcal{a}$ e $\mathcal{b}$. Estas noções são estendidas de maneira natural à interseção e à soma de um conjunto qualquer de ideais de A.

O **produto** $\mathcal{a}.\mathcal{b}$ definido como o conjunto das somas finitas $a_1.b_1 + \cdots + a_r.b_r$, onde $a_j \in \mathcal{a}$, $b_j \in \mathcal{b}$, $1 \leq j \leq r$, é um ideal de A contido em $\mathcal{a} \cap \mathcal{b}$. Mais geralmente, define-se o produto $\prod \mathcal{a}_j$ de um número finito de ideais $\mathcal{a}_j$, e em particular, as potências $\mathcal{a}^0 = A$, $\mathcal{a}^1 = \mathcal{a}$, $\mathcal{a}^2 = \mathcal{a}\mathcal{a}, \cdots, \mathcal{a}^m = \mathcal{a}.\cdots\mathcal{a}$ (m vezes) de um ideal $\mathcal{a}$.

Dizemos que os ideais $\mathcal{a}$ e $\mathcal{b}$ de A são **comaximais** se $\mathcal{a} + \mathcal{b} = A$. Obviamente, isto ocorrerá quando $\mathcal{a}$ e $\mathcal{b}$ forem ideais maximais distintos. Temos que a inclusão $\mathcal{a}.\mathcal{b} \subseteq \mathcal{a} \cap \mathcal{b}$ é sempre verdadeira, mas a igualdade somente é verificada quando os ideais forem comaximais.

Sejam $\mathcal{a}_1, \cdots, \mathcal{a}_r$ ideais de A. Consideramos o produto cartesiano $\mathcal{A} = A/\mathcal{a}_1 \times \cdots \times A/\mathcal{a}_r$ (que é um anel comutativo em relação às operações componente a componente) e o homomorfismo $\phi : A \to \mathcal{A}$ definido por $a \to (a+\mathcal{a}_1, \cdots, a+\mathcal{a}_r)$ $(a \in A)$. Este homomorfismo tem obviamente $\mathcal{a}_1 \cap \cdots \cap \mathcal{a}_r$ como núcleo e, portanto, induz um homomorfismo injetivo de $A/(\mathcal{a}_1 \cap \cdots \cap \mathcal{a}_r)$ em $\mathcal{A}$. Em geral, temos o seguinte teorema.

Teorema 2.1.3 *(***Teorema Chinês do Resto***) Para quaisquer ideais* $\mathcal{a}_1, \cdots, \mathcal{a}_r$ *de* A, *com* $r \geq 2$, *as seguintes condições são equivalentes.*

- $\mathcal{a}_1, \cdots, \mathcal{a}_r$ *são comaximais dois a dois.*
- *O homomorfismo* $\phi : A \to A/\mathcal{a}_1 \times \cdots \times A/\mathcal{a}_r$ *é sobrejetivo. Neste caso, este homomorfismo induz um isomorfismo de* $A/(\mathcal{a}_1 \cap \cdots \cap \mathcal{a}_r) = A/(\mathcal{a}_1 \cdots \mathcal{a}_r)$ *sobre* $A/\mathcal{a}_1 \times \cdots \times A/\mathcal{a}_r$. ■

No caso particular em que $A = \mathbb{Z}$, concluimos do Teorema 2.1.3 que, para quaisquer números $m_1, \cdots, m_r \in \mathbb{N}$, primos dois a dois, o anel quociente $\frac{\mathbb{Z}}{<m_1 \cdots m_r>}$ é isomorfo ao produto direto $\frac{\mathbb{Z}}{m_1\mathbb{Z}} \times \cdots \times \frac{\mathbb{Z}}{m_r\mathbb{Z}}$.

Exemplo 2.1.1 *Sejam* $A = \mathbb{Z}[i]$ *e os ideais primos* $< 1+i >$ *e* $< 2+i >$ *do anel* A. *Como os ideais* $< 1+i >$ *e* $< 2+i >$ *são comaximais, segue do Teorema 2.1.3 que* $\frac{A}{< 1+i > \cap < 2+i >} = \frac{A}{< 1+i >< 2+i >} = \frac{A}{< 1+3i >}$ *é isomorfo ao produto direto* $\frac{A}{< 1+i >} \times \frac{A}{< 2+i >}$. ■

Teorema 2.1.4 *Sejam* $\dot{a}_1, \cdots, \dot{a}_r$ *ideais e* P *um ideal primo de um anel* A. *Se* $\dot{a}_1 \cdots \dot{a}_r \subseteq P$ *(respectivamente =) então* $\dot{a}_j \subseteq P$ *(respectivamente =) para algum* $j = 1, \cdots, r$. ■

Definição 2.1.12 *Um* **corpo** $\mathbb{K}$ *é um anel comutativo com identidade tal que todo elemento não nulo é uma unidade. Um* **subcorpo** *de* $\mathbb{K}$ *é um subconjunto não vazio de* $\mathbb{K}$ *que considerado com as operações de* $\mathbb{K}$ *é um corpo.* ■

Observação 2.1.1 *Um ideal* $\dot{a}$ *é* **primo** *(respectivamente,* **maximal***) se, e somente se,* $A/\dot{a}$ *for um domínio (respectivamente corpo). Em particular,* A *é um domínio se, e somente se,* $< 0 >$ *for um ideal primo;* A *é um corpo se, e somente se,* $< 0 >$ *for um ideal maximal e, portanto, o único ideal de* A *diferente de* A. ■

Definição 2.1.13 *Seja* $\mathbb{K}$ *um corpo. Um* **espaço vetorial** V *sobre* $\mathbb{K}$ *é um grupo abeliano em relação a adição junto com uma operação de multiplicação (por escalar) de* $\mathbb{K} \times V$ *em* V, *ou seja, a cada par* (α, u) *de* $\mathbb{K} \times V$ *está associado um único elemento de* V *que se indica por* αu, *e para essa multiplicação tem-se que:*

- $\alpha(\beta u) = (\alpha\beta)u$
- $(\alpha + \beta)u = \alpha u + \beta u$
- $\alpha(u + v) = \alpha u + \alpha v$
- $1u = u$

para quaisquer $u, v \in V$ *e* $\alpha, \beta \in \mathbb{K}$. ■

Definição 2.1.14 *Seja* $\mathbb{K}$ *um corpo. Uma* **álgebra** V *sobre* $\mathbb{K}$ *é um espaço vetorial* V *sobre* $\mathbb{K}$, *junto com uma operação binária de multiplicação sobre* V *tal que para todo* $\alpha \in \mathbb{K}$ *e* $u, v, w \in V$, *as seguints condições são satisfeitas:*

- $(\alpha u)v = \alpha(uv) = u(\alpha v)$
- $(u+v)w = uw + vw$
- $u(v+w) = uv + uw$.

Além disso, se $(uv)w = u(vw)$, *para todo* $u, v, w \in V$, *dizemos que* **V é uma álgebra associativa** *sobre* $\mathbb{K}$. ■

Sejam $\mathbb{K}$ um corpo e V um espaço vetorial sobre o corpo $\mathbb{K}$. Dizemos que o conjunto de vetores $S = \{v_i \mid i \in I\}$, onde I é um conjunto de índices de V, **gera** V se para todo $u \in V$ tem-se que $u = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n}$ onde $\alpha_j \in \mathbb{K}$ e $v_{i_j} \in S$, $j = 1, 2, \cdots, n$. Neste caso, o vetor u é chamado uma **combinação linear** dos v_{i_j}. Dizemos que o conjunto de vetores $S = \{v_i \mid i \in I\}$ é **linearmente independente** sobre $\mathbb{K}$ se a igualdade $\sum_{j=1}^{n} \alpha_j v_{i_j} = 0$ implicar que $a_j = 0$, $j = 1, 2, \cdots, n$. Se o conjunto de vetores não for linearmente independente, dizemos que é **linearmente dependente.**

Definição 2.1.15 *Seja* V *um espaço vetorial sobre um corpo* $\mathbb{K}$. *Dizemos que o conjunto de vetores* $\mathcal{B} = \{v_i \mid i \in I\}$ *de* V *forma uma* **base** *para* V *sobre* $\mathbb{K}$ *se gera* V *e se for linearmente independente. O número de vetores de* $\mathcal{B}$ *é chamado* **dimensão** *de* V *sobre* $\mathbb{K}$. ■

Teorema 2.1.5 *(***Hammel***) Todo* $\mathbb{K}$*-espaço vetorial* V *possui uma base.* ■

A definição de A-módulo, para um anel A qualquer, é obtida como estensão natural da noção de $\mathbb{K}$-espaço vetorial.

Definição 2.1.16 *Entendemos por um* A**-módulo** *um grupo abeliano* M, *escrito aditivamente e munido de uma operação externa (usualmente escrita como multiplicação), satisfazendo os axiomas* $a.(x+y) = a.x + a.y$, $(a+b).x = a.x + b.x$, $(a.b).x = a.(b.x)$, $1.x = x$, *para quaisquer* $a, b \in A$ *e* $x, y \in M$. ■

Evidentemente, todo grupo abeliano aditivo G é um $\mathbb{Z}$-módulo, sendo $(\pm n).x$ definido por $\pm(x + \cdots + x)$ (n vezes) para quaisquer $n \in \mathbb{N}$ e $x \in G$.

O **produto cartesiano** de A-módulos é um A-módulo, sendo as operações definidas componente a componente. Um subconjunto $N \subseteq M$ é um A**-submódulo** se considerado com as operações de

M é um A-módulo. Também temos o módulo quociente M/N definido de maneira análoga para anéis.

Se M é um A-módulo com submódulos $N_1, N_2, \cdots, N_r$ então M é chamado **produto (soma) direto** de $N_1, N_2, \cdots, N_r$ se todo elemento $m \in M$ pode ser expresso unicamente como $m = n_1 + n_2 + \cdots + n_r$ com $n_i \in N_i$. Equivalentemente, $M = N_1 + N_2 + \cdots + N_r$ e $N_i \cap (N_1 + N_2 + \cdots + N_{i-1} + N_{i+1} + \cdots + N_r) = 0$ para $1 \leq i \leq r$. Neste caso, M é escrito como $M = N_1 \times N_2 \times \cdots \times N_r$.

Um A-módulo M é **finitamente gerado** se existem $x_1, \cdots, x_s \in M$ tais que $M = A.x_1 + \cdots + A.x_s$; neste caso, dizemos que $x_1, \cdots, x_s$ formam um **sistema de geradores** de M. Dizemos que os elementos $y_1, \cdots, y_s \in M$ são **linearmente independentes** (sobre A) se, para quaisquer $a_1, \cdots, a_s \in A$, a igualdade $\sum_{i=1}^{s} a_j.y_j = 0$ implicar que $a_1 = \cdots = a_s = 0$. Se, além disso, $y_1, \cdots, y_s$ formarem um sistema de geradores de M, então este conjunto é chamado uma **base** de M. É importante notar que nem todo módulo finitamente gerado possui uma base. Em particular, a caracterização das bases de um $\mathbb{K}$-espaço, como sistemas minimais de geradores ou sistemas maximais de elementos linearmente independentes, não se estende a módulos. Um A-módulo que possua uma base é chamado **livre**.

Muitas vezes consideramos módulos cuja operação externa é definida por uma multiplicação interna. Em particular, dado um anel A, todo anel B que contiver A é considerado como A-módulo, sendo o produto externo $a.x$, com $a \in A$ e $x \in B$, definido como o produto $a.x$ tomado em B. Em particular, o próprio A é considerado como um A-módulo, e os submódulos de A são exatamente os ideais de A. Note que um ideal $\mathcal{a}$ de A será um A-módulo livre se, e somente se, $\mathcal{a}$ for um ideal principal, gerado por um elemento que não seja um divisor de zero de A. Tal elemento forma então uma base de $\mathcal{a}$.

Sejam $\mathbb{F}$ um corpo e $\mathbb{K}$ um subcorpo de $\mathbb{F}$. Considerando $\mathbb{F}$ como $\mathbb{K}$-espaço vetorial (em relação à adição em $\mathbb{F}$ e a multiplicação com elementos de $\mathbb{K}$), dizemos que $\mathbb{F}$ é uma **extensão** de $\mathbb{K}$ e se uma (e logo toda) base do $\mathbb{K}$-espaço $\mathbb{F}$ tiver n elementos, dizemos que $\mathbb{F}$ é uma **extensão finita** de $\mathbb{K}$, de grau $[\mathbb{F} : \mathbb{K}] = n$. Se tivermos a inclusão de corpos $\mathbb{K} \subseteq \mathbb{F} \subseteq \mathbb{E}$ decorre que $[\mathbb{E} : \mathbb{K}] = [\mathbb{E} : \mathbb{F}][\mathbb{F} : \mathbb{K}]$, e que $[\mathbb{F} : \mathbb{K}] = 1$ se, e somente se, $\mathbb{F} = \mathbb{K}$.

Dado um subconjunto $\mathcal{A}$ de $\mathbb{F}$, denotamos por $\mathbb{K}(\mathcal{A})$ o corpo obtido de $\mathbb{K}$ pela adjunção de $\mathcal{A}$, ou seja, o menor corpo entre $\mathbb{K}$ e $\mathbb{F}$ que contém $\mathcal{A}$. No caso de $\mathcal{A} = \{\alpha_1, \cdots, \alpha_s\}$ escrevemos $\mathbb{K}(\alpha_1, \cdots, \alpha_s)$.

Um elemento $\alpha \in \mathbb{F}$ é chamado **algébrico** sobre $\mathbb{K}$ se for raiz de um polinômio mônico com coeficientes em $\mathbb{K}$. Caso contrário, é chamado **transcendente**. Dizemos que $\mathbb{F}$ é uma **extensão algébrica** de $\mathbb{K}$ se todo elemento $\alpha \in \mathbb{F}$ for algébrico sobre $\mathbb{K}$. O corpo $\mathbb{K}(\mathcal{A})$, obtido de $\mathbb{K}$ pela adjunção de um conjunto $\mathcal{A}$ de elementos algébricos sobre $\mathbb{K}$, é chamada uma **extensão algébrica** sobre $\mathbb{K}$. Uma extensão $\mathbb{F}$ de $\mathbb{K}$ será finita se existirem elementos $\alpha_1, \cdots, \alpha_s$ algébricos sobre $\mathbb{K}$ tais que $\mathbb{F} = \mathbb{K}(\alpha_1, \cdots, \alpha_s)$. Neste caso, $\mathbb{F}$ é uma extensão algébrica de $\mathbb{K}$.

Um polinômio $f(x) \in \mathbb{K}[x]$ é dito **irredutível** sobre $\mathbb{K}$, se sempre que $f(x) = a(x)b(x)$, com $a(x)$, $b(x) \in \mathbb{K}[x]$, implicar que um dos $a(x)$ ou $b(x)$ tem grau zero (isto é, uma constante). Um polinômio mônico (não necessariamente irredutível) $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{K}[x]$ é chamado **separável** se $f(x)$ e sua derivada $f'(x) = nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$ forem primos entre si, isto é, se $mdc(f(x), f'(x)) = 1$. Um polinômio irredutível e mônico $f(x) \in \mathbb{K}[X]$ será separável se, e somente se, $f'(x) \neq 0$. Um elemento α algébrico sobre $\mathbb{K}$ é chamado **separável** sobre $\mathbb{K}$ se for raiz de um polinômio separável $f(x) \in \mathbb{K}[x]$.

Uma extensão algébrica $\mathbb{F}$ de $\mathbb{K}$ é chamada **separável** se todo $\alpha \in \mathbb{F}$ for separável sobre $\mathbb{K}$; para isto é suficiente que $\mathbb{F} = \mathbb{K}(\mathcal{A})$ para algum conjunto $\mathcal{A}$ de elementos separáveis sobre o corpo $\mathbb{K}$.

Dados um anel B, um subanel A e um subconjunto $\mathcal{A}$ de B, denotamos por $A[\mathcal{A}]$ o anel obtido de A pela adjunção de $\mathcal{A}$, ou seja, o menor anel entre A e B que contém $\mathcal{A}$. Se $\mathcal{A} = \{\alpha_1, \cdots, \alpha_s\}$ escrevemos $A(\alpha_1, \cdots, \alpha_s)$ em vez de $A(\mathcal{A})$.

Todo subanel A de um corpo $\mathbb{F}$ é um domínio, e $\mathbb{K} = \{a.b^{-1} : a, b \in A, b \neq 0\}$ é o menor subcorpo de $\mathbb{F}$ que contém A, chamado **corpo de frações** de A em $\mathbb{F}$.

Sejam A um domínio e $\mathbb{F}$ uma extensão do corpo de frações $\mathbb{K}$ de A. Os elementos $\alpha_1, \cdots, \alpha_s \in \mathbb{F}$ serão linearmente independentes sobre A se, e somente se, o forem sobre $\mathbb{K}$. Temos ainda que se $[\mathbb{F} : \mathbb{K}] = n$ e se B é um subanel de $\mathbb{F}$ que contém A e que, como A-módulo, possua uma base $\beta_1, \cdots, \beta_q$, então $q \leq n$, e que $q = n$ se, e somente se, $\beta_1, \cdots, \beta_q$ formam uma base de $\mathbb{F}$ sobre $\mathbb{K}$.

Para finalizar esta seção, veremos alguns fatos sobre corpos finitos. Seja n um inteiro positivo fixo e sejam h e k dois inteiros quaisquer. O resto r quando $h + k$ (hk) é dividido por n é a soma (produto) de h e k módulo n. O conjunto $\{0, 1, \cdots, n-1\}$ é um anel sob estas operações, denotado por $\mathbb{Z}_n$. Quando n é um primo p temos que $\mathbb{Z}_p$ é um corpo de p elementos e denotado por $GF(p)$. Seja $f(x) \in GF(p)$ um polinômio irredutível de grau m. O quociente de $GF(p)[x]$

módulo $f(x)$ é um corpo, chamado **corpo de Galois** de ordem p^m, e denotado por $GF(p^m)$. O corpo $GF(p^m)$ é um $GF(p)$-espaço vetorial de dimensão m e, portanto, possui p^m elementos. A unicidade de $GF(p^m)$ é garantida pelo seguinte teorema.

Teorema 2.1.6 *Se $\mathbb{E}$ e $\mathbb{F}$ são dois corpos finitos de ordem p^m, então os corpos $\mathbb{E}$ e $\mathbb{F}$ são isomorfos, ou seja, existe um isomorfismo entre $\mathbb{E}$ e $\mathbb{F}$ que preserva a adição e a multiplicação.* ■

Teorema 2.1.7 *O conjunto dos elementos inversíveis de um corpo finito com q elementos, onde $q = p^m$, é um grupo cíclico multiplicativo de ordem $q - 1$.* ■

2.2 Inteiros Algébricos sobre um Anel

A noção principal na Teoria dos Números Algébricos é a de anel dos inteiros algébricos de um corpo numérico. Este anel será introduzido nesta seção num contexto mais geral, através da noção de elemento inteiro sobre um anel A, que estende a noção de elemento algébrico sobre um corpo $\mathbb{K}$. O fato do anel dos inteiros algébricos sobre $\mathbb{Z}$ ser sempre um $\mathbb{Z}$-módulo livre também será visto. Como exemplos mais simples de corpos numéricos serão vistos os corpos quadráticos. A teoria apresentada nesta seção segue de [43].

Por um **número algébrico** entendemos qualquer $\alpha \in \mathbb{C}$ (corpos dos números complexos) que é algébrico sobre $\mathbb{Q}$ (corpo dos números racionais). Um subcorpo $\mathbb{F}$ de $\mathbb{C}$ é chamado um **corpo numérico** (ou **corpo de números algébricos**) se for uma extensão finita de $\mathbb{Q}$.

Veremos que todo corpo numérico $\mathbb{F} = \mathbb{Q}(\alpha)$ possui um subanel com corpo de frações igual a $\mathbb{F}$. Este anel, em geral, não coincide com o anel $\mathbb{Z}(\alpha)$, o qual depende da escolha do elemento primitivo α da extensão $\mathbb{F}$ de $\mathbb{Q}$. Este anel é definido como o conjunto dos elementos de $\mathbb{F}$ que são "inteiros" sobre $\mathbb{Z}$ nos termos da seguinte definição, que vamos introduzir num contexto mais geral.

Definição 2.2.1 *Sejam B um anel e A um subanel de B. Um elemento $\alpha \in B$ é dito* **inteiro** *sobre A se α é raiz de um polinômio mônico com coeficientes em A. Em particular, todo elemento de A é inteiro sobre A.* ■

No caso de $B = \mathbb{C}$ e $A = \mathbb{Z}$ os elementos inteiros sobre $\mathbb{Z}$ são chamados **inteiros algébricos**. Por exemplo, os números $i = \sqrt{-1}$ e $\sqrt{2}$ são inteiros algébricos pois são raízes dos polinômios $x^2 + 1$ e $x^2 - 2$, respectivamente.

É claro que no caso de corpos $B = \mathbb{F}$ e $A = \mathbb{K}$, um elemento $\alpha \in \mathbb{F}$ é inteiro sobre $\mathbb{K}$ se, e somente se, for algébrico sobre $\mathbb{K}$.

Veremos, a seguir, que o conjunto dos elementos de B que são inteiros sobre A forma um subanel $I_B(A)$ de B. Para provar que a diferença (respectivamente, o produto) de dois elementos α, $\beta \in B$ inteiros sobre A é também um inteiro, podemos tentar construir, a partir de polinômios mônicos $f, g \in A[x]$ tais que $f(\alpha) = g(\beta) = 0$, um polinômio mônico $h \in A[x]$ com raiz $\alpha - \beta$ (respectivamente, $\alpha.\beta$). Tal construção, entretanto, é viável somente em casos bem simples. Em geral, é mais prático usar as condições equivalentes do próximo teorema, que relaciona a propriedade de um elemento ser inteiro sobre A com a propriedade de um certo A-módulo ser finitamente gerado.

Teorema 2.2.1 *Seja B um anel e $A \subseteq B$ um subanel. Para qualquer elemento $\alpha \in B$ as seguintes condições são equivalentes.*

- *α é inteiro sobre A.*
- *$A(\alpha)$ é um A-módulo finitamente gerado.*
- *Existe um subanel S de B que contém A e α que é um A-módulo finitamente gerado.* ■

Proposição 2.2.1 *Sejam B um anel, A um subanel de B, e seja $(x_i)_{1 \leq i \leq n}$ um conjunto finito de elementos de B. Se, para todo i, x_i é inteiro sobre $A[x_1, x_2, \cdots, x_n]$ (em particular se todos os x_i's são inteiros sobre A), então $A[x_1, x_2, \cdots, x_n]$ é um A-módulo finitamente gerado.* ■

Denotando por $I_B(A)$ o conjunto dos elementos de B que são inteiros sobre A, temos o seguinte corolário.

Corolário 2.2.1 *$I_B(A)$ é um subanel de B que contém A.* ■

O anel $I_B(A)$ é chamado o **fecho inteiro** de A em B. Quando $I_B(A) = A$ dizemos que A é **integralmente fechado** em B. Quando $I_B(A) = B$ dizemos que B é **inteiro** sobre A. Quando $B = \mathbb{F}$ é um corpo numérico e $A = \mathbb{Z}$, o anel $I_{\mathbb{F}}(\mathbb{Z})$ é chamado **anel dos inteiros algébricos** de $\mathbb{F}$.

Definição 2.2.2 *Dizemos que um domínio A é* **integralmente fechado** *se for integralmente fechado no seu corpo de frações $\mathbb{K}$, isto é, se $I_{\mathbb{K}}(A) = A$.* ■

Quando $B = \mathbb{Q}$ e $A = \mathbb{Z}$, o anel $I_B(\mathbb{Z})$ coincide com o anel $\mathbb{Z}$ dos números inteiros. Isto é uma consequência imediata do seguinte teorema.

Teorema 2.2.2 *Todo domínio fatorial A é integralmente fechado no seu corpo de frações.* ■

O seguinte teorema fornece uma caracterização da relação entre os ideais primos de A e os ideais primos de B.

Teorema 2.2.3 *Sejam A e B anéis, com $A \subseteq B$, tal que B é inteiro sobre A.*

- *Se $Q \subseteq B$ é um ideal primo e $P = Q \cap A$, então Q é maximal se, e somente se, P é maximal.*
- *Se Q_1 e Q_2 são ideais primos de B tais que $Q_1 \subseteq Q_2$ e $Q_1 \cap A = Q_2 \cap A = P$, então $Q_1 = Q_2$.*
- *Se P é um ideal primo de A, então existe um ideal primo Q de B tal que $Q \cap A = P$.* ■

Exemplo 2.2.1 *Sejam os anéis $A = \mathbb{Z}$ e $B = \mathbb{Z}[i]$, onde $i^2 = -1$. Tomando os ideais primos $< 2 >$ e $< 5 >$ do anel A, e os ideais primos $< 1+i >$ e $< 2+i >$ do anel B, é de verificação imediata que $< 2 >=< 1+i > \cap \mathbb{Z}$ e $< 5 >=< 2+i > \cap \mathbb{Z}$.* ■

Um **corpo quadrático** é um subcorpo $\mathbb{F}$ de $\mathbb{C}$ tal que $[\mathbb{F} : \mathbb{Q}] = 2$. Observando que qualquer $\alpha \in \mathbb{F} - \mathbb{Q}$ é um elemento primitivo da extensão $\mathbb{Q} \subseteq \mathbb{F}$ (uma vez que de $1 < [\mathbb{Q}(\alpha) : \mathbb{Q}] = [\mathbb{F} : \mathbb{Q}] = 2$ decorre que $\mathbb{Q}(\alpha) = \mathbb{F}$), temos que os elementos $1, \alpha$ formam uma base desta extensão. Para cada $r \in \mathbb{Q} - \{0\}$ denotamos por $\sqrt{r}$ uma das raízes do polinômio $x^2 - r$, a outra raiz é $-\sqrt{r}$.

No próximo teorema veremos que todo corpo quadrático possui um elemento primitivo da forma $\sqrt{d}$, univocamente determinado a menos do sinal, onde d é um inteiro "livre de quadrados".

Teorema 2.2.4 *Todo corpo quadrático é da forma $\mathbb{Q}(\sqrt{d})$, onde d é um número inteiro livre de quadrados.* ■

O seguinte teorema fornece o anel dos inteiros algébricos de um corpo quadrático.

Teorema 2.2.5 *Seja $\mathbb{F} = \mathbb{Q}(\sqrt{d})$ um corpo quadrático onde d é um inteiro livre de quadrados.*

- *Se* $d \equiv 2$ *ou* $d \equiv 3\ (mod\ 4)$, *o anel* $I_{\mathbb{F}}(\mathbb{Z})$ *dos inteiros de* $\mathbb{F}$ *consiste de todos os elementos da forma* $a + b\sqrt{d}$ *com* $a, b \in \mathbb{Z}$.

- *Se* $d \equiv 1\ (mod\ 4)$, *o anel* $I_{\mathbb{F}}(\mathbb{Z})$ *consiste de todos os elementos da forma* $\frac{1}{2}(a + b\sqrt{d})$ *com* $a, b \in \mathbb{Z}$ *de mesma paridade.* ■

Exemplo 2.2.2 *Sejam os corpo* $\mathbb{F} = \mathbb{Q}(i)$, *onde* $i^2 = -1$, *e* $\mathbb{K} = \mathbb{Q}(\sqrt{5})$. *Pelo Teorema 2.2.5 temos que* $\mathbb{Z}[i]$ *é o anel dos inteiros algébricos de* $\mathbb{F}$ *e o conjunto dos elementos da forma* $\frac{1}{2}(a + b\sqrt{5})$, *com* $a, b \in \mathbb{Z}$ *de mesma paridade, é o anel dos inteiros algébricos do corpo* $\mathbb{K}$, *pelo fato de que* $-1 \equiv 3\ (mod\ 4)$ *e* $5 \equiv 1\ (mod\ 4)$. ■

2.2.1 Domínios de Dedekind

O exemplo do corpo quadrático $\mathbb{Q}(\sqrt{-5})$ mostra que o anel dos inteiros algébricos de um corpo de números algébricos nem sempre é fatorável. Para consertar esta falha, foi introduzido por Kummer a noção de "número ideal", que deu origem à noção de "ideal", devido a Dedekind. Em lugar da fatoração única em potências de elementos irredutíveis, válida apenas no caso de domínios fatoráveis, prova-se que, em qualquer anel de inteiros algébricos, todo ideal não-nulo possui uma fatoração única em potências de ideais primos. Esta e outras propriedades do anel dos inteiros algébricos generalizam-se facilmente a uma classe de domínios chamados "domínios de Dedekind".

Definição 2.2.3 *Um domínio* A *é chamado* **domínio de Dedekind** *se for integralmente fechado, todo ideal de* A *for finitamente gerado como* A*-módulo, e todo ideal primo não nulo de* A *for maximal.* ■

Obviamente, todo corpo e todo domínio principal é de Dedekind. O anel de polinômios $\mathbb{Q}[x, y]$ é um exemplo de um domínio fatorial que não é de Dedekind, pois o ideal gerado por x é primo mas não é maximal.

O próximo teorema estabelece que a propriedade de um domínio A ser de Dedekind é preservada na passagem de seu corpo de frações $\mathbb{K}$ a qualquer extensão separável finita $\mathbb{F}$. Mais precisamente temos o seguinte.

Teorema 2.2.6 *Sejam* A *um domínio de Dedekind e* $\mathbb{F}$ *uma extensão finita e separável de seu corpo de frações* $\mathbb{K}$. *Então o anel dos inteiros algébricos de* $\mathbb{F}$ *é um domínio de Dedekind e um* A*-módulo finitamente gerado.* ■

Corolário 2.2.2 *Seja $\mathbb{F}$ um corpo de números algébricos. Então o anel de inteiros algébricos de $\mathbb{F}$ é um domínio de Dedekind.* ■

Exemplo 2.2.3 *Pelo Teorema 2.2.6 temos que os anéis $\mathbb{Z}[i]$ e $\mathbb{Z}(\sqrt{2})$ são domínios de Dedekind, pelo fato de serem anéis de inteiros algébricos dos corpos $\mathbb{Q}(i)$ e $\mathbb{Q}(\sqrt{2})$, respectivamente.* ■

O nosso intuito, a seguir, é verificar, em domínios de Dedekind A, a existência e a unicidade da fatoração de ideais não nulos em ideais primos, já que a fatoração de elementos de $A - \{0\}$ em elementos irredutíveis, embora sempre exista, nem sempre é unica. Por exemplo, em $\mathbb{Z}[\sqrt{-5}]$ temos $6 = 2.3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$. Da mesma maneira que a fatoração em elementos irredutíveis é estendida aos elementos não nulos do corpo de frações $\mathbb{K}$ de A, admitindo-se potências negativas dos elementos irredutíveis, convém estender a fatoração em ideais primos.

Teorema 2.2.7 *Num domínio de Dedekind todo ideal não nulo tem uma única fatoração como um produto de ideais primos.* ■

O próximo corolário será de grande utilidade na determinação dos ideais num anel quociente e também fornece o conceito de divisibilidade entre ideais de um anel.

Corolário 2.2.3 *Seja A um domínio de Dedekind e sejam $\mathfrak{a}$ e $\mathfrak{b}$ ideais de A. Então $\mathfrak{a} \subseteq \mathfrak{b}$ se, e somente se, $\mathfrak{b}$ divide $\mathfrak{a}$, ou seja, se e somente se, existe um ideal $\mathbf{c}$ de A tal que $\mathfrak{a} = \mathfrak{b}\mathbf{c}$.* ■

Exemplo 2.2.4 *Seja o anel dos inteiros Gaussianos $\mathbb{Z}[i]$. Pelo Teorema 2.2.7 temos a seguinte fatoração única $2\mathbb{Z}[i] =< 1 + i >^2$.* ■

Se A é um domínio de Dedekind e $\mathfrak{a}$ um ideal de A, nem sempre o anel quociente $A/\mathfrak{a}$ é finito. Veremos, a seguir, quando $A/\mathfrak{a}$ é finito.

Sejam $\mathbb{F}$ um corpo de números algébricos tal que $[\mathbb{F} : \mathbb{Q}] = n$ e A o anel dos inteiros algébricos de $\mathbb{F}$. Sendo $A = I_{\mathbb{F}}(\mathbb{Z})$, todo ideal primo não nulo P de A é maximal. Logo, A/P é um corpo chamado **corpo residual** de A módulo P.

Teorema 2.2.8 *Seja P um ideal primo de A. Então,*

- $P \cap \mathbb{Z} = p\mathbb{Z} =< p >$, *sendo p o único número primo em P, isto é, $< p >= P \cap \mathbb{Z}$.*
- A/P *é uma extensão finita do corpo $\mathbb{Z}_p = GF(p)$, de grau $[A/P : \mathbb{Z}_p] \leq n$.* ■

O grau $[A/P : \mathbb{Z}_p]$ é chamado o **grau de inércia** de P e denotado pelo símbolo $f = f(P)$.

Definição 2.2.4 *Seja $\mathfrak{a}$ um ideal não nulo de A. A* **norma do ideal** *$\mathfrak{a}$ é o número de elementos de $A/\mathfrak{a}$, e será denotada por $N(\mathfrak{a})$.* ■

No próximo teorema veremos que este número é sempre finito.

Teorema 2.2.9 *Com as notações acima temos:*

- *Para todo ideal primo não nulo P de A temos que $N(P) = p^f$, onde f é o grau de inércia de P e p é o único primo em P.*
- *Para todo ideal não nulo $\mathfrak{a}$ de A temos que $N(\mathfrak{a}) \in \mathbb{N} - \{0\}$. Em particular, $N(\mathfrak{a}) = 1$ se, e somente se, $\mathfrak{a} = A$.*
- *Para quaisquer ideais não nulos $\mathfrak{a}, \mathfrak{b}$ de A temos que $N(\mathfrak{a}.\mathfrak{b}) = N(\mathfrak{a}).N(\mathfrak{b})$.*
- *Se $N(\mathfrak{a})$ for um número primo então $\mathfrak{a}$ é um ideal primo.*
- *Para todo $m \in \mathbb{N} - \{0\}$ existe somente um número finito de ideais não nulos $\mathfrak{a}$ do anel A tais que $N(\mathfrak{a}) = m$.* ■

Obervamos que a partir do Teorema 2.2.9 é possível encontrar um grande número de de anéis finitos através do conceito de norma de um ideal. Por exemplo, tomando $A = \mathbb{Z}[i]$ e $p = 5$, temos que $\frac{A}{<2+i>}$ é um anel com 5 elementos, $\frac{A}{<2-i>^2}$ é um anel com 25 elementos, e de modo análogo podemos obter outros anéis finitos.

Sejam A um domínio de Dedekind, $\mathbb{F}$ uma extensão finita e separável do corpo de frações $\mathbb{K}$ de A e $B = I_{\mathbb{F}}(A)$. Veremos no próximo teorema a decomposição de um ideal primo P de A em $\mathbb{F}$, isto é, a fatoração do ideal estendido $P.B$ em ideais primos Q do anel B.

Teorema 2.2.10 *Sejam A, B, $\mathbb{K}$, $\mathbb{F}$ como definidos acima, $n = [\mathbb{F} : \mathbb{K}]$ e P um ideal primo não nulo de A. Então,*

- *$P.B = Q_1^{e_1}.\cdots.Q_r^{e_r}$, onde os Q_i's são ideais primos distintos de B e e_i, $i = 1, \cdots, r$, são inteiros maiores ou iguais a 1.*
- *Q_i, $i = 1, \cdots, r$ são os ideais primos que estão acima de P, isto é, $P = Q_i \cap A$.*

- *Para todo ideal* $\mathbb{b}$ *do anel* B *tal que* $\mathbb{b} \cap A = P$, *o* A/P*-espaço vetorial* $B/\mathbb{b}$ *tem dimensão finita, e em particular,* $B/Q_i^{l_i}$, $1 \leq l_i \leq e_i$, $i = 1, \cdots, r$, *é um espaço vetorial sobre* A/P. ■

O corpo A/P pode ser identificado com a imagem de A sob o homomorfismo canônico de B sobre $B/\mathbb{b}$ e, portanto, pode ser considerado como subcorpo do anel $B/\mathbb{b}$. Em particular, para cada ideal primo Q de B que está acima de P, o corpo B/Q pode ser considerado como uma extensão do corpo A/P, de grau $[B/Q : A/P] = dim(B/Q)$, o qual é chamado **grau de inércia** de Q e denotado por $f(Q \mid P)$. O maior número e tal que Q^e divida $P.B$ é chamado de **índice de ramificação** de Q e denotado por $e(Q \mid P)$.

O próximo teorema fornece uma igualdade fundamental de muita utilidade.

Teorema 2.2.11 *Sejam* A *um domínio de Dedekind,* $B = I_{\mathbb{F}}(A)$, *onde* $\mathbb{F}$ *é uma extensão finita e separável do corpo de frações* $\mathbb{K}$ *de* A, P *um ideal primo não nulo do anel* A *e* $Q_1, \cdots, Q_r$ *os ideais primos distintos do anel* B *que estão acima do ideal primo* P. *Então temos a seguinte igualdade* $\sum_{i=1}^{r} e(Q_i \mid P).f(Q_i \mid P) = [B/P.B : A/P] = [\mathbb{F} : \mathbb{K}]$. ■

Pelo Teorema Chinês do Resto, Teorema 2.1.3, existe um isomorfismo canônico de $B/P.B = B/(Q_1^{e_1}.\cdots.Q_r^{e_r})$ sobre o produto cartesiano $B/Q_1^{e_1} \times \cdots \times B/Q_r^{e_r}$, que induz a aplicação diagonal de A/P sobre $A/P \times \cdots \times A/P$ (r vezes).

A igualdade fundamental do Teorema 2.2.11 permite vários tipos de decomposição de P. Convém dar nomes aos casos extremos, e no caso de $n = 2$ são obviamente os únicos. Dizemos que o ideal primo P de A é

- **Totalmente decomposto** em $\mathbb{F}$, quando $r = n$, ou seja, $e(Q \mid P) = f(Q \mid P) = 1$ para todo ideal primo Q de B que está acima de P.
- **Totalmente inerte** em $\mathbb{F}$, quando $r = 1$ e $e(Q \mid P) = 1$, ou seja, $f(Q \mid P) = n$ para algum ideal primo Q de B que está acima de P.
- **Totalmente ramificado** em $\mathbb{F}$, quando $r = 1$ e $f(Q \mid P) = 1$, ou seja, $e(Q \mid P) = n$ para algum ideal primo Q de B que está acima de P.

A seguir veremos um teorema de decomposição de um número primo num corpo quadrático.

Teorema 2.2.12 *Seja* $\mathbb{F} = \mathbb{Q}(\sqrt{d})$, *o corpo quadrático associado com o inteiro livre de quadrados* d. *Então,*

- *Os primos impares* p, *onde* d *é um resíduo quadrático módulo* p, *decompõe em* $\mathbb{F}$. *Assim como 2, se* $d \equiv 1 \ (mod\ 8)$

- *Os primos impares* p, *onde* d *não é um resíduo quadrático módulo* p, *permanece primo em* $\mathbb{F}$. *Assim como 2, se* $d \equiv 5 \ (mod\ 8)$.

- *Todo primo impar* p *divisor de* d *ramifica em* $\mathbb{F}$. *Assim como 2, se* $d \equiv 2$ *ou* $3 \ (mod\ 4)$. ■

Exemplo 2.2.5 *Seja a extensão* $\mathbb{Q} \subseteq \mathbb{Q}(i)$. *Esta extensão é de grau 2 e* $A = \mathbb{Z}[i]$ *é o anel dos inteiros algébricos. Assim, pelo Teorema 2.2.12 temos que*

- *O número primo 2 ramifica em* $\mathbb{Q}(i)$, *isto é,* $< 2 >=< 1+i >^2$.

- *Os primos da forma* $4k+1$, $k \in \mathbb{N}$ *decompõem em* $\mathbb{Q}(i)$, *como por exemplo,* $< 5 >=< 2+i >< 2-i >$ *e* $< 13 >=< 3+2i >< 3-2i >$.

- *Os primos da forma* $4k+3$, $k \in \mathbb{N}$ *permanecem primos em* $\mathbb{Q}(i)$. ■

Dados um número primo impar p e qualquer $c \in \mathbb{Z}$, não divisível por p, dizemos que c é um **resíduo quadrático (ou um resto quadrático)** módulo p (respectivamente, c **não é um resíduo quadrático** módulo p) se $c \equiv x^2 \ (mod\ p)$ (respectivamente, $c \not\equiv x^2 \ (mod\ p)$). O **símbolo de Legendre** é definido da seguinte maneira

$$\left(\frac{c}{p}\right) = \begin{cases} 1 \text{ quando } c \text{ é um resto quadrático módulo } p \\ \\ -1 \text{ quando } c \text{ não é um resto quadrático módulo } p. \end{cases}$$

Para finalizar esta seção veremos um teorema que fornece um método de calcular o símbolo de Legendre.

Teorema 2.2.13 *Dados* p, q *números primos impares e* a, $b \in \mathbb{Z}$ *não divisíveis por* p, *temos que:*

- $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right) \cdot \left(\frac{b}{p}\right)$.
- $\left(\frac{a}{p}\right) = a^{(p-1)/2} \ (mod\ p)$.
- $\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}$.
- $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$ *e* $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$. ■

2.3 Códigos de Bloco

Esta seção tem a finalidade de familiarizar o leitor com os conceitos básicos e essenciais a respeito de códigos de bloco. O ponto de partida é um conjunto A, que pode ser finito ou infinito, chamado **alfabeto**.

Definição 2.3.1 *Sejam A um alfabeto e n um número natural.*

- *Um* **espaço de seqüências** A^I *é o conjunto de todas as seqüências* $\mathbf{c} = \{c_i : i \in I\}$ *de elementos* c_i *sobre* A, *onde* I *é um conjunto de índices.*
- *Um* **código** $\mathcal{C}$ *sobre* A *é qualquer subconjunto não vazio do espaço de seqüências* A^I.
- *Um* **código de bloco** $\mathcal{C}$ **de comprimento** n *sobre* A *é qualquer subconjunto não vazio do conjunto* A^n *de todas as seqüências* $\mathbf{c} = \{c_i : 1 \leq i \leq n\}$ *(que agora chamamos "palavras-código").*
- *Dados dois elementos* $\mathbf{u}, \mathbf{v} \in A^n$, *definimos a* **distância de Hamming** *entre* $\mathbf{u}$ *e* $\mathbf{v}$ *como sendo* $d(\mathbf{u}, \mathbf{v}) = \#\{i : u_i \neq v_i, 1 \leq i \leq n\}$. ■

Seja $\mathcal{C}$ um código sobre um alfabeto A.

- As coordenadas de uma palavra-código são chamadas de **símbolos**.
- A cardinalidade de um código finito é chamado **tamanho** do código.
- Quando o conjunto de índices I é finito temos o que chamamos de **código de bloco**. Neste caso, a cardinalidade de I, denotada por n, é denominada **comprimento** do código.
- Quando I é infinito temos um **código convolucional** ou um **código de treliça**.

Neste trabalho, consideramos somente o caso em que os alfabetos são finitos. Em geral, é conveniente que o alfabeto A tenha uma estrutura algébrica que pode ser grupo, anel ou corpo, de modo que a codificação e a decodificação sejam simplificadas. Neste trabalho, estaremos trabalhando somente com códigos de bloco, onde I é um conjunto finito.

O desempenho de um código pode ser avaliado através de uma distância. Em geral, quando pensamos em distância logo lembramos o conceito de métrica que é a estrutura matemática que mais

se aproxima do conceito intuitivo de medir distâncias num determinado conjunto. Assim, notamos que para quaisquer elementos $\mathbf{u}, \mathbf{v}$ e $\mathbf{w}$ em A^n, temos que a distância de Hamming é uma métrica, ou seja, são válidas as seguintes propriedades:

- $d(\mathbf{u}, \mathbf{v}) \geq 0$, valendo a igualdade se, e somente se, $\mathbf{u} = \mathbf{v}$ (positividade).
- $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$ (simetria).
- $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ (desigualdade triangular).

Definição 2.3.2 *Seja $\mathcal{C}$ um código de bloco. A* **distância de Hamming mínima** *do código $\mathcal{C}$ é definida como o número $d = min\{d(\mathbf{u}, \mathbf{v}) : \mathbf{u}, \mathbf{v} \in \mathcal{C} \text{ e } \mathbf{u} \neq \mathbf{v}\}$.* ■

Observamos que para um código de bloco $\mathcal{C}$ de comprimento n tal que $|\mathcal{C}| \geq 2$ temos que $1 \leq d \leq n$. Se $|\mathcal{C}| = 1$ adotamos por convenção que $d = \infty$

Aparentemente, para calcular d é necessário calcular $\begin{pmatrix} M \\ 2 \end{pmatrix}$ distâncias, onde M é o número de palavras do código, o que tem custo computacional elevado. Veremos adiante como se pode calcular d de modo mais econômico em códigos com estrutura de espaço vetorial.

Existem outras distâncias tais como as distâncias de Lee [34], de Mannheim [30] e Euclidiana [40]. A distância Euclidiana é usada quando associamos um conjunto de pontos do $\mathbb{R}^n$ ao código. Contudo, nos Capítulos 3, 4 e Seção 5.1 do Capítulo 5 deste trabalho, usamos somente a distância de Hamming e na Seção 5.2 do Capítulo 5 faremos uso da distância de Lee.

2.3.1 Códigos sobre Corpos Finitos

Nesta seção, faremos uma breve revisão de códigos sobre corpos finitos necessários para o desenvolvimento deste trabalho. A teoria apresentada nesta seção segue de [35].

Definição 2.3.3 *Se o alfabeto A for um corpo finito $\mathbb{F}_q = GF(q)$, onde $q = p^m$, com p um número primo e m um inteiro positivo, dizemos que um código de bloco $\mathcal{C}$ sobre $\mathbb{F}_q$ é* **linear** *se for um subespaço vetorial de $\mathbb{F}_q^n$.* ■

Na Definição 2.2.1, de maneira implícita foi definido o parâmetro n, o comprimento do código. A seguir, definimos dimensão e taxa, que são outros dois parâmetros importantes de um código de bloco.

Definição 2.3.4 *Seja* $\mathcal{C}$ *um código de bloco linear.*

- *A* **dimensão** *do código* $\mathcal{C}$ *é definida como* $k = log_{|A|} \mid \mathcal{C} \mid$ *símbolos por bloco, onde* $|\cdot|$ *denota a cardinalidade do conjunto.*
- *A* **taxa** *do código* $\mathcal{C}$ *é definida como* $R_{\mathcal{C}} = k/n$, *onde* k *é a dimensão e* n *é o comprimento do código.*
- *Um código de bloco* $\mathcal{C}$ *de comprimento* n, *dimensão* k *e distância de Hamming mínima* d *é chamado um* **(n,k,d)-código**. ■

Quando k é inteiro, dizemos que a palavra-código contém k dígitos de **informação** e $n-k$ dígitos de **redundância**.

Há várias maneiras de descrever um código, podemos por exemplo, dar uma base formada por k vetores, isto é, um conjunto de vetores (palavras-código) linearmente independentes sobre $GF(q)$. Assim, é possível gerar q^k vetores distintos (o número total de palavras-código), e que a dimensão do código é dada por $log_{|A|} \mid \mathcal{C} \mid = log_q q^k = k$. Consequentemente, k vetores linearmente independentes de $\mathbb{F}_q^n$ geram um (n,k)-código de bloco linear sobre $\mathbb{F}_q$, onde q é uma potência de primo.

Se $\mathbf{g_1}, \mathbf{g_2}, \cdots, \mathbf{g_k}$ são vetores de $\mathbb{F}_q^n$ que geram um (n,k)-código linear $\mathcal{C}$, onde $\mathbf{g_i} = (g_{i1}, g_{i2}, \cdots, g_{in})$, $i = 1, 2, \cdots, k$, então a aplicação $f : \mathbb{F}_q^k \to \mathbb{F}_q^n$ definida por $f(u) = f(a_1, a_2, \cdots, a_k)$ $= \sum_{i=1}^{k} a_i \mathbf{g_i} = (\sum_{i=1}^{k} a_i g_{i1}, \cdots, \sum_{i=1}^{k} a_i g_{in})$ é um "codificador", isto é, pensando $\mathbb{F}_q^k$ como o conjunto das palavras numa linguagem "natural", a função f nos diz como codificar as palavras. Deste modo, qualquer palavra $\mathbf{v} \in \mathcal{C}$ pode ser expressa na forma $\mathbf{v} = \mathbf{u}G$, onde $\mathbf{u} = (a_1, \cdots, a_k) \in \mathbb{F}_q^k$ e

$$G = \begin{bmatrix} \mathbf{g_1} \\ \mathbf{g_2} \\ \vdots \\ \mathbf{g_k} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k1} & g_{k2} & \cdots & g_{kn} \end{bmatrix}.$$

A matriz $G = [g_{ij}]$, que descreve a aplicação linear f, é chamada **matriz geradora** do código $\mathcal{C}$. Dizemos que G está na forma sistemática se $G = [I_k \mid P_{k \times n-k}]$ para alguma matriz P, isto é, $g_{ij} = \delta_{ij}$ para $i, j = 1, 2, \cdots, k$ (onde $\delta_{ij} = 0$ se $i \neq j$ e $\delta_{ii} = 1$). Neste caso, dizemos que o código é **sistemático**, e que os primeiros k símbolos ou coordenadas de $\mathbf{c} \in \mathcal{C}$ são os símbolos de informação, e os demais símbolos de paridade.

Os símbolos de controle tem a seguinte finalidade. Dado um vetor $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ em $\mathbb{F}_q^n$, para checar se $\mathbf{x} \in \mathcal{C}$ basta verificar se $\mathbf{x} = f(\mathbf{u})$ para algum $\mathbf{u} \in \mathbb{F}_q^k$ e neste caso (estamos supondo $\mathcal{C}$ na forma sistemática), $x_j = a_j$, $1 \leq k$ e $x_j = \sum_{i=1}^{k} a_i g_{ij} = \sum_{i=1}^{k} x_i g_{ij}$, para $j > k$. Logo, para verificar se $\mathbf{x} \in \mathcal{C}$ basta verificar se $x_j = \sum_{i=1}^{k} x_i v_{ij}$, para $j = k+1, \cdots, n$.

Note que a matriz G não é univocamente determinada por $\mathcal{C}$, pois ela depende da escolha da base. Duas matrizes geradoras de um mesmo código $\mathcal{C}$ podem ser obtidas uma da outra por uma seqüência de operações elementares do tipo:

- Troca de duas linhas,
- Multiplicação de uma linha por um elemento não nulo do corpo $\mathbb{F}_q$,
- Adição de qualquer múltiplo de uma linha à outra.

Se além disso, efetuarmos seqüências de operações sobre G do tipo

- Permutação de duas colunas.
- Multiplicação de uma coluna por um escalar não nulo,

obtemos uma matriz G' de um código equivalente a $\mathcal{C}$. Inversamente, podemos construir códigos a partir de matrizes geradoras G. Para isto basta tomar uma matriz cujas linhas são linearmente independentes.

Note que pelo fato de G ter posto k é sempre possível reduzí-la na forma escalonada através das operações elementares sobre as suas linhas.

Dizemos que dois códigos são **equivalentes** se um pode ser obtido do outro através de permutações de suas cordenadas. Pode-se provar que todo código é equivalente a um código que pode ser gerado por uma matriz na forma sistemática (Teorema 3.2 de [39]).

Pelo que vimos acima, a informação contida numa palavra $\mathbf{c} \in \mathcal{C}$ depende de k de suas coordenadas e as demais coordenadas representam a redundância usada para a verificação da paridade. A taxa do código $\mathcal{C}$, denotada por $R_{\mathcal{C}} = k/n$, é a razão entre o número de coordenadas "informativas" e o número total de coordenadas.

Uma outra maneira de descrever um código é estabelecer uma aplicação linear sobrejetora $h : \mathbb{F}_q^n \to \mathbb{F}_q^k$ tal que o núcleo de h é o código $\mathcal{C}$. Neste caso, temos que $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \mathcal{C}$

se, e somente se, $h(x_1, x_2, \cdots, x_n) = \mathbf{0}$. Se temos uma matriz H de posto k e ordem $n-k$ por n dada por $H = [h_{ij}]$, então esta última equação se escreve como $\sum_{i=1}^{n} h_{ij}x_j = 0$, $i = 1, 2, \cdots, k$, isto é, $\mathbf{x}H^T = \mathbf{0}$, onde H^T significa a transposta da matriz H. A matriz H é chamada **matriz verificação de paridade** do código $\mathcal{C}$ e gera um $(n, n-k)$-código sobre $\mathbb{F}_q$ chamado **código dual** de $\mathcal{C}$.

Note que num código $\mathcal{C}$ com matriz geradora G, para verificar que um determinado vetor $\mathbf{c} \in \mathbb{F}_q^n$ pertence ou não a $\mathcal{C}$, é preciso verificar se o sistema de equações com k incógnitas $\mathbf{x} = (x_1, x_2, \cdots, x_k)$ dado por $\mathbf{x}G = \mathbf{c}$, admite solução. Em geral, esta questão para ser respondida, requer um custo computacional elevado. No entanto, trabalhando com uma matriz verificação de paridade H, a questão pode ser respondida bem mais rapidamente. É só verificar se é nulo o vetor $\mathbf{c}H^T$, o que é fácil de realizar computacionalmente.

Dados um código $\mathcal{C}$ com matriz verificação de paridade H e um vetor $\mathbf{c} \in \mathbb{F}_q^n$, chamamos o vetor $\mathbf{c}H^T$ de **síndrome** de $\mathbf{c}$.

Os códigos devem ser tomados de tal modo que duas palavras-código devam ser bem diferentes, de modo que ao recebermos uma mensagem com possíveis erros possamos decodificá-la como a palavra-código mais parecida com a mensagem recebida. Para formalizar este conceito revemos a seguinte definição.

Definição 2.3.5 *O* **peso de Hamming** *de um vetor $\mathbf{v} \in \mathbb{F}_q^n$, denotado por $w_H(\mathbf{v})$, é o número de componentes não nulas de $\mathbf{v}$. O* **peso de Hamming mínimo** *de um código $\mathcal{C}$ é o número $w_H(\mathcal{C}) = min\{w_H(\mathbf{c}) : \mathbf{u} \in \mathcal{C} \ e \ \mathbf{u} \neq \mathbf{0}\}$.* ■

Como a distância de Hamming entre dois vetores $\mathbf{u}$ e $\mathbf{v}$ é o número de coordenadas em que eles diferem, segue, portanto, que a distância entre estes vetores é dada por $d(\mathbf{u}, \mathbf{v}) = w_H(\mathbf{u} - \mathbf{v})$. Além disso, se $\mathbf{u}$ e $\mathbf{v}$ são palavras de um código linear $\mathcal{C}$ então $\mathbf{u} - \mathbf{v} \in \mathcal{C}$, assim, a distância entre quaisquer duas palavras é igual ao peso de uma outra palavra e, portanto, a distância de Hamming mínima de um código linear é igual ao peso de Hamming mínimo de suas palavras não nulas. Portanto, em um código de bloco linear, para determinar a distância de Hamming mínima do código é equivalente a determinar seu peso de Hamming mínimo.

Observação 2.3.1 *Do ponto de vista da Álgebra Linear não se faz nenhuma distinção entre dois subespaços vetoriais de mesma dimensão de um dado espaço vetorial, pois sempre existe um automor-*

fismo linear do espaço ambiente levando um subespaço no outro. No entanto, do ponto de vista da Teoria de Códigos, dois subespaços de mesma dimensão podem ter pesos distintos. ■

Passamos, agora, a análise da correção e deteção de erros. Dizemos que um código $\mathcal{C}$ **corrige até** t **erros** se para todo $\mathbf{y} \in \mathbb{F}_q^n$ existe no máximo um único $\mathbf{x} \in \mathcal{C}$ com $d(\mathbf{x}, \mathbf{y}) \leq 2t + 1$. Isto significa que ao recebermos uma mensagem $\mathbf{y}$ com no máximo t erros, devemos ter que $\mathbf{y}$ difere de algum elemento $\mathbf{x} \in \mathcal{C}$ no máximo em t coordenadas. Sendo $\mathbf{x}$ a mensagem enviada, então $\mathbf{x}$ é o único elemento de $\mathcal{C}$ mais próximo de $\mathbf{y}$ e, portanto, podemos recuperar $\mathbf{x}$ a partir de $\mathbf{y}$. Entretanto, em geral, o número de palavras-código é muito grande e este método torna-se complexo e impraticável. Deste modo, um dos objetivos da teoria de códigos corretores de erros é o de propor métodos mais eficientes de decodificação.

Dizemos que um decodificador **detecta** erros quando ele determina que a palavra recebida não é uma palavra-código. Se, além disso, o decodificador ainda for capaz de determinar corretamente a palavra-código transmitida, dizemos que ele **corrige** erros. Um decodificador que somente detecta erros é usado em sistemas de comunicações onde é possível a retransmissão de dados, enquanto que um decodificador que corrige erros é usado mais frequentemente em sistemas de comunicações onde não é possível essa retransmissão.

Dado um código $\mathcal{C}$ com distância mínima d, definimos $t = [\frac{d-1}{2}]$, onde $[x]$ representa a parte inteira de um número real x. A importância da distância mínima d de um código é revelada no seguinte resultado, que é válido para qualquer distância.

Teorema 2.3.1 *Seja $\mathcal{C}$ um código com distância mínima d. Então $\mathcal{C}$ pode corrigir todos os padrões de erros de peso menor ou igual a t e detectar todos os padrões de erros de peso menor ou igual a $d - 1$.* ■

Note que em virtude do Teorema 2.3.1, um código terá maior capacidade de correção de erros quanto maior for a sua distância mínima. Portanto, é fundamental para a Teoria de Códigos poder calcular d ou pelo menos determinar uma cota inferior.

O Teorema 2.3.1 também permite traçar uma estratégia para a detecção e a correção de erros. Seja $\mathcal{C}$ um código com distância mínima d e seja t como acima. Quando o receptor recebe uma palavra $\mathbf{v}$, uma das seguintes situações é verificada.

- A palavra $\mathbf{v}$ encontra-se numa esfera de raio t em torno de uma palavra-código $\mathbf{c}$ (esta palavra é única, pelo Teorema 2.3.1). Neste caso, devemos substituir $\mathbf{v}$ pela palavra $\mathbf{c}$.

- A palavra $\mathbf{v}$ não se encontra em nenhuma esfera de raio t em torno de uma palavra $\mathbf{c}$ do código. Neste caso, não é possível decodificar $\mathbf{v}$ com boa margem de segurança.

Observe que no primeiro item não se pode ter certeza absoluta de que $\mathbf{c}$ tenha sido a palavra transmitida, pois poderíamos ter cometido mais do que t erros, afastando assim $\mathbf{v}$ da palavra transmitida e aproximando-a de outra palavra-código. A questão deve ser encarada em termos probabilísticos. Observe também que a hipótese no segundo item não ocorre em códigos perfeitos.

O teorema seguinte é de grande importância para que possamos calcular a distância de Hamming mínima de um código linear sobre $\mathbb{F}_q$.

Teorema 2.3.2 *Seja $\mathcal{C}$ um código com matriz verificação de paridade H e peso de Hamming mínimo w_H. Então quaisquer $w_H - 1$ ou menos colunas de H são linearmente independentes e existem w_H colunas de H linearmente dependentes.* ■

O próximo corolário fornece um limitante superior para a distância de Hamming mínima de um código $\mathcal{C}$ em função dos parâmetros n e k.

Corolário 2.3.1 (Limitante de Singleton) *Se $\mathcal{C}$ é um (n, k, d)-código, então a distância de Hamming mínima d satisfaz $d \leq n - k + 1$.* ■

Códigos com distância mínima $d = n - k + 1$ são chamados **códigos de máxima distância de separação** (maximum distance separable). Estes códigos têm uma descrição interessante como conjuntos satisfazendo certas propriedades geométricas em espaços projetivos sobre corpos finitos.

Passamos agora a ver a importância prática dos parâmetros distância mínima, capacidade de correção e taxa de um código corretor de erros.

Suponha que a mensagem $\mathbf{u} = (u_1, \cdots, u_k)$ corresponda a palavra-código $\mathbf{c} = (c_1, c_2, \cdots, c_n)$ que é enviada através do canal. Devido ao efeito do ruído introduzido pelo canal, o vetor recebido $\mathbf{b} = (b_1, b_2, \cdots, b_n)$ pode ser diferente de $\mathbf{c}$. Definimos então o **vetor erro** como $\mathbf{e} = \mathbf{b} - \mathbf{c} = (e_1, e_2, \cdots, e_n)$.

O decodificador deve decidir a partir de $\mathbf{b}$ qual foi a palavra-código enviada. Da natureza aleatória do ruído, o decodificador não é capaz de determinar com precisão qual foi o vetor erro

que ocorreu. Deste modo, ele escolherá o vetor erro que tenha ocorrido com maior probabilidade, dado que $\mathbf{b}$ foi recebido. Isto minimizará a probabilidade de erro de palavra. Supondo que todas as palavras sejam igualmente prováveis, este método é ótimo no sentido em que minimiza a probabilidade de que o decodificador cometa um erro. Este processo é chamado **decodificação por máxima verossimilhança**.

Dizemos que não ocorreram erros durante a transmissão em uma determinada palavra-código quando o vetor erro for todo nulo. Agora, se X é uma variável aleatória que representa o número de erros ocorridos em uma transmissão, isto é, X é o número de componentes não nulas de $\mathbf{e}$. Em geral, se os erros forem independentes, tem-se que $P(X = 0) > P(X = 1) > \cdots > P(X = n)$. Assim, o método de decodificação por máxima verossimilhança será selecionar o vetor erro de menor peso de Hamming. Este tipo de decodificação é chamado **decodificação pelo vizinho mais próximo**.

Um teste comum realizado para se detectar a presença de erros em códigos lineares sobre $I\!F_q$ é o cálculo do **vetor síndrome** $\mathbf{s}$ de um vetor recebido $\mathbf{b}$, que é definido como $\mathbf{s} = \mathbf{b}H^T$. Assim, $\mathbf{s} = \mathbf{0}$ se, e somente se, $\mathbf{b} \in \mathcal{C}$. Mas, o contrário não é verdadeiro, pois dependendo do vetor erro, o vetor recebido poderá ser uma outra palavra-código diferente da enviada. Isso nos conduz a seguinte definição.

Definição 2.3.6 *A* **probabilidade de erro da palavra-código**, *denotada por* P_e, *para um determinado esquema de decodificação é a probabilidade de que a saída do decodificador seja uma palavra-código diferente da enviada.* ■

Se o código corretor de erros (linear ou não) possuir M palavras $\mathbf{c_1}, \mathbf{c_2}, \cdots, \mathbf{c_M}$ as quais são usadas com igual probabilidade, temos que

$$P_{\mathrm{e}} = (1/M) \sum_{i=1}^{M} \mathrm{Prob}\{\text{saída do decodificador } \neq \mathbf{c_i} \mid \mathbf{c_i} \text{ foi enviada}\}.$$

O princípio básico da teoria de códigos corretores de erros é a introdução de redundâncias nas mensagens enviadas. Quanto mais redundância for introduzida nas palavras de um código, maior serão a distância mínima e a sua capacidade de correção de erros e com isto a probabilidade de erro diminui. Contudo, a taxa do código diminui a medida que a redundância aumenta. O que se espera na prática é construir bons códigos que apresentem boas taxas e baixas probabilidades de erro. Deste modo, podemos perceber que estes objetivos são conflitantes, mas ainda assim a teoria de Shannon garante a existência de bons códigos.

2.4 Conclusões

Através deste capítulo fizemos uma breve revisão dos conceitos básicos de álgebra abstrata, teoria dos números algébricos e teoria de códigos lineares sobre corpos finitos, enfocando seus principais parâmetros e propriedades, e de anéis dos inteiros algébricos de um corpo numérico, com a finalidade de encontrar anéis finitos com identidade.

Códigos lineares podem ser definidos com símbolos escolhidos de um conjunto de tamanho arbitrário. Contudo, quase todos os resultados principais da teoria da codificação e decodificação são obtidos assumindo que o alfabeto do código são elementos de um corpo finito. Em vista disto, nos próximos capítulos passamos a procura de extensões dos códigos de Hamming, Reed Solomon, BCH e alternantes sobre anéis finitos com identidade.

Capítulo 3

Códigos sobre Anéis Comutativos Finitos com Identidade

Neste capítulo temos dois objetivos básicos. O primeiro é estender o conceito de códigos via anéis de grupo sobre anéis comutativos finitos com identidade, e o segundo é estender os códigos de Hamming e Reed-Solomon sobre anéis comutativos finitos locais com identidade.

Na Seção 3.1 veremos o conceito de códigos cíclicos sobre anéis finitos e que códigos lineares sobre anéis finitos locais possuem propriedades de distâncias tão boas quanto, ou melhores que códigos lineares definidos sobre anéis finitos quaisquer. Na Seção 3.2, estendemos o conceito de códigos via anéis de grupo sobre anéis finitos. Na Seção 3.3, apresentamos uma técnica de construção de códigos de Hamming sobre anéis finitos locais provenientes do anel dos inteiros algébricos de um corpo numérico (provenientes dos anéis de Dedekind), e também um algoritmo de decodificação para estes códigos. Na Seção 3.4, apresentamos uma técnica de construção de códigos Reed-Solomon sobre anéis finitos locais quaisquer. Um algoritmo eficiente de decodificação para estes códigos será apresentado na Seção 4.3. Na Seção 3.5, faremos uma proposta alternativa, baseada nas Seções 4.6 e 4.7 de [31], através da qual códigos lineares sobre anéis finitos podem ser obtidos via concatenação generalizada de códigos sobre anéis finitos locais.

3.1 Códigos Cíclicos

Seja A um anel finito. Nesta seção, apresentamos definições e teoremas envolvendo códigos cíclicos sobre A, os quais serão de grande utilidade na construção de códigos BCH sobre o anel A.

A seguir veremos que códigos cíclicos de comprimento n sobre A são ideais no anel quociente $A_n = \frac{A[x]}{<x^n-1>}$, onde 1 é a identidade de A. Assim, os principais teoremas que caracterizam os polinômios geradores para códigos cíclicos sobre $GF(p)$ podem ser estendidos em muitas situações para códigos cíclicos sobre A. Como uma consequência, em geral, códigos cíclicos sobre A podem também ser descritos pelas matrizes geradora e verificação de paridade.

Definição 3.1.1 *Sejam A um anel finito e n um número inteiro positivo.*

- *Um* **código linear** *$\mathcal{C}$ de comprimento n sobre A é o conjunto das n-uplas de A^n que formam um A-módulo.*
- *Um* **(n,k)-código linear** *$\mathcal{C}$ sobre A é um A-módulo livre de dimensão k no espaço das n-uplas A^n.*
- *Um código linear $\mathcal{C}$ sobre A é* **cíclico** *se $\mathbf{c}^{(1)} = (c_{n-1}, c_0, \cdots, c_{n-2}) \in \mathcal{C}$ sempre que $\mathbf{c} = (c_0, c_1, \cdots, c_{n-1}) \in \mathcal{C}$ com $c_i \in A$, $i = 0, \cdots, n-1$. Isto é, $\mathcal{C}$ é invariante com respeito as permutações cíclicas das coordenadas.* ■

Se o código linear $\mathcal{C}$ possuir uma base, ou seja, é um A-módulo livre, sua dimensão é o menor número de vetores que o geram. Assim, códigos que possuem uma base podem ser descritos através das matrizes geradora e verificação de paridade. Se o código linear for apenas finitamente gerado tomamos o menor conjunto de vetores que o geram, e assim podem ser descritos através da matriz geradora, e neste caso, pode acontecer que o número de geradores é maior que o comprimento do código. Contudo, sempre podemos definir códigos lineares sobre A a partir de matrizes verificação de paridade. Além disso, o interesse fundamental dos códigos cíclicos reside no fato de admitirem uma representação polinomial de suas palavras-código e com isso apresentam um procedimento simples de decodificação.

Seja o anel $A_n = \frac{A[x]}{<x^n-1>}$, quociente do anel de polinômios $A[x]$ pelo ideal gerado por $x^n - 1$. Seja $\mathbf{v} = (v_0, v_1, \cdots, v_{n-1})$ uma palavra de um código cíclico $\mathcal{C}$ sobre A. De maneira análoga, quando estes códigos são definidos sobre corpos finitos, representamos esta palavra na forma polinomial, isto

é, pelo polinômio $v(x) = v_0 + \cdots + v_{n-1}x^{n-1}$, cujo grau máximo é $n-1$. Fazendo o produto $xv(x)(mod\ x^n - 1)$, obtemos $v'(x) = v_{n-1} + v_0x + \cdots + v_{n-2}x^{n-1}$ que corresponde à palavra-código $\mathbf{v}' = (v_{n-1}, v_0, \cdots, v_{n-2}) = \mathbf{v}^{(1)}$. Portanto, $v^{(1)}(x)$ é obtida calculando o produto $xv(x)$ no anel quociente A_n, onde a soma de duas palavras-código é realizada em $A[x]$. Deste modo, códigos cíclicos sobre A podem ser caracterizados como os subconjuntos $\mathcal{C}$ de A_n tais que $c(x) \in \mathcal{C}$ acarreta em $xc(x) \in \mathcal{C}$.

Sendo $\mathcal{C}$ um código cíclico e tomando $c(x) \in \mathcal{C}$ temos que $xc(x) \in \mathcal{C}$, $x^2c(x) \in \mathcal{C}$, $\cdots$, logo, $b(x)a(x) \in \mathcal{C}$ para todo $b(x) \in A_n$. Mas, isto quer dizer que $\mathcal{C}$ é um ideal de A_n. Reciprocamente, se $\mathcal{C}$ é um ideal de A_n então $\mathcal{C}$ é um subconjunto de A_n e $xc(x) \in \mathcal{C}$ sempre que $c(x) \in \mathcal{C}$. Logo, $\mathcal{C}$ é um código cíclico. Resumindo, códigos cíclicos são os ideais de A_n, onde provamos o seguinte teorema.

Teorema 3.1.1 *Um conjunto $\mathcal{C} \subseteq A_n$ é um código cíclico se, e somente se, $\mathcal{C}$ é um ideal.* ■

Exemplo 3.1.1 *Considere o anel $A = \mathbb{Z}_4$ e o anel quociente $A_2 = \frac{A[x]}{<x^2-1>}$. Pelo Teorema 3.1.1 temos que o ideal $\mathcal{C} = \{0, 2, 2x, 2+2x, 1+3x, 3+x, 3+3x, 1+x\}$ de A_2 é um código cíclico.* ■

A questão que agora se coloca naturalmente é a seguinte: "Será que os códigos cíclicos definidos desta maneira possuem um polinômio gerador ?"

A resposta a esta questão foi apresentada por Interlando e Palazzo em [32] e Calderbank e Sloane em [13]. Interlando e Palazzo provaram que um código cíclico $\mathcal{C}$ no anel $A_n = \frac{\mathbb{Z}_m[x]}{<x^n-1>}$, onde m é um inteiro positivo, possui um polinômio gerador se o coeficiente dominante do polinômio de menor grau em $\mathcal{C}$ for uma unidade de $\mathbb{Z}_m$, e este polinômio, por sua vez, gera o código e divide $x^n - 1$. Neste caso, as matrizes geradora e verificação de paridade são obtidas de maneira análoga quando estes códigos são definidos sobre corpos finitos. Calderbank e Sloane provaram, usando números p-ádicos, que o anel $A_n = \frac{\mathbb{Z}_q[x]}{<x^n-1>}$, onde q é uma potência de primo, é principal. Assim, como todo código cíclico em A_n é um ideal, segue portanto, que todo código cíclico em A_n possui um polinômio gerador.

Os resultados obtidos por Interlando e Palazzo foram sobre o anel $A = \mathbb{Z}_m$, isto é, A sendo o anel dos inteiros módulo m. Com demonstrações análogas, os resultados ainda valem num contexto mais geral, isto é, sobre qualquer anel comutativo com identidade e, em particular, para anéis comutativos finitos com identidade, que é o caso de nosso interesse. A relevância do trabalho de Interlando e

Palazzo reside no fato de fornecerem um método de construir códigos cíclicos sobre anéis finitos, que é bem similar a construção destes códigos sobre corpos finitos. O procedimento para tal, passa a ser o seguinte: fatora-se o polinômio $x^n - 1$ num anel apropriado e, então seleciona-se um fator como o polinômio gerador. Na construção de códigos BCH, apresentada no Capítulo 4, veremos como realizar esta fatoração quando n divide $p^r - 1$, onde p é um número primo e r um inteiro positivo.

Observação 3.1.1 *Seja A um anel finito. Sejam a matriz*

$$M = \begin{bmatrix} \alpha_1^0 & \alpha_1^1 & \cdots & \alpha_1^{n-2} & \alpha_1^{n-1} \\ \alpha_2^0 & \alpha_2^1 & \cdots & \alpha_2^{n-2} & \alpha_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_r^0 & \alpha_r^1 & \cdots & \alpha_r^{n-2} & \alpha_r^{n-1} \end{bmatrix} \tag{3.1}$$

e o polinômio $f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} \in A[x]$. É de fácil verificação que $f(x)$ admite $\alpha_1, \alpha_2, \cdots, \alpha_r$ como raízes se, e somente se, o vetor correspondente $(a_0, a_1, \cdots, a_{n-1})$ pertence ao espaço nulo da matriz M. Deste modo, temos que o conjunto de todos os polinômios que tem $\alpha_1, \alpha_2, \cdots, \alpha_r$ como raízes é o espaço nulo da matriz M. ■

A seguir revemos um teorema, Teorema VI.2 de [37], que será de grande utilidade no desenvolvimento do restante desta seção e da Seção 4.2. Este teorema afirma que todo anel comutativo finito pode ser expresso, de maneira única, como um produto (soma) direto de anéis comutativos finitos locais.

Teorema 3.1.2 (Teorema de Estrutura para um Anel Comutativo Finito) *Seja A um anel comutativo finito. Então A pode ser decomposto de maneira única (a menos da ordem dos somandos) como o produto (soma) direto de anéis locais, ou seja, $A = A_1 \times A_2 \times \cdots \times A_r$, onde A_i, $i = 1, 2, \cdots, r$, são anéis comutativos finitos locais.* ■

O isomorfismo do Teorema 3.1.2 é definido como

$$\begin{aligned} \phi : A &\rightarrow A_1 \times A_2 \times \cdots \times A_r \\ a &\rightarrow \phi(a) = (\phi_1(a), \phi_2(a), \cdots, \phi_r(a)) = (a + P_1^m, a + P_2^m, \cdots, a + P_r^m) \end{aligned} \tag{3.2}$$

onde $\phi_i : A \rightarrow A_i$, $A_i = A/P_i^m$, $i = 1, 2, \cdots, r$ são os homomorfismos canônicos, P_i, $i = 1, 2, \cdots, r$, são os ideais primos de A e m o menor inteiro positivo tal que $\cap_{i=1}^r P_i^m = 0$.

Definição 3.1.2 *Se* $\dot{a}_i$ *são ideais em* A_i, *para* $i = 1, 2, \cdots, r$, *definimos o* **produto direto** $\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r$ *em* $A_1 \times A_2 \times \cdots \times A_r$ *como o conjunto dos elementos* $\{(a_1, a_2, \cdots, a_r) : a_i \in \dot{a}_i\}$, *onde todas as possíveis combinações de elementos são consideradas e, portanto, a adição e a multiplicação são definidas componente por componente. O produto direto* $\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r$ *é um ideal do produto direto* $A_1 \times A_2 \times \cdots \times A_r$. ■

Observação 3.1.2 *Seja* A *um anel comutativo finito com decomposição em anéis locais dada por* $A = A_1 \times A_2 \times \cdots \times A_r$. *Seja* $\dot{a}$ *um ideal de* A. *Então* $\phi_i(\dot{a}) \subseteq A_i$, $i = 1, 2, \cdots, r$, *é um ideal, pois se* $b_1, b_2 \in \phi_i(\dot{a})$, *então existe* $a_1, a_2 \in \dot{a}$ *tais que* $\phi_i(a_1) = b_1$ *e* $\phi_i(a_2) = b_2$, *e assim,* $b_1 + b_2 = \phi_i(a_1 + a_2) \in \phi_i(\dot{a})$ *e se* $\alpha \in A_i$ *e* $b \in \phi_i(\dot{a})$ *então existem* $\beta \in A$ *e* $a \in \dot{a}$ *tais que* $\phi_i(\beta) = \alpha$ *e* $\phi_i(a) = b$, *e assim* $\alpha b = \phi_i(\beta)\phi_i(a) = \phi_i(\beta a) \in \phi_i(\dot{a})$ *(como* $\dot{a}$ *é um ideal temos que* $a_1 + a_2$ *e* βa *pertencem a* $\dot{a}$*). Concluímos, deste modo, que* $\phi(\dot{a})$ *é um produto direto de ideais em* $A = A_1 \times A_2 \times \cdots \times A_r$, *ou seja, todo ideal de* A *é isomorfo a um produto direto de ideais* $\dot{a}_i$ *de* A_i, $i = 1, 2, \cdots, r$. *Por outro lado, se* $\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r$ *é um produto direto de ideais em* $A = A_1 \times A_2 \times \cdots \times A_r$, *então* $\phi^{-1}(\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r) \subseteq A$ *é um ideal. De fato, se* $a_1, a_2 \in \phi^{-1}(\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r)$, *então* $\phi(a_1), \phi(a_2) \in \dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r$ *e, portanto,* $\phi(a_1 + a_2) = \phi(a_1) + \phi(a_2) \in \dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r$, *ou seja,* $a_1 + a_2 \in \phi^{-1}(\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r)$. *Agora, se* $\alpha \in A$ *e* $a \in \phi^{-1}(\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r)$, *então* $\phi(a) \in \dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r$ *e, portanto,* $\phi(\alpha a) = \alpha\phi(a) \in \dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r$, *ou seja,* $\alpha a \in \phi^{-1}(\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r)$. *Logo,* $\phi^{-1}(\dot{a}_1 \times \dot{a}_2 \times \cdots \times \dot{a}_r)$ *é um ideal. Concluímos deste modo que o produto direto de ideais* $\dot{a}_i$ *de* A_i *é isomorfo a um ideal de* A. *Assim, existe uma correspondência biunívoca entre os ideais de* A *e o produto direto de ideais em* $A_1 \times A_2 \times \cdots \times A_r$, *e uma vez que os ideais de* A_i *são conhecidos, todo ideal de* A *pode ser obtido através de um produto direto apropriado e aplicar a inversa* ϕ^{-1}. ■

Exemplo 3.1.2 *Seja* A *o anel dos inteiros Gaussianos, isto é,* $A = \mathbb{Z}[i]$. *Do Exemplo 2.2.5 temos que* $< 5 >=< 2 + i >< 2 - i >$ *e* $< 2 + i > \cap < 2 - i >=< 0 >$ *e pelo Teorema 2.1.3 temos o isomorfismo* $\mathbb{Z}_5(i) \simeq \frac{\mathbb{Z}[i]}{<2+i>} \times \frac{\mathbb{Z}[i]}{<2-i>} \simeq \mathbb{Z}_5 \times \mathbb{Z}_5$. *Assim, os ideais de* $\mathbb{Z}_5[i]$ *são cópias isomorfas dos ideais* $\{0\} \times \{0\}, \{0\} \times \mathbb{Z}_5, \mathbb{Z}_5 \times \{0\}$ *e* $\mathbb{Z}_5 \times \mathbb{Z}_5$ *do anel* $\mathbb{Z}_5 \times \mathbb{Z}_5$. ■

A seguir veremos um teorema, cuja demonstração é bastante simples através do uso do Teorema 3.1.2, e que é válido para qualquer distância.

Teorema 3.1.3 *Seja* A *um anel comutativo finito com decomposição em anéis locais dada por*

$A = A_1 \times A_2 \times \cdots \times A_r$. *Se* $\mathcal{C}$ *é um código linear sobre* A, *então* $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_r$, *isto é,* $\mathcal{C}$ *é um produto (soma) direto de subcódigos lineares* $\mathcal{C}_i$ *sobre os anéis locais* A_i, $i = 1, 2, \cdots, r$. *Além disso, temos ainda que*

- $d_{min}(\mathcal{C}) = min_{1 \leq i \leq r}\{d_{min}(\mathcal{C}_i)\}$, *onde* $d_{min}(\mathcal{C}_i)$ *é a distância mínima do i-ésimo código* $\mathcal{C}_i$.
- $| \mathcal{C} | = \prod_{i=1}^{r} | \mathcal{C}_i |$.

Demonstração: A aplicação ϕ da equação (3.2) é um isomorfismo de A-módulos, pois para todo $\alpha, a \in A$ temos que $\phi(\alpha a) = (\alpha a + P_1^m, \alpha a + P_2^m, \cdots, \alpha a + P_r^m) = \alpha\phi(a)$. Agora, se $\mathcal{C}$ é um código linear sobre A, ou seja, um A-módulo, temos pelo isomorfismo ϕ que $\phi(\mathcal{C}) \subseteq A_1 \times A_2 \times \cdots \times A_r$ é um código linear sobre A. Assim, $\phi(\mathcal{C}) = \mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_r$, onde cada $\mathcal{C}_i$, $i = 1, 2, \cdots, r$, são A_i-módulos (para todo $\alpha = a + P_i^m \in A_i$ e para todo $c_i = c + P_i^m \in \mathcal{C}_i$, temos que $\alpha c_i = ac + P_i^m \in A_i$ e $\phi(ca) = (ac + P_1^m, ac + P_2^m, \cdots, ac + P_r^m)$, portanto $ac + P_i^m \in \mathcal{C}_i$, para todo $i = 1, 2, \cdots, r$, ou seja, $\mathcal{C}_i$ é um A_i-módulo para todo $i = 1, 2, \cdots, r$). Reciprocamente, se $\mathcal{C}_i$ são A_i-módulos para $i = 1, 2, \cdots, r$, então $\mathcal{C}_i$ são também A-módulos através da operação por escalar dada por $ac_i =: \phi_i(a)c_i$ para todo $a \in A$ e $c_i \in \mathcal{C}_i$. Deste modo, temos que $\mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_r$ é um A-módulo. Logo, sendo ϕ um isomorfismo, existe um A-módulo $\mathcal{C}$ tal que $\phi(\mathcal{C}) = \mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_r$, ou seja, $\mathcal{C} \simeq \mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_r$. ■

O próximo corolário, segue diretamente do Teorema 3.1.3.

Corolário 3.1.1 *Se* $\mathcal{C}$ *é um código cíclico sobre* A, *então* $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_r$, *onde cada* $\mathcal{C}_i$ *é um código cíclico sobre* A_i, $i = 1, 2, \cdots, r$. ■

3.2 Anéis de Grupos

Um (n, k)-código sobre o anel dos inteiros módulo m, $\mathbb{Z}_m$, pode ser visto como um ideal no anel de grupo $\mathbb{Z}_m G$, onde G é um grupo abeliano finito de ordem n. O parâmetro k é definido em termos da noção de posto de um submódulo sobre um anel de ideais principais. Nesta seção, estendemos o conceito de códigos para A um anel finito e G um grupo abeliano finito, onde veremos que um código sobre A pode ser visto como um ideal em AG, o anel de grupo de G sobre A. Antes, porém, veremos o conceito de anel de grupo.

Sejam G um grupo e A um anel com identidade. Denotamos por AG o conjunto de todas as combinações lineares formais do tipo

$$\sum_i a_i g_i, \quad a_i \in A, \quad g_i \in G$$

com adição, multiplicação por escalar e multiplicação dadas por

$$\begin{aligned} &\sum_i a_i g_i + \sum_i b_i g_i = \sum_i (a_i + b_i) g_i \qquad \text{(soma)} \\ &a(\sum_i a_i g_i) = \sum_i (a a_i) g_i \qquad \text{(produto por escalar)} \\ &(\sum_i a_i g_i)(\sum_j b_j g_j) = \sum_i \sum_j a_i b_j g_i g_j = \sum_k t_k g_k \qquad \text{(produto)} \end{aligned} \tag{3.3}$$

onde $a_i, b_i \in A$, $g_i \in G$, $g_i g_j = g_k$ e $t_k = \sum_{g_i g_j = g_k} a_i b_j$. Se o grupo G é infinito consideramos apenas combinações lineares tais que (a_i) é uma sequência quase nula, isto é, tal que apenas um número finito de coeficientes a_i é diferente de zero.

Definição 3.2.1 *Com as operações de soma e produto como em (3.3), AG é um anel chamado* **anel de grupo** *de G sobre A. Se além disso, A é comutativo, considerando as três operações, temos que AG é uma A-álgebra chamada* **álgebra de grupo** *de G sobre A. Em outras palavras, AG é um A-módulo livre cuja base é formada pelos elementos de G.* ■

Observação 3.2.1 *Considerando G, um grupo finito de ordem n e que seus elementos são escritos na ordem $g_0, g_1, \cdots, g_{n-1}$, temos que a todo elemento $a = \sum_{i=0}^{n-1} a_i g_i \in AG$ podemos associar (de maneira única) a n-upla $(a_0, a_1, \cdots, a_{n-1})$.* ■

3.2.1 Códigos via Anéis de Grupos

Blake em [8], [9] e Spiegel em [48], [49], propuseram construções de códigos via anéis de grupos sobre anéis de inteiros residuais módulo m, $\mathbb{Z}_m$. Nesta seção, propomos uma construção de códigos via anéis de grupos sobre anéis comutativos finitos com identidade.

Definição 3.2.2 *Sejam A um anel finito e G um grupo abeliano finito de ordem n.*

- *Um* **código linear** *de comprimento n sobre A é o conjunto das n-uplas associadas com os elementos de um ideal de AG.*

- *Se G é um grupo cíclico de ordem n, um* **código cíclico linear** *de comprimento n sobre A é o conjunto das n-uplas associadas com os elementos de um ideal de AG.* ■

Observação 3.2.2 *Os ideais de AG são do tipo $\dot{a}G$ com $\dot{a}$ um ideal de A. Com efeito, se $I \subseteq AG$ é um ideal, então $I = \{\sum_i a_i g_i : a_i \in \dot{a}\}$, onde $\dot{a} \subseteq A$ é um subconjunto. Tomando $x, y \in \dot{a}$ temos que $x.1$ e $y.1$ pertencem a I por definição. Como I é um ideal segue que $x.1 + y.1 = (x+y).1 \in I$, e portanto, $x + y \in \dot{a}$. Analogamente, prova-se que $\alpha x \in \dot{a}$ para todo $x \in \dot{a}$ e $\alpha \in A$. Por outro lado, temos claramente que $\dot{a}G$, onde $\dot{a} \subseteq A$ é um ideal, é um ideal de AG.* ■

Em vista do Teorema 3.1.2 temos o seguinte resultado.

Teorema 3.2.1 *Se A é um anel finito tal que $A = A_1 \times A_2 \times \cdots \times A_r$, onde os $A_i's$, são anéis locais, com $i = 1, 2, \cdots, r$, e se G é um grupo finito de ordem n, com elementos escritos na ordem $g_0, g_1, \cdots, g_{n-1}$, então AG é isomorfo a $A_1G \times A_2G \times \cdots \times A_rG$.*

Demonstração: O isomorfismo ϕ da equação (3.2) induz a aplicação ψ definida como

$$\begin{aligned} \psi : AG &\rightarrow A_1G \times A_2G \times \cdots \times A_rG \\ \sum_{i=0}^{n-1} a_i g_i &\rightarrow \psi(\sum_{i=0}^{n-1} a_i g_i) = \sum_{i=0}^{n-1} \phi(a_i) g_i = \sum_{i=0}^{n-1} (\phi_1(a_i), \phi_2(a_i), \cdots, \phi_r(a_i)) g_i \\ &= (\sum_{i=0}^{n-1} \phi_1(a_i) g_i, \sum_{i=0}^{n-1} \phi_2(a_i) g_i, \cdots, \sum_{i=0}^{n-1} \phi_r(a_i) g_i), \end{aligned} \tag{3.4}$$

onde as duas representações de elementos em $A_1G \times A_2G \times \cdots \times A_rG$ são equivalentes e usadas quando conveniente. A adição e a multiplicação em $A_1G \times A_2G \times \cdots \times A_rG$ são induzidas de AG mediante a aplicação ϕ, isto é, se

$$a = \sum_{i=0}^{n-1} \phi(a_i) g_i \quad \text{e} \quad b = \sum_{i=0}^{n-1} \phi(b_i) g_i$$

são dois elementos de $A_1G \times A_2G \times \cdots \times A_rG$ então

$$a + b = \sum_{i=0}^{n-1} (\phi(a_i) + \phi(b_i)) g_i \quad \text{e} \quad ab = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \phi(a_i)\phi(b_j) g_i g_j.$$

Agora, se $x = \sum_{i=0}^{n-1} a_i g_i$ e $y = \sum_{i=0}^{n-1} y_i g_i$ são dois elementos de AG então $\psi(x+y) = \sum_{i=0}^{n-1} \phi(a_i + b_i) g_i = \sum_{i=0}^{n-1} (\phi(a_i) + \phi(b_i)) g_i = \sum_{i=0}^{n-1} \phi(a_i) g_i + \sum_{i=0}^{n-1} \phi(b_i) g_i = \psi(x) + \psi(y)$ e $\psi(xy) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \phi(a_i b_j) g_i g_j =$

$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\phi(a_i)\phi(b_j)g_ig_j = \sum_{i=0}^{n-1}\phi(a_i)g_i\sum_{j=0}^{n-1}\phi(b_j)g_j = \psi(x)\psi(y)$, ou seja, ψ é um homomorfismo de anéis. Sendo AG um A-módulo livre gerado pelos elementos de G, é claro que ψ é injetora. Para mostrar que é sobrejetora, seja $b = (b_1, b_2, \cdots, b_r) \in A_1G \times A_2G \times \cdots \times A_rG$, onde $b_j = \sum_{i=0}^{n-1} a_{i_j}g_i$, para $j = 1, 2, \cdots, r$. Sendo ϕ sobrejetora e como $(a_{i_1}, a_{i_2}, \cdots, a_{i_r})$ é um elemento de $A_1 \times A_2 \times \cdots \times A_r$, para todo $i = 0, 1, \cdots, n-1$, existe $a_i \in A$ tal que $\phi(a_i) = (\phi_1(a_i), \phi_2(a_i), \cdots, \phi_r(a_i)) = (a_{i_1}, a_{i_2}, \cdots, a_{i_r})$. Tomando $a = \sum_{i=0}^{n-1} a_ig_i \in AG$ temos que $\psi(x) = \sum_{i=0}^{n-1}\phi(a_i)g_i = (\sum_{i=0}^{n-1}\phi_1(a_i)g_i, \sum_{i=0}^{n-1}\phi_2(a_i)g_i, \cdots, \sum_{i=0}^{n-1}\phi_r(a_i)g_i) = (\sum_{i=0}^{n-1}a_{i_1}g_i, \sum_{i=0}^{n-1}a_{i_2}g_i, \cdots, \sum_{i=0}^{n-1}a_{i_r}g_i) = (b_1, b_2, \cdots, b_r) = b$, ou seja, ψ é sobrejetora. Sendo ψ um homomorfismo injetor e sobrejetor, é portanto um isomorfismo. ■

Seja $I \subseteq AG$ um ideal. Pela Observação 3.2.2 temos que $I = \dot{a}G = \{\sum_{i=0}^{n-1} a_ig_i : a_i \in \dot{a}\}$, onde $\dot{a} \subseteq A$ é um ideal. Pela Observação 3.1.2 temos que a imagem deste ideal mediante a aplicação ψ é dada por

$$\psi(I) = (\{\sum_{i=0}^{n-1}\phi_1(a_i)g_i : a_i \in \dot{a}\} \times \{\sum_{i=0}^{n-1}\phi_2(a_i)g_i : a_i \in \dot{a}\} \times \cdots \times \{\sum_{i=0}^{n-1}\phi_r(a_i)g_i : a_i \in \dot{a}\})$$

o qual esta contida em $A_1G \times A_2G \times \cdots \times A_rG$. Como $I \subseteq AG$ é um ideal, ou seja, é fechado em relação a adição e a multiplicação por elementos de AG, segue que $I_l = \{\sum_{i=0}^{n-1}\phi_l(a_i)g_i : a_i \in \dot{a}\}$ é um ideal de A_lG para $l = 1, 2, \cdots, r$, pois contém o elemento zero, e é fechado em relação a adição e a multiplicação por elementos de A_lG. Assim, se I é um ideal de AG então $\psi(I)$ é um produto direto de ideais I_l de A_lG, $l = 1, 2, \cdots, r$, e reciprocamente, temos que um produto direto de ideais I_l em A_lG é isomorfo a um ideal de AG. Deste modo, existe uma correspondência bijetora entre os ideais de AG e o produto direto de ideais em $A_1G \times A_2G \times \cdots \times A_rG$, e uma vez que os ideais de A_lG são conhecidos, todo ideal de AG pode ser obtido através de um produto apropriado e aplicando a função inversa ψ^{-1}.

Deste modo, do Teorema 3.2.1, temos que conhecer os ideais (códigos) de AG é equivalente a conhecer os ideais (códigos) de A_iG, $i = 1, 2, \cdots, r$. Se, além disso, G é um grupo cíclico finito de ordem n gerado por um elemento $g \in G$, então conhecer os ideais e portanto, os códigos cíclicos sobre A, é equivalente a conhecer os códigos cíclicos sobre A_i, $i = 1, \cdots, r$. Temos também, pelo Teorema 3.1.3, que a distância mínima de um código em AG é igual ao $min_{1 \leq i \leq r}\{d_i\}$, onde d_i é

a distância mínima de um código em A_iG, $i = 1, 2, \cdots, r$.

Pela Observação 3.2.2 temos que os ideais de A_iG são da forma $\dot{a}_iG$, onde $\dot{a}_i$ é um ideal de A_i, para todo $i = 1, 2, \cdots, r$. Para encontrarmos os ideais de A_i, $i = 1, 2, \cdots, r$, necessitamos do seguinte teorema (Proposição 1.1 de [4]).

Teorema 3.2.2 *Seja A um anel e $\dot{a}$ um ideal de A. Existe uma bijeção preservando a ordem entre os ideais de A que contém $\dot{a}$ e os ideais de $A/\dot{a}$.* ■

De modo a exibirmos a construção de um código sobre A mais explicitamente, tomamos $a_l = \sum_{i=0}^{n-1} \phi_l(a_i)g_i \sim (a_0^{(l)}, a_1^{(l)}, \cdots, a_{n-1}^{(l)})$ uma palavra-código em A_l, $l = 1, 2, \cdots, r$, tal que o correspondente elemento em $A_1G \times A_2G \times \cdots \times A_rG$ é $(\sum_{i=0}^{n-1} a_i^{(1)}g_i, \sum_{i=0}^{n-1} a_i^{(2)}g_i, \cdots, \sum_{i=0}^{n-1} a_i^{(r)}g_i)$, o qual é identificado, mediante a aplicação inversa ψ^{-1}, com o elemento $\sum_{i=0}^{n-1} \alpha_i g_i \in AG$, onde $\alpha_i = \phi^{-1}(a_i^{(1)}, a_i^{(2)}, \cdots, a_i^{(r)})$, e com este elemento em AG identificamos a n-upla $(\alpha_0, \alpha_1, \cdots, \alpha_{n-1})$, como sendo uma palavra-código sobre A. Uma maneira mais clara deste processo é como segue: listamos as n-uplas correspondentes aos elementos a_l de AG, $l = 1, 2, \cdots, r$, como segue

$$a_1 = (a_0^{(1)}, a_1^{(1)}, \cdots, a_{n-1}^{(1)}) \in A_1G$$

$$a_2 = (a_0^{(2)}, a_1^{(2)}, \cdots, a_{n-1}^{(2)}) \in A_2G$$

$$\vdots \quad \vdots$$

$$a_r = (a_0^{(r)}, a_1^{(r)}, \cdots, a_{n-1}^{(r)}) \in A_rG$$

Assim a primeira posição da n-upla correspondente em AG é obtida pelas primeiras posições destas n-uplas como $\phi^{-1}(a_0^{(1)}, a_0^{(2)}, \cdots, a_0^{(r)})$ e, similarmente, para as outras coordenadas.

Exemplo 3.2.1 *Considere o anel de grupo $\mathbb{Z}_5[i]G$, onde G é um grupo cíclico de ordem 4. Do Exemplo 3.1.2 temos que $\mathbb{Z}_5[i]$ é isomorfo a $\mathbb{Z}_5 \times \mathbb{Z}_5$ e do Teorema 3.2.1 temos que $\mathbb{Z}_5[i]G$ é isomorfo a $\mathbb{Z}_5G \times \mathbb{Z}_5G$. Considere o (4,2) código $\mathcal{C}_1$ sobre $\mathbb{Z}_5$ gerado por $g_1(x) = 2+3x+x^2$ com distância de Hamming mínima 3, e o (4,1) código $\mathcal{C}_2$ sobre $\mathbb{Z}_5$ gerado por $g_2(x) = 1+x+x^2+x^3$ com distância de Hamming mínima 4. Portanto, o código $\mathcal{C} \simeq \mathcal{C}_1 \times \mathcal{C}_2$ sobre $\mathbb{Z}_5[i]$ tem distância de Hamming mínima 3. Assim, se $\mathbf{c}_1 = (2, 3, 1, 0)$ é uma palavra do código $\mathcal{C}_1$ e se $\mathbf{c}_2 = (1, 1, 1, 1)$ é uma palavra do código $\mathcal{C}_2$, a primeira coordenada da palavra $\mathbf{c}$ do código $\mathcal{C}$ é dada por $\phi^{-1}(2, 1) =$*

$4+4i$, *a segunda por* $\phi^{-1}(3,1) = 2+3i$, *a terceira por* $\phi^{-1}(1,1) = 1$ *e a quarta e última por* $\phi^{-1}(0,1) = 3+i$. ■

Sejam A um domínio de Dedekind e $\mathfrak{a} \subseteq A$ um ideal. Pelo Teorema 2.2.7, o ideal $\mathfrak{a}$ pode ser expresso como um produto de ideais primos, isto é, $\mathfrak{a} = \prod_{i=1}^{r} P_i^{e_i}$, onde P_i são ideais primos de A e e_i, $i = 1, \cdots, r$, são inteiros maiores ou iguais a 1. Nem sempre o anel quociente $A/\mathfrak{a}$ é finito. Mas se A for o anel dos inteiros algébricos de um corpo numérico, segue do Teorema 2.2.9 que o anel quociente $A/\mathfrak{a}$ é sempre finito e o número de elementos é definido como a norma do ideal $\mathfrak{a}$.

Pelo Teorema Chinês do Resto (Teorema 2.1.3) a aplicação $\phi : A/\mathfrak{a} \to A/P_1^{e_1} \times \cdots \times P_r^{e_r}$, definida por $\phi(a+\mathfrak{a}) = (\phi_1(a+\mathfrak{a}), \cdots, \phi_r(a+\mathfrak{a})) = (a+P_1^{e_1}, \cdots, a+P_r^{e_r})$ é um isomorfismo, onde $a \in A$ e $\phi_i : A/\mathfrak{a} \to A/P_i^{e_i}$, $i = 1, \cdots, r$, são os homorfismos canônicos. Se I é um ideal de $A/\mathfrak{a}$, então pela Observação 3.1.2 temos que $\phi_i(I) = \{\phi_i(a+\mathfrak{a}) : a \in A\}$ é um ideal de $A/P_i^{e_i}$, $i = 1, \cdots, r$. Reciprocamente, se $\mathfrak{a}_i \subseteq A/P_i^{e_i}$, $i = 1, \cdots, r$, são ideais, então $\phi^{-1}\{(a_1, \cdots, a_r) : a_i \in \mathfrak{a}_i,\ i = 1, \cdots, r\}$ é um ideal de $A/\mathfrak{a}$. Assim, conhecer os ideais de $A/\mathfrak{a}$ é equivalente a conhecer os ideais de $A/P_i^{e_i}$, $i = 1, \cdots, r$.

Se $\mathfrak{a}$ é um produto de ideais primos distintos, isto é, $\mathfrak{a} = \prod_{i=1}^{r} P_i$ onde P_i são ideais primos distintos de A, então A/P_i, $i = 1, \cdots, r$, são corpos, e portanto, possuem apenas os ideais triviais, e portanto, em $A/\mathfrak{a}$ existem 2^r ideais, que são facilmente determinados pelo isomorfismo ϕ. Por outro lado, se $\mathfrak{a} = \prod_{i=1}^{r} P_i^{e_i}$, onde P_i, $i = 1, \cdots, r$, são ideais primos de A e se $e_i > 1$ para algum $i = 1, \cdots, r$, para calcularmos os ideais de $A/\mathfrak{a}$ veremos inicialmente quais são os ideais de $A/P_i^{e_i}$, $i = 1, \cdots, r$. Pelo Teorema 3.2.2 temos que os ideais de A/P^e estão em bijeção com os ideais $\mathfrak{a} \subseteq A$ tal que $\mathfrak{a} \supseteq P^e$. Assim, pelo Corolário 2.2.3 existe um ideal $\mathfrak{b} \subseteq A$ tal que $P^e = \mathfrak{a}\mathfrak{b}$. Pela unicidade da fatoração de ideais (Teorema 2.2.7) segue que $\mathfrak{a} = P^k$, $0 \leq k \leq e$. Assim, os ideais de A/P^e são completamente determinados.

Exemplo 3.2.2 *Considerando* $\mathbb{F} = \mathbb{Q}(i)$ *temos pelo Exemplo 2.2.2 que o anel dos inteiros algébricos de* $\mathbb{F}$ *é* $\mathbb{Z}[i]$. *Sendo* $\mathbb{Q} \subseteq \mathbb{F}$ *uma extensão de grau 2, segue pelo Teorema 2.2.11 que* $\sum_{i=1}^{2} e_i f_i = 2$. *Tomando* $p = 5$ *temos que* $p\mathbb{Z}[i] =< 2+i >< 2-i >= P_1 P_2$ *onde* $P_i \cap \mathbb{Z} =< 5 >$, $i = 1, 2$, *e* $f_1 = f_2 = e_1 = e_2 = 1$. *Assim,* $\mathbb{Z}[i]/(p\mathbb{Z}[i]) \simeq \mathbb{Z}_5 \times \mathbb{Z}_5$ *pois* $[\mathbb{Z}[i]/(5\mathbb{Z}[i]) : \mathbb{Z}/(5\mathbb{Z})] = 2$. *Como* $\mathbb{Z}_5$ *possui apenas os ideais triviais, segue que* $\mathbb{Z}[i]/(5\mathbb{Z}[i])$ *tem* $2^2 = 4$ *ideais e são cópias isomorfas dos ideais* $\{0\} \times \{0\}$, $\{0\} \times \mathbb{Z}_5$, $\mathbb{Z}_5 \times \{0\}$, $\mathbb{Z}_5 \times \mathbb{Z}_5$ *de* $\mathbb{Z}_5 \times \mathbb{Z}_5$. *Analogamente, se*

$\dot{a} =< 2+i >< 3+2i >=< 4+7i >$ *temos que* $\frac{A}{\dot{a}} \simeq \frac{A}{<2+i>} \times \frac{A}{<3+2i>}$, *onde* $< 2+i > \cap \mathbb{Z} = < 5 >$ *e* $< 3+2i > \cap \mathbb{Z} =< 13 >$. *Como* $\frac{A}{<2+i>}$ *e* $\frac{A}{<3+2i>}$ *são corpos de 5 e 13 elementos, respectivamente, segue que* $A/\dot{a}$ *tem* $2^2 = 4$ *ideais e que seus ideais são cópias isomorfas dos ideais* $\{0\} \times \{0\}$, $\{0\} \times \mathbb{Z}_{13}$, $\mathbb{Z}_5 \times \{0\}$, $\mathbb{Z}_5 \times \mathbb{Z}_{13}$ *de* $\mathbb{Z}_5 \times \mathbb{Z}_{13}$ ■

3.3 Códigos de Hamming

Blake em [9] propôs uma construção de códigos de Hamming sobre anéis de inteiros residuais do tipo $\mathbb{Z}_{p^l}$ onde p é um número primo e l é um inteiro maior ou igual a 1. Nesta seção, descrevemos a generalização destes códigos para anéis finitos locais do tipo $R = A/P^e$, onde e é um inteiro maior ou igual a 1, A é o anel dos inteiros algébricos de um corpo numérico $\mathbb{K}$ (uma extensão finita do corpo dos números racionais) e P é um ideal primo de A tal que $< p >= P \cap \mathbb{Z}$. O ideal maximal de R é $\mathcal{M} = P/P^e$. Segue do Teorema 2.1.2 que este ideal é composto por todos os divisores de zero de R.

A construção sendo proposta garante que estes códigos tenham distâncias de Hamming mínimas iguais a três e, portanto, capacidade de correção de um erro aleatório. De modo análogo ao caso de códigos de Hamming sobre corpos, a construção da matriz verificação de paridade será realizada a partir do parâmetro $m \geq 2$ representando o número de linhas da matriz e o comprimento dado em função de m.

3.3.1 Construção da Matriz Verificação de Paridade

Sendo $\mathcal{M}$ um subgrupo aditivo de R, como a ordem (norma do ideal P^e), isto é, o número de elementos do anel R é p^{fe}, onde $f = [A/P : \mathbb{Z}_p]$ é o grau da extensão, segue que a ordem de $\mathcal{M}$ é p^s, com $s < fe$, donde concluímos que o conjunto dos divisores de zero deste anel possui p^s elementos.

Sejam $\mathbf{a} = (a_1, \cdots, a_m)$ e $\mathbf{b} = (b_1, \cdots, b_m)$ pertencentes a R^m. Tomando as matrizes

$$M = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_m & b_m \end{bmatrix} \quad \text{e} \quad M_{ij} = \begin{bmatrix} a_i & b_i \\ a_j & b_j \end{bmatrix},$$

definimos $d_{ij} = a_i b_j - b_i a_j$, para todo par (i,j) tal que $1 \leq i \neq j \leq m$.

Lema 3.3.1 *Dois elementos* $\mathbf{a}, \mathbf{b} \in R^m$ *são linearmente dependentes se, e somente se,* d_{ij} *é um divisor de zero para todo par* (i,j) *tal que* $1 \leq i \neq j \leq m$.

Demonstração: Se um dos d_{ij}, $i \neq j$ for uma unidade então, pelo Teorema 5.17 de [12], o sistema linear homogêneo

$$\begin{cases} \alpha a_i + \beta b_i = 0 \\ \\ \alpha a_j + \beta b_j = 0 \end{cases} \qquad \text{onde} \qquad X = \begin{bmatrix} \alpha \\ \\ \beta \end{bmatrix}$$

com $\alpha, \beta \in R$, somente admite a solução trivial, isto é, $\alpha = \beta = 0$, contradizendo o fato de $\mathbf{a}$ e $\mathbf{b}$ serem linearmente dependentes. Reciprocamente, assumindo que cada $d_{ij} = a_i b_j - b_i a_j$ é um divisor de zero para todo par (i,j) tal que $1 \leq i \neq j \leq m$, pela Propriedade 4.11(e) de [12] temos que para todo $i \neq j$ o posto da matriz M_{ij} é menor que 2. Como o posto da matriz M é menor ou igual o máximo dos postos das matrizes M_{ij}, segue que o posto da matriz M é menor que 2. Assim, pelo Teorema 5.3 de [12] temos que o sistema linear homogêneo $M\mathbf{X} = \mathbf{0}$, tem uma solução não trivial, e portanto, os vetores $\mathbf{a}$ e $\mathbf{b}$ são linearmente dependentes. ■

Seja B o conjunto dos p^{sm} elementos em R^m cujas componentes são elementos divisores de zero. Dizemos que $\mathbf{a}$ e $\mathbf{b}$ em $R^m - B$ são **equivalentes** se forem linearmente dependentes, isto é, se existirem α, β em R, não todos nulos, tais que $\alpha\mathbf{a} + \beta\mathbf{b} = \mathbf{0}$. Este fato define claramente uma relação de equivalência.

A fim de construirmos a matriz verificação de paridade H de um código que tenha distância de Hamming mínima 3, não podemos permitir que:

1. Uma coluna de H pertença a B, isto é, todas as suas componentes são elementos divisores de zero;

2. Em H existam colunas equivalentes.

Por (1), estamos evitando que o código tenha distância 1 e por (2) que tenha distância 2. O teorema seguinte fornece quantos elementos podemos escolher em R^m, satisfazendo estas condições. Sua demonstração é bastante análoga a apresentada em [9]. Contudo, para efeito de referências futuras e também para uma maior clareza deste trabalho, apresentamos a demonstração.

Teorema 3.3.1 *O número das classes de equivalência de* $R^m - B$ *é dado por* $(q^m - 1)/(q - 1)$, *onde* $q = p^{v-s}$ *e* $v = fe$.

Demonstração: Seja $\mathbf{a} = (a_1, \cdots, a_i, 0, \cdots, 0)$ um elemento com i unidades e $m - i$ zeros. Seja $\mathbf{b} = (b_1, \cdots, b_i, b_{i+1}, \cdots, b_m)$ um elemento de $R^m - B$. Temos pelo Lema 3.3.1 que $\mathbf{b}$ é equivalente a $\mathbf{a}$ se, e somente se,

(i) $a_j b_l$ é um divisor de zero para $1 \leq j \leq i$, $l > i$ e

(ii) $a_j b_l - a_l b_j$ é um divisor de zero para $1 \leq j$, $l \leq i$.

Sendo a_j uma unidade, (i) implica que b_l, $l > i$, é um divisor de zero. De (ii) temos que $b_l = a_i^{-1} a_l b_i + a_i^{-1} \times$ (divisor de zero), $l = 1, \cdots, i-1$, e se b_i é um divisor de zero então b_l, $l = 1, \cdots, i$, é um divisor de zero, o que implica que $\mathbf{b} \in B$, contradizendo $\mathbf{b}$ dado acima. Agora, se $b_i \in R^*$ então b_l, $l = 1, \cdots, i-1$, é a soma de uma unidade e um divisor de zero, que é sempre uma unidade. Para $b_i \in R^*$ fixo, existem exatamente p^s soluções para b_l, $l = 1, \cdots, i-1$, correspondendo na escolha de divisor de zero. Como b_i pode ser escolhido de $p^v - p^s$ maneiras e cada um dos b_l, $l > i$, em p^s maneiras, segue que existem $(p^{v-s} - 1)p^{sm}$ elementos de $R^m - B$ equivalentes a $\mathbf{a}$. Este argumento é estendido de maneira natural para $\mathbf{a}$ com i unidades e $m - i$ divisores de zero, e portanto, é independente de i. Assim, o número das classes de equivalência de $R^m - B$ é $\dfrac{\mid R^m - B \mid}{p^{sm}(p^{v-s} - 1)} = \dfrac{p^{vm} - p^{sm}}{p^{sm}(p^{v-s} - 1)} = \dfrac{p^{sm}(p^{(v-s)m} - 1)}{p^{sm}(p^{v-s} - 1)} = \dfrac{p^{(v-s)m} - 1}{p^{v-s} - 1} = \dfrac{q^m - 1}{q - 1}$, onde $q = p^{v-s}$ e $v = fe$. ■

Observamos pelo Teorema 3.3.1 que o número das classes de equivalência de $R^m - B$ depende de p, m, v e s e que o número máximo de colunas na matriz verificação de paridade de um código sobre R com distância de Hamming mínima 3 é $n = (p^{(v-s)m} - 1)/(p^{v-s} - 1)$, onde m representa o número de linhas de H. Isto implica que dado um parâmetro m, com $m \geq 2$, podemos sempre construir um código sobre R com os seguintes parâmetros:

- comprimento $n = (p^{(v-s)m} - 1)/(p^{v-s} - 1)$,
- distância de Hamming mínima $d_{min} = 3$.

O próximo teorema fornece quais elementos devem ser escolhidos em R de tal modo a satisfazer as condições (1) e (2), dadas acima.

Teorema 3.3.2 *Sejam* $\mathbf{a} = (a_1, \cdots, a_m)$ *e* $\mathbf{b} = (b_1, \cdots, b_m)$ *elementos de* R^m *e* $\mathbf{a}' = (a_1', \cdots, a_m')$ *e* $\mathbf{b}' = (b_1', \cdots, b_m')$, *com* $a_i' \equiv a_i (mod\ p)$ *e* $b_i' \equiv b_i (mod\ p)$, $1 \leq i \leq m$. *Se* $\mathbf{a}'$ *e* $\mathbf{b}'$ *são linearmente dependentes sobre* $\mathbb{Z}_p$, *então* $\mathbf{a}$ *e* $\mathbf{b}$ *são linearmente dependentes sobre* R.

Demonstração: Suponha que $\mathbf{a}'$ e $\mathbf{b}'$ sejam linearmente dependentes sobre $\mathbb{Z}_p$, isto é, que existam α e β em $\mathbb{Z}_p$, não todos nulos, tais que $\alpha\mathbf{a}' + \beta\mathbf{b}' = \mathbf{0}$. Consequentemente, $\alpha a_i' + \beta b_i' = 0$, para todo $1 \leq i \leq m$. Como $a_i \equiv a_i' (mod\ p)$ e $b_i \equiv b_i' (mod\ p)$, para todo $1 \leq i \leq m$, temos que $a_i = r_i p + a_i'$ e $b_i = s_i p + b_i'$, onde r_i e s_i, para todo $1 \leq i \leq m$, são inteiros. Assim, $\alpha(a_i - r_i p) + \beta(b_i - s_i p) = 0$ para todo $1 \leq i \leq m$. Agora, multiplicando ambos os membros desta última igualdade por p^{u-1}, obtemos que $\alpha' a_i + \beta' b_i = 0$, para todo $1 \leq i \leq m$ onde $\alpha' = \alpha p^{u-1}$ e $\beta' = \beta p^{u-1}$ pertencem a R e p^u é a característica de R. Assim, $\mathbf{a}$ e $\mathbf{b}$ são linearmente dependentes sobre R. ■

Corolário 3.3.1 *A matriz verificação de paridade de "algum" código de Hamming sobre* R *pode ser tomada como sendo a matriz verificação de paridade de um código de Hamming sobre* $GF(p)$, *onde as colunas são* m*-uplas (não nulas) pertencentes a* $\mathbb{Z}_p$ *e tais que quaisquer duas delas são linearmente independentes sobre* $\mathbb{Z}_p$. ■

Note que para um dado parâmetro m, os códigos de Hamming sobre $GF(p)$, que também são códigos de Hamming sobre R, possuem o mesmo comprimento e com número diferente de palavras-código.

Exemplo 3.3.1 *Vamos construir um código de Hamming sobre* $R = \mathbb{Z}_2[i]$. *Pela notação usada temos* $p = 2$, $f = s = 1$, $e = 2$, $v = fe = 2$ *e* $q = p^{v-s} = 2$. *Adotamos* $m = 3$. *Deste modo, temos que* $n = (q^m - 1)/(q - 1) = 7$. *A matriz verificação de paridade* H *do código de Hamming sobre* $GF(2)$ *com comprimento* $n = 7$ *e dimensão* $k = 4$ *é dada por*

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

Pelo Corolário 3.3.1, temos que a matriz H *é também a matriz verificação de paridade de um código de Hamming sobre* $\mathbb{Z}_2[i]$ *com comprimento* $n = 7$ *e distância de Hamming mínima* $d_{min} = 3$. ■

3.3.2 Algoritmo de Decodificação

Para finalizar o estudo dos códigos de Hamming, apresentamos um algoritmo de decodificação para estes códigos. Esta proposta é uma extensão simples do método para a decodificação destes códigos quando definidos sobre os corpos finitos $GF(q)$.

Interlando em [31] descreveu um método para a decodificação de códigos de Hamming sobre anéis de inteiros residuais. Aquí, veremos que este método também pode ser adaptado para a decodificação de códigos de Hamming definidos sobre anéis locais provenientes de anéis de inteiros algébricos, que é bastante similar ao caso destes códigos quando definidos sobre os corpos $GF(q)$, onde q é uma potência de primo.

Suponha que as palavras-código tenham comprimento n, e que H seja a matriz verificação de paridade. Suponha ainda que $\mathbf{c}$ seja a palavra transmitida, $\mathbf{e}$ o vetor erro introduzido pelo canal, $\mathbf{b} = \mathbf{c} + \mathbf{e}$ o vetor recebido e $\mathbf{s} = \mathbf{b}H^T$ o vetor síndrome. Quando ocorrer apenas um erro, o vetor síndrome será suficiente para detectar a posição do erro na palavra-código e a magnitude do mesmo.

Sendo $\mathbf{b} = \mathbf{c}+\mathbf{e}$, com $\mathbf{c}$ uma palavra-código, segue que $\mathbf{s} = \mathbf{b}H^T = (\mathbf{c}+\mathbf{e})H^T = \mathbf{c}H^T + \mathbf{e}H^T = \mathbf{e}H^T$. Assim, quando ocorrer apenas um erro na posição j, o vetor síndrome será a j-ésima coluna da matriz H multiplicada pela componente e_j do vetor erro. Deste modo, devemos procurar na matriz H uma coluna de modo que quando multiplicada por um elemento de R^* (elementos invertíveis de R) obtenhamos o vetor síndrome. Consequentemente, a coluna encontrada representará a posição j na palavra-código em que ocorreu o erro e o elemento de R^* será a magnitude do mesmo.

Exemplo 3.3.2 *Considere o código do Exemplo 3.3.1. Suponha que a palavra transmitida seja a palavra toda nula* $\mathbf{c} = (0000000)$ *e que o vetor erro introduzido pelo canal seja* $\mathbf{e} = (000000i)$. *Multiplicando o vetor recebido* $\mathbf{b} = \mathbf{c} + \mathbf{e}$ *pela matriz* H, *obtemos* $\mathbf{s} = (iii)$, *que é a última coluna da matriz* H *multiplicada por* i. *Deste modo, podemos dizer que ocorreu um erro de magnitude* i *na posição 7 da palavra-código.* ■

Devido ao fato de que quaisquer duas colunas de H são sempre linearmente independentes, quando ocorrer apenas um erro, este método determinará de maneira única a posição e a magnitude do erro. Deste modo, não há possibilidade de se determinar dois ou mais padrões de erros para uma mesma síndrome.

Mas, quando estiver fora da capacidade de correção do código, isto é, ocorrer dois ou mais erros,

pode acontecer que não exista uma coluna de H tal que quando multiplicada por um elemento de R^* dê o vetor síndrome. Neste caso, o decodificador declara apagamento por falta de dados para a decodificação correta. Caso se encontre tal coluna, um erro de decodificação ocorrerá.

3.4 Códigos Reed-Solomon

Blake em [9] propôs a construção de códigos Reed-Solomon sobre anéis de inteiros residuais do tipo $\mathbb{Z}_{p^l}$ onde p é um número primo e l é um inteiro maior ou igual a 1. Nesta seção, apresentamos duas propostas de construções de códigos Reed-Solomon sobre anéis finitos locais. Primeiramente, tratamos o caso de anéis finitos locais quaisquer e em seguida anéis finitos locais do tipo A/P^e, onde e é um inteiro maior ou igual a 1, $< p >= P \cap \mathbb{Z}$ é um ideal primo de $\mathbb{Z}$, $N(P^e) = p^{fe}$ (norma), onde $f = [A/P : \mathbb{Z}_p]$, A é o anel dos inteiros algébricos de um corpo numérico $\mathbb{K}$ e P é um ideal primo de A. Estas construções são muito semelhantes a construção destes códigos sobre o anel dos inteiros residuais módulo q ou sobre corpos $GF(q)$, o corpo de Galois com q elementos.

Seja R um anel finito local e $\mathcal{M}$ seu ideal maximal. Pelo Teorema 2.1.2 temos que $\mathcal{M}$ é formado por todas as não unidades do anel R. Seja p um número primo e $\alpha \in R$ tal que $\alpha^{p-1} \equiv 1 \ (mod\ \mathcal{M})$ e $\alpha^i \not\equiv 1 \ (mod\ \mathcal{M})$ se $1 \leq i \leq p-2$. Sejam $s = p-1$, $G_s = \{1, \alpha, \alpha^2, \cdots, \alpha^{s-1}\}$, o **vetor localizador** $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ consistindo de elementos distintos de G_s e m um inteiro não negativo. Seja a matriz H definida como

$$H = \begin{bmatrix} \alpha_1^m & \alpha_2^m & \cdots & \alpha_n^m \\ \alpha_1^{m+1} & \alpha_2^{m+1} & \cdots & \alpha_n^{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{m+d-2} & \alpha_2^{m+d-2} & \cdots & \alpha_n^{m+d-2} \end{bmatrix}. \tag{3.5}$$

Teorema 3.4.1 *O espaço nulo da matriz H sobre R é um código $\mathcal{C}(n, \eta)$ de comprimento $n \leq s$ e distância de Hamming mínima $d_{min}(\mathcal{C}) \geq d$.*

Demonstração: Pelo Teorema 2.3.2 temos que a distância de Hamming mínima deste código é maior ou igual a d se, e somente se, quaisquer $d-1$ colunas de H são linearmente independentes sobre R. Mostraremos que a matriz H possui posto $d-1$, isto é, qualquer conjunto de $d-1$ colunas

são linearmente independentes sobre R. Seja a submatriz H_1 de H dada por

$$H_1 = \begin{bmatrix} \alpha^{j_1 m} & \alpha^{j_2 m} & \cdots & \alpha^{j_{d-1} m} \\ \alpha^{j_1 (m+1)} & \alpha^{j_2 (m+1)} & \cdots & \alpha^{j_{d-1} (m+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{j_1 (m+d-2)} & \alpha^{j_2 (m+d-2)} & \cdots & \alpha^{j_{d-1} (m+d-2)} \end{bmatrix}.$$

Vamos mostrar que esta matriz é não singular, onde $j_i = 0, \cdots, n-1$, $i = 1, \cdots, d-1$, com $j_1 < j_2 < \cdots < j_{d-1}$. O determinante desta matriz pode ser expresso na seguinte forma

$$det(H_1) = \alpha^{(j_1 + \cdots + j_{d-1})m} det(H_2)$$

onde a matriz H_2 é definida como

$$H_2 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{j_1} & \alpha^{j_2} & \cdots & \alpha^{j_{d-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{j_1 (d-2)} & \alpha^{j_2 (d-2)} & \cdots & \alpha^{j_{d-1} (d-2)} \end{bmatrix}.$$

Note que $det(H_2)$ é o determinante de Vandermonde, donde concluímos que o determinante da matriz H_1 é dado por

$$det(H_1) = \alpha^{(j_1 + \cdots + j_{d-1})m} \prod_{i>l} (\alpha^{j_i} - \alpha^{j_l}). \tag{3.6}$$

Assim, H_1 será não singular (terá posto $d-1$) se, e somente se, a equação (3.6) for uma unidade em R. As potências de α são sempre unidades em R pelo fato de que $\alpha^{p-1} \equiv 1 \ (mod \ \mathcal{M})$, ou seja, $\alpha^{p-1} \in 1 + \mathcal{M} \subseteq R - \mathcal{M}$. Assim, pelo Teorema 2.1.2 temos que α e suas potências são unidades em R. Os termos que fazem parte do produtório em (3.6) são da forma $\alpha^{j_i} - \alpha^{j_l} = \alpha^{j_i}(1 - \alpha^{j_l - j_i})$, onde estamos supondo $j_l > j_i$. Ambos os termos deste produto são unidades, pois o primeiro é uma potência de α e o segundo sendo da forma $1 - \alpha^j$, $j = 1, \cdots, p-2$. Deste modo, para $1 \leq j \leq p-2$, $\alpha^j \not\equiv 1 \ (mod \ \mathcal{M})$, isto é, $\alpha^j - 1 \notin \mathcal{M}$, $j = 1, \cdots, p-2$. Pelo Teorema 2.1.2 temos que $\alpha^j - 1$, $j = 1, \cdots, p-2$ são unidades em R. Concluímos deste modo que $det(H_1)$ é uma unidade em R e, portanto, H tem posto $d-1$. Assim $d_{min}(\mathcal{C}) \geq d$. ■

Exemplo 3.4.1 *Seja o anel finito local* $R = \mathbb{Z}_{25}$ *cujo ideal maximal é* $\mathcal{M} = \{0, 5, 10, 15, 20\}$. *Vamos construir um código Reed-Solomon* $\mathcal{C}(n, \eta)$ *sobre* $\mathbb{Z}_{25}$ *com* $d_{min}(\mathcal{C}) \geq 3$. *Sejam* $p = 5$, $n = p - 1 = 4$, $\alpha = 2$ *(elemento do anel) e também vamos supor que* $m = 1$. *É de fácil verificação que* $\alpha^4 \equiv 1 \ (mod\ \mathcal{M})$ *e que* $\alpha^j \not\equiv 1 \ (mod\ \mathcal{M})$, $j = 1, 2, 3$. *Seja* $\eta = (1, \alpha, \alpha^2, \alpha^3)$ *o vetor localizador. A matriz verificação de paridade deste código é dada por*

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 & 8 \\ 1 & 4 & 16 & 14 \end{bmatrix}.$$

Pelo Teorema 3.4.1, a matriz H *descreve um código* $\mathcal{C}(n, \eta)$ *sobre* $\mathbb{Z}_{25}$ *de comprimento 4 e* $d_{min}(\mathcal{C}) \geq 3$. *Este código tem capacidade de correção de um erro aleatório.* ■

Agora, iremos propor a construção de códigos Reed-Solomon sobre anéis finitos locais do tipo $R = A/P^e$. Seja $\alpha = a + P^e \in A/P^e$ um elemento tal que $a^{p-1} \equiv 1 \ (mod\ P)$ e $a^i \not\equiv 1 \ (mod\ P)$ se $1 \leq i \leq p-2$. Sejam $s = p-1$, $G_s = \{1, \alpha, \alpha^2, \cdots, \alpha^{s-1}\}$, o **vetor localizador** $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ consistindo de elementos distintos de G_s e m um inteiro não negativo. Seja a matriz H definida como

$$H = \begin{bmatrix} \alpha_1^m & \alpha_2^m & \cdots & \alpha_n^m \\ \alpha_1^{m+1} & \alpha_2^{m+1} & \cdots & \alpha_n^{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{m+d-2} & \alpha_2^{m+d-2} & \cdots & \alpha_n^{m+d-2} \end{bmatrix}. \tag{3.7}$$

Observação 3.4.1 *Com as notações acima temos que* $R/\mathcal{M}$ *é isomorfo a* A/P, *onde* $\mathcal{M} = P/P^e$ *é o único ideal maximal de* R. ■

Teorema 3.4.2 *O espaço nulo da matriz* H *sobre* R *é um código* $\mathcal{C}(n, \eta)$ *de comprimento* $n \leq s$ *e distância de Hamming mínima* $d_{min}(\mathcal{C}) \geq d$.

Demonstração: De maneira análoga à demonstração do Teorema 3.4.1 temos que o determinante da submatriz H_1 é dado por

$$det(H_1) = \alpha^{(j_1 + \cdots + j_{d-1})m} \prod_{i>l} (\alpha^{j_i} - \alpha^{j_l}). \tag{3.8}$$

Assim, H_1 é não singular (terá posto $d-1$) se, e somente se, a equação (3.8) for uma unidade em R. As potências de α são sempre unidades em R pelo fato de que $a^{p-1} \equiv 1 \ (mod\ P)$, ou

seja, $a^{p-1}-1 \in P$, e assim $\alpha^{p-1}-1+P=0+P$ em R. Pela Observação 3.4.1 temos que $a^{p-1}-1+P^e+\mathcal{M}=0+P^e+\mathcal{M}$ em $R/\mathcal{M}$. Portanto, $\overline{a^{p-1}}-\overline{1} \in \mathcal{M}$ não é uma unidade em R. Pelo Teorema 2.1.2 segue que $\overline{a^{p-1}}=a^{p-1}+P^e$ é uma unidade em R. Consequentemente, as potências de α também são unidades em R. Agora, os termos que fazem parte do produtório em (3.8) são da forma $\alpha^{j_l}-\alpha^{j_i}=\alpha^{j_i}(1-\alpha^{j_l-j_i})$, onde estamos supondo $j_l > j_i$. Ambos os termos deste produto são unidades, pois o primeiro é uma potência de α e o segundo sendo da forma $1-\alpha^j$, $j=1,2,\cdots,p-2$. Deste modo, para $j=1,2,\cdots,p-2$, temos que $a^j \not\equiv 1 \ (mod\ P)$, isto é, $a^j-1 \notin P$. Portanto, a^j-1+P é uma unidade em A/P (corpo) e pela Observação 3.4.1 temos que $a^j-1+P+\mathcal{M}$ é uma unidade em $R/\mathcal{M}$. Assim, existe $\overline{c}+\mathcal{M} \in R/\mathcal{M}$ tal que $(\overline{b}+\mathcal{M})(\overline{c}+\mathcal{M})=\overline{1}+\mathcal{M}=\overline{bc}+\mathcal{M}$, onde $\overline{b}=a^j-1+P^r$. Logo, $\overline{1}-\overline{bc} \in \mathcal{M}$ não é uma unidade em R. Pelo Teorema 2.1.2 temos que $\overline{bc}$ é uma unidade em R e, portanto, $\overline{b}=a^j-1+P^r=\alpha^j-1$ é uma unidade em R. Concluímos deste modo que o $det(H_1)$ é uma unidade em R. Assim, H tem posto $d-1$ e, portanto $d_{min}(\mathcal{C}) \geq d$. ■

Exemplo 3.4.2 *Seja A o anel dos inteiros algébricos de $\mathbb{Q}(i)$, isto é, $A=\mathbb{Z}[i]$. Seja $p=7$ um número primo. Temos que $R=\mathbb{Z}[i]/(7\mathbb{Z}[i])=\mathbb{Z}_7[i]=\{a+bi : a,b \in \mathbb{Z}_7\}$. Vamos construir um código Reed-Solomon $\mathcal{C}(n,\eta)$ sobre $\mathbb{Z}_7[i]$ com $d_{min}(\mathcal{C}) \geq 5$. Para isso, sejam $p=7$, $n=p-1=6$, $\alpha=3$ (elemento primitivo de $\mathbb{Z}_7[i]$) e vamos supor que $m=1$. Seja $\eta=(1,\alpha,\alpha^2,\cdots,\alpha^5)$ o vetor localizador. Deste modo, temos que a matriz verificação de paridade H, é dada por*

$$H=\begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha^8 & \alpha^{10} \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & \alpha^{15} \\ 1 & \alpha^4 & \alpha^8 & \alpha^{12} & \alpha^{16} & \alpha^{20} \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 & 6 & 4 & 5 \\ 1 & 2 & 4 & 1 & 2 & 4 \\ 1 & 6 & 1 & 6 & 1 & 6 \\ 1 & 4 & 2 & 1 & 4 & 2 \end{bmatrix}. \tag{3.9}$$

Desta forma, pelo Teorema 3.4.2, a matriz H descreve um código $\mathcal{C}(n,\eta)$ sobre $\mathbb{Z}_7[i]$ de comprimento 6 e $d_{min}(\mathcal{C}) \geq 5$. Este código tem capacidade de correção de dois erros aleatórios. ■

Observação 3.4.2 *Observamos que estas classes de códigos Reed-Solomon sobre anéis locais não são formadas por códigos cíclicos. Por uma análise direta, podemos verificar que $\mathbf{c}=(4\ 2\ 3\ 6\ 1\ 0)$ é uma palavra-código do Exemplo 3.4.2, pois $\mathbf{c}H^T=\mathbf{0}$, onde H é a matriz verificação de paridade dada por (3.9). Entretanto, a palavra $\mathbf{c}^{(2)}=(1\ 0\ 4\ 2\ 3\ 6)$ não pertence a este código, pelo fato de*

que $\mathbf{c}^{(2)}H^T \neq \mathbf{0}$. *Concluímos deste modo que, em geral, os códigos Reed-Solomon sobre anéis locais não são cíclicos.* ■

Por outro lado, podemos afirmar que todo polinômio $v(x)$ de um código Reed-Solomon é múltiplo de um determinado polinômio mônico $g(x)$, também pertencente ao código. Este fato pode ser visto da seguinte maneira. Como $\mathbf{v}H^T = \mathbf{0}$ temos que $\alpha^m, \alpha^{m+1}, \cdots, \alpha^{m+d-2}$ são raízes de $v(x)$. Temos também, que os coeficientes de $v(x)$ pertencem ao anel R (que possui identidade multiplicativa), donde podemos concluir que $f_1(x) = x - \alpha^m$, $f_2(x) = x - \alpha^{m+1}$, $f_3(x) = x - \alpha^{m+2}, \cdots, f_{d-1}(x) = x - \alpha^{m+d-2}$ e $g(x) = mmc\{f_1(x), f_2(x), \cdots, f_{d-1}(x)\}$ são fatores de $v(x)$. Isto implica que todo polinômio código $v(x)$ é múliplo de $g(x)$. Por outro lado, nem todo múltiplo de $g(x)$ é um polinômio código.

Para finalizar esta seção veremos alguns fatos sobre o problema de definir códigos BCH sobre anéis comutativos finitos A com identidade. A definição de códigos BCH sobre anéis finitos segue os mesmos passos quando estes códigos são definidos sobre corpos finitos. Deste modo, sejam $A(\delta)$, onde δ é raiz de um polinômio irredutível sobre o anel A, uma extensão do anel A e $\alpha \in A(\delta)^*$ de ordem s tais que $1 - \alpha^j$, $j = 1, \cdots, s-1$ são unidades em $A(\delta)$. Definimos um código BCH sobre o anel A, de comprimento $n \leq s$ e distância projetada d, como sendo o espaço nulo da matriz H definida por

$$H = \begin{bmatrix} \alpha_1^m & \alpha_2^m & \cdots & \alpha_n^m \\ \alpha_1^{m+1} & \alpha_2^{m+1} & \cdots & \alpha_n^{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{m+d-2} & \alpha_2^{m+d-2} & \cdots & \alpha_n^{m+d-2} \end{bmatrix},$$

onde $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ é o vetor localizador.

O problema de determinar os elementos $\alpha \in A(\delta)^*$ tais que $1 - \alpha^j$, $j = 1, 2, \cdots, n-1$, são unidades está diretamente relacionado com a determinação da estrutura de $A(\delta)^*$. No próximo capítulo faremos um estudo mais detalhado de códigos BCH sobre anéis finitos, onde veremos que encontrar o anel em que o polinômio $x^n - 1$ se decompõe, é de extrema importância.

3.5 Proposta de Construção e Decodificação de Códigos sobre Anéis Finitos

No decorrer desta seção faremos um estudo a respeito da construção e da decodificação de códigos lineares sobre anéis comutativos finitos com identidade. Este mesmo problema foi tratado nas seções 4.6 e 4.7 de [31] para o caso de grupos abelianos finitos, que pode ser escrito como o produto direto de subgrupos cíclicos cujas ordens são potências de primos. Nesta seção, usaremos o fato de que todo anel finito comutativo com identidade pode ser expresso, de maneira única, como o produto direto de anéis finitos locais (Teorema 3.1.2).

Faremos, agora, a construção de um código linear sobre um anel comutativo finito com identidade tal que $A = A_1 \times A_2 \times \cdots \times A_r$, onde A_i, $i = 1, 2, \cdots, r$, são anéis locais. Para isto, selecionamos r códigos lineares sobre $A_1, A_2, \cdots, A_r$, respectivamente, e tal que a distância de Hamming mínima de cada um desses códigos seja pelo menos d. Se $d \leq 3$ os códigos de Hamming são suficientes e se $d > 3$ recorremos aos códigos de múltipla correção, isto é, Reed-Solomon, BCH ou alternantes.

Para codificarmos uma seqüência de informação $(c_1, c_2, \cdots, c_k)$ com $c_i \in A$, $i = 1, 2, \cdots, k$, primeiramente expressamos cada elemento c_i, $i = 1, 2, \cdots, k$, na forma de um produto direto, isto é, $c_i = (a_{1i}, a_{2i}, \cdots, a_{ri})$ para $i = 1, 2, \cdots, k$ e $a_{li} \in A_l$ com $l = 1, 2 \cdots, r$. Deste modo, toda seqüência de informação é processada da seguinte maneira

$$
\begin{array}{ccccc|c|ccccccc}
(a_{11} & a_{12} & \cdots & a_{1,k}) & \rightarrow & \text{C} & \rightarrow & (a_{11} & \cdots & a_{1,k} & a_{1,k+1} & \cdots & a_{1,n}) \\
(a_{21} & a_{22} & \cdots & a_{2,k}) & \rightarrow & \text{O} & \rightarrow & (a_{21} & \cdots & a_{2,k} & a_{2,k+1} & \cdots & a_{2,n}) \\
 & & \vdots & & & \text{D} & & & \vdots & & & & \\
(a_{r,1} & a_{r,2} & \cdots & a_{r,k}) & \rightarrow & \text{I} & \rightarrow & (a_{r,1} & \cdots & a_{r,k} & a_{r,k+1} & \cdots & a_{r,n}) \\
\uparrow & \uparrow & \cdots & \uparrow & & \text{F} & & \downarrow & \cdots & \downarrow & \downarrow & \cdots & \downarrow \\
(c_1 & c_2 & \cdots & c_k) & & \text{I} & & (c_1 & \cdots & c_k & c_{k+1} & \cdots & c_n)
\end{array}
\tag{3.10}
$$

No lado esquerdo da equação (3.10) os elementos de cada linha são da forma a_{ij} com i fixo e $j = 1, 2, \cdots, k$, e todos pertencem ao mesmo anel A_i. Deste modo, a seqüência de elementos $(a_{i1}, a_{i2}, \cdots, a_{ik})$ pode ser considerada como os símbolos de informação de um código sobre A_i, $i = 1, 2, \cdots, r$. O lado direito é formado pelas palavras-código correspondentes a cada uma dessas

seqüências de informação. De modo análogo, fazemos o produto direto de cada coluna no lado direito para obtermos os correspondentes elementos do anel A que formarão a palavra-código.

A distância de Hamming mínima do código $\mathcal{C}$ é dada por $d_{min}(\mathcal{C}) = min_{1\leq i\leq r}\{d_i\}$ onde d_i é a distância de Hamming mínima do i-ésimo código definido sobre A_i, $i = 1, 2, \cdots, r$.

A decodificação destes códigos é realizada da seguinte maneira. Tendo recebido uma palavra de comprimento n (com elementos em A), o receptor decompõe a mesma num produto direto (como no lado direito da equação (3.10)) para obter r palavras, também de comprimento n, cada uma sobre A_i, $i = 1, 2, \cdots, r$. A partir daí, r decodificadores (um para cada código em (3.10)) decodificam as palavras correspondentes.

3.6 Conclusões

Neste capítulo vimos que códigos cíclicos de comprimento n sobre um anel comutativo finito A com identidade são ideais no anel $A_n = \frac{A[x]}{<x^n-1>}$, e assim os principais teoremas caracterizando o polinômio gerador para códigos cíclicos sobre o corpo $GF(q)$ podem ser generalizados em muitas situações para códigos cíclicos sobre o anel A. Vimos que códigos lineares sobre anéis finitos locais possuem propriedades de distâncias tão boas quanto ou melhores que a dos códigos lineares definidos sobre anéis finitos quaisquer.

Em seguida, apresentamos construções de códigos sobre anéis finitos via anéis de grupo, fazendo uso do Teorema de Estrutura para um Anel Comutativo Finito. Vimos que estes códigos são facilmente determinados quando o anel finito é obtido através do anel dos inteiros algébricos de um corpo numérico.

Foram apresentadas construções dos códigos de Hamming e Reed-Solomon sobre anéis finitos locais. Os códigos de Hamming para correção de um erro aleatório e os códigos Reed-Solomon para múltipla correção de erros. Os códigos Reed-Solomon foram estendidos sobre quaisquer anéis finitos locais enquanto que os de Hamming sobre anéis finitos locais provenientes dos anéis dos inteiros algébricos de um corpo numérico. Um algoritmo para a decodificação dos códigos de Hamming também foi apresentado. A decodificação dos códigos Reed-Solomon será apresentada na Seção 4.3 do Capítulo 4, pelo fato de necessitarmos de uma teoria mais elaborada.

Finalmente, com o objetivo de sistematizar um outro método para a decodificação de códigos de múltipla correção de erros nos levou a estabelecer uma proposta alternativa, que consiste em

concatenar códigos definidos sobre anéis comutativos finitos locais com identidade para obtermos um código linear sobre um anel comutativo finito com identidade.

No próximo capítulo faremos o estudo dos códigos BCH (para múltipla correção de erros) sobre anéis comutativos finitos com identidade. Neste caso, encontrar o anel em que o polinômio $x^n - 1$ se decompõe em fatores lineares, é de extrema importância.

Capítulo 4

Códigos BCH sobre Anéis Comutativos Finitos com Identidade

Códigos algébricos lineares podem ser definidos através de alfabetos escolhidos em um conjunto de tamanho arbitrário. Contudo muitos dos resultados em Teoria da Codificação tem sido obtidos assumindo que os símbolos são elementos de um corpo finito. Neste capítulo veremos uma importante classe de códigos algébricos, os códigos BCH, que podem ser estendidos para anéis arbitrários comutativos finitos com identidade.

Os códigos BCH (Bose-Chadhuri-Hocquenghem) são uma generalização extraordinária dos códigos de Hamming para múltipla correção de erros e formam a classe dos melhores códigos construtivos para canais em que os erros afetam os símbolos de maneira independente. Os códigos BCH foram descobertos por R.C. Bose e D.K. Ray-Chaudhuri [10], [11] e independentemente por A. Hocquenghem [29], os quais levam as iniciais de seus nomes. Os códigos BCH são definidos a partir da matriz verificação de paridade e projetados para corrigir um certo número de erros.

O estudo dos códigos BCH é importante nos seguintes aspectos:

- Quando o comprimento da palavra-código não é grande, existem bons códigos nesta classe;
- As técnicas de codificação e decodificação são relativamente simples;
- A subclasse não binária dos códigos Reed-Solomon tem certas propriedades de otimalidade e uma estrutura de distância bem entendida;

- O conhecimento destes códigos é, provavelmente, o ponto inicial para estudar outras classes de códigos cíclicos.

Blake em [8] propôs uma construção de códigos cíclicos sobre anéis de inteiros residuais do tipo $\mathbb{Z}_m$, onde m é um produto de primos distintos, isto é, $m = \prod_{i=1}^{r} p_i$, com p_i, $i = 1, 2, \cdots, r$, números primos, a partir de códigos cíclicos sobre $\mathbb{Z}_{p_i}$. Blake em um trabalho posterior [9], propôs a construção de matrizes verificação de paridade para códigos sobre $\mathbb{Z}_m$ de maneira análoga aos códigos de Hamming e Reed-Solomon. Algumas questões deixadas em aberto por Blake foram resolvidas por Spiegel em [48] e [49]. Em [48] Spiegel propôs uma construção de códigos sobre $\mathbb{Z}_m$, para m da forma $m = \prod_{i=1}^{r} p_i^{k_i}$, a partir de códigos sobre $\mathbb{Z}_{p_i^{k_i}}$, que por sua vez são obtidos a partir de códigos sobre $\mathbb{Z}_{p_i}$. Em 1978, Spiegel em [49], propôs uma construção de códigos BCH sobre os anéis $\mathbb{Z}_{p^k}$ a partir de códigos BCH sobre corpos p-ádicos e seus corpos de classes residuais. Shankar em [46] propôs uma construção de códigos BCH com símbolos em anéis de inteiros residuais.

Nas Seções 4.1 e 4.2 estendemos os códigos BCH para símbolos sobre anéis comutativos finitos com identidade. A obtenção destes códigos está baseada na fatoração de $x^s - 1$ sobre o grupo das unidades de uma extensão conveniente destes anéis. Esta construção será feita de maneira bastante similar aos códigos BCH sobre corpos finitos. Na Seção 4.3 apresentamos um algoritmo de decodificação eficiente que servirá para a decodificação dos códigos Reed-Solomon e BCH definidos sobre anéis comutativos finitos locais com identidade. O algoritmo apresentado nesta seção é uma adaptação do algoritmo de Berlekamp-Massey modificado proposto por Interlando, Palazzo e Elia em [33].

A Seção 4.4 tem como objetivo complementar a teoria desenvolvida sobre o algoritmo de Berlekamp-Massey modificado, onde veremos que este algoritmo pode ser aplicado para geração de seqüências através do uso de registradores de deslocamento, cujos elementos pertencem a um anel comutativo finito local com identidade.

4.1 Códigos BCH sobre Anéis Finitos Locais

Incialmente recordamos que sobre corpos $GF(p)$, define-se um código BCH cíclico de comprimento s, com $mdc(p, s) = 1$, e distância projetada d, como sendo um código cujo polinômio gerador tem como raízes os elementos $\alpha^b, \alpha^{b+1}, \cdots, \alpha^{b+d-2}$, onde $\alpha \in GF(p^m)$ é uma raiz primitiva de $x^s - 1$, b é um inteiro não negativo e m é tal que s divide $p^m - 1$. Quando $s = p^m - 1$ o código é chamado **BCH primitivo**.

Nesta seção iremos descrever a construção de códigos BCH sobre anéis comutativos finitos locais A com identidade, a qual é bastante semelhante a de códigos BCH sobre corpos finitos. A diferença é que os elementos da matriz verificação de paridade estão num anel conveniente.

Do capítulo anterior recordamos que um código cíclico de comprimento s sobre um anel comutativo finito A com identidade é um ideal no anel de polinômios $A[x]$ módulo o polinômio $x^s - 1$. Nosso primeiro passo será identificar uma estrutura algébrica sobre A que contenha as raízes de $x^s - 1$. O conhecimento da estrutura do grupo multiplicativo das unidades de tal anel nos levará a identificação do subgrupo cíclico consistindo das raízes de $x^s - 1$.

Com o intuito de uma maior clareza possível, julgamos necessário fazer algumas considerações iniciais sobre a notação e a terminologia que utilizamos ao longo deste capítulo.

Começamos por observar que daqui em diante a letra A denotará sempre um anel comutativo finito local com identidade, a qual denotamos por 1. O ideal maximal de A será denotado por $\mathcal{M}$ e seu corpo residual por $I\!K = \frac{A}{\mathcal{M}} = GF(p^m)$, para algum $m \geq 1$ (esta última igualdade segue do Teorema 2.1.6).

Seja $A[x]$ o anel de polinômios na variável x sobre o anel A. Denotamos a projeção natural $A[x] \to I\!K[x]$ por μ, e assim $\mu(f(x)) = \mu(a_0 + a_1x + \cdots + a_nx^n) = \overline{f}(x) = \overline{a_0} + \overline{a_1}x + \cdots + \overline{a_n}x^n$, onde $\overline{a_i} = a_i + \mathcal{M}$, $i = 1, \cdots, n$, e $f(x) = a_0 + a_1x + \cdots + a_nx^n \in A[x]$. Assim, o homomorfismo natural $A \to A/\mathcal{M} = I\!K$ é simplesmente a restrição de μ aos polinômios constantes.

A seguir veremos alguns conceitos básicos envolvendo elementos do anel de polinômios $A[x]$.

Definição 4.1.1 *Seja $f(x)$ um polinômio em $A[x]$. Dizemos que*

- *$f(x)$ é uma* **unidade** *se existe um polinômio $a(x) \in A[x]$ tal que $a(x)f(x) = 1$.*
- *$f(x)$ é um* **divisor de zero** *se existe $a(x) \in A[x]$, $a(x) \neq 0$, tal que $a(x)f(x) = 0$.*
- *$f(x)$ é* **regular** *se não é um divisor de zero em $A[x]$.* ■

Os três seguintes teoremas são de [37], Teoremas XIII.2, XV.1 e XIII.7, pelo fato de estarmos considerando A um anel local com $\mathcal{M}$ seu ideal maximal e a aplicação μ como sendo a projeção natural.

Teorema 4.1.1 *Seja $f(x) = a_0 + a_1x + \cdots + a_nx^n \in A[x]$. As seguintes condições são equivalentes.*

- *$f(x)$ é regular.*

- $< a_0, a_1, \cdots, a_n >= A$.
- a_i *é uma unidade para algum* i, $0 \leq i \leq n$.
- $\mu(f(x)) \neq 0$. ■

Teorema 4.1.2 *Seja* $f(x)$ *um polinômio regular em* $A[x]$ *e suponha que* $\mu(f(x))$ *tenha uma raiz simples* $\overline{\alpha} \in I\!K$. *Então* $f(x)$ *tem uma e somente uma raiz* α *tal que* $\mu(\alpha) = \overline{\alpha}$. ■

Teorema 4.1.3 *Seja* $f(x)$ *um polinômio regular em* $A[x]$. *Se* $\mu(f(x))$ *é irredutível em* $I\!K[x]$, *então* $f(x)$ *é irredutível em* $A[x]$. ■

Seja $f(x)$ um polinômio mônico de grau h em $A[x]$ tal que $\mu(f(x))$ seja irredutível em $I\!K[x]$. Segue do Teorema 4.1.3 que $f(x)$ também é irredutível em $A[x]$. Seja o anel quociente $\frac{A[x]}{<f(x)>}$ das classes residuais dos polinômios em x sobre A, módulo o polinômio $f(x)$. Este anel, o qual denotamos por R, é um anel comutativo finito local com identidade e é chamado uma **extensão de Galois** de A, cujo ideal maximal é dado por $\overline{\mathcal{M}_1} = \frac{\mathcal{M}_1}{<f(x)>}$, onde $\mathcal{M}_1 =< \mathcal{M}, f(x) >$ e seu corpo residual é dado por $I\!K_1 = \frac{R}{\mathcal{M}_1} = \frac{A[x]/<f(x)>}{<\mathcal{M},f(x)>/<f(x)>} = \frac{A[x]}{<\mathcal{M},f(x)>} = \frac{(A/\mathcal{M})[x]}{<\mu(f(x))>}$, que possui p^{mh} elementos. O corpo residual $I\!K_1$ será denotado por $GF(p^{mh})$.

Observação 4.1.1 *O anel* R *é local, isto é, os seus elementos divisores de zero formam um grupo abeliano aditivo e consistem daqueles polinômios de grau menor ou igual a* $h-1$ *cujos coeficientes são todos divisores de zero no anel* A. *Um polinômio* $f(x) \in R$ *que possui pelo menos um coeficiente que seja uma unidade em* A *não é um divisor de zero em* R *e, portanto, pertence a* R^* *(o grupo multiplicativo das unidades de* R*), ou seja, é possível encontrar um polinômio* $g(x) \in R[x]$, *tal que* $f(x)g(x) = 1$. ■

Vamos introduzir agora algumas notações das quais faremos uso durante esta seção. R^* denotará o grupo multiplicativo das unidades de R e $I\!K_1^*$ o grupo cíclico multiplicativo de $GF(p^{mh})$ de ordem $p^{mh} - 1$.

Desde que R^* é um grupo abeliano multiplicativo, o mesmo pode ser expresso como um produto direto de grupos cíclicos [37]. Estamos interessado no grupo cíclico de R^*, que será denotado por G_s, cujos elementos são as raízes de $x^s - 1$ para algum s tal que $mdc(s, p) = 1$ (isto garante que $x^s - 1$ não tenha fatores repetidos). Uma vez identificado este grupo, o problema da construção dos

códigos reduz-se a escolher certos elementos deste grupo para serem as entradas da matriz verificação de paridade do códigos BCH.

O próximo teorema segue de [37], Teorema XVIII.2, e fornece a ordem do subgrupo de interesse para que possamos definir os códigos BCH sobre o anel A.

Teorema 4.1.4 *Existe um único subgrupo cíclico maximal de* R^* *cuja ordem é relativamente prima com* p. *Este subgrupo tem ordem* $p^{mh}-1$. ∎

O Teorema 4.1.4 afirma que R^* tem um e somente um subgrupo cíclico cuja ordem $p^{mh}-1$ é relativamente prima com p. O próximo teorema fornece um método de obter este subgrupo cíclico.

Teorema 4.1.5 *Suponha que* α *gera um subgrupo de ordem* s *em* R^*, *onde* $mdc(s,p)=1$. *Então o polinômio* x^s-1 *pode ser fatorado como* $x^s-1=(x-\alpha)(x-\alpha^2)\cdots(x-\alpha^s)$ *se, e somente se,* $\overline{\alpha}$ *tem ordem* s *em* $I\!K_1^*$.

Demonstração: Suponha, primeiramente, que x^s-1 possa ser fatorado sobre R^* como $x^s-1=(x-\alpha)(x-\alpha^2)\cdots(x-\alpha^s)$ (segue do Teorema 4.1.4 que s divide $p^{mh}-1$ pois $mdc(p,s)=1$). Assim, através da aplicação μ temos que $x^s-1=(x-\overline{\alpha})(x-\overline{\alpha^2})\cdots(x-\overline{\alpha^s})$ sobre $I\!K_1^*$, onde $\overline{\alpha}=\mu(\alpha)=\alpha+\overline{\mathcal{M}_1}$. Sendo s um divisor de $p^{mh}-1$ temos que x^s-1 não tem raiz múltipla em $I\!K_1^*$, e portanto, $\overline{\alpha}$ tem ordem s em $I\!K_1^*$.

Reciprocamente, suponha que $\overline{\alpha}$ tem ordem s em $I\!K_1^*$. Então x^s-1 pode ser fatorado em $I\!K_1^*$, isto é, $x^s-1=(x-\overline{\alpha})(x-\overline{\alpha^2})\cdots(x-\overline{\alpha^s})$. Consideramos o conjunto $\overline{F}=\{\overline{\alpha},\overline{\alpha^2},\cdots,\overline{\alpha^s}\}$. Como o $mdc(s,p)=1$, segue que x^s-1 não tem raiz múltipla em $I\!K_1^*$. Daí pelo Teorema 4.1.2 temos que para cada $\overline{\alpha^i}$, $i=1,2,\cdots,s$, existe um único elemento $\alpha_i\in R^*$ tal que $\mu(\alpha_i)=\overline{\alpha^i}$ e α_i é uma raiz de x^s-1 em R^*. Devemos, agora, mostrar que x^s-1 pode ser fatorado como $x^s-1=(x-\alpha)(x-\alpha^2)\cdots(x-\alpha^s)$, onde $\alpha^i=\alpha_i$, $i=1,2,\cdots,s$, isto é, devemos mostrar que $F=\{\alpha,\alpha^2,\cdots,\alpha^s\}$ são todas as raízes de x^s-1 em R^*. Seja β uma raiz de x^s-1 tal que $\beta\neq\alpha^i$, $i=1,2,\cdots,s$. Então $\mu(\beta)$ é uma raiz de x^s-1 em $I\!K_1^*$. Mas $\mu(\beta)\notin\overline{F}$ pelo Teorema 4.1.2, e como x^s-1 tem exatamente s raízes em $I\!K_1^*$ segue que $\mu(\beta)$ não é um novo elemento de $\overline{F}$. Portanto, x^s-1 pode ser fatorado de maneira única sobre R^* como $x^s-1=(x-\alpha)(x-\alpha^2)\cdots(x-\alpha^s)$. ∎

Em vista do Teorema 4.1.5, denotamos por G_s o subgrupo cíclico de ordem $s=p^{mh}-1$ de

R^* o qual tem como elementos todas as raízes de $x^s - 1$. Além disso, todo polinômio $h(x)$ que divide $x^s - 1$ pode ser fatorado de maneira única sobre $\mathbb{K}_1^*$. Assim, pelo Teorema 4.1.5, temos que a fatoração de $h(x)$ sobre G_s também é única. Deste modo, temos o seguinte corolário.

Corolário 4.1.1 *Um polinômio $h(x)$ que divide $x^s - 1$ com coeficientes em A pode ser fatorado sobre G_s como $h(x) = (x - \alpha^{e_1})(x - \alpha^{e_2}) \cdots (x - \alpha^{e_l})$ se, e somente se, $\overline{h}(x)$ pode ser fatorado sobre $\mathbb{K}_1$ como $\overline{h}(x) = (x - \overline{\alpha}^{e_1})(x - \overline{\alpha}^{e_2}) \cdots (x - \overline{\alpha}^{e_l})$.* ■

O próximo teorema, cuja a demonstração é bem elementar, é de grande utilidade na obtenção do gerador de G_s.

Teorema 4.1.6 *Seja $\overline{\alpha}$ o gerador de um subgrupo cíclico de ordem s em $\mathbb{K}_1^*$. Então α gera um subgrupo cíclico de ordem sd em R^*, onde d é um inteiro maior ou igual a 1 e α^d gera o subgrupo cíclico G_s de R^*.*

Demonstração: Suponha que a ordem de α é k, isto é, $\alpha^k = 1$ e $\alpha^j \neq 1$ para $1 \leq j \leq k-1$. Portanto, $(\overline{\alpha})^k = \overline{\alpha^k} = 1$. Assim, devemos ter que s é um divisor de k, isto é, $k = sd$. ■

Definição 4.1.2 *Um* **código BCH** *$\mathcal{C}(n, \eta)$ de comprimento reduzido $n \leq s$ sobre A tem matriz verificação de paridade definida como*

$$H = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{2t} & \alpha_2^{2t} & \cdots & \alpha_n^{2t} \end{bmatrix} \tag{4.1}$$

para algum $t \geq 1$, onde $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ é o **vetor localizador**, *consistindo de elementos distintos de G_s.* ■

Os códigos BCH definidos desta maneira, em geral, não são cíclicos. Na Seção 4.1.1 faremos o estudo de códigos BCH cíclicos através de uma escolha particular do vetor localizador. Uma palavra $\mathbf{c} = (c_1, c_2, \cdots, c_n) \in A^n$ está em $\mathcal{C}(n, \eta)$ se, e somente se, satisfaz as $2t$ equações de verificação de paridade sobre R, isto é,

$$\sum_{i=1}^{n} c_i \alpha_i^l = 0, \quad l = 1, 2, \cdots, 2t.$$

Para $l \geq 1$, cada uma dessas equações de verificação de paridade pode ser reescrita em h equações sobre A.

O código $\mathcal{C}(n, \eta)$ para o qual $n = p^{mh} - 1$ será chamado **primitivo**. Neste caso, η é único, a menos de permutação de suas coordenadas.

Uma matriz verificação de paridade H_1 com elementos sobre A pode ser obtida substituindo cada elemento em H pelo correspondente vetor coluna de comprimento h sobre A. É possível obter uma estimativa da distância de Hamming mínima, d, diretamente da matriz verificação de paridade.

Para a demonstração do próximo teorema necessitamos do seguinte lema.

Lema 4.1.1 *Seja α um elemento de G_s de ordem s. Então as diferenças $\alpha^{l_1} - \alpha^{l_2}$ são unidades em R se $0 \leq l_1 \neq l_2 \leq s-1$.*

Demonstração: Sem perda de generalidade podemos assumir que $l_1 > l_2$. Temos que $\alpha^{l_1} - \alpha^{l_2}$ pode ser escrito como $\alpha^{l_2}(\alpha^{l_1 - l_2} - 1)$. O primeiro termo α^{l_2} no último produto é uma unidade. Agora, o segundo termo pode ser escrito como $1 - \alpha^j$, $j = 1, 2, \cdots, s-1$. Assim, para $1 \leq j \leq s-1$, temos pelo Teorema 4.1.5 que $1 - \alpha^j$ é uma unidade em R. ■

Teorema 4.1.7 *A distância de Hamming mínima de um código BCH satisfaz $d \geq 2t + 1$.*

Demonstração: Suponha $\mathbf{c}$ uma palavra-código não nula de $\mathcal{C}(n, \eta)$ tal que $w_H(\mathbf{c}) \leq 2t$. Então $\mathbf{c}H^T = \mathbf{0}$. Deletando as $n - 2t$ colunas da matriz H, que corresponde aos zeros da palavra-código, obtemos uma nova matriz H_1 que é Vandermonde. Pelo Lema 4.1.1 temos que o determinante da matriz H_1 é uma unidade em R. Assim, a única possibilidade para a palavra-código $\mathbf{c}$ é ser a palavra toda nula. ■

Exemplo 4.1.1 *O polinômio $f(x) = x^3 + x + 1$ é irredutível sobre $\mathbb{Z}_2$ e sobre $A = \mathbb{Z}_2[i]$. Seja R o anel $\frac{A[x]}{<f(x)>}$ e o corpo $\mathbb{K}_1 = \frac{\mathbb{Z}_2[x]}{<f(x)>}$. O elemento α tal que $f(\alpha) = 0$ tem ordem 7 em R^*. Assim, os elementos de G_7 são $1, \alpha, \alpha^2, \cdots, \alpha^6$, isto é, os elementos de G_7 são raízes de $x^7 - 1$. Tomando $\eta = (1, \alpha, \alpha^2, \cdots, \alpha^6)$, então se $t = 1$ o código BCH $\mathcal{C}(n, \eta)$ tem matriz verificação de paridade igual a*

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

onde a distância de Hamming mínima é maior ou igual a 3. ■

4.1.1 Códigos BCH como Códigos Cíclicos

Nesta seção veremos que por uma escolha particular na ordem de $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ a geração dos códigos BCH torna-se simples. Com isso, será possível especificar uma estimativa para a distância de Hamming mínima.

Lembramos a definição de um código BCH: $\mathbf{c} = (c_1, c_2, \cdots, c_n)$ é uma palavra-código se, e somente se, $\sum_{i=1}^{n} c_i \alpha_i^l = 0$ para $l = 1, 2, \cdots, 2t$, onde $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ é um vetor arbitrário, mas ordenado, de elementos distintos de G_s. A mudança de tal ordenação não afeta a propriedade da correção de erros do código. Assim, tomando $s = n$ e $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, onde α é um elemento primitivo de G_n e $\alpha_i = \alpha^{i-1}$, temos que a matriz H em (4.1) define um código BCH que tem um polinômio gerador.

O Teorema 4.1.6 é útil na determinação do gerador de G_n, e o Corolário 4.1.1 é útil na determinação dos polinômios minimais de α^i. Assim, temos que $M_i(x)$, o polinômio minimal de α^i sobre R^* (onde α é primitivo em G_n), terá como raízes todos os elementos distintos da sequência $\alpha^i, (\alpha^i)^q, (\alpha^i)^{q^2}, \cdots, (\alpha^i)^{q^{h-1}}$, onde $q = p^m$. Portanto, $M_i(x)$ pode ser construído de maneira idêntica à construção de $m_i(x)$, o polinômio minimal de $\overline{\alpha}^i$ sobre $GF(p^m)$.

Através do conhecimento do grupo cíclico G_n, a construção dos códigos BCH cíclicos sobre o anel A reduz-se ao problema de escolher elementos deste grupo para serem raízes do polinômio gerador $g(x)$. Deste modo, se α é um elemento primitivo de G_n e se $\alpha^{e_1}, \alpha^{e_2}, \cdots, \alpha^{e_t}$ são raízes do polinômio $g(x)$ sobre A dado por

$$g(x) = mmc\{M_{e_1}(x), M_{e_2}(x), \cdots, M_{e_t}(x)\},$$

onde $M_{e_i}(x)$ é o polinômio minimal de α^{e_i} para $i = 1, 2, \cdots, t$, temos que os polinômios $v(x)$ tal que $v(x) = c(x)g(x)\ (mod\ x^n - 1)$, onde $c(x) \in A[x]$, define um código cíclico de comprimento n sobre A.

Definição 4.1.3 *Seja α um elemento primitivo de G_n. Um código* **BCH** *definido sobre o anel A é um código cíclico de comprimento n gerado pelo polinômio de grau mínimo $g(x)$ cujas raízes são $\alpha^{b+1}, \alpha^{b+2}, \cdots, \alpha^{b+2t}$, para algum $b \geq 0$, e $t \geq 1$, isto é, $g(x) = mmc\{M_1(x), M_2(x), \cdots, M_{2t}(x)\}$, onde $M_i(x)$ é o polinômio minimal de α^{b+i}, $i = 1, 2, \cdots, 2t$, sobre A. Além disso, $\overline{g}(x) = mmc\{m_1(x), m_2(x), \cdots, m_{2t}(x)\}$ onde $m_i(x)$ é o polinômio minimal de $\overline{\alpha}^i$, $i = 1, \cdots, 2t$, gera um código BCH sobre $GF(p^m)$.* ■

Segue da Definição 4.1.3 que $v(x) = v_0 + v_1 x + v_2 x^2 + \cdots + v_{n-1} x^{n-1} \in A[x]$ é uma palavra-código se, e somente se, $v(\alpha^{b+e_i}) = 0$, para $i = 1, 2, \cdots, 2t$. Deste modo, a matriz verificação de paridade H para o código BCH tendo $g(x)$ como polinômio gerador é dada por

$$H = \begin{bmatrix} 1 & \alpha^{b+1} & \alpha^{2(b+1)} & \cdots & \alpha^{(n-1)(b+1)} \\ 1 & \alpha^{b+2} & \alpha^{2(b+2)} & \cdots & \alpha^{(n-1)(b+2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{b+2t} & \alpha^{2(b+2t)} & \cdots & \alpha^{(n-1)(b+2t)} \end{bmatrix}. \tag{4.2}$$

Assim, $v(x)$ é uma palavra-código se, e somente se, $\mathbf{v}H^T = \mathbf{0}$.

Observação 4.1.2

- *O método sistemático para a determinação do mmc de um conjunto de polinômios* $\{p_1(x), p_2(x), \cdots, p_n(x)\}$, *consiste em calcular primeiramente o mdc (usando o algorítmo de Euclides) e em seguida calcular o mmc que é dado pelo quociente* $mmc\{p_1(x), p_2(x), \cdots, p_n(x)\} = (\prod_{i=1}^{n} p_i(x))/(mdc\{p_1(x), p_2(x), \cdots, p_n(x)\})$.
- *Esta construção de códigos BCH conduz somente a códigos cíclicos, enquanto que a construção de códigos Reed-Solomon do capítulo anterior não pode ser derivada desta construção pelo fato de que estes códigos não são cíclicos.* ■

No próximo teorema estabelecemos um limitante inferior para a distância de Hamming mínima do código obtido. Este limitante também se aplica aos códigos cíclicos em geral. Mas em códigos BCH cíclicos os polinômios geradores são escolhidos de modo que a distância mínima seja garantida por este limitante.

Teorema 4.1.8 *Seja* $g(x)$ *o polinômio gerador de um código cíclico de comprimento* n *com símbolos em* A *e sejam* $\alpha^{e_1}, \alpha^{e_2}, \cdots, \alpha^{e_{n-k}}$ *raízes de* $g(x)$ *em* G_n, *onde* α *é um elemento de ordem* n. *Então a distância de Hamming mínima do código é maior do que o número máximo de inteiros consecutivos módulo* n *no conjunto* $E = \{e_1, e_2, \cdots, e_{n-k}\}$.

Demonstração: Sejam $m, m+1, \cdots, m+d-2$ o conjunto máximo de inteiros consecutivos módulo n no conjunto E. Sabemos que um código cíclico com raízes $\alpha^{e_1}, \alpha^{e_2}, \cdots, \alpha^{e_{n-k}}$, é o espaço nulo

da matriz

$$H = \begin{bmatrix} 1 & \alpha^{e_1} & \cdots & (\alpha^{e_1})^{n-2} & (\alpha^{e_1})^{n-1} \\ 1 & \alpha^{e_2} & \cdots & (\alpha^{e_2})^{n-2} & (\alpha^{e_2})^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \alpha^{e_{n-k}} & \cdots & (\alpha^{e_{n-k}})^{n-2} & (\alpha^{e_{n-k}})^{n-1} \end{bmatrix}.$$

Agora, se nenhuma combinação linear de $d-1$ colunas da matriz

$$H_1 = \begin{bmatrix} 1 & \alpha^{m} & \cdots & (\alpha^{m})^{n-2} & (\alpha^{m})^{n-1} \\ 1 & \alpha^{m+1} & \cdots & (\alpha^{m+1})^{n-2} & (\alpha^{m+1})^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \alpha^{m+d-2} & \cdots & (\alpha^{m+d-2})^{n-2} & (\alpha^{m+d-2})^{n-1} \end{bmatrix}$$

é nula, então claramente nenhuma combinação linear de $d-1$ colunas de H é nula. Assim pelo Teorema 2.3.2 o código tem distância de Hamming mínima maior ou igual a d.

Para verificarmos que a matriz H_1 tem esta propriedade, iremos observar se para a matriz H_2, formada por qualquer conjunto de $d-1$ de suas colunas,

$$H_2 = \begin{bmatrix} (\alpha^{m})^{j_1} & \cdots & (\alpha^{m})^{j_{d-2}} & (\alpha^{m})^{j_{d-1}} \\ (\alpha^{m+1})^{j_1} & \cdots & (\alpha^{m+1})^{j_{d-2}} & (\alpha^{m+1})^{j_{d-1}} \\ \vdots & \ddots & \vdots & \vdots \\ (\alpha^{m+d-2})^{j_1} & \cdots & (\alpha^{m+d-2})^{j_{d-2}} & (\alpha^{m+d-2})^{j_{d-1}} \end{bmatrix},$$

o $det(H_2) \neq 0$. Vamos mostrar que a matriz H_2 é não singular. O determinante desta matriz pode ser expresso na seguinte formula

$$det(H_2) = \alpha^{m(j_1+j_2+\cdots+j_{d-1})} det(H_3)$$

onde a matriz H_3 é dada por

$$H_3 = \begin{bmatrix} 1 & \cdots & 1 & 1 \\ \alpha^{j_1} & \cdots & \alpha^{j_{d-2}} & \alpha^{j_{d-1}} \\ (\alpha^{j_1})^2 & \cdots & (\alpha^{j_{d-2}})^2 & (\alpha^{j_{d-1}})^2 \\ \vdots & \ddots & \vdots & \vdots \\ (\alpha^{j_1})^{d-2} & \cdots & (\alpha^{j_{d-2}})^{d-2} & (\alpha^{j_{d-1}})^{d-2} \end{bmatrix}.$$

Note que $det(H_3)$ é o determinante de Vandermonde. Disso, concluímos que o determinante da matriz H_2 é dado por

$$det(H_2) = \alpha^{m(j_1+\cdots+j_{d-1})} \prod_{i>k} (\alpha^i - \alpha^k). \tag{4.3}$$

Assim a matriz H_2 é não singular (tem posto $d-1$) se, e somente se, (4.3) é uma unidade em R. Pelo Lema 4.1.1, segue que o determinante em (4.3) é uma unidade em R. Concluímos deste modo, que o determinante de H_2 é uma unidade em R. Portanto, qualquer combinação de $d-1$, ou menos, colunas de H é linearmente independente. Assim, pelo Teorema 2.3.2, o código, que é o espaço nulo de H, tem distância de Hamming mínima maior ou igual a d. ■

Exemplo 4.1.2 *Referindo ao Exemplo 4.1.1, temos que se* α *é uma raiz de* $g(x)$, *então* $g(x) = (x-\alpha)(x-\alpha^2)(x-\alpha^4)$. *Portanto,* $g(x)$ *gera um código BCH de comprimento 7 sobre* $\mathbb{Z}_2[i]$ *com distância de Hamming mínima maior ou igual a 3.* ■

4.2 Códigos BCH sobre Anéis Finitos

Seja A um anel comutativo finito com identidade, o qual denotaremos por 1. Lembramos do Teorema 3.1.2 que o anel A pode ser expresso de maneira única como um produto (soma) direto de aneis comutativos finitos locais, ou seja, existe um isomorfismo ϕ definido da seguinte maneira

$$\begin{aligned} \phi : A &\rightarrow A_1 \times A_2 \times \cdots \times A_r \\ a &\rightarrow \phi(a) = (\phi_1(a), \phi_2(a), \cdots, \phi_r(a)) = (a + P_1^m, a + P_2^m, \cdots, a + P_r^m) \end{aligned}$$

onde $\phi_i : A \rightarrow A_i$, $A_i = A/P_i^m$, $i = 1, 2, \cdots, r$ são os homomorfismos canônicos, P_i, $i = 1, 2, \cdots, r$, são os ideais maximais do anel A e m o menor inteiro positivo tal que $\cap_{i=1}^r P_i^m = 0$.

Teorema 4.2.1 *A aplicação*

$$\varphi : A[x] \to A_1[x] \times A_2[x] \times \cdots \times A_r[x]$$

definida por $\varphi(a(x)) = (\varphi_1(a(x)), \varphi_2(a(x)), \cdots, \varphi_r(a(x)))$ *onde* $a(x) = a_0 + a_1x + \cdots + a_nx^n$ *e* $\varphi_i(a(x)) = \phi_i(a_0) + \phi_i(a_1)x + \cdots + \phi_i(a_n)x^n$, $i = 1, 2, \cdots, r$, *é um isomorfismo.*

Demonstração: Claramente temos que φ é um homomorfismo injetor. Para verificarmos a sobrejetividade tomamos $g(x) = (g_1(x), g_2(x), \cdots, g_r(x)) \in A_1[x] \times A_2[x] \times \cdots \times A_r[x]$. Disso temos que

$$\begin{aligned} g(x) &= (a_0^1 + a_1^1x + \cdots + a_n^1x^n, \cdots, a_0^r + a_1^rx + \cdots + a_n^rx^n) \\ &= (a_0^1, a_0^2, \cdots, a_0^r) + (a_1^1, a_1^2, \cdots, a_1^r)x + \cdots + (a_r^1, a_r^2, \cdots, a_r^r)x^n. \end{aligned}$$

Como ϕ é um isomorfismo temos que existe $a(x) = a_0 + a_1x + \cdots + a_nx^n \in A[x]$ onde $\phi(a_i) = (a_i^1, a_i^2, \cdots, a_i^r)$, $i = 1, 2, \cdots, r$ e tal que $\varphi(a(x)) = g(x)$. Portanto, φ é um isomorfismo. ■

Nosso objetivo agora é obter um polinômio irredutível em $A[x]$ a partir de polinômios irredutíveis em $A_i[x]$, $i = 1, 2, \cdots, r$. É isto que veremos no teorema a seguir. Sejam $\mathcal{M}_i$ os ideais maximais dos anéis locais A_i, $\mathbb{K}_i = \frac{A_i}{\mathcal{M}_i}$ os corpos residuais e $\mu_i : A_i \to \mathbb{K}_i$ as projeções naturais para todo $i = 1, 2, \cdots, r$. Do Teorema 4.1.3 temos que um polinômio mônico $g_i(x)$ tal que $\mu_i(g_i(x))$ é irredutível sobre $\mathbb{K}_i$, é também irredutível sobre A_i, para todo $i = 1, 2, \cdots, r$.

Teorema 4.2.2 *Se os* $g_i(x)$*'s são polinômios mônicos de grau* h *tal que* $\mu_i(g_i(x))$ *são irredutíveis sobre* $\mathbb{K}_i$, *para* $i = 1, 2, \cdots, r$, *então o polinômio* $f(x) = \varphi^{-1}(g(x)) \in A[x]$, *onde* $g(x) = (g_1(x), g_2(x), \cdots, g_r(x)) \in A_1[x] \times \cdots \times A_r[x]$, *é mônico e irredutível de grau* h *sobre* A.

Demonstração: Claramente $f(x)$ é um polinômio mônico de grau h. Agora, suponha que $f(x)$ seja redutível em $A[x]$, isto é, $f(x) = a(x)b(x)$, onde $a(x), b(x) \in A[x]$ e $grau(a(x)) = u \geq 1$, $grau(b(x)) = v \geq 1$. Pelo isomorfismo φ, Teorema 4.2.1, temos que existe $f(x) \in A[x]$ tal que

$$\begin{aligned} g(x) &= (g_1(x), g_2(x), \cdots, g_r(x)) = (\varphi(f(x)) \\ &= \varphi(a(x))\varphi(b(x)) \\ &= (\varphi_1(a(x)), \varphi_2(a(x)), \cdots, \varphi_r(a(x)))(\varphi_1(b(x)), \varphi_2(b(x)), \cdots, \varphi_r(b(x))) \\ &= (\varphi_1(a(x))\varphi_1(b(x)), \varphi_2(a(x))\varphi_2(b(x)), \cdots, \varphi_r(a(x))\varphi_r(b(x))). \end{aligned}$$

Portanto, $g_i(x) = \varphi_i(a(x))\varphi_i(b(x))$, $i = 1, 2, \cdots, r$. Como o $grau(a(x)) = grau(\varphi_i(a(x))$, $i = 1, 2, \cdots, r$, temos que os $g_i(x)'s$ são redutíveis em $A_i[x]$, $i = 1, 2, \cdots, r$, contradizendo a hipótese. Portanto, $f(x)$ é mônico e irredutível em $A[x]$. ■

Exemplo 4.2.1 *Do Exemplo 3.1.2 temos que* $\mathbb{Z}_5[i] \simeq \frac{\mathbb{Z}[i]}{\mathcal{M}_1} \times \frac{\mathbb{Z}[i]}{\mathcal{M}_2} \simeq \mathbb{Z}_5 \times \mathbb{Z}_5$, *onde* $\mathbb{Z}_5[i] = \{a+bi : a, b \in \mathbb{Z}_5\}$, $\mathcal{M}_1 =< 2+i >$ *e* $\mathcal{M}_2 =< 2-i >$. *Este isomorfismo é definido por*

0	→	(0,0)	3i	→	(1,4)	2+2i	→	(1,3)	4+3i	→	(0,3)
1	→	(1,1)	4i	→	(3,2)	4+i	→	(1,2)	3+i	→	(0,1)
2	→	(2,2)	1+i	→	(3,4)	2+3i	→	(3,1)	1+3i	→	(2,0)
3	→	(3,3)	3+2i	→	(2,4)	4+4i	→	(2,1)	2+i	→	(4,0)
4	→	(4,4)	1+4i	→	(4,3)	1+2i	→	(0,2)	4+2i	→	(3,0)
i	→	(2,3)	3+3i	→	(4,2)	2+4i	→	(0,4)	3+4i	→	(1,0)
2i	→	(4,1)									

e é um isomorfismo entre $\mathbb{Z}_5[i]$ *e* $\mathbb{Z}_5 \times \mathbb{Z}_5$. ■

Teorema 4.2.3 *A aplicação*

$$\psi : A[x] \to \frac{A_1[x]}{< g_1(x) >} \times \frac{A_2[x]}{< g_2(x) >} \times \cdots \times \frac{A_r[x]}{< g_r(x) >}$$

definida por

$$\psi(a(x)) = (\varphi_1(a(x)) + < g_1(x) >, \varphi_2(a(x)) + < g_2(x) >, \cdots, \varphi_r(a(x)) + < g_r(x) >)$$

é um homomorfismo sobrejetor cujo núcleo é $< f(x) >$, *onde os* $g_i(x)'s$ *são definidos como anteriormente.*

Demonstração: Claramente temos que a aplicação ψ é um homomorfismo. Para a sobrejetividade, seja $h(x) \in \frac{A_1[x]}{<g_1(x)>} \times \frac{A_2[x]}{<g_2(x)>} \times \cdots \times \frac{A_r[x]}{<g_r(x)>}$. Assim,

$$\begin{aligned} h(x) &= (h_1(x) + < g_1(x) >, h_2(x) + < g_2(x) >, \cdots, h_r(x) + < g_r(x) >) \\ &= (h_1(x), h_2(x), \cdots, h_r(x)) + (< g_1(x) >, < g_2(x) >, \cdots, < g_r(x) >). \end{aligned}$$

Como φ é sobrejetora, existe $f(x) \in A[x]$ tal que $\varphi(f(x)) = (\varphi_1(f(x)), \varphi_2(f(x)), \cdots, \varphi_r(f(x))) = (h_1(x), h_2(x), \cdots, h_r(x))$, e portanto $\psi(f(x)) = h(x)$, ou seja, ψ é sobrejetora. Para o núcleo temos claramente que $< f(x) > \subseteq Ker(\psi)$. Por outro lado, do Teorema 4.2.2 temos que existe $f(x) \in A[x]$ tal que $\varphi(f(x)) = (g_1(x), g_2(x), \cdots, g_r(x)) = (\varphi_1(f(x)), \varphi_2(f(x)), \cdots, \varphi_r(f(x)))$. Portanto, $\varphi_i(f(x)) = g_i(x)$ para todo $i = 1, 2, \cdots, r$. Considerando $h(x) \in ker(\psi)$, isto é, $\psi(h(x)) = (\varphi_1(h(x)) + < g_1(x) >, \varphi_2(h(x)) + < g_2(x) >, \cdots, \varphi_r(h(x)) + < g_r(x) >) = \overline{0}$, temos que $\varphi_i(h(x)) = a_i(x)g_i(x)$ com $a_i(x) \in A_i[x]$, $i = 1, 2, \cdots, r$. Tomando $a(x) = (a_1(x), a_2(x), \cdots, a_r(x))$, existe $b(x) \in A[x]$ tal que $\varphi(b(x)) = a(x)$, isto é, $\varphi_i(b(x)) = a_i(x)$, $i = 1, 2, \cdots, r$. Portanto, $\varphi_i(h(x)) = \varphi_i(b(x))\varphi_i(f(x)) = \varphi_i(b(x)f(x))$, $i = 1, 2, \cdots, r$. Assim, $\varphi(h(x)) = (\varphi_1(h(x)), \varphi_2(h(x)), \cdots, \varphi_r(h(x)))$ $= \varphi(b(x)f(x))$. Mas como φ é injetora temos que $h(x) = b(x)f(x)$, isto é, $h(x) \in < f(x) >$. Portanto, $k(\psi) = < f(x) >$. ■

Observação 4.2.1 *Segue do teorema do homomorfismo de anéis (Teorema 3.4.1 de [27]) e do Teorema 4.2.3 que a aplicação*

$$\psi' : \frac{A[x]}{< f(x) >} \to \frac{A_1[x]}{< g_1(x) >} \times \frac{A_2[x]}{< g_2(x) >} \times \cdots \times \frac{A_r[x]}{< g_r(x) >}$$

definida por

$$\psi'(h(x) + < f(x) >) = (\varphi_1(h(x)) + < g_1(x) >, \varphi_2(h(x)) + < g_2(x) >, \cdots, \varphi_r(h(x)) + < g_r(x) >),$$

onde o polinômio $h(x) \in A[x]$, é um isomorfismo. ■

Exemplo 4.2.2 *Usando as notações do Exemplo 4.2.1 temos que os polinômios $x^2 + 2$ e $x^2 + 3$ são mônicos e irredutíveis sobre o anel $\mathbb{Z}_5$. Assim, pelo Teorema 4.2.2, temos que o polinômio $x^2 + i$ é mônico e irredutível sobre o anel $\mathbb{Z}_5[i]$. Neste caso, temos que $\frac{\mathbb{Z}_5[i][x]}{<x^2+i>}$ é isomorfo a $\frac{\mathbb{Z}_5[x]}{<x^2+2>} \times \frac{\mathbb{Z}_5[x]}{<x^2+3>}$.* ■

O próximo teorema segue de [37], Theorem XVIII.1, e é fundamental na decomposição do polinômio $x^n - 1$ em fatores lineares sobre R^*.

Teorema 4.2.4 *Seja $R = R_1 \times R_2 \times \cdots \times R_r$ onde R_i, $1 \leq i \leq r$, são anéis finitos locais. Então R^* é o produto direto de grupos, isto é, $R^* = R_1^* \times R_2^* \times \cdots \times R_r^*$.* ■

Observação 4.2.2 *O Teorema 4.2.4 assegura que para cada $\alpha \in R^*$ existe um único elemento $\beta = (\beta_1, \beta_2, \cdots, \beta_r) \in R_1^* \times R_2^* \times \cdots \times R_r^*$.* ■

Considerando $R = \frac{A[x]}{<f(x)>}$, $R_i = \frac{A_i[x]}{<g_i(x)>}$, $\mathcal{M}_i$ como os ideais maximais de A_i, $I\!K_i = \frac{A_i}{\mathcal{M}_i} = GF(p_i^{m_i})$ como os corpos residuais de A_i e $\mu_i : A_i[x] \to I\!K_i[x]$ como as projeções naturais, temos que $\overline{\mathcal{M}_i} = \frac{<\mathcal{M}_i, g_i(x)>}{<g_i(x)>}$ são os ideais maximais de R_i e $\overline{I\!K_i} = \frac{A_i[x]/<g_i(x)>}{<\mathcal{M}_i, g_i(x)>/<g_i(x)>} = \frac{A_i[x]}{<\mathcal{M}_i, g_i(x)>} = \frac{(A_i/\mathcal{M}_i)[x]}{<\mu_i(g_i(x))>} = GF(p_i^{m_i h})$, como seus corpos residuais, para todo $i = 1, 2, \cdots, r$.

O próximo teorema fornece a condição de modo que o polinômio $x^n - 1$ possa ser fatorado em fatores lineares sobre R^*.

Teorema 4.2.5 *O polinômio $x^n - 1$ pode ser fatorado sobre o grupo multiplicativo R^* como $x^n - 1 = (x - \alpha)(x - \alpha^2)\cdots(x - \alpha^n)$ se, e somente se, $\overline{\beta_i}$ tem ordem n em $\overline{I\!K_i^*}$, onde $mdc(n, p_i) = 1$, $i = 1, 2, \cdots, r$.*

Demonstração: Suponha, primeiramente, que $x^n - 1$ possa ser fatorado sobre R^* como $x^n - 1 = (x - \alpha)(x - \alpha^2)\cdots(x - \alpha^n)$. Então $x^n - 1$ pode ser fatorado sobre R_i^*, $i = 1, 2, \cdots, r$, da seguinte maneira $x^n - 1 = (x - \beta_i)(x - \beta_i^2)\cdots(x - \beta_i^n)$. Agora, segue pelo Teorema 4.1.5 que $\overline{\beta_i}$ tem ordem n em $\overline{I\!K_i^*}$, $i = 1, 2, \cdots, r$. Reciprocamente, suponha que $\overline{\beta_i}$ tem ordem n em $\overline{I\!K_i^*}$, $i = 1, 2, \cdots, r$. Assim, pelo Teorema 4.1.5 temos que $x^n - 1$ pode ser fatorado de maneira única sobre R_i^* como $x^n - 1 = (x - \beta_i)(x - \beta_i^2)\cdots(x - \beta_i^n)$. Tomando $\alpha \in R^*$ como sendo a imagem inversa do elemento $\beta = (\beta_1, \beta_2, \cdots, \beta_r) \in R_1^* \times R_2^* \times \cdots \times R_r^*$, é fácil verificar que $x^n - 1 = (x - \alpha)(x - \alpha^2)\cdots(x - \alpha^n)$. ■

Em vista do Teorema 4.2.5, denotamos por $H_{\alpha,n}$ o subgrupo cíclico de R^* de ordem n gerado por α, ou seja, $H_{\alpha,n}$ contém todas as raízes de $x^n - 1$ sobre R^*. Deste modo, a fim de construirmos um código BCH cíclico sobre R^*, devemos escolher certos elementos de $H_{\alpha,n}$ para serem as raízes do polinômio gerador $g(x)$. Assim, se $\alpha^{e_1}, \alpha^{e_2}, \cdots, \alpha^{e_{n-k}}$ são raízes de $g(x)$ em $H_{\alpha,n}$, construímos o polinômio gerador $g(x)$ como

$$g(x) = mmc\{M_{e_1}(x), M_{e_2}(x), \cdots, M_{e_{n-k}}(x)\}$$

onde $M_i(x)$ são os polinômios mínimos de α^i, $i = e_1, e_2, \cdots, e_{n-k}$. Além disso, pelo Teorema 4.2.4, podemos obter (de uma maneira bem semelhante ao Teorema 4.1.8) uma estimativa para a distância de Hamming mínima do código BCH construído desta maneira.

O próximo teorema fornece um método de construção dos $M_i(x)$, $i = 1, 2, \cdots, r$, isto é, o polinômio minimal de α^i sobre A.

Teorema 4.2.6 *Seja* $M_j(x)$ *o polinômio minimal de* α^j *sobre* A, *onde* α *gera* $H_{\alpha,n}$. *Então* $M_j(x) = \prod_{\lambda \in S_j} (x - \lambda)$, *onde* $S_j = \{(\alpha^j)^m : m = \prod_{i=1}^{r} q_i^{s_i}$ *para todo* $0 \leq s_i \leq h-1\}$, *com* $q_i = p_i^{m_i}$, $i = 1, 2, \cdots, h-1$.

Demonstração: Sejam $\overline{M_j}(x)$ a projeção de $M_j(x)$ sobre o corpo $\mathbb{K}_i$ e $M_j^{(i)}(x)$ o polinômio minimal de $\overline{\alpha^j}$ sobre $\mathbb{K}_i^*$, para todo $i = 1, 2, \cdots, r$. É fácil verificar que $\overline{M_j}(x)$ é divisível por $M_j^{(i)}(x), i = 1, 2, \cdots, r$. Assim, as raízes de $\overline{M_j}(x)$ estão entre os elementos distintos da sequência $\overline{\alpha^j}, \overline{\alpha^{jq_i}}, \overline{\alpha^{jq_i^2}}, \cdots, \overline{\alpha^{jq_i^{h-1}}}$ (i-ésima projeção). Portanto, $M_j(x)$ tem entre suas raízes elementos distintos da sequência $\alpha^j, \alpha^{jq_i}, \alpha^{jq_i^2}, \cdots, \alpha^{jq_i^{h-1}}$, $i = 1, 2, \cdots, r$. Logo, todo elemento γ da forma $\gamma = (\alpha^j)^{q_i^s}$ é uma raiz de $M_j(x)$, $1 \leq i \leq r$, $0 \leq s \leq h-1$. Agora, seja k tal que $1 \leq k \leq r$ e $k \neq i$. Então, se $\overline{\gamma}$ é uma raiz de $\overline{M_j}(x)$, segue que $(\overline{\gamma})^{q_k^l}$ é uma raiz de $\overline{M_j}(x)$ (com coeficientes em $GF(q_k)$), para todo $l = 0, 1, \cdots, h-1$. Portanto, $\gamma^{q_k l} = (\alpha^j)^{q_i^s q_k^l}$ é um raiz de $M_j(x)$. Procedendo desta maneira, podemos mostrar que $M_j(x)$ tem necessariamente como raízes todos os elementos distintos de S_j. Mas o polinômio $\prod_{\lambda \in S_j} (x - \lambda)$ por construção tem coeficientes em A. Portanto, o polinômio minimal de α^j é dado por $M_j(x) = \prod_{\lambda \in S_j} (x - \lambda)$. ■

Concluímos esta seção com o exemplo de um código BCH sobre o anel $\mathbb{Z}_5[i]$ para ilustrarmos os resultados apresentados nesta seção.

Exemplo 4.2.3 *Vamos construir um código de comprimento 24 e distância de Hamming mínima maior ou igual a 5 sobre* $A = \mathbb{Z}_5[i]$. *Claramente 24 divide* $mdc(5^2 - 1, 5^2 - 1)$. *O elemento* $\alpha = \alpha_1 + 1 = \psi^{-1}(\beta_1, \beta_2)$ *gera* $H_{\alpha,24}$, *onde* $\beta_i = 1 + \gamma_i$ *são elementos primitivos de* R_i, $i = 1, 2$, *respectivamente. Sejam* $\alpha^4, \alpha^5, \alpha^6, \alpha^7$ *raízes de* $g(x)$. *Então* $M_4(x)$ *tem como raízes todos os elementos distintos no conjunto* $S_4 = \{\alpha^4, \alpha^{20}\}$; $M_5(x)$ *tem como raízes todos os elementos distintos no conjunto* $S_5 = \{\alpha, \alpha^5\}$; $M_6(x)$ *tem como raiz o elemento* α^6; *e* $M_7(x)$ *tem como raízes todos os elementos distintos no conjunto* $S_7 = \{\alpha^7, \alpha^{11}\}$. *Assim, o polinômio* $g(x) = mmc\{M_4(x), M_5(x), M_6(x), M_7(x)\} = (x-\alpha)(x-\alpha^4)(x-\alpha^5)(x-\alpha^6)(x-\alpha^7)(x-\alpha^{11})(x-\alpha^{20})$ *gera um código BCH de comprimento 24 sobre* A *e distância de Hamming mínima maior ou igual a 5.* ■

4.3 Algoritmo de Decodificação

Nesta seção apresentamos um algoritmo eficiente de decodificação que servirá para a decodificação dos códigos Reed-Solomon e BCH definidos sobre anéis comutativos finitos locais com identidade. Quando estes códigos são definidos sobre corpos $GF(q)$, o primeiro passo da decodificação é localizar as posições dos erros na palavra recebida, através do uso do algoritmo de Berlekamp-Massey [39]. Em seguida, aplica-se o procedimento proposto por Forney [15] para a determinação das magnitudes dos erros. A teoria apresentada nesta seção é uma adaptação da apresentada por Interlando, Palazzo e Elia [33].

Lembramos que quando estivermos decodificando os códigos Reed-Solomon o anel R em questão será um anel comutativo finito local com identidade e para os códigos BCH o anel R será um anel comutativo finito com identidade definido por $R = \frac{A[x]}{<f(x)>}$ onde A é um anel comutativo finito local com identidade e $f(x)$ um polinômio mônico de grau h, tal que $\mu(f(x))$ é irredutível sobre $\mathbb{K} = \frac{A}{\mathcal{M}}$ (que por sua vez é irredutível sobre A, pelo Teorema 4.1.3), onde $\mathcal{M}$ é o ideal maximal de A.

O nome algoritmo de Berlekamp-Massey se deve ao fato de primeiramente ter sido desenvolvido em [5] por Berlekamp e posteriormente melhorado por Massey em [36].

Interlando, Palazzo e Elia em [33], estenderam o algoritmo de Berlekamp-Massey, chamando-o de algoritmo de Berlekamp-Massey modificado. Este algoritmo é um método de decodificação dos códigos Reed-Solomon e BCH sobre $\mathbb{Z}_{p^k}$, onde p é um número primo e k é um inteiro maior ou igual a 1. Com demonstrações análogas, este algoritmo também vale num contexto mais geral, ou seja, pode ser utilizado para a decodificação de códigos Reed-Solomon e BCH definidos sobre anéis comutativos finitos R com identidade. Além disso, o procedimento proposto por Forney em [15] ainda continua válido. Esta adaptação não altera os fundamentos do algoritmo original, e deste modo a complexidade permanece praticamente inalterada.

No Capítulo 3, Seção 3.4, e no Capítulo 4, Seção 4.2, caracterizamos os códigos Reed-Solomon e BCH, respectivamente, através de suas matrizes verificação de paridade. Seja $\mathcal{C}(n, \eta)$ tal código. Analisando as equações (3.5), (3.7), (4.1) e (4.2) podemos observar que estes códigos quando projetados

para a correção de até t erros, possuem matrizes verificação de paridade semelhantes e da forma

$$H = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{2t} & \alpha_2^{2t} & \alpha_3^{2t} & \cdots & \alpha_n^{2t} \end{bmatrix} \tag{4.4}$$

onde $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ é o vetor localizador do código e $\alpha_j = \alpha^{k_j}$, $j = 1, 2, \cdots, n$. Recordamos que no caso de códigos Reed-Solomon sobre R, temos que α é um elemento primitivo sobre R e que no caso de códigos BCH, temos que α é uma raiz de $x^s - 1$. Assim, devido a esta semelhança, o procedimento da correção de erros que descrevemos nesta seção (baseado na informação dada pelo vetor síndrome) servirá para a decodificação de ambos os códigos, e que será sempre capaz de corrigir qualquer combinação de até t erros.

Denotamos por $\mathbf{c} = (c_1, c_2, \cdots, c_n)$ a palavra-código transmitida e por $\mathbf{b} = (b_1, b_2, \cdots, b_n)$ a palavra recebida, com vetor erro introduzido pelo canal dado por $\mathbf{e} = (e_1, e_2, \cdots, e_n) = \mathbf{b} - \mathbf{c}$ e vetor síndrome dado por $\mathbf{s} = (s_1, s_2, \cdots, s_{2t}) = \mathbf{b}H^T$.

Suponhamos que $\nu \leq t$ erros tenham ocorridos nas localizações $x_1 = \alpha_{i_1}, x_2 = \alpha_{i_2}, \cdots, x_\nu = \alpha_{i_\nu}$ com valores (ou magnitudes) $y_1 = e_{i_1}, y_2 = e_{i_2}, \cdots, y_\nu = e_{i_\nu}$.

Sendo $\mathbf{b} = \mathbf{c} + \mathbf{e}$ temos que $\mathbf{s} = \mathbf{b}H^T = (\mathbf{c}+\mathbf{e})H^T = \mathbf{e}H^T$. Assim, em termos de pares (x_i, y_i), as componentes s_j de $\mathbf{s}$ são dadas por

$$s_j = \sum_{i=1}^{\nu} y_i x_i^j \tag{4.5}$$

onde $j = 1, 2, \cdots, 2t$. Portanto, um método para corrigir erros é resolver o sistema de equações (4.5), dando como soluções os pares (x_i, y_i) que representam as posições e as magnitudes dos erros. Observamos que os números localizadores de erros de cada componente do padrão de erro devem ser distintos pelo fato de α ser um elemento primitivo.

Inicialmente veremos o problema da localização dos erros e em seguida resolvemos o problema da determinação das magnitudes dos mesmos. Lembramos que no caso de códigos binários, encontrar a localização dos erros implica necessariamente na determinação das magnitudes dos mesmos. Deste modo, considere o seguinte conjunto de equações

$$(x - x_1)(x - x_2)\cdots(x - x_\nu) = x^\nu + \sigma_1 x^{\nu-1} + \cdots + \sigma_{\nu-1}x + \sigma_\nu, \tag{4.6}$$

onde os coeficientes $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ são chamados **funções simétricas elementares** dos $x_i's$, $i = 1, 2, \cdots, \nu$.

O primeiro passo é no sentido de obtermos uma relação entre os s_j e os σ_i, e em seguida analisar a existência de soluções. Assim, multiplicando ambos o lados de (4.6) por $y_i x_i^j$ e em seguida substituindo x por x_i, $i = 1, 2, \cdots, \nu$, obtemos o seguinte conjunto de equações

$$y_i x_i^{j+\nu} + y_i x_i^{j+\nu-1}\sigma_1 + \cdots + y_i x_i^{j+1}\sigma_{\nu-1} + y_i x_i^j \sigma_\nu = 0. \tag{4.7}$$

Agora somando estas equações para $i = 1, 2, \cdots, \nu$ e usando a equação (4.5) obtemos a seguinte relação entre σ_i e s_j,

$$s_{j+\nu} + s_{j+\nu-1}\sigma_1 + \cdots + s_{j+1}\sigma_{\nu-1} + s_j\sigma_\nu = 0 \tag{4.8}$$

e todos os s_j são conhecidos se $j = 1, 2, \cdots, 2t - \nu$. Deste modo, o cálculo dos σ_i a partir do vetor síndrome é feito resolvendo o sistema linear (4.8) de modo que ν tenha o menor valor possível (isto é exigido pois sempre estaremos assumindo que o vetor erro que ocorre é aquele que possui o menor peso de Hamming possível). Por construção, o sistema linear (4.8) sempre possui uma solução. Veremos no próximo teorema que se os y_i, $i = 1, 2, \cdots, \nu$, são unidades, a solução de (4.8) é única.

Teorema 4.3.1 *[31] O sistema linear (4.8), nas incógnitas $\sigma_1, \sigma_2, \cdots, \sigma_\nu$, dado por*

$$\begin{bmatrix} s_1 & s_2 & \cdots & s_\nu \\ s_2 & s_3 & \cdots & s_{\nu+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{2t-\nu} & s_{2t-\nu+1} & \cdots & s_{2t-1} \end{bmatrix} \begin{bmatrix} \sigma_\nu \\ \sigma_{\nu-1} \\ \vdots \\ \sigma_1 \end{bmatrix} = \begin{bmatrix} -s_{\nu+1} \\ -s_{\nu+2} \\ \vdots \\ -s_{2t} \end{bmatrix} \tag{4.9}$$

possui solução única se, e somente se, todas as magnitudes y_i, $i = 1, 2, \cdots, \nu$, dos erros ocorridos forem unidades no anel sobre o qual o código está definido. ■

Em vista destes fatos temos que o procedimento eficiente de decodificação de códigos Reed-Solomon e BCH compreende os seguintes passos.

- **Passo 1** Cálculo do vetor síndrome $\mathbf{s} = (s_1, s_2, \cdots, s_{2t})$ a partir do vetor $\mathbf{b}$;
- **Passo 2** Cálculo das funções simétricas elementares $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ a partir do vetor $\mathbf{s}$;

- **Passo 3** Cálculo dos números localizadores de erros $x_1, x_2, \cdots, x_\nu$ a partir das funções simétricas elementares $\sigma_i, i = 1, 2, \cdots, \nu$;
- **Passo 4** Cálculo das magnitudes dos erros $y_i's$ a partir dos $x_i's$, $i = 1, 2, \cdots, \nu$, e do vetor síndrome $\mathbf{s}$.

A seguir passaremos a caracterizar cada um dos quatros passos mencionados acima.

Passo 1. *Cálculo do Vetor Síndrome.* $\mathbf{s} = \mathbf{b}H^T$.

Passo 2. *Cálculo das Funções Simétricas Elementares.*

O problema aqui é determinar, a partir do sistema linear (4.8), os σ_i, $1 \leq i \leq \nu \leq t$, conhecendo os elementos $s_1, s_2, \cdots, s_{2t}$, as componentes do vetor síndrome $\mathbf{s}$, os quais pertencem a um anel comutativo finito local R, de tal maneira que ν seja mínimo. Pelo Teorema 4.3.1 temos que a solução de (4.8) será única somente quando as magnitudes de todos os erros forem unidades em R.

Este algorítmo é iterativo de modo que no n-ésimo passo o decodificador determina um conjunto de l_n vetores $\sigma_i^{(n)}$ tal que as $n - l_n$ equações

$$\begin{cases} s_n\sigma_0^{(n)} + s_{n-1}\sigma_1^{(n)} + \cdots + s_{n-l_n}\sigma_{l_n}^{(n)} = 0 \\ s_{n-1}\sigma_0^{(n)} + s_{n-2}\sigma_1^{(n)} + \cdots + s_{n-l_n-1}\sigma_{l_n}^{(n)} = 0 \\ \qquad \vdots \\ s_{l_n+1}\sigma_0^{(n)} + s_{l_n}\sigma_1^{(n)} + \cdots + s_1\sigma_{l_n}^{(n)} = 0 \end{cases} \tag{4.10}$$

chamadas **somas de potências**, sejam satisfeitas com l_n o menor possível e $\sigma_0^{(n)} = 1$. É conveniente representar o conjunto dos σ_i na seguinte forma polinomial $\sigma^{(n)}(x) = \sigma_0^{(n)} + \sigma_1^{(n)}x + \cdots + \sigma_{l_n}^{(n)}x^{l_n}$. Este polinômio tem grau menor ou igual a l_n e representa a solução do n-ésimo estágio.

Agora, suponha que no n-ésimo estágio o decodificador tenha determinado $\sigma^{(n)}(x)$, com l_n mínimo tal que o sistema (4.8) seja satisfeita. No $(n+1)$-ésimo estágio o decodificador procura encontrar o polinômio $\sigma^{(n+1)}(x)$ de menor grau tal que as equações

$$\sum_{i=0}^{l_{n+1}} s_{j-i}\sigma_i^{(n+1)} = 0, \quad l_{n+1} + 1 \leq j \leq n+1 \tag{4.11}$$

sejam satisfeitas. Definimos a n-ésima **discrepância** d_n como $d_n = s_{n+1}\sigma_0^{(n)} + s_n\sigma_1^{(n)} + \cdots + s_{n+1-l_n}\sigma_{l_n}^{(n)}$. Se $d_n = 0$, a equação (4.11) vale com $\sigma^{(n+1)}(x) = \sigma^{(n)}(x)$. E como $\sigma^{(n)}(x)$ é uma solução mínima no n-ésimo estágio, é certamente uma solução mínima no $(n+1)$-ésimo estágio. Contudo, se $d_n \neq 0$ a determinação de $\sigma^{(n+1)}(x)$ a partir de $\sigma^{(n)}(x)$ não é trivial. Em vista disto, temos dois lemas que estão diretamente relacionados com a determinação de $\sigma^{(n+1)}(x)$ (não necessariamente com o menor valor de l_{n+1} possível) a partir de $\sigma^{(n)}(x)$.

Lema 4.3.1 *[33] Suponha que $\sigma^{(n)}(x)$ seja um polinômio solução minimal para as n primeiras somas de potências (isto é, tal que as equações (4.10) sejam satisfeitas com l_n mínimo) e tenha próxima discrepância $d_n \neq 0$. Seja $\sigma^{(m)}(x) = 1 + \sigma_1^{(m)}x + \sigma_2^{(m)}x^2 + \cdots + \sigma_{l_m}^{(m)}x^{l_m}$ qualquer polinômio solução para as m primeiras somas de potências com $1 \leq m < n$ e tal que a equação $d_n - yd_m = 0$ admita uma solução y sobre o anel R (note que mesmo sendo d_m um divisor de zero em R, pode ocorrer que esta equação ainda admita uma solução). Então o polinômio $\sigma^{(n)}(x) - yx^{n-m}\sigma^{(m)}(x) = \sigma^{(n+1)}(x)$ é uma solução para as $(n+1)$ primeiras somas de potências. Além disso, $l_{n+1} = max\{l_n, l_m + n - m\}$.* ■

Lema 4.3.2 *[33] Sejam $\sigma^{(n)}(x)$, l_n e $d_n \neq 0$ como definidos no Lema 4.3.1. Suponha que $\sigma^{(n+1)}(x)$ seja uma solução polinomial das equações em (4.10) satisfazendo $n + 1 - l_{n+1}$ equações e que $\sigma^{(n+1)}(x) - \sigma^{(n)}(x) = ax^{n-m}\sigma^{(m)}(x)$, onde a é uma unidade em R e $\sigma_0^{(m)} = 1$. Então o polinômio $\sigma^{(m)}(x)$ é um polinômio solução para as $m - l_m$ primeiras equações em (4.10), e tendo próxima discrepância $d_m \neq 0$, satisfazendo $d_n + ad_m = 0$ e $l_m = l_{n+1} - (n - m)$.* ■

Em vista dos Lemas 4.3.1 e 4.3.2 temos o seguinte teorema.

Teorema 4.3.2 *[33] Sejam $\sigma^{(n)}(x)$ um polinômio solução minimal no n-ésimo estágio e $\sigma^{(m)}(x)$, $1 \leq m < n$ uma das soluções minimais anteriores tal que a equação $d_n - yd_m = 0$ admita uma solução em y e $m - l_m$ tenha o máximo valor possível. Então uma solução no $(n+1)$-ésimo estágio é $\sigma^{(n+1)}(x)$ onde*

$$\sigma^{(n+1)}(x) = \sigma^{(n)}(x) \quad e \quad l_{n+1} = l_n \quad para \quad d_n = 0 \tag{4.12}$$

$$\begin{cases} \sigma^{(n+1)}(x) = \sigma^{(n)}(x) - yx^{n-m}\sigma^{(m)}(x) \quad e \\ l_{n+1} = max\{l_n, l_m + n - m\} \end{cases} \quad para\ d_n \neq 0. \tag{4.13}$$

■

Observação 4.3.1 *A solução* $\sigma^{(n+1)}(x)$ *dada no Teorema 4.3.2 não é necessariamente a resposta almejada, pelo fato de que o mesmo não garante a minimalidade quando o coeficiente da menor potência da variável* x *em* $\sigma^{(n+1)}(x) - \sigma^{(n)}(x)$ *não for uma unidade em* R. *Contudo, em muitos casos este teorema ja aponta corretamente a solução minimal no* $(n+1)$*-ésimo estágio.* ■

Para que possamos discutir o emprego da generalização do algoritmo de Berlekamp-Massey na determinação iterativa da solução minimal, iremos assumir as seguintes condições iniciais

$$\sigma^{(-1)}(x) = 1, \quad l_{-1} = 0, \quad d_{-1} = 1,$$

$$\sigma^{(0)} = 1, \quad l_0 = 0, \quad d_0 = s_1.$$

onde 1 é o elemento identidade do anel em questão e s_1 é a primeira componente não nula do vetor síndrome.

Agora, pelo Lema 1 de [36] temos que se $\sigma^{(n)}(x)$ satisfaz $n - l_n$ equações em (4.10), mas não satisfaz $n + 1 - l_n$ equações em (4.10), então a solução $\sigma^{(n+1)}(x)$ deverá satisfazer $n + 1 - l_{n+1}$ equações em (4.10), onde $l_{n+1} \geq max\{l_n, n + 1 - l_n\}$.

Segue da Seção III de [36] que se a equação $d_{n'} - yd_{m'} = 0$, na variável y, sempre admite uma solução para todo m' tal que $1 \leq m' < n'$, então $l_{n+1} = max\{l_n, n+1-l_n\} = max\{l_n, l_m+n-m\}$. Por outro lado, se existir n' tal que $d_{n'} - yd_{m'} = 0$ não admita solução para todo m' tal que $1 \leq m' < n'$, então as soluções $\sigma^{(n)}(x)$, $n \geq n'$, dadas pelo Teorema 4.3.2 (equações (4.12) e (4.13)) não necessariamente serão minimais. Para este caso, suponha então que $\sigma^{(n)}(x)$ seja uma solução minimal no n-ésimo estágio e $\sigma^{(n+1)}(x)$ seja uma solução no $(n+1)$-ésimo estágio, dadas a partir do Teorema 4.3.2. Deste modo temos dois casos a considerar.

a) Se $l_{n+1} = max\{l_n, n + 1 - l_n\}$, então $\sigma^{(n+1)}(x)$ já é a solução minimal procurada no $(n + 1)$-ésimo estágio.

b) Se $l_{n+1} > max\{l_n, n+1-l_n\}$, é possível que exista uma solução polinomial minimal $D^{(n+1)}(x)$ com grau l tal que $max\{l_n, n+1-l_n\} \leq l < l_{n+1}$. Assim, um polinômio $D^{(n+1)}(x)$ com menor grau possível l, onde $max\{l_n, n + 1 - l_n\} \leq l < l_{n+1}$, será uma solução minimal no $(n + 1)$-ésimo estágio se, e somente se, o polinômio $\sigma^{(m)}(x)$ definido pela relação $D^{(n+1)}(x) - \sigma^{(n)}(x) = x^{n-m}\sigma^{(m)}(x)$ for uma solução para as m primeiras somas de potências, com $d_m = -d_n$ e com $\sigma_0^{(m)}$ sendo um divisor de zero no anel R.

Observação 4.3.2 *Seja* R *um anel comutativo com identidade. Seja a equação linear* $ax - b = 0$,

na variável x, *sobre o anel* R. *Temos três casos a considerar. Se* a *for uma unidade em* R, *temos que a equação linear* $ax - b = 0$ *possui solução única dada por* $x = a^{-1}b$. *Se* b *for uma unidade e* a *for um divisor de zero em* R, *temos que a equação linear* $ax - b = 0$ *não possui solução em* R. *Finalmente, se* a *e* b *forem ambos divisores de zero em* R, *então a equação linear* $ax - b = 0$ *poderá ter ou não solução em* R. ■

A seguir descrevemos o **Algoritmo de Berlekamp-Massey Modificado para anéis comutativos finitos locais com identidade** que resolve o problema original, isto é, as equações em (4.8). As entradas destas equações são as componentes do vetor síndrome $\mathbf{s} = (s_1, s_2, \cdots, s_{2t})$ pertencentes ao anel R. Este algoritmo nos dará como saída um conjunto de vetores σ_i, $i = 1, 2, \cdots, \nu$, tal que as equações em (4.8) sejam satisfeitas com o valor mínimo possível de ν.

Inicie com a seguinte tabela

n	$\sigma^{(n)}(x)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	s_1	0	0
1				
2				
$\vdots$				
$2t$				

e em seguida complete-a da seguinte maneira:

1. $n \leftarrow 0$,

2. Se $d_n = 0$, então $\sigma^{(n+1)}(x)$ e l_{n+1} são dados por (4.12). Vá para o item 5,

3. Se $d_n \neq 0$, então encontre m tal que a equação $d_n - yd_m = 0$, na variável y, tenha solução sobre o anel R e $m - l_m$ tenha máximo valor possível. Neste caso, $\sigma^{(n+1)}(x)$ e l_{n+1} são dados por (4.13),

4. Se $l_{n+1} = max\{l_n, n+1-l_n\}$ vá para o item 5. Caso contrário, devemos procurar uma solução $D^{(n+1)}(x)$ com grau mínimo possível l no intervalo $max\{l_n, n+1-l_n\} \leq l < max\{l_n, l_m + n - m\}$

tal que o polinômio $\sigma^{(m)}(x)$ definido pela equação $D^{(n+1)}(x) - \sigma^{(n)}(x) = x^{n-m}\sigma^{(m)}(x)$ seja uma solução para as m primeiras somas de potências tal que $d_m = -d_n$ e $\sigma_0^{(m)}$ seja um divisor de zero em R. Se este polinômio for encontrado, então faça $\sigma^{(n+1)}(x) \leftarrow D^{(n+1)}(x)$ e $l_{n+1} \leftarrow l$,

5. Se $n < 2t-1$ calcule $d_{n+1} = s_{n+2} + s_{n+1}\sigma_1^{(n+1)} + \cdots + s_{n+2-l_{n+1}}\sigma_{l_{n+1}}^{(n+1)}$,

6. $n \leftarrow n+1$, se $n < 2t$ vá para o item 2. Caso contrário, fim.

Deste modo a resposta desejada será dada pelo polinômio $\sigma^{(2t)}(x)$, isto é, os seus coeficientes formam uma solução para as equações em (4.8).

Observação 4.3.3 *Notamos que a principal diferença deste algoritmo de Berlekamp-Massey modificado para o original está no item 4, onde é verificado se a solução $\sigma^{(n)}(x)$ calculada no item 3 é minimal. Caso negativo, devemos encontrar uma solução minimal testando se um polinômio $\sigma^{(m)}(x)$ (satisfazendo certas condições) é uma solução para as m primeiras somas de potência. Observamos também que o número de polinômios $\sigma^{(m)}(x)$ a serem testados não é muito grande e a complexidade do algoritmo modificado é a mesma comparando com o algoritmo original.* ■

Finalizado este passo da determinação dos σ_i, $i = 1, 2, \cdots, \nu$, que satisfaçam as equações (4.8), passamos a analizar o terceiro passo para a decodificação dos códigos Reed-Solomon e BCH. Este passo tem como finalidade calcular os números localizadores de erros corretos $x_i's$ a partir dos $\sigma_i's$, $i = 1, 2, \cdots, \nu$.

Passo 3. *Cálculo dos Números Localizadores de Erros.*

O objetivo neste passo é resolver a equação polinomial $\rho(z) = 0$ sobre o anel R onde $\rho(z) = z^{\nu}\sigma^{(2t)}(z^{-1}) = z^{\nu} + \sigma_1 z^{\nu-1} + \cdots + \sigma_{\nu-1}z + \sigma_{\nu}$. Sendo R um anel local devemos ter que as raízes de $\sigma^{(2t)}(z) = 1 + \sigma_1 z + \cdots + \sigma_{\nu}z^{\nu}$ estão em R^*, e assim as raízes de $\rho(z)$ (o polinômio recíproco de $\sigma^{(2t)}(z)$) são as inversas das raízes de $\sigma^{(2t)}(z)$.

A solução do sistema (4.8), em geral, não é única, e deste modo os valores de σ_i, $i = 1, 2, \cdots, \nu$, produzidos pelo algoritmo de Berlekamp-Massey modificado para anéis finitos locais pode não ser os mesmos definidos pela equação (4.6). Assim, as raízes do polinômio $\rho(z)$ (com os coeficientes σ_i, $i = 1, 2, \cdots, \nu$, obtidos pelo algoritmo de Berlekamp-Massey modificado) podem não ser os números localizadores de erros corretos. Contudo veremos, a seguir, que de fato estes números podem ser obtidos

a partir das raízes de $\rho(z)$, desde que os coeficientes σ_i, $i = 1, 2, \cdots, \nu$, sejam uma solução de (4.8).

Com o intuito de analisarmos a **relação entre as raízes de $\rho(z)$ e os números localizadores de erros** suponha que $\rho(z)$ tenha pelo menos ν raízes distintas, $z_1, z_2, \cdots, z_\nu$, sobre o anel R. Assim, podemos escrever $\rho(z) = (z - z_1)(z - z_2) \cdots (z - z_\nu)$. Observamos que pelo menos uma solução de $\sigma(z)$ produzida pelo algoritmo de Berlekamp-Massey modificado terá esta propriedade.

No sentido de converter as raízes de $\rho(z)$ nos números localizadores de erros corretos, suponhamos que estes números sejam $x_1, x_2, \cdots, x_\nu$, que as magnitudes dos erros sejam $y_1, y_2, \cdots, y_\nu$ e que as raízes de $\rho(z)$ sejam $z_1, z_2, \cdots, z_\nu$. Assim, podemos escrever

$$y_i x_i^j \rho(z) = y_i x_i^j (z - z_1)(z - z_2) \cdots (z - z_\nu) \tag{4.14}$$

para $i = 1, 2, \cdots, \nu$ e $j = 1, 2, \cdots, 2t - \nu$. Substituindo z por x_i e somando o primeiro membro desta equação para $i = 1, 2, \cdots, \nu$, obtemos

$$s_{j+\nu} + s_{j+\nu-1}\sigma_1 + \cdots + s_{j+1}\sigma_{\nu-1} + s_j\sigma_\nu. \tag{4.15}$$

Esta equação se anula para todo $j = 1, 2, \cdots, 2t - \nu$ (pois os $\sigma_i's$ formam um conjunto solução de (4.8)) e deste modo

$$\sum_{i=1}^{\nu} y_i x_i^j (x_i - z_1)(x_i - z_2) \cdots (x_i - z_\nu) = 0 \tag{4.16}$$

para $j = 1, 2, \cdots, 2t - \nu$. Escrevendo na forma matricial temos o seguinte sistema homogêneo

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_\nu \\ x_1^2 & x_2^2 & \cdots & x_\nu^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{2t-\nu} & x_2^{2t-\nu} & \cdots & x_\nu^{2t-\nu} \end{bmatrix} \begin{bmatrix} y_1 p_1 \\ y_2 p_2 \\ \vdots \\ y_\nu p_\nu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.17}$$

onde $p_i = \prod_{l=1}^{\nu} (x_i - z_l)$ para $i = 1, 2, \cdots, \nu$. Sendo $2t - \nu$ maior ou igual a ν (pois $\nu \leq t$) e pelo Teorema 4.3.1 o posto (maior número r tal que existe uma submatriz $r \times r$ tal que o determinante é uma unidade em R) da matriz $\nu \times \nu$ em (4.17) é ν (que é igual ao número de incógnitas), segue do Teorema 5.3 de [12] que este sistema possui somente a solução trivial, isto é, $y_i p_i = 0$, $i = 1, 2, \cdots, \nu$.

Como consequência deste fato temos que cada p_i, $i = 1, 2, \cdots, \nu$, é um divisor de zero em R. Logo, em cada p_i existe pelo menos um l-ésimo fator $(x_i - z_l)$ que é um divisor de zero em R.

Além disso, se o l_1-ésimo fator em p_i, $i-1,2,\cdots,\nu$, é um divisor de zero d_1 e o l_2-ésimo fator em p_k é também um divisor de zero d_2, então $l_1 \neq l_2$ para $i \neq k$. De fato, se $l_1 = l_2$ para $i \neq k$, temos que $x_i - x_k$, onde $x_i - z_{l_1} = d_1$ e $x_k - x_{l_2} = d_2$, é um divisor de zero em R (pelo fato de R ser um anel local), o que é uma contradição pelas demonstrações dos Teoremas 3.4.1 e 3.4.2 e do Lema 4.1.1, já que $i \neq k$. Concluímos deste modo que para cada z_i existe um único número localizador de erro x_i, $i = 1,2,\cdots,\nu$, correspondente.

Resumindo, temos o seguinte procedimento para o cálculo dos números localizadores de erros corretos.

- Calcule as raízes $z_1, z_2, \cdots, z_\nu$ do polinômio $\rho(z) = z^\nu \sigma^{(2t)}(z^{-1})$, o recíproco do polinômio produzido pelo algoritmo de Berlekamp-Massey modificado.

  Observação 4.3.4 *O método usado para encontrar a solução da equação polinomial* $f(x) = 0$, *onde* $f(x) = a_0 + a_x + a_2x^2 + \cdots + a_nx^n$ *sobre o anel* R, *é o de fazer uma busca exaustiva em* R *das raízes de* $f(x)$. *Isto é análogo ao que se faz quando da decodificação de códigos BCH sobre corpos finitos* $GF(q)$, *onde* q *é uma potência de um número primo.* ■

- Entre os $x_i = \alpha^l$, $l = 1,2,\cdots,s$, selecionamos aqueles $x_i's$ tal que as diferenças $x_i - z_i$ sejam divisores de zero em R. Os $x_i's$ selecionados serão os números localizadores de erros corretos e cada $x_i = \alpha^{k_j}$, $j = 1,2,\cdots,n$, indica a posição j do erro na palavra-código.

Finalizado este passo de localização dos erros, passamos ao quarto e último passo do procedimento de decodificação dos códigos Reed-Solomon e BCH.

Passo 4. *Determinação das Magnitudes dos Erros.*

A fim de completarmos o processo de decodificação do códigos Reed-Solomon e BCH descrevemos, a seguir, um método para a determinação das magnitudes y_i, $i = 1,2,\cdots,\nu$, dos erros. Neste sentido, primeiramente iremos mostrar que estas magnitudes ficam determinadas de maneira única após o conhecimento dos x_i, $i = 1,2,\cdots,\nu$, que são os números localizadores de erros corretos determinados no passo anterior.

As ν primeiras equações em (4.5) (que relacionam as componentes do vetor síndrome com os x_i e y_i, $i = 1, 2, \cdots, \nu$) podem ser escritas na forma

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_\nu \\ x_1^2 & x_2^2 & \cdots & x_\nu^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^\nu & x_2^\nu & \cdots & x_\nu^\nu \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_\nu \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_\nu \end{bmatrix}. \tag{4.18}$$

Pelo Teorema 4.3.1 segue que a matriz $\nu \times \nu$ em (4.18) é não singular, isto é, seu determinante é uma unidade em R, acarretando que o vetor $\mathbf{y} = (y_1, y_2, \cdots, y_\nu)$ é determinado de maneira única.

Para determinar o vetor $\mathbf{y}$ usamos o procedimento proposto por Forney [15]. Este método requer o conhecimento dos números localizadores de erros corretos $x_1, x_2, \cdots, x_\nu$ e de suas funções simétricas elementares $\sigma_1, \sigma_2, \cdots, \sigma_\nu$, que são calculadas a partir das equações (4.6).

Inicialmente, definimos as funções simétricas elementares σ_{jl} dos números localizadores de erros $x_1, x_2, \cdots, x_{j-1}, x_{j+1}, \cdots, x_\nu$ através da seguinte equação

$$\prod_{i \neq j}(x - x_i) = \sum_{l=0}^{\nu-1} \sigma_{jl} x^{\nu-1-l}. \tag{4.19}$$

Fazendo uso da equação (4.6) obtemos

$$\prod_{i=1}^{\nu}(x - x_i) = \sum_{i=0}^{\nu} \sigma_i x^{\nu-i}, \tag{4.20}$$

onde $\sigma_0 = \sigma_{j0} = 1$, o elemento identidade do anel R. Das equações (4.19) e (4.20) chegamos a

$$(x - x_j) \sum_{l=0}^{\nu-1} \sigma_{jl} x^{\nu-1-l} = \sum_{i=0}^{\nu} \sigma_i x^{\nu-i}. \tag{4.21}$$

Desenvolvendo o primeiro membro da equação (4.21) obtemos

$$\sum_{l=0}^{\nu-1} \sigma_{jl} x^{\nu-l} - \sum_{l=0}^{\nu-1} \sigma_{jl} x_j x^{\nu-1-l} = \sum_{i=0}^{\nu} \sigma_i x^{\nu-i}. \tag{4.22}$$

Da equação (4.22) concluímos que os coeficientes σ_{jl} podem ser obtidos de maneira recursiva a partir dos $x_i's$ e $\sigma_i's$, $i = 1, 2, \cdots, \nu$ (que são conhecidos), mediante a seguinte relação

$$\sigma_{ji} = \sigma_i + x_j \sigma_{j,i-1} \quad \text{para } 1 \leq i \leq \nu - 1 \quad \text{e com} \quad \sigma_0 = \sigma_{j0} = 1. \tag{4.23}$$

Denotando a magnitude de cada erro por y_j, temos então que

$$\sum_{l=0}^{\nu-1} \sigma_{jl} s_{\nu-l} = \sum_{l=0}^{\nu-1} \sigma_{jl} \sum_{i=1}^{\nu} y_i x_i^{\nu-l} = \sum_{i=1}^{\nu} y_i x_i \sum_{l=0}^{\nu-1} \sigma_{jl} x_i^{\nu-1-l}. \tag{4.24}$$

Fazendo uso da equação (4.19) chegamos que

$$\sum_{l=0}^{\nu-1} \sigma_{jl} s_{\nu-l} = \sum_{i=1}^{\nu} y_i x_i \prod_{m \neq j} (x_i - x_m) = y_j x_j \prod_{m \neq j} (x_j - x_m), \tag{4.25}$$

onde esta última igualdade segue do fato que o somatório em questão só não se anula se $i = j$. Da equação (4.25) concluímos então que

$$\sum_{l=0}^{\nu-1} \sigma_{jl} s_{\nu-l} = y_j \sum_{l=0}^{\nu-1} \sigma_{jl} x_j^{\nu-l}, \tag{4.26}$$

e consequentemente, cada y_j, $j = 1, 2, \cdots, \nu$, é dado por

$$y_j = \frac{\sum_{l=0}^{\nu-1} \sigma_{jl} s_{\nu-l}}{\sum_{l=0}^{\nu-1} \sigma_{jl} x_j^{\nu-l}}. \tag{4.27}$$

Observação 4.3.5 *O denominador da equação (4.27) é uma unidade no anel R pois é igual a $x_j \prod_{m \neq j} (x_j - x_m)$, onde cada fator é do tipo $\alpha^i - \alpha^j$ com $0 \leq i \neq j \leq n-1$. Assim, segue das demonstrações dos Teoremas 3.4.1, 3.4.2 e Lema 4.1.1 que as diferenças $(\alpha^i - \alpha^j)$, com $i \neq j$, são sempre unidades no anel R, o que implica que o referido denominador é uma unidade no anel R.* ■

Concluímos deste modo o quarto passo de decodificação dos códigos Reed-Solomon e BCH, que é a determinação das magnitudes y_i, $i = 1, 2, \cdots, \nu$, dos erros através da equação (4.27). Assim, finalizamos o processo de decodificação destes códigos definidos sobre anéis comutativos finitos locais com identidade. Finalizamos esta seção com dois exemplos ilustrando este procedimento de decodificação dos códigos Reed-Solomon e BCH, respectivamente. Sendo que no primeiro exemplo faremos a decodificação de um código Reed-Solomon e no segundo de um código BCH.

Exemplo 4.3.1 (Decodificação de código Reed-Solomon) *Seja o código do Exemplo 3.4.2. Os parâmetros deste código são $p = 7$, $n = p - 1 = 6$, $\alpha = 3$, $d_{min} \geq 5$ e $m = 1$. A matriz verificação de paridade é dada por*

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha^8 & \alpha^{10} \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & \alpha^{15} \\ 1 & \alpha^4 & \alpha^8 & \alpha^{12} & \alpha^{16} & \alpha^{20} \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 & 6 & 4 & 5 \\ 1 & 2 & 4 & 1 & 2 & 4 \\ 1 & 6 & 1 & 6 & 1 & 6 \\ 1 & 4 & 2 & 1 & 4 & 2 \end{bmatrix}.$$

Suponha que a palavra toda nula $\mathbf{c} = (000000)$ *seja transmitida através do canal e que o mesmo introduza o vetor erro* $\mathbf{e} = (0, 0, i, 0, 2i, 0)$. *O receptor terá como vetor recebido* $\mathbf{b} = \mathbf{c} + \mathbf{e} = (0, 0, i, 0, 2i, 0)$. *O processo de decodificação é dado da seguinte maneira:*

1. *Cálculo do vetor síndrome:* $\mathbf{s} = \mathbf{b}H^T = (3i, i, 3i, 3i)$.

2. *Cálculo das variáveis* $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ *que satisfaçam o sistema (4.8). Neste passo vamos aplicar o algoritmo de Berlekamp-Massey modificado com as entradas* $s_1 = 3i$, $s_2 = i$, $s_3 = 3i$ *e* $s_4 = 3i$. *Deste modo obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	$3i$	0	0
1	$1 + 4iz$	$i + 2$	1	0
2	$1 + 2z$	$5i$	1	1
3	$1 + 2z + 3z^2$	$5i$	2	1
4	$1 + z + z^2$	$-$	2	2

3. *Determinação das raízes de* $\rho(z) = z^2 + z + 1$ *(o polinômio recíproco de* $\sigma^{(4)}(z)$*). Através do cálculo direto vemos que* $z_1 = 2$ *e* $z_2 = 4$ *são raízes de* $\rho(z)$. *Dentre os elementos* $\alpha^0 = 1, \alpha = 3, \alpha^2 = 2, \alpha^3 = 6, \alpha^4 = 4$ *e* $\alpha^5 = 5$ *temos que* $x_1 = \alpha^2 = 2$ *e* $x_2 = \alpha^4 = 4$ *são tais que* $x_1 - z_1 = x_2 - z_2 = 0$ *(divisor de zero em* $\mathbb{Z}_7(i)$*). Assim,* $x_1 = \alpha^{k_3}$ *e* $x_2 = \alpha^{k_5}$ *são os números localizadores de erros corretos e que os erros ocorreram nas posições 3 e 5 da palavra toda nula. As funções simétricas elementares* σ_1 *e* σ_2 *satisfazem a relação* $(x - x_1)(x - x_2) = x^2 + \sigma_1 x + \sigma_2$. *Donde concluímos que* $\sigma_1 = \sigma_2 = 1$.

4. *Cálculo da magnitude dos erros. Este cálculo é feito através do uso da equação (4.27). Contudo, devemos primeiramente determinar os coeficientes* σ_{jl} *por intermédio da equação (4.23). Assim,* $\sigma_{11} = \sigma_1 + x_1\sigma_{10} = 1 + 2.1 = 3$ *e* $\sigma_{21} = \sigma_1 + x_2\sigma_{20} = 1 + 4.1 = 5$, *e substituindo esses valores na equação (4.27) obtemos* $y_1 = i$ *e* $y_2 = 2i$.

Deste modo concluímos que ocorreram dois erros. Um na posição 3 da palavra-código de magnitude i *e o outro na posição 5 com magnitude* $2i$. ■

Exemplo 4.3.2 (Decodificação de código BCH) *Seja* $\mathcal{C}(n,\eta)$ *o código BCH sobre* $\mathbb{Z}_2(i)$ *gerado pelo polinômio* $g(x) = x^8 + x^4 + x^2 + x + 1$. *Sejam* $R = \frac{\mathbb{Z}_2(i)}{<x^4+x+1>}$, *onde* $f(x) = x^4 + x + 1$ *é irredutível sobre* $\mathbb{Z}_2$, G_{15} *o subgrupo cíclico de* R *que contém todas as raízes de* $x^{15} - 1$ *e* α *um elemento primitivo de* G_{15}. *Como* $\alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^6, \alpha^8, \alpha^9$ *e* α^{12} *são as raízes de* $g(x)$ *segue que* $d_{min}(\mathcal{C}) \geq 5$, *e este código tem capacidade de correção de até* $t = 2$ *erros. Seja* $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_{15}) = (\alpha^0, \alpha^1, \cdots, \alpha^{14}) = (\alpha^{k_1}, \alpha^{k_2}, \cdots, \alpha^{k_{15}})$ *o vetor localizador e a matriz verificação de paridade dada por*

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 & \alpha^8 & \alpha^9 & \alpha^{10} & \alpha^{11} & \alpha^{12} & \alpha^{13} & \alpha^{14} \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha^8 & \alpha^{10} & \alpha^{12} & \alpha^{14} & \alpha & \alpha^3 & \alpha^5 & \alpha^7 & \alpha^9 & \alpha^{11} & \alpha^{13} \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} \\ 1 & \alpha^4 & \alpha^8 & \alpha^{12} & \alpha & \alpha^5 & \alpha^9 & \alpha^{13} & \alpha^2 & \alpha^6 & \alpha^{10} & \alpha^{14} & \alpha^3 & \alpha^7 & \alpha^{11} \end{bmatrix}.$$

Supondo que a palavra toda nula $\mathbf{c} = (000000000000000)$ *seja transmitida através do canal, e que o mesmo introduza o seguinte vetor erro* $\mathbf{e} = (0i0000000000010)$. *Deste modo, o vetor recebido será* $\mathbf{b} = \mathbf{c} + \mathbf{e}$. *O procedimento de decodificação é da seguinte maneira:*

1. *Cálculo da síndrome:* $\mathbf{s} = \mathbf{b}H^T = (\alpha i + \alpha^{13}, \alpha^2 i + \alpha^{11}, \alpha^3 i + \alpha^9, \alpha^4 i + \alpha^7)$.

2. *Cálculo das variáveis* $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ *que satisfaçam o sistema (4.8). Para este caso, vamos aplicar o algoritmo de Berlekamp-Massey modificado com síndromes* $s_1 = \alpha i + \alpha^{13}$, $s_2 = \alpha^2 i + \alpha^{11}$, $s_3 = \alpha^3 i + \alpha^9$ *e* $s_4 = \alpha^4 i + \alpha^7$. *Assim, através do algoritmo de Berlekamp-Massey modificado,*

obtemos a seguinte tabela

n	$\sigma^{(n)}(z)$	d_n	l_n	$n-l_n$
-1	1	1	0	-1
0	1	$\alpha i+\alpha^{13}$	0	0
1	$1+(\alpha i+\alpha^{13})z$	$\alpha^2+\alpha^2 i$	1	0
2	$1+(\alpha^2+\alpha^7 i)z$	$\alpha^4 i+\alpha^{13}$	1	1
3	$1+(\alpha^2+\alpha^7 i)z+(\alpha^4 i+\alpha^9)z^2$	$\alpha^3 i+\alpha^3$	2	1
4	$1+\alpha^{12}z+\alpha^{14}z^2$	-	2	2

3. Determinação das raízes de $\rho(z)=z^2+\alpha^{12}z+\alpha^{14}$ *(o polinômio recíproco de* $\sigma^{(4)}(z)$*). Por inspeção direta, temos que* $z_1=\alpha$ *e* $z_2=\alpha^{13}$ *são as raízes de* $\rho(z)$*. Dentre os elementos* $\alpha^0=1,\alpha,\alpha^2,\alpha^3,\alpha^4=\alpha+1,\alpha^5=\alpha^2+\alpha,\alpha^6=\alpha^3+\alpha^2,\alpha^7=\alpha^3+\alpha+1,\alpha^8=\alpha^2+1,\alpha^9=\alpha^3+\alpha,\alpha^{10}=\alpha^2+\alpha+1,\alpha^{11}=\alpha^3+\alpha^2+\alpha,\alpha^{12}=\alpha^3+\alpha^2+\alpha+1,\alpha^{13}=\alpha^3+\alpha^2+1,\alpha^{14}=\alpha^3+1$*, temos que* $x_1=\alpha$ *e* $x_2=\alpha^{13}$ *são tais que* $x_1-z_1=x_2-z_2=0$ *(divisor de zero em R). Assim,* $x_1=\alpha^{k_2}$ *e* $x_2=\alpha^{k_{14}}$ *são os números localizadores de erros corretos e que os erros ocorreram nas posições 2 e 14 da palavra toda nula. As funções simétricas elementares* σ_1 *e* σ_2 *satisfazem a relação* $(x-x_1)(x-x_2)=x^2+(x_1+x_2)x+x_1x_2=x^2+\sigma_1x+\sigma_2$*. Deste modo, temos que* $\sigma_1=\alpha^{12}$ *e* $\sigma_2=\alpha^{14}$.

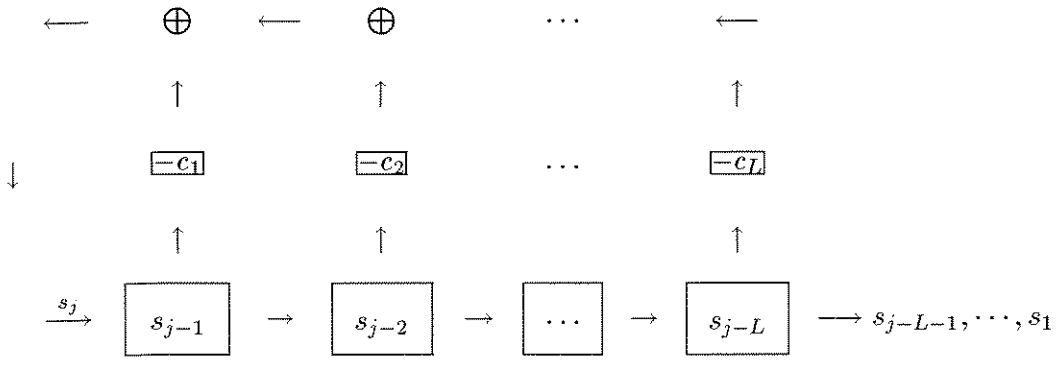
4. Cálculo da magnitude dos erros. De modo análogo ao exemplo anterior, inicialmente determinaremos os coeficientes σ_{jl} *por intermédio da equação (4.23). Assim,* $\sigma_{11}=\sigma_1+x_1\sigma_{10}=\alpha^{12}+\alpha=\alpha^{13}$ *e* $\sigma_{21}=\sigma_1+x_2\sigma_{20}=\alpha^{12}+\alpha^{13}=\alpha$*, e aplicando esses valores na equação (4.27), obtemos* $y_1=i$ *e* $y_2=1$.

Podemos então afirmar que ocorreram dois erros. Um na posição 2 da palavra código com magnitude i *e outro na posição 14 com magnitude 1.* ■

4.4 Aplicação do Algoritmo de Berlekamp-Massey Modificado

O problema de circuitos lineares de deslocamentos com realimentação para dígitos pertencentes a um corpo finito foi tratado em [36]. Este mesmo problema foi estendido em [31] para o caso de anéis locais da forma $\mathbb{Z}_{p^k}$. Nesta seção, veremos uma extensão deste mesmo problema para o caso em que a seqüência pertence a um anel comutativo finito local com identidade, fazendo uso do algoritmo de Berlekamp-Massey modificado.

Um circuito linear de deslocamentos com relimentação (LFSR - linear feedback register) de comprimento L (ver figura abaixo) consiste de uma cascata de L atrasadores (registros de deslocamentos) e alguns multiplicadores e somadores capazes de gerar uma combinação linear dos conteúdos destes registros.



LFSR de comprimento L

O conteúdo do último registro é a saída do LFSR. Os conteúdos iniciais $s_1, s_2, \cdots, s_L$ dos L atrasadores coincidem com os L primeiros dígitos de saída, e os digítos subsequentes de saída são obtidos através da seguinte relação de recorrência $s_j = -\sum_{i=1}^{L} c_i s_{j-i}$ para $j \geq L+1$. Os dígitos de saída e os coeficientes de realimentação $c_1, c_2, \cdots, c_L$ são elementos do anel R. Quando $c_L = 0$ o LFSR é dito **singular**.

Um LFSR gera uma seqüência finita de dígitos $s_1, s_2, \cdots, s_N$ quando esta seqüência coincide com os N primeiros dígitos de saída do mesmo para algum conteúdo inicial. Se $L \geq N$ o LFSR sempre gera a seqüência e se $L < N$ temos que o LFSR gera a seqüência se, e somente se,

$$s_j + s_{j-1}c_1 + \cdots + s_{j-L+1}c_{L-1} + s_{j-L}c_L = 0 \quad \text{para} \quad L+1 \leq j \leq N. \tag{4.28}$$

Em [36] foi mostrado que o algoritmo usado para decodificação de códigos BCH também pode ser usado para sintetizar um LFSR de comprimento mínimo L que gera uma seqüência prescrita. Isto é, o problema de geração de um LFSR e a decodificação de um código BCH são equivalentes.

De uma maneira bem semelhante, o algoritmo de Berlekamp-Massey modificado pode ser aplicado para sintetizar um LFSR de comprimento mínimo que gera uma sequência $s_1, s_2, \cdots, s_N$ de elementos do anel R, em [31] isto foi mostrado para o anel $\mathbb{Z}_{p^k}$. Comparando as equações (4.8) e (4.28) podemos constatar que este fato faz sentido pois em ambos os casos o objetivo é encontrar a menor quantidade de variáveis (ν ou L) que satisfaçam os respectivos conjuntos de equações.

Deste modo, temos que as entradas do algoritmo serão os elementos $s_1, s_2, \cdots, s_N$ que formam a seqüência dada e a saída do mesmo será o polinômio $c(x) = 1 + c_1 x + \cdots + c_L x^L$ na variável x, onde os coeficientes são os coeficientes de realimentação do LFSR mínimo de comprimento L que gera $s_1, s_2, \cdots, s_N$. Este LFSR mínimo será único se, e somente se, $2L \leq N$ e em cada estágio do algoritmo a equação linear $d_n - y d_m = 0$ na variável y tiver solução única, onde d_n e d_m são as n-ésima e m-ésima discrepâncias, respectivamente. Caso contrário, haverá mais de um LFSR mínimo de comprimento L que gera $s_1, s_2, \cdots, s_N$.

Observamos que, conforme discutido em [36], o algoritmo de Berlekamp-Massey modificado pode ainda ser aplicado como parte de um codificador de fonte ou compressor de dados.

Exemplo 4.4.1 *Encontrar o LFSR mínimo que gera a seqüência* $s_1 = 1, s_2 = i, s_3 = i, s_4 = 1$ *e* $s_5 = i$ *sobre o anel* $\mathbb{Z}_2[i]$. *Através da aplicação do algoritmo de Berlekamp-Massey modificado*

encontramos

n	$\sigma^{(n)}(x)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	1	0	0
1	$1 + x$	$i + 1$	1	0
2	$1 + ix$	$i + 1$	1	1
3	$1 + ix + (i + 1)x^2$	$i + 1$	2	1
4	$1 + (i + 1)x + x^2$	$i + 1$	2	2
5	$1 + (i + 1)x + ix^3$	$-$	3	2

Portanto, $c(x) = 1 + (i + 1)x + ix^3$ *e o LFSR mínimo que gera a seqüência dada tem suas saídas determinadas através da seguinte relação* $s_n + (i+1)s_{n-1} + is_{n-3} = 0$ *para* $n \geq 4$, *onde as operações são realizadas em* $\mathbb{Z}_2[i]$.

4.5 Conclusões

Neste capítulo apresentamos uma proposta de extensão dos códigos BCH sobre anéis comutativos finitos com identidade. A construção dos códigos BCH cíclicos foi baseada na fatoração do polinômio $x^s - 1$ sobre o grupo das unidades de uma extensão do anel. Este procedimento é bastante análogo a construção destes códigos sobre corpos finitos. O processo de obtenção dos mesmos não exige conhecimento de códigos sobre o corpo $GF(p)$, com p um número primo.

Também apresentamos um algoritmo de decodificação para os códigos Reed-Solomon e BCH definidos sobre um anel comutativo finito local com identidade. O algoritimo apresentado mantém as características essenciais dos correspondentes algorítmos para decodificação desses códigos definidos sobre anéis de inteiros residuais $\mathbb{Z}_q$, onde q é uma potência de um número primo, e este por sua vez mantem as mesmas características essenciais dos correspondentes algoritmos de decodificação para códigos sobre os corpos finitos $GF(q)$, não implicando deste modo em um aumento considerável de complexidade.

Vimos também que o algoritmo de Berlekamp-Massey modificado, usado na decodificação de códigos Reed-Solomon e BCH definidos sobre anéis comutativos finitos locais com identidade, pode ainda ser aplicado para sintetizar um LFSR mínimo que gera seqüências de elementos pertencentes a estes anéis.

Capítulo 5

Códigos Alternantes e BCH sobre Anéis Comutativos Finitos Locais com Identidade

Códigos lineares sobre anéis de inteiros tem recentemente dispertado um grande interesse devido a sua nova função em teoria de codificação algébrica e suas bem sucedidas aplicações em modulação e codificação. Um trabalho notável de Hammons et. al. [22], mostrou que certos códigos não lineares binários com boa capacidade de correção de erros podem ser vistos, através do mapeamento de Gray, como códigos lineares sobre $\mathbb{Z}_4$. Trabalhos anteriores por Calderbenk e Sloane [13], Blake [8] e [9], e Shankar [46] estendem a noção de códigos cíclicos, códigos Reed-Solomon (RS) e códigos BCH sobre $GF(q)$ para as classes de códigos sobre os anéis finitos $\mathbb{Z}_q$, com q uma potência de um primo. Examinando esquemas BCM (block coded modulation) como códigos sobre grupos, Biglieri e Elia [6], mostraram que códigos sobre grupos abelianos podem ser estudados via códigos lineares sobre anéis finitos.

É conhecido que códigos alternantes formam uma grande e poderosa família de códigos que podem ser obtidos por uma simples modificação na matriz verificação de paridade de um código BCH.

Por outro lado, a maioria dos códigos não binários são determinados sob a métrica de Hamming. Códigos com a métrica de Hamming são ideais para canais balanceados em que a probabilidade de erro para cada símbolo é a mesma. A métrica de Lee foi desenvolvida como uma alternativa para a métrica de Hamming para transmissão de sinais não binários sobre certos canais ruidosos. Ulrich [51]

considerou códigos que podem corrigir um erro de $+1$ ou -1 numa palavra-código. Códigos com a métrica de Lee sobre corpos finitos, foram descritos por Lee [34], Berlekamp [5], Chiang e Wolf [14], Satyanarayana [44], e Roth e Siegel [42].

Neste capítulo, analisamos o problema da construção de códigos alternantes sobre anéis comutativos finitos locais com identidade através da métrica de Hamming e códigos BCH sobre $\mathbb{Z}_q$, com q uma potência de um primo p, através da métrica de Lee. Também analisamos o problema da decodificação destes códigos. O algoritmo de decodificação sendo proposto pode também ser usado para corrigir apagamentos por uma simples modificação do método para corpos finitos descritos em [7] quando códigos alternantes são considerados. Contudo, o algoritmo de decodificação para os códigos BCH sob a métrica de Lee sobre $\mathbb{Z}_q$ combina o algoritmo de Berlekamp-Massey modificado proposto por Interlando, Palazzo, e Elia [33] com o método proposto por Roth e Siegel [42].

Este capítulo considera os seguintes aspectos. Na Seção 5.1, apresentamos uma técnica de construção de códigos alternantes sobre anéis comutativos finitos locais com identidade sob a métrica de Hamming. Baseado no algoritmo de Berlekamp-Massey modificado, um processo eficiente de decodificação é estabelecido. Na Seção 5.2, a construção de códigos BCH sobre o anel dos inteiros residuais $\mathbb{Z}_q$ sob a métrica de Lee é considerado. Propomos um algoritmo de decodificação alternativo para esses códigos onde também fazemos uso do algoritmo de Berlekamp-Massey modificado. A questão relativa à existência de um algoritmo de decodificação simples sob a métrica de Lee permanece aberta. Um processo sistemático para decodificação de códigos sob a métrica de Lee tem sido a motivação das pesquisas recentes.

5.1 Códigos Alternantes

Nesta seção apresentamos um processo de construção de códigos alternantes sobre anéis comutativos finitos locais com identidade e um processo eficiente de decodificação, baseado no algoritmo de Berlekamp-Massey modificado, para correção de até $r-1$ erros, mediante o número de erros de Hamming garantido pela distância de Hamming mínima projetada $r+1$. Primeiramente, iremos rever as propriedades importantes de extensão de Galois sobre anéis, as quais caracterizam os códigos alternantes.

Ao longo desta seção, assumimos que A é um anel comutativo finito local com identidade com ideal maximal $\mathcal{M}$ e corpo residual $\mathbb{K} = \frac{A}{\mathcal{M}} = GF(p^m)$, onde m é um inteiro positivo e p é

um primo. Seja $A[x]$ o anel de polinômios na variável x sobre A. Denotamos a projeção natural $A[x] \rightarrow I\!K[x]$ por μ. Seja $f(x)$ um polinômio mônico sobre o anel A de grau h, tal que $\mu(f(x))$ seja irredutível em $I\!K[x]$. Então pelo Teorema 4.1.3 temos que $f(x)$ é também irredutível em $A[x]$. Seja $\frac{A[x]}{<f(x)>}$ o conjunto das classes residuais de polinômios em x sobre A módulo $f(x)$. Este anel, denotado por R, é um anel comutativo finito local com identidade cujo ideal maximal é dado por $\overline{\mathcal{M}_1} = \frac{\mathcal{M}_1}{<f(x)>}$, onde $\mathcal{M}_1 =< \mathcal{M}, f(x) >$ e seu corpo residual é dado por $I\!K_1 = \frac{R}{\mathcal{M}_1} = \frac{A[x]/<f(x)>}{<\mathcal{M},f(x)>/<f(x)>} = \frac{A[x]}{<\mathcal{M},f(x)>} = \frac{(A/\mathcal{M})[x]}{<\mu(f(x))>}$, o qual tem ordem p^{mh}. O corpo residual $I\!K_1$ será denotado por $GF(p^{mh})$.

Seja R^* o grupo multiplicativo das unidades de R. Segue que R^* é um grupo abeliano e, portanto, pode ser expresso como um produto direto de grupos cíclicos [37]. Estamos interessados no grupo cíclico de R^*, denotado por G_s, cujos elementos são as raízes de $x^s - 1$ para algum inteiro positivo s. O Teorema 4.1.4 estabelece a existência do grupo G_s, onde $s = p^{mh} - 1$.

Definição 5.1.1 *Seja* $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ *o* **vetor localizador**, *consistindo de elementos distintos de* G_s, *e seja* $\mathbf{w} = (w_1, w_2, \cdots, w_n)$ *um vetor consistindo de elementos de* G_s. *Seja a matriz* H *definida como*

$$H = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \\ w_1\alpha_1 & w_2\alpha_2 & \cdots & w_n\alpha_n \\ w_1\alpha_1^2 & w_2\alpha_2^2 & \cdots & w_n\alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ w_1\alpha_1^{r-1} & w_2\alpha_2^{r-1} & \cdots & w_n\alpha_n^{r-1} \end{bmatrix}, \tag{5.1}$$

onde r *é um inteiro positivo. Esta matriz define um* **código alternante** *de comprimento* $n \leq s$ *sobre* A, *e será denotado por* $C(n, \eta, \mathbf{w})$. ■

Observação 5.1.1 *Os códigos alternantes foram definidos e estudados por Helgert nos trabalhos [23]-[26]. O nome código alternante esta baseado no fato de que a matriz ou determinante da forma*

$$\begin{bmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_{n-1}(x_0) \\ f_0(x_1) & f_1(x_1) & \cdots & f_{n-1}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{r-1}) & f_1(x_{r-1}) & \cdots & f_{n-1}(x_{r-1}) \end{bmatrix}$$

é chamado um **alternante.** ■

A matriz H definida em (5.1) pode ser decomposta como

$$H = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{r-1} & \alpha_2^{r-1} & \cdots & \alpha_n^{r-1} \end{bmatrix} \begin{bmatrix} w_1 & 0 & 0 & \cdots & 0 \\ 0 & w_2 & 0 & \cdots & 0 \\ 0 & 0 & w_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & w_n \end{bmatrix} = XY. \tag{5.2}$$

Uma palavra $\mathbf{c} \in A^n$ está em $\mathcal{C}(n, \eta, \mathbf{w})$ se, e somente se, $\mathbf{c}H^T = \mathbf{0}$. O peso de Hamming de uma palavra $\mathbf{c}$ em A^n será denotado por $w_H(\mathbf{c})$. É possível obter uma estimativa da distância de Hamming mínima, d, diretamente da matriz verificação de paridade. Para isto,

Teorema 5.1.1 *$\mathcal{C}(n, \eta, \mathbf{w})$ tem distância de Hamming mínima $d \geq r + 1$.*

Demonstração: Suponhamos que exista uma palavra $\mathbf{c} \in \mathcal{C}(n, \eta, \mathbf{w})$ tal que $w_H(\mathbf{c}) < r$. Então $\mathbf{c}H^T = \mathbf{c}(XY)^T = \mathbf{0}$. Tomando $\mathbf{b} = \mathbf{c}Y^T$ temos que $w_H(\mathbf{b}) = w_H(\mathbf{c})$ pelo fato de Y ser uma matriz diagonal inversível. Assim, $\mathbf{b}X^T = \mathbf{0}$. Deletando $n - r$ colunas da matriz X que correspondam a zeros da palavra, então a nova matriz X' é Vandermonde. Segue pelo Lema 4.1.1 que este determinante é uma unidade em R. Assim, a única possibilidade para $\mathbf{c}$ é a palavra-código toda nula. Deste modo, temos que $d \geq r + 1$. ■

Exemplo 5.1.1 *Considerando $\eta = (1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6)$, onde α é o gerador do grupo cíclico G_7 do Exemplo 4.1.1 e $\mathbf{w} = (1, 1, 1, 1, 1, 1, 1)$. Então se $r = 2$ a seguinte matriz*

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

é a matriz verificação de paridade de um código alternante $\mathcal{C}(n, \eta, \mathbf{w})$. ■

Exemplo 5.1.2 *Outro exemplo de código alternante é um código BCH. Pela equação (4.1) temos que a matriz verificação de paridade de um código BCH é dada por*

$$H = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{2t} & \alpha_2^{2t} & \cdots & \alpha_n^{2t} \end{bmatrix},$$

onde $\alpha_i = \alpha^{k_i}$, $i = 1, 2, \cdots, n$, *a qual podemos decompor como*

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{k_1} & \alpha^{k_2} & \cdots & \alpha^{k_n} \\ (\alpha^{k_1})^2 & (\alpha^{k_2})^2 & \cdots & (\alpha^{k_n})^2 \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha^{k_1})^{2t-1} & (\alpha^{k_2})^{2t-1} & \cdots & (\alpha^{k_n})^{2t-1} \end{bmatrix} \begin{bmatrix} \alpha^{k_1} & 0 & \cdots & 0 \\ 0 & \alpha^{k_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha^{k_n} \end{bmatrix}.$$

Portanto, um código alternante. ■

5.1.1 Algoritmo de Decodificação

Nesta seção apresentamos um algoritmo de decodificação eficiente para os códigos alternantes $\mathcal{C}(n, \eta, \mathbf{w})$. Este algoritmo, baseado no algorítmo de Berlekamp-Massey modificado, corrige todos os erros de peso de Hamming $t \leq r/2$, isto é, quando a distância de Hamming mínima é maior ou igual a $r + 1$.

Inicialmente estabelecemos algumas notações. Seja $\mathcal{C}(n, \eta, \mathbf{w})$ um código alternante sobre A definido pela equação (5.1). Seja R o anel comutativo finito local com identidade e α um elemento primitivo do grupo cíclico G_s, onde $s = p^{mh} - 1$. Seja $\mathbf{c} = (c_1, c_2, \cdots, c_n)$ a palavra-código transmitida. Seja $\mathbf{b} = (b_1, b_2, \cdots, b_n)$ o vetor recebido, onde o vetor erro é dado por $\mathbf{e} = (e_1, e_2, \cdots, e_n) = \mathbf{b} - \mathbf{c}$.

Dado um vetor localizador $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha^{k_1}, \alpha^{k_2}, \cdots, \alpha^{k_n})$ sobre G_s, definimos os **valores síndromes** s_l de um vetor erro $\mathbf{e} = (e_1, e_2, \cdots, e_n)$ como

$$s_l = \sum_{j=1}^{n} e_j w_j \alpha_j^l, \quad l \geq 0.$$

Suponhamos que $\nu \leq t$ é o número de erros nas localizações $x_1 = \alpha_{i_1}, x_2 = \alpha_{i_2}, \cdots, \cdots, x_\nu = \alpha_{i_\nu}$ com valores $y_1 = e_{i_1}, y_2 = e_{i_2}, \cdots, y_\nu = e_{i_{nu}}$.

Quando $\mathbf{c}$ em $C(n, \eta, \mathbf{w})$ é a palavra-código transmitida, os primeiros r valores síndromes s_l podem ser determinados pelo vetor recebido $\mathbf{b}$ como segue

$$s_l = \sum_{j=1}^{n} e_j w_j \alpha_j^l = \sum_{j=1}^{n} b_j w_j \alpha_j^l, \quad l = 0, 1, \cdots, r-1.$$

O algoritmo sendo proposto consiste dos seguintes quatro passos fundamentais:

Passo 1 - Cálculo do vetor síndrome $\mathbf{s} = \mathbf{b}H^T$, onde $\mathbf{b}$ é o vetor recebido;

Passo 2 - Cálculo das funções simétricas elementares a partir de $\mathbf{s}$;

Passo 3 - Cálculo dos números localizadores de erros $x_1, x_2, \cdots, x_\nu$ a partir de $\sigma_1, \sigma_2, \cdots, \sigma_\nu$;

Passo 4 - Cálculo das magnitudes dos erros $y_1, y_2, \cdots, y_\nu$ a partir dos $x_i's$ e $\mathbf{s}$.

A seguir passamos a caracterizar cada um destes passos.

O primeiro passo, **Passo 1**, do algoritmo de decodificação consiste em encontrar o vetor síndrome $\mathbf{s} = (s_0, s_1, \cdots, s_{r-1})$, o qual é bastante simples.

O conjunto dos números localizadores de erros consiste dos elementos $\alpha^0, \alpha^1, \cdots, \alpha^{s-1}$. As funções simétricas elementares $\sigma_1, \sigma_2, \cdots, \sigma_\nu$ (onde ν denota o número de erros introduzido pelo canal) são definidas como os coeficientes do polinômio

$$(x - x_1)(x - x_2)\cdots(x - x_\nu) = x^\nu + \sigma_1 x^{\nu-1} + \cdots + \sigma_{\nu-1}x + \sigma_\nu.$$

No **Passo 2**, o cálculo das funções simétricas elementares é equivalente a encontrar uma solução $\sigma_1, \sigma_2, \cdots, \sigma_\nu$, com o menor ν possível, a partir do seguinte conjunto de equações lineares recorrentes sobre R

$$s_{j+\nu} + s_{j+\nu-1}\sigma_1 + \cdots + s_{j+1}\sigma_{\nu-1} + s_j\sigma_\nu = 0, \quad j = 0, 1, 2, \cdots, (r-1) - \nu \tag{5.3}$$

onde $s_0, s_1, \cdots, s_{r-1}$ são as componentes do vetor síndrome. Fazemos uso do algoritmo de Berlekamp-Massey modificado para encontrar as soluções de (5.3). Isto é possível pois as provas apresentadas em [33] também valem para anéis commutativos locais com identidade. Neste caso, a discrepância d_n é dada por $d_n = s_n\sigma_0^{(n)} + s_{n-1}\sigma_1^{(n)} + \cdots + s_{n-l_n}\sigma_{l_n}^{(n)}$.

No **Passo 3** a solução em (5.3) geralmente não é única e o recíproco do polinômio $\sigma^{(r)}(z)$ (saída do algoritmo de Berlekamp-Massey modificado), a saber $\rho(z)$, pode não ser o polinômio localizador de erros correto

$$(z - x_1)(z - x_2)\cdots(z - x_\nu)$$

onde $x_i = \alpha^{k_j}$ são os números localizadores de erros corretos (j varia entre $1 \leq j \leq n$ e indica a posição j do erro na palavra-código), ν é o número de erros, e α é o gerador de G_s. Assim, o método para o cálculo dos números localizadores de erros corretos é o seguinte

- Calcule as raízes de $\rho(z)$ (o recíproco de $\sigma^{(r)}(z)$), digamos, $z_1, z_2, \cdots, z_\nu$;
- Entre os $x_i = \alpha^l$, $l = 1, 2, \cdots, s$, selecionamos aqueles $x_i's$ tais que as diferenças $x_i - z_i$ são divisores de zero em R. Os $x_i's$ selecionados serão os números localizadores de erros corretos e cada $x_i = \alpha^{k_j}$, $j = 1, 2, \cdots, n$, indica a posição do erro na palavra-código.

No **Passo 4**, o cálculo das magnitudes dos erros é baseado no método de Forney [15]. As magnitudes dos erros $y_1, y_2, \cdots, y_\nu$, são dadas por

$$y_j = \frac{\sum_{l=0}^{\nu-1} \sigma_{jl} s_{\nu-1-l}}{E_j \sum_{l=0}^{\nu-1} \sigma_{jl} x_j^{\nu-1-l}}, \qquad j = 1, 2, \cdots, \nu \tag{5.4}$$

onde os coeficientes $\sigma_{ji}'s$ são calculados recursivamente por

$$\sigma_{ji} = \sigma_i + x_j\sigma_{j,i-1}, \quad i = 1, 2, \cdots, \nu - 1,$$

iniciando com $\sigma_0 = \sigma_{j0} = 1$. Os $E_j = w_{i_j}$, $j = 1, 2, \cdots, \nu$, $i = 1, 2, \cdots, n$ são as localizações correspondentes dos erros no vetor $\mathbf{w}$. Segue pelo Lema 4.1.1, que o denominador da equação (5.4) é uma unidade em R.

Exemplo 5.1.3 *Seja o grupo G_7 do Exemplo 4.1.1. Considerando $\eta = (\alpha^5, \alpha, \alpha^4, \alpha^2)$, $\mathbf{w} = (\alpha^4, \alpha, \alpha^4, \alpha)$ e $r = 2$ temos que a seguinte matriz*

$$H = \begin{bmatrix} \alpha^4 & \alpha & \alpha^4 & \alpha \\ \alpha^2 & \alpha^2 & \alpha & \alpha^3 \end{bmatrix}$$

define um código alternante de comprimento 4 e distância de Hamming mínima maior ou igual a 3. Este código tem capacidade de correção de erro de Hamming de peso no máximo 1. Supondo que a palavra-código toda nula $\mathbf{c} = (0000)$ *seja transmitida, e o vetor* $\mathbf{b} = (00i0)$ *recebido. Então o vetor síndrome é* $\mathbf{s} = (i\alpha^4, i\alpha)$. *Aplicando o algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	$i\alpha^4$	0	0
1	$1 + i\alpha^4 z$	$\alpha(i+1)$	1	0
2	$1 + \alpha^4 z$	-	1	1

A raiz de $\rho(z) = z + \alpha^4$ *(o recíproco de* $\sigma^{(2)}(z)$*) é* α^4. *Entre os elementos* $\alpha^0, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5$ *e* α^6 *temos que* $x_1 = \alpha^4$ *é tal que* $x_1 - z_1 = 0$ *(divisor de zero em* R*). Portanto,* x_1 *é o número localizador de erro correto e como* $k_3 = 4$ *temos que um erro ocorreu na posição 3 da palavra-código toda nula. A função simétrica elementar* σ_1 *é obtida de* $x - x_1 = x - \sigma_1 = x - \alpha^4$. *Finalmente, aplicando o método de Forney a* $\mathbf{s}$ *e* $\sigma_1 = \alpha^4$ *obtemos* $y_1 = i$. *Portanto, o padrão de erro é* $\mathbf{e} = (00i0)$. ■

Exemplo 5.1.4 *Seja o grupo cíclico* G_{15} *do Exemplo 4.3.2. Considerando o vetor localizador* $\eta = (\alpha^{14}, 1, \alpha^5, \alpha^3, \alpha^{10}, \alpha^{11}, \alpha, \alpha^9, \alpha^{13})$, $\mathbf{w} = (\alpha, \alpha^2, \alpha, \alpha^3, \alpha^2, \alpha^9, \alpha, \alpha^{10}, \alpha^5)$ *e* $r = 4$ *temos que a matriz* H *definida por*

$$H = \begin{bmatrix} \alpha^{14} & \alpha^2 & \alpha & \alpha^3 & \alpha^2 & \alpha^9 & \alpha & \alpha^{10} & \alpha^5 \\ 1 & \alpha^2 & \alpha^6 & \alpha^6 & \alpha^{12} & \alpha^5 & \alpha^2 & \alpha^4 & \alpha^3 \\ \alpha^{14} & \alpha^2 & \alpha^{11} & \alpha^9 & \alpha^7 & \alpha & \alpha^3 & \alpha^{13} & \alpha \\ \alpha^{13} & \alpha^2 & \alpha & \alpha^{12} & \alpha^2 & \alpha^{12} & \alpha^4 & \alpha^7 & \alpha^{14} \end{bmatrix}$$

define um código alternante de comprimento 9 e distância de Hamming miníma maior ou igual a 5. Este código tem capacidade de correção de erros de Hamming de peso até $t = 2$. *Temos também que* $k_1 = 14, k_2 = 0, k_3 = 5, k_4 = 3, k_5 = 10, k_6 = 11, k_7 = 1, k_8 = 9$ *e* $k_9 = 13$. *Supondo que a palavra toda nula* $\mathbf{c} = (000000000)$ *seja transmitida, e o vetor* $\mathbf{b} = (00001000i)$ *é recebido. Então o vetor síndrome é* $\mathbf{s} = (\alpha^2 + i\alpha^5, \alpha^{12} + i\alpha^3, \alpha^7 + i\alpha, \alpha^2 + i\alpha^{14})$. *Aplicando o algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	$\alpha^2 + i\alpha^5$	0	0
1	$1 + (\alpha^2 + i\alpha^5)z$	$\alpha^7 + i\alpha^3$	1	0
2	$1 + (\alpha^4 + i\alpha^{14})z$	$\alpha^{13} + i\alpha^{10}$	1	1
3	$1 + (\alpha^4 + i\alpha^{14})z + i\alpha^8 z^2$	$\alpha^8(1 + i)$	2	1
4	$1 + \alpha^9 z + \alpha^8 z^2$	-	2	2

Assim, o polinômio $\rho(z) = z^{\nu}\sigma^{(4)}(z^{-1}) = z^2 + \alpha^9 z + \alpha^8$ *possui como raízes* $z_1 = \alpha^{10}$ *e* $z_2 = \alpha^{13}$. *Entre os elementos* $1, \alpha, \alpha^2, \cdots, \alpha^{14}$ *temos que* $x_1 = \alpha^{10}$ *e* $x_2 = \alpha^{13}$ *são tais que* $z_1 - x_1$ *e* $z_2 - x_2$ *são divisores de zero em* R. *Portanto,* x_1 *e* x_2 *são os números localizadores de erros corretos e indicam que dois erros ocorreram. Como* $k_5 = 10$ *e* $k_9 = 13$ *temos que os erros ocorreram nas posições 5 e 9, respectivamente, da palavra-código toda nula. Finalmente, aplicando o método de Forney a* $\sigma_1 = \alpha^9$ *e* $\sigma_2 = \alpha^8$ *obtemos as magnitudes dos erros* $y_1 = 1$ *e* $y_2 = i$. *Deste modo, concluímos que ocorreram dois erros. Um erro ocorreu na posição 5 da palavra-código toda nula com magnitude 1 e o outro na posição 9 com magnitude* i. ■

5.2 Códigos BCH sobre $\mathbb{Z}_q$ para a Métrica de Lee

Nesta seção apresentamos uma técnica de construção de códigos BCH sobre anéis de inteiros residuais $\mathbb{Z}_q$, onde q é uma potência de um primo. Baseado no algoritmo de Berlekamp-Massey modificado e na proposta em [42], apresentamos um algoritmo de decodificação alternativo para correção de erros de Lee garantido pela distância de Lee mínima projetada. Para isto,

Definição 5.2.1 *Seja* $\mathbb{Z}_m$ *o anel dos inteiros módulo* m, *onde* m *é um inteiro positivo, isto é,* $\mathbb{Z}_m = \{0, 1, 2, \cdots, m-1\}$.

- *Para um elemento* $\alpha \in \mathbb{Z}_m$, *definimos o* **valor de Lee** $|\ \alpha\ |$ *por*

$$|\ \alpha\ | = \begin{cases} \alpha & se \quad 0 \leq \alpha \leq [\frac{m}{2}] \\ m - \alpha & se \quad [\frac{m}{2}] < \alpha \leq m-1 \end{cases}$$

  onde $[\frac{m}{2}]$ *significa o maior inteiro menor ou igual* $\frac{m}{2}$.

- *A* **distância de Lee** *entre dois vetores* $\mathbf{u} = (a_1, a_2, \cdots, a_n)$ *e* $\mathbf{v} = (b_1, b_2, \cdots, b_n)$ *sobre* $\mathbb{Z}_m^n$ *é definida por*

$$d_L(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} d_L(a_i, b_i),$$

  onde $d_L(a_i, b_i) = min\{a_i - b_i, b_i - a_i\}\ (mod\ m)$, $i = 1, \cdots, n$. *O* **peso de Lee** *de um vetor* $\mathbf{u} = (a_1, a_2, \cdots, a_n)$ *sobre* $\mathbb{Z}_m^n$ *é definido como*

$$w_L(\mathbf{u}) = d_L(\mathbf{u}, \mathbf{0}) = \sum_{i=1}^{n} |\ a_i\ |\ .$$

- *A* **distância de Lee mínima**, $d_L(X)$, *de um subconjunto* X *de* $\mathbb{Z}_m^n$ *é a menor distância de Lee entre quaisquer dois vetores distintos em* X. ■

Observação 5.2.1

- *Os elementos* $0, 1, 2, \cdots, [\frac{m}{2}]$ *de* $\mathbb{Z}_m$ *são definidos como os* **elementos positivos** *do anel para o qual* $|\ \alpha\ | = \alpha$. *Os demais elementos são chamados* **negativos**.

- *Se* X *for um subgrupo aditivo de* $\mathbb{Z}_m^n$, *sua distância de Lee mínima é o peso de Lee mínimo de qualquer vetor não nulo de* X *(a diferença de quaisquer dois vetores em* X *é também um vetor em* X*). Assim, a distância de Lee mínima de um código linear é igual ao peso de Lee mínimo de suas palavras-código não nulas.*

- *A distância de Lee define uma métrica sobre* $\mathbb{Z}_m$.

- *Para* $m = 2$ *ou 3, as distâncias de Hamming e a de Lee coincidem. Para* $m > 3$, *a distância de Lee entre dois vetores é maior ou igual a distância de Hamming entre os mesmos. Assim, a distância de Lee mínima de um código é maior ou igual a distância de Hamming mínima do mesmo código, e menor ou igual a distância de Lee entre os dois vetores que definem a distância de Hamming mínima. Portanto,* $d_H \leq d_L \leq [\frac{m}{2}] d_H$.

- *O problema de escolher uma métrica para um dado canal ou encontrar um canal com uma dada métrica depende do esquema de decodificação.* ■

Seja $\mathbb{Z}_q[x]$ o anel de polinômios na variável x sobre $\mathbb{Z}_q$, onde q é uma potência de um primo. Seja $f(x)$ um polinômio mônico e irredutível sobre $\mathbb{Z}_p$ de grau h e que, consequentemente, é irredutível sobre $\mathbb{Z}_q$. Seja R o anel quociente $\frac{\mathbb{Z}_q[x]}{<f(x)>}$ o qual é formado pelas classes residuais de polinômios em $\mathbb{Z}_q[x]$ módulo $f(x)$. Este anel consiste de todos os polinômios de grau menor ou igual a $h-1$ cujas operações de adição e multiplicação são sempre tomadas módulo $f(x)$. Além disso, R é um anel comutativo finito local com identidade e é chamado uma extensão de Galois de grau h de $\mathbb{Z}_q$. Similarmente à Seção 5.1, temos que R^* é um grupo multiplicativo de elementos de R, o qual contém um subgrupo cíclico G_s, cuja ordem é $s = p^h - 1$.

Definição 5.2.2 *Um* **código BCH** *de comprimento reduzido $n \leq s$ sobre $\mathbb{Z}_q$ é definido pela seguinte matriz verificação de paridade*

$$H = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{r-1} & \alpha_2^{r-1} & \cdots & \alpha_n^{r-1} \end{bmatrix} \tag{5.5}$$

onde r é um inteiro positivo, e $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ é o **vetor localizador**, *consistindo de elementos distintos de G_s. O código será denotado por $\mathcal{C}(n, \eta)$.* ■

Pela Definição 5.2.2 temos que uma palavra $\mathbf{c} = (c_1, c_2, \cdots, c_n) \in \mathbb{Z}_q^n$ pertence ao código se, e somente se, satisfaz as r equações de verificação de paridade

$$\sum_{j=1}^{n} c_j \alpha_j^l = 0, \quad l = 0, 1, \cdots, r-1 \tag{5.6}$$

Para $l \geq 1$, temos que para cada equação de verificação de paridade em (5.6) obtemos h equações sobre $\mathbb{Z}_q$.

Observação 5.2.2 *Um código para o qual $n = p^h - 1$, será chamado* **primitivo**. *Neste caso, η é único, a menos de permutação de suas coordenadas. Por outro lado, quando $h = 1$ temos que o*

código correspondente é **Reed-Solomon** *de comprimento* $n \leq s$ *sobre* $\mathbb{Z}_q$. *Neste caso, os* α_j*'s são unidades distintas de* $\mathbb{Z}_q$. ■

Dadas uma palavra transmitida $\mathbf{c}$ e uma palavra recebida $\mathbf{b} \in \mathbb{Z}_q^n$, o vetor erro é definido como $\mathbf{e} = \mathbf{b} - \mathbf{c}$. **O número de erros de Lee** é dado por $w_L(\mathbf{e})$, isto é, o número de erros de Lee é o menor número de adições de ± 1 nas coordenadas da palavra transmitida $\mathbf{c}$ que produz a palavra recebida $\mathbf{b}$. Como o peso de Lee satisfaz a desigualdade do triângulo, usando um código de distância de Lee mínima d_L podemos corrigir todos os padrões de erros de peso de Lee até $(d_L - 1)/2$ (Teorema 2.3.1).

Definição 5.2.3 *Dados um vetor localizador* $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ *de um código* $\mathcal{C}(n, \eta)$ *e um vetor* $\mathbf{c} = (c_1, c_2, \cdots, c_n) \in \mathbb{Z}_q^n$, *definimos o* **polinômio localizador** *associado com* $\mathbf{c}$ *como o polinômio* $\sigma(x)$ *sobre* R *dado por*

$$\sigma(x) = \prod_{j=1}^{n} (1 - \alpha_j x)^{w_L(c_j)}.$$

Também definimos os **valores síndromes** s_l *de um vetor erro* $\mathbf{e} = (e_1, e_2, \cdots, e_n)$ *como*

$$s_l = \sum_{j=1}^{n} e_j \alpha_j^l,$$

para todo $l \geq 0$. *A série* $s(x)$ *definida por*

$$s(x) = \sum_{j=1}^{\infty} s_l x^l,$$

é a **série síndrome formal**. ■

Dada uma palavra $\mathbf{c} \in \mathcal{C}(n, \eta)$, definimos a palavra $\mathbf{c}^+ = (c_1^+, c_2^+, \cdots, c_n^+)$ por

$$c_j^+ = \begin{cases} c_j & \text{se} \quad c_j \in \{1, 2, \cdots, [\frac{q}{2}]\} \\ 0 & \text{caso contrário} \end{cases}$$

e $\mathbf{c}^-$ por $\mathbf{c}^- = \mathbf{c}^+ - \mathbf{c}$. Isto é, $\mathbf{c}^+$ é igual a $\mathbf{c}$ nas entradas positivas e zero nas demais, ao passo que as entradas de $\mathbf{c}^-$ tomam em consideração os valores de Lee das entradas negativas de $\mathbf{c}$, e zero nas demais. Sejam $\sigma^+(x)$ e $\sigma^-(x)$ os polinômios localizadores de $\mathbf{c}^+$ e $\mathbf{c}^-$, respectivamente. Sejam $s^+(x) = \sum_{l=1}^{\infty} s_l^+ x^l$ e $s^-(x) = \sum_{l=1}^{\infty} s_l^- x^l$ as séries síndromes formais sobre R associadas com $\sigma^+(x)$ e $\sigma^-(x)$, respectivamente.

Sendo $\mathbf{c} \in C(n,\eta)$ uma palavra-código, temos que $\mathbf{c}H^T = \mathbf{0}$, isto é, $\mathbf{c}^+H^T = \mathbf{c}^-H^T$. A primeira equação desta última igualdade fornece que $w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-)\ (mod\ q)$, ou equivalentemente, $w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-) \pm lq$, para algum inteiro positivo l, ao passo que as demais $r-1$ equações podem ser reescritas como $s_l^+ = s_l^-$, $l = 1, 2, \cdots, r-1$, ou equivalentemente, $s^+(x) \equiv s^-(x)\ (mod\ x^r)$.

Lema 5.2.1 *Se* $w_L(\mathbf{c}^+) \neq w_L(\mathbf{c}^-)$ *então* $w_L(\mathbf{c}) \geq q$.

Demonstração: Como $w_L(\mathbf{c}^+) \neq w_L(\mathbf{c}^-)$ temos que $w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-) \pm lq$ para algum inteiro positivo l. Portanto, $w_L(\mathbf{c}) = w_L(\mathbf{c}^+) + w_L(\mathbf{c}^-) \geq q$, o que prova o lema. ■

Através do Lema 5.2.1 obtemos um limitante inferior para a distância de Lee mínima quando $w_l(\mathbf{c}^+) \neq w_L(\mathbf{c}^-)$. No próximo teorema mostramos que para qualquer valor de r, o limitante inferior pode ser melhorado. Sua demonstração é bastante análoga à apresentada em [42]. Contudo, para efeito de referências futuras e também para uma maior clareza deste trabalho, apresentamos a demonstração.

Teorema 5.2.1 *Seja* $C(n,\eta)$ *um código com distância de Lee mínima* $d_L(C)$. *Então o limitante inferior da distância de Lee mínima é* $d_L(C) \geq \begin{cases} 2r & se \quad 1 \leq r \leq \left[\frac{q}{2}\right] \\ q & se \quad \left[\frac{q}{2}\right] < r \leq q-1 \end{cases}$

Demonstração: Do Lema 5.2.1, resta apenas mostrar que se $\mathbf{c} \in C(n,\eta)$ é tal que $w_L(\mathbf{c}) < 2r$ e $w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-)$ então $\mathbf{c} = \mathbf{0}$. Deste modo, assumindo que $w_L(\mathbf{c}) < 2r$ e $w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-)$ obtemos $w_L(\mathbf{c}) = 2w_L(\mathbf{c}^+) = 2w_L(\mathbf{c}^-)$. Portanto, $grau(\sigma^+(x)) = w_L(\mathbf{c}^+) = w_L(\mathbf{c}^-) = grau(\sigma^-(x)) \leq r-1$. Agora, usando a congruência $s^+(x) \equiv s^-(x)\ (mod\ x^r)$ e o algoritmo para a decodificação de códigos Reed-Solomon e BCH, dado na Seção 4.3, obtemos $\sigma^+(x) = \sigma^-(x)$. Contudo, como os suportes (o conjunto das coordenadas não nulas) de $\mathbf{c}^+$ e $\mathbf{c}^-$, são disjuntos, os polinômios $\sigma^+(x)$ e $\sigma^-(x)$ são relativamente primos. Portanto, devemos ter $\sigma^+(x) = \sigma^-(x) = 1$, acarretando que a palavra-código deva ser a palavra toda nula, ou seja, $\mathbf{c} = \mathbf{0}$. ■

5.2.1 Algoritmo de Decodificação

Nesta seção apresentamos um processo alternativo de decodificação dos códigos BCH $C(n,\eta)$ quando a métrica de Lee é considerada. Para um código com distância de Lee mínima d_L que corrige até $t = \left[\frac{d_L-1}{2}\right]$ erros de Lee, o método requer o conhecimento de $2r-2$ síndromes, onde

as primeiras r síndromes são conhecidas e as restantes são não conhecidas. A decodificação é então acompanhada pela determinação do polinômio localizador de erros através do algoritmo de Berlekamp-Massey modificado que corrige todos os erros de Lee até $r-1$, garantido pela distância de Lee mínima projetada $2r$ (se r varia entre 1 e $[\frac{q}{2}]$) e que também corrige todos os erros Lee até $[\frac{q}{2}]$ garantido pela distância de Lee mínima projetada q (se $[\frac{q}{2}] < r \leq q-1$).

Seja $\mathcal{C}(n,\eta)$ um código BCH definido sobre $\mathbb{Z}_q$, onde q é uma potência de um primo p, com matriz verificação de paridade H dado pela equação (5.5) e com distância de Lee mínima $d_L(C)$.

Sendo G_s um grupo cíclico multiplicativo com gerador α, temos que o vetor localizador η é dado por $\eta = (\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha^{k_1}, \alpha^{k_2}, \cdots, \alpha^{k_n})$, onde k_i, $i = 1, 2, \cdots, n$ são inteiros positivos.

Seja $\mathbf{c} = (c_1, c_2, \cdots, c_n)$ a palavra-código transmitida e $\mathbf{b} = (b_1, b_2, \cdots, b_n)$ a palavra recebida. O vetor erro é dado por $\mathbf{e} = (e_1, e_2, \cdots, e_n) = \mathbf{b} - \mathbf{c}$. Os valores das síndromes são dados por $s_l = \sum_{i=1}^{n} b_i \alpha_i^l$, $l \geq 0$. Suponha que $\nu \leq t$ é o número de erros nas localizações $x_1 = \alpha_{i_1}, x_2 = \alpha_{i_2}, \cdots, x_\nu = \alpha_{i_\nu}$ com valores $y_1 = e_{i_1}, y_2 = e_{i_2}, \cdots, y_\nu = e_{i_\nu}$.

O processo de decodificação sendo proposto (muito similar ao dos códigos alternantes) consiste dos seguintes quatro passos fundamentais:

Passo 1 - Cálculo dos valores síndromes (conhecidos e não conhecidos) $s_l = \sum_{i=1}^{n} b_i \alpha_i^l$, $l = 0, 1, \cdots, 2r-3$ se $1 \leq r \leq [\frac{q}{2}]$ ou $l = 0, 1, \cdots, 2[\frac{q}{2}] - 1$ se $[\frac{q}{2}] < r \leq q-1$, do vetor recebido;

Passo 2 - Cálculo das funções simétricas elementares $\sigma_1, \sigma_2, \cdots, \sigma_\nu$, a partir da aplicação do algoritmo de Berlekamp-Massey modificado;

Passo 3 - Cálculo dos números localizadores de erros $x_1, x_2, \cdots, x_\nu$ a partir das funções simétricas elementares $\sigma_1, \sigma_2, \cdots, \sigma_\nu$. O processo para o cálculo dos números localizadores de erros é dado por:

- Calcule as raízes de $\rho(z)$ (o recíproco de $\sigma^{(r)}(z)$), digamos, $z_1, z_2, \cdots, z_\nu$, onde o polinômio $\sigma^{(r)}(z)$ é a saída do algoritmo de Berlekamp-Massey modificado;
- Entre os $x_i = \alpha^{k_j}$, $j = 1, 2, \cdots, n$, selecione aqueles $x_i's$ tais que $z_i - x_i$ são divisores de zero em R. Os $x_i's$ selecionados são os números localizadores de erros corretos e cada $x_i = \alpha^{k_j}$, $j = 1, 2, \cdots, n$, indica a posição j do erro na palavra-código.

Passo 4 - Cálculo das magnitudes dos erros $y_1, y_2, \cdots, y_\nu$ a partir dos $x_i's$, $i = 1, 2, \cdots, \nu$ e s_l através do método de Forney.

A seguir, passaremos a caracterizar cada um destes passos.

O **Passo 1** do algoritmo de decodificação tem como objetivo encontrar os valores das síndromes, isto é, encontrar $s_l = \sum_{i=1}^{n} b_i \alpha_i^l$, $l = 0, 1, \cdots, 2r-3$, quando $1 \leq r \leq [\frac{q}{2}]$ ou $l = 0, 1, \cdots, 2[\frac{q}{2}] - 1$, quando $[\frac{q}{2}] < r \leq q-1$.

O **Passo 2** faz uso dos valores das síndromes s_l, $l \geq 0$, para obter o polinômio $\sigma(z) = \prod_{i=1}^{\nu}(1 - x_i z) = \sum_{i=0}^{\nu} \sigma_i z^i$, $\sigma_0 = 1$, através do algoritmo de Berlekamp-Massey modificado.

No **Passo 3** os números localizadores de erros são encontrados calculando as raízes recíprocas de $\sigma^{(r)}(z)$, ou seja, calculando as raízes $z_1, z_2, \cdots, z_\nu$ do polinômio $\rho(z) = z^\nu \sigma^{(r)}(z^{-1})$ e em seguida selecionar aqueles $x_i = \alpha^i$, $i = 0, 1, \cdots, n-1$ que tornam as diferenças $x_i - z_j$, $j = 1, 2, \cdots, \nu$, elementos divisores de zero em R. Deste modo, temos que os $x_i's$ selecionados são os números localizadores de erros corretos e que $x_i = \alpha^{k_j}$, $j = 1, 2, \cdots, n$, significa que ocorreu um erro na posição j da palavra-código.

Uma vez que os localizadores de erros x_i, $i = 1, 2, \cdots, \nu$, são conhecidos, o **Passo 4** faz uso do procedimento proposto por Forney para determinar as magnitudes y_i, $i = 1, 2, \cdots, \nu$ dos erros. Estas magnitudes são calculadas como os quocientes

$$y_j = \frac{\sum_{l=0}^{\nu-1} \sigma_{jl} s_{\nu-1-l}}{\sum_{l=0}^{\nu-1} \sigma_{jl} x_j^{\nu-1-l}}, \qquad j = 1, 2, \cdots, \nu, \tag{5.7}$$

onde os coeficientes $\sigma_{ji}'s$ são calculados recursivamente por $\sigma_{ji} = \sigma_i + x_j \sigma_{j,i-1}$ para $i = 1, 2, \cdots, \nu-1$, iniciando com $\sigma_0 = \sigma_{j0} = 1$. Segue do Lema 4.1.1 que o denominador da equação (5.7) é uma unidade no anel R.

O algoritmo de decodificação dos códigos BCH definidos sobre anéis de inteiros residuais módulo q mediante a métrica de Lee fica então estabelecido.

Para finalizar esta seção apresentamos dois exemplos de decodificação de códigos BCH através da métrica de Lee, sendo que o primeiro é projetado para a correção de erros de Lee de peso 1 e o segundo de erros de Lee de peso até 3.

Exemplo 5.2.1 *Seja o anel* $R = \frac{\mathbb{Z}_4[x]}{<f(x)>}$, *onde* $f(x) = x^3 + x + 1$ *é irredutível sobre* $\mathbb{Z}_2$ *e portanto sobre* $\mathbb{Z}_4$. *Temos que* $\alpha = \beta^2$, *com* β *uma raiz de* $f(x)$, *é o gerador do grupo cíclico de ordem* $s = 2^3 - 1 = 7$. *Tomando* $r = 2$ *e* $\eta = (\alpha^5, \alpha, 1, \alpha^3, \alpha^6, \alpha^2, \alpha^4)$ *temos que a seguinte matriz*

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

define um código BCH de comprimento 7. Este código corrige todos os erros de Lee de peso 1, alguns erros de peso 2 e detecta todos os erros de peso 2. Assim, supondo que a palavra toda nula $\mathbf{c} = (0000000)$ *seja transmitida através do canal, e que o mesmo introduza o vetor erro* $\mathbf{e} = (0300000)$, *temos que o vetor recebido é então dado por* $\mathbf{b} = \mathbf{c} + \mathbf{e}$, *cujos valores das síndromes são* $s_0 = 3$ *e* $s_1 = 3\alpha$. *Aplicando o algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	3	0	0
1	$1 - 3z$	$3\alpha - 1$	1	0
2	$1 - \alpha z$	$-$	1	1

Assim, $\rho(z) = z - \alpha$, *cuja raiz é* $z_1 = \alpha$. *Como* $\alpha - \alpha = 0$ *é um divisor de zero em* R *e que* $\alpha_2 = \alpha$, *segue que o erro ocorreu na posição 2 da palavra-código toda nula. Finalmente, aplicando o procedimento proposto por Forney a* $\sigma_1 = \alpha$ *temos que a magnitude do erro é dado por* $y_1 = 3$. *Assim, o padrão de erro é dado por* $\mathbf{e} = (0300000)$ *e o polinômio localizador é dado por* $\sigma(x) = 1 - \alpha x$.

Por outro lado, supondo que a palavra toda nula $\mathbf{c} = (0000000)$ *seja transmitida através do canal, e que o mesmo introduza o seguinte vetor erro* $\mathbf{e} = (0101000)$. *Deste modo o vetor recebido é dado por* $\mathbf{b} = \mathbf{c} + \mathbf{e}$, *cujos valores das síndromes são* $s_0 = 2$ *e* $s_1 = \alpha + \alpha^3$. *Aplicando o algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n-l_n$
-1	1	1	0	-1
0	1	2	0	0
1	$1+2z$	$\alpha+\alpha^3$	1	0
2	$1+2z-(\alpha+\alpha^3)z^2$	$-$	2	0

Assim, $\rho(z) = z^2 + 2z - (\alpha + \alpha^3)$ *é de grau dois, indicando que ocorreu um erro cujo peso de Lee é pelo menos dois. Todavia, não é possível corrigi-lo pelo fato de que este polinômio é irredutível sobre* R. *Portanto, apenas detecta.*

Por outro lado, suponha que a palavra toda nula $\mathbf{c} = (0000000)$ *é transmitida através do canal e que o mesmo introduza o seguinte vetor erro* $\mathbf{e} = (0200000)$. *Deste modo, o vetor recebido é dado por* $\mathbf{b} = \mathbf{c} + \mathbf{e}$, *cujos valores das síndromes são* $s_0 = 2$ *e* $s_1 = 2\alpha$. *Aplicando o algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n-l_n$
-1	1	1	0	-1
0	1	2	0	0
1	$1+2z$	2α	1	0
2	$1+(2-\alpha)z$	$-$	1	1

Assim, $\rho(z) = z - (2 - \alpha)$, *cuja raiz é* $z_1 = 2 + \alpha$. *Como* $(2+\alpha) - \alpha = 2$ *é um divisor de zero em* R *e que* $\alpha_2 = \alpha$, *segue que o erro ocorreu na posição 2 da palavra-código toda nula. Finalmente, aplicando o procedimento proposto por Forney a* $\sigma_1 = \alpha$ *temos que a magnitude do erro é dado por* $y_1 = 2$. *Assim, o padrão de erro é dado por* $\mathbf{e} = (0200000)$ *e o polinômio localizador é dado por* $\sigma(x) = (1 - (2 + \alpha)x)^2$. ■

Exemplo 5.2.2 *Seja* R *o anel* $\frac{\mathbb{Z}_9[x]}{<f(x)>}$, *onde* $f(x) = x^2 + 1$ *é irredutível sobre* $\mathbb{Z}_9$. *Temos que* α $(= 2\beta + 2$, *com* β *uma raiz de* $f(x))$, *é o gerador do grupo cíclico de ordem* $s = 3^2 - 1 = 8$. *Considerando* $r = 4$ *e* $\eta = (\alpha^2, \alpha, 1, \alpha^3, \alpha^6, \alpha^5, \alpha^4, \alpha^7)$ *temos que a seguinte matriz*

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \alpha^2 & \alpha & 1 & \alpha^3 & \alpha^6 & \alpha^5 & \alpha^4 & \alpha^7 \\ \alpha^4 & \alpha^2 & 1 & \alpha^6 & \alpha^4 & \alpha^2 & 1 & \alpha^6 \\ \alpha^6 & \alpha^3 & 1 & \alpha & \alpha^2 & \alpha^7 & \alpha^4 & \alpha^5 \end{bmatrix}$$

define um código BCH de comprimento 8. Este código corrige todos os erros de Lee de peso até 3, corrige alguns erros de Lee de peso 4 e detecta todos os erros de Lee de peso 4. Assim, supondo que a palavra toda nula $\mathbf{c} = (00000000)$ *é transmitida e o vetor* $\mathbf{b} = (01010800)$ *é recebido. Então os valores das síndromes são* $s_0 = 1$, $s_1 = 2\alpha + 6$, $s_2 = \alpha$, $s_3 = 7\alpha + 6$, $s_4 = -1$ *e* $s_5 = 7\alpha + 3$. *Através do algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	1	0	0
1	$1 - z$	$2\beta + 5$	1	0
2	$1 - (2\beta + 6)z$	$4\beta + 4$	1	1
3	$1 - (-\beta + 2)z + (5 + 6\beta)z^2$	$6\beta + 5$	2	1
4	$1 - 7\beta z + (4 + 5\beta)z^2$	$7\beta + 7$	2	2
5	$1 - (2 + 6\beta)z + (\beta + 7)z^2 - (7\beta - 2)z^3$	$7\beta + 2$	3	2
6	$1 - (2 + 7\beta)z + \beta z^2 - (2\beta - 7)z^3$	-	3	3

As raízes de $\rho(z) = z^3 - (2 + 7\beta)z^2 + \beta z + 7\beta + 7$ *(o polinômio recíproco de* $\sigma^{(4)}(z)$*) são* $z_1 = 2\beta + 2$, $z_2 = 7\beta + 2$ *e* $z_3 = 7\beta + 7$. *Entre os elementos* $\alpha^0, \alpha^1, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6$ *e* α^7 *temos que as diferenças* $\alpha - z_1 = \alpha^3 - z_2 = \alpha^5 - z_3 = 0$ *são divisores de zero em* R. *Portanto,* $x_1 = \alpha_2 = \alpha$, $x_2 = \alpha_4 = \alpha^3$ *e* $x_3 = \alpha_6 = \alpha^5$, *são os números localizadores de erros corretos e indica que ocorreu um erro de Lee de peso no mínimo 3, nas posições 2, 4 e 6 da palavra toda nula, respectivamente. As funções simétricas elementares corretas* σ_1, σ_2 *e* σ_3 *são obtidas de*

$$(x - x_1)(x - x_2)(x - x_3) = x^3 + \sigma_1 x^2 + \sigma_2 x + \sigma_3 = x^3 + (7 + 2\beta)x^2 + \beta x + 7\beta + \beta.$$

Finalmente, através do método de Forney aplicado a $\sigma_1 = 7 + 2\beta$, $\sigma_2 = \beta$ *e* $\sigma_3 = 7\beta + 7$ *temos que as magnitudes dos erros são dadas por* $y_1 = 1$, $y_2 = 1$ *e* $y_3 = 8$. *Assim, o padrão de erro é dado por* $\mathbf{e} = (01010800)$ *e o polinômio localizador é dado por* $\sigma(x) = (1 - \alpha x)(1 - \alpha^3 x)(1 - \alpha^5 x)$.

Por outro lado, supondo que a palavra toda nula $\mathbf{c} = (00000000)$ *é transmitida através do canal e que o mesmo introduza o seguinte vetor erro de peso de Lee 4,* $\mathbf{e} = (02020000)$, *temos que o vetor recebido é dado por* $\mathbf{b} = \mathbf{c} + \mathbf{e}$, *cujos valores das síndromes são* $s_0 = 4$, $s_1 = 8$, $s_2 = 0$, $s_3 = 8$, $s_4 = -4$ *e* $s_5 = 1$. *Aplicando o algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n - l_n$
-1	1	1	0	-1
0	1	4	0	0
1	$1 - 42z$	1	1	0
2	$1 - 2z$	2	1	1
3	$1 - 4z + 8z^2$	0	2	1
4	$1 - 4z + 8z^2$	0	2	2
5	$1 - 4z + 8z^2$	0	2	3
6	$1 - 4z + 8z^2$	-	2	4

Assim, $\rho(z) = z^2 - 4z + 8$, *cujas raízes são* $z_1 = \alpha$ *e* $z_2 = \alpha^3$. *Como as diferenças* $\alpha - z_1 = \alpha^3 - z_2 = 0$ *são divisores de zero em* R *e que* $\alpha_2 = \alpha$ *e* $\alpha_4 = \alpha^3$, *segue que os erros ocorreram nas posições 2 e 4, respectivamente, da palavra toda nula. Finalmente, aplicando o procedimento proposto por Forney a* $\sigma_1 = 5$ *e* $\sigma_2 = 8$ *temos que as magnitudes dos erros são dadas por* $y_1 = y_2 = 2$. *Assim, o padrão de erro é dado por* $\mathbf{e} = (02020000)$ *e o polinômio localizador é dado por* $\sigma(x) = (1 - \alpha x)^2(1 - \alpha^3 x)^2$.

Agora, suponha que a palavra toda nula $\mathbf{c} = (00000000)$ *seja transmitida através do canal e que o mesmo introduza o seguinte vetor erro de peso de Lee 4* $\mathbf{e} = (01010101)$ *temos que o vetor recebido é dado por* $\mathbf{b} = \mathbf{c} + \mathbf{e} = \mathbf{e}$, *cujos valores das síndromes são* $s_0 = 4$, $s_1 = s_2 = s_3 = 0$, $s_4 = -4$ *e* $s_5 = 0$. *Aplicando o algoritmo de Berlekamp-Massey modificado obtemos a seguinte tabela*

n	$\sigma^{(n)}(z)$	d_n	l_n	$n-l_n$
-1	1	1	0	-1
0	1	4	0	0
1	$1-4z$	2	1	0
2	1	0	1	1
3	1	0	1	2
4	1	-4	1	3
5	$1+2z^3-8z^4$	0	4	1
6	$1+2z^3-8z^4$	-	4	2

Assim, $\rho(z) = z^4+2z+1 = (z-8)(z^3+8z^2+z+1)$, *onde o polinômio* x^3+8z^2+z+1 *é irredutível sobre o anel* R. *Portanto, neste caso, detectamos um padrão de erro de peso de Lee maior ou igual a 4, mas através deste método não é possível corrigi-lo.* ■

5.3 Conclusões

Neste capítulo foram apresentadas técnicas de construções de códigos alternantes e BCH. Na primeira técnica de construção, códigos alternantes sobre anéis comutativos finitos locais sob a métrica de Hamming foi apresentado. Um processo de decodificação eficiente baseado no algoritmo de Berlekamp-Massey modificado foi fornecido. Na segunda técnica de construção, códigos BCH sobre anéis de inteiros residuais do tipo $\mathbb{Z}_q$, onde q é uma potência de um primo, sob a métrica de Lee foi proposto. Um processo de decodificação alternativo também baseado no algoritmo de Berlekamp-Massey modificado foi fornecido.

Capítulo 6

Conclusões e Sugestões

6.1 Conclusões

Neste trabalho foram estudados e apresentados métodos de construção e de decodificação de códigos de bloco lineares sobre anéis finitos. A principal motivação do presente trabalho é que códigos lineares sobre anéis finitos tem recentemente despertado um grande interesse pela sua nova função em teoria de codificação algébrica e pelas suas bem sucedidas aplicações em modulação e codificação. Como exemplo temos o trabalho de Hammons et al. [22], onde mostram que certos códigos não lineares binários do tipo Kerdoch, Preparata e Goethals com boa capacidade de correção de erros podem ser vistos, através do mapeamento de Gray, como códigos lineares sobre $\mathbb{Z}_4$.

Códigos algébricos lineares podem ser definidos através de alfabetos escolhidos em um conjunto de tamanho arbitrário. Contudo, em muitos dos resultados em Teoria da Codificação tem sido considerado que os símbolos são elementos de um corpo finito. Deste modo, no presente trabalho, apresentamos métodos de construção e de decodificação de códigos lineares sobre anéis comutativos finitos com identidade.

O estudo de códigos lineares sobre anéis não é recente. Códigos lineares sobre anéis de inteiros residuais foram considerados por Blake [8]-[9], Spiegel [48]-[49], Shankar [46], Interlando et al. [33], dentre outros. Por outro lado, a maioria dos códigos sobre anéis são determinados sob a métrica de Hamming. A métrica de Lee foi desenvolvida como uma alternativa para a métrica de Hamming para transmissão de sinais não binários sobre certos canais ruidosos. Deste modo, códigos lineares sob a métrica de Lee tem recentemente despertado um grande interesse em teoria da decodificação, mas

estas classes se ressentem de uma sistematização na determinação dos melhores códigos e também da existência de algoritmos de decodificação simples para códigos definidos sobre anéis finitos e, portanto, a obtenção de um processo sistemático para a decodificação de códigos sob a métrica de Lee tem sido a motivação das pesquisas recentes. Neste sentido, no presente trabalho, damos continuidade a caracterização da construção de códigos lineares sobre anéis arbitrários comutativos finitos. Assim, as principais contribuições podem ser resumidas como se segue:

No Capítulo 3, apresentamos construções de códigos lineares sobre anéis finitos via anéis de grupo, fazendo uso do Teorema de Estrutura para um anel comutativo finito. Também apresentamos construções de códigos de Hamming e Reed-Solomon sobre anéis finitos locais. Um algoritmo de decodificação dos códigos de Hamming também foi apresentado. Finalmente, baseado na proposta alternativa de codificação e decodificação sobre grupos abelianos apresentada por Interlando [31], apresentamos uma extensão desta mesma construção para anéis finitos. O processo de obtenção de códigos sobre anéis finitos é através da concatenação generalizada de códigos sobre anéis finitos locais, e o processo de decodificação consiste basicamente em construir codificadores para cada código componente na concatenação.

No Capítulo 4, apresentamos uma proposta de extensão dos códigos BCH sobre anéis comutativos finitos arbitrários. A construção destes códigos foi baseada na obtensão do grupo cíclico $G - S$ que contêm as raízes do polinômio $x^s - 1$ sobre o grupo das unidades de uma extensão do anel. Este procedimento é similar a construção destes códigos sobre corpos finitos. Também apresentamos um algoritmo de decodificação para os códigos Reed-Solomon e BCH definidos sobre um anel comutativo finito local com identidade, fazendo uso do algoritmo de Berlekamp-Massey modificado.

No Capítulo 5, apresentamos técnicas de construções de códigos alternantes e BCH. Na primeira técnica de construção, códigos alternantes sobre anéis comutativos finitos locais sob a métrica de Hamming foi apresentado. Um processo de decodificação eficiente baseado no algoritmo de Berlekamp-Massey modificado foi fornecido. Na segunda técnica de construção, códigos BCH sobre os anéis dos inteiros residuais $\mathbb{Z}_q$, onde q é uma potência de um primo, sob a métrica de Lee foi proposto. Um processo de decodificação alternativo também baseado no algoritmo de Berlekamp-Massey modificado foi fornecido.

6.2 Sugestões para Futuras Pesquisas

Como direções alternativas para posterior prosseguimento deste trabalho, podemos sugerir algumas extensões e aplicações dos resultados como se segue.

- determinar o polinômio gerador (caso exista) de códigos cíclicos definidos sobre anéis finitos.
- extensões de outras classes de códigos com símbolos em corpos finitos, como por exemplo códigos de Goppa, para símbolos em anéis finitos;
- determinação da distância mínima verdadeira dos códigos alternantes e BCH.
- estender o conceito da métrica de Lee a anéis finitos e consequentemente gerar e decodificar códigos sobre estes anéis sob esta métrica.
- decodificar códigos cíclicos e BCH com símbolos em anéis comutativos finitos até a distância mínima atual através do algoritmo Iterativo Fundamental modificado.
- utilização de códigos de bloco lineares sobre anéis comutativos com identidade em codificação de fontes, bem como em codificação conjunta fonte - canal.

Bibligrafia

[1] A.A. Andrade and R. Palazzo Jr., Applications of local finite rings in coding theory, submetido para publicação.

[2] A.A. Andrade and R. Palazzo Jr., BCH codes over arbitrary finite commutative rings, submetido ao 12th Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes, 23 à 27/06/97, França (Toulouse).

[3] A.A. Andrade, R. Palazzo Jr. and J.C. Interlando, Alternant and BCH codes over certain local finite rings, submetido ao International Symposium on Information Theory, 29/06 à 04/07/97, Alemanha (Ulm).

[4] M.F. Atiyah and I.G. MacDonald, *Introduction to Commutative Algebra*, Massachusetts, Addison-Wesley, 1969.

[5] E.R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.

[6] E. Biglieri, and M. Elia, "On the construction of group block codes", *Annales des Telecommunications*, Tome 50, No. 9-10, pp. 817-823, Sept-Oct 1995.

[7] R.E. Blahut, *Theory and Practice of Error Control Codes*, Reading: Addison-Wesley, 1983.

[8] I.F. Blake, "Codes over certain rings", *Inform. and Control*, No. 20, pp. 396-404, 1972.

[9] I.F. Blake, "Codes over integer residue rings", *Inform. and Control*, No. 29, pp. 295-300, 1975.

[10] R.C. Bose and D.K. Ray-Chaudhuri, "On class of error-correcting binary group codes", *Inform. and Control*, Vol. 3, pp. 68-79, March 1960.

[11] R.C. Bose and D.K. Ray-Chaudhuri, "Further results on error-correcting binary group codes", *Inform. and Control*, Vol. 3, pp. 279-290, September 1960.

[12] W.C. Brown, *Matrices over Commutative Rings*, New York: Marcel Dekker, 1993.

[13] A.R. Calderband and N.J.A. Sloane, "Modular and p-adic cyclic codes", *Designs, Codes, and Crystography*, submetido para publicação.

[14] J. Chung-Yaw Chiang and J.K. Wolf, "On channels and codes for the Lee metric", *Inform. and Control*, Vol. 19, No. 2, pp. 159-173, 1971.

[15] G.D. Forney Jr, "On decoding BCH codes", *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 549-557, October 1965.

[16] G.D. Forney Jr, "Coset codes I: Introduction and geometrical classification", *IEEE Trans. Inform. Theory*, vol. 34, pp. 1123-1151, September 1988.

[17] G.D. Forney Jr, "Coset codes II: Binary lattices and related codes", *IEEE Trans. Inform. Theory*, vol. 34, pp. 1152-1187, September 1988.

[18] G.D. Forney Jr., "Geometrically uniform codes", *IEEE Trans. Inform. Theory*, Vol. 37, No. 5, pp. 1241-1260, September 1991.

[19] M.J.E. Golay, "Notes on digital coding", *Proc. IEEE*, 37 (1949), pp. 657.

[20] V.D. Goppa, "Algebraico - geometic codes", *Math. URSS Izvestiya*, 21 (1983), pp. 75-91.

[21] R.W. Hamming, "Error detecting and error correcting codes", *Bell Syst. Tech. J.*, 29 (1950), pp. 147-160.

[22] A.R. Hammons, Jr., P.V. Kumar, A.R. Calderbank, N.J.A. Sloane e P. Solé, "The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals and related codes", *IEEE Trans. Inform. Theory*, Vol. 40, No. 2, pp. 301-319, March, 1994.

[23] H.J. Helgert, "Srivastava codes", *IEEE Trans. Inform. Theory*, Vol. IT-18, No. 2, pp. 292-297, March 1972.

[24] H.J. Helgert, "Noncyclic generalizations of BCH and Srivastava codes", *Inform. and Control*, No. 21, pp. 280-290, 1972.

[25] H.J. Helgert, "Alternant codes", *Inform. and Control*, No. 26, pp. 369-380, 1974.

[26] H.J. Helgert, "Binary primitive alternant codes", *Inform. and Control*, No. 27, pp. 101-108, 1975.

[27] I.N. Herstein, *Topics in Algebra*, John Wiley and Sons, New York, 1975.

[28] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes", *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 371-377, March 1977.

[29] A. Hocquenghem, "Codes correcteurs d'erreurs", *Chiffres*, Vol. 2, pp. 147-156, September 1959.

[30] K. Huber, "Codes over Gaussian integers", *IEEE Trans. Inform. Theory*, Vol. 40, No. 1, pp. 207-216, January 1994.

[31] J.C. Interlando, *Uma Contribuição à Construção e Decodificação de Códigos Lineares sobre Grupos Abelianos via Concatenação de Códigos sobre Anéis de Inteiros Residuais*, Tese de Doutorado, Fee-Unicamp, Dezembro de 1994.

[32] J.C. Interlando and R. Palazzo, Jr., "A note on cyclic codes over $\mathbb{Z}_m$", *VI Reunión de Trabajo en Processamiento de la Información y Control*, Bahia Blanca, Argentina, November 1995.

[33] J.C. Interlando, R. Palazzo, Jr., and M. Elia, "On the decoding of Reed-Solomon and BCH codes over integer residue rings", *IEEE Trans. Inform. Theory*, aceito para publicação.

[34] C.Y. Lee, "Some properties of nonbinary error-correcting codes", *IEEE Trans. Inform. Theory*, Vol. pp. 77-82, June 1958.

[35] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error Correcting Codes*, North-Holland Publishing Company, 1977.

[36] J.L. Massey, "Shift register synthesis and BCH decoding", *IEEE Trans. Inform. Theory*, Vol. IT-15, pp. 122-127, January 1969.

[37] B.R. McDonald, *Finite Rings with Identity*, New York, Marcel Dekker, 1974.

[38] E.E. Nemirovskii and S.L. Portnoi, "Matching block codes to a channel with phase shifts", *Probl. Peredachii Inform.*, Vol. 22, pp. 27-34, 1986.

[39] W.W. Peterson and E.J. Weldon, Jr., *Error Correcting Codes*, 2nd. ed., MIT Press, Cambridge, Mass., 1972.

[40] P.M. Piret, "Algebraic construction of cyclic codes over $\mathbb{Z}_8$ with a good Euclidean minimum distance", *IEEE Trans. Inform. Theory*, Vol. 41, No. 3, pp. 815-818, May 1995.

[41] E. Prange, "Cyclic error-correcting codes in two symbols", *TN*-57-103, September 1957.

[42] R.M. Roth and P.H. Siegel, "Lee-metric BCH codes and their application to constrained and partial-response channels", *IEEE Trans. Inform. Theory*, Vol. 40, No. 4, pp. 1083-1096, July 1994.

[43] P. Samuel, *Algebraic Theory of Numbers*, Paris, Hermann, 1970.

[44] C. Satyanarayana, "Lee metric codes over integer residue rings", *IEEE Trans. Inform. Theory*, Vol. IT-25, No. 2, pp. 250-254, March 1979.

[45] S.J. Sayegh, "A class of optimum block codes in signal space", *IEEE Trans. Commun.*, vol. COM-34, pp. 1043-1045, October 1986.

[46] P. Shankar, "On BCH codes over arbitrary integer rings", *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 480-483, July 1979.

[47] C.E. Shannon, "A Mathematical Theory of Communication", *Bell Syst., Tech. J.*, Vol. 27 (1948), pp. 397-423 and 623-656.

[48] E. Spiegel, "Codes over $\mathbb{Z}_m$", *Inform. and Control*, No. 35, pp. 48-51, 1977.

[49] E. Spiegel, "Codes over $\mathbb{Z}_m$, Revisited", *Inform. and Control*, No. 37, pp. 100-104, 1978.

[50] H. Stichtenoth, *Algebraic Function Fields and Codes*, Springer Verlag, 1993.

[51] W. Ulrich, "Non-binary error correction codes", *Bell Syst. Tech. J.*, Vol. 36, No. 6, pp. 1341-1387, November 1957.