Dane Marques de Ávila

O Segundo Peso de Hamming do Código de Reed-Muller Generalizado

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA
2016

Dane Marques de Ávila

O Segundo Peso de Hamming do Código de Reed-Muller Generalizado

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Uberlândia, como parte dos requisitos para obtenção do título de **MESTRE EM MATEMÁTICA.**

Área de Concentração: Matemática.
Linha de Pesquisa: Álgebra.

Orientador: Prof. Dr. Cícero Fernandes de Carvalho.

UBERLÂNDIA - MG
2016

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

A958s 2016

Ávila, Dane Marques de, 1987-
O segundo peso de Hamming do código de Reed-Muller generalizado / Dane Marques de Ávila. - 2016.
53 f.

Orientador: Cícero Fernandes de Carvalho.
Dissertação (mestrado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Matemática.
Inclui bibliografia.

1. Matemática - Teses. 2. Corpos finitos (Álgebra) - Teses. 3. Álgebra comutativa - Teses. 4. Bases de Gröbner - Teses. I. Carvalho, Cícero Fernandes de. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Matemática. III. Título.

CDU: 51

SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Ata da defesa de DISSERTAÇÃO DE MESTRADO junto ao Programa de Pós-Graduação em Matemática da Faculdade de Matemática da Universidade Federal de Uberlândia

Defesa de: Dissertação de Mestrado Acadêmico, nº. 56, PPGMAT;
Data: 29 de fevereiro de 2016;
Discente: 11412MAT004 – Dane Marques de Ávila;
Título do Trabalho: O segundo peso de Hamming do código de Reed-Muller generalizado;
Linha de Pesquisa: Geometria Algébrica;
Projeto de Pesquisa de Vinculação: Estudo de parâmetros de códigos de avaliação do tipo Reed-Muller e seus duais.

Às 10:00 horas do dia 29 de fevereiro do ano de 2016 na Sala 1F 119 - Campus Santa Mônica da Universidade Federal de Uberlândia, reuniu-se a Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Matemática em 29 de fevereiro de 2016, assim composta: Professores(as) Doutores(as) Paulo Roberto Brumatti - IMECC-UNICAMP – Campinas/SP; Victor Gonzalo Lopez Neumann - FAMAT/UFU e Cícero Fernandes de Carvalho - FAMAT/UFU, orientador do candidato, sendo que a participação do primeiro foi feita de forma virtual via Skype.

Iniciando os trabalhos, o presidente da mesa, Prof. Dr. Cícero Fernandes de Carvalho, agradeceu a presença do público, apresentou à Comissão Examinadora o candidato, concedendo ao mesmo a palavra para exposição de seu trabalho de dissertação. A duração da apresentação do candidato e o tempo de arguição e resposta foram cumpridos conforme as normas do Programa.

A seguir, o senhor presidente concedeu a palavra, pela ordem acima estabelecida, sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, avaliou a dissertação, a apresentação e as respostas do candidato à arguição.

Em face do resultado obtido, a Banca Examinadora considerou o candidato aprovado.

Esta defesa de Dissertação de Mestrado Acadêmico é parte dos requisitos necessários à obtenção do título de Mestre em Matemática. O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, legislação e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos às 11:30 horas. Foi lavrada a presente ata que, após lida e achada conforme, foi assinada pela Banca Examinadora.

Prof. Dr. Paulo Roberto Brumatti
IMECC/UNICAMP

Prof. Dr. Victor Gonzalo Lopez Neumann
FAMAT/UFU

Prof. Dr. Cícero Fernandes de Carvalho
FAMAT/UFU - Orientador

Dedicatória

Dedico esse trabalho à minha esposa Mariana Mendes Silva, que em momentos de fraqueza, sempre me deu força, amor, carinho e apoio para continuar.

Agradecimentos

Primeiramente, agradeço a Deus, pois sem Ele, eu nada teria alcançado.

Agradeço ao meu orientador, Cícero Fernandes de Carvalho pelos ensinamentos dados e aos professores Paulo Roberto Brumatti e Victor Gonzalo Lopez Neumann por terem aceito o convite para fazerem parte da minha banca.

Agradeço também à minha esposa Mariana Mendes Silva, não só por esta etapa concluída, mas pelos nossos mais de oito anos de amor, paciência, compreensão e companheirismo.

Agradeço à minha família (meu pai Jair Rodrigues de Ávila, minhas mães Leni Marques Gonçalves de Ávila e Milta Pereira Menezes de Ávila e ao meu irmão Daniel Menezes de Ávila) pelo amor, pelo apoio e incentivo aos estudos.

E por fim agradeço à agência CAPES pelo fornecimento da bolsa de pesquisa ao longo da Pós-Graduação.

ÁVILA, D. M. *O Segundo Peso de Hamming do Código de Reed-Muller Generalizado.* 2016. 53 p. Dissertação de Mestrado, Universidade Federal de Uberlândia, Uberlândia-MG.

Resumo

Nesse trabalho apresentamos o cálculo do segundo peso de Hamming de códigos de Reed-Muller generalizados na maioria dos casos (v. Teorema 4.6). Nossa referência principal será [13], embora tenhamos utilizado também resultados de [3] e [5]. No primeiro capítulo descrevemos os corpos finitos e mostramos como podem ser construídos. No capítulo 2 apresentamos os conceitos básicos da teoria de códigos. Nele, definimos o que são os códigos corretores de erros, a métrica de Hamming, os parâmetros de um código, a equivalência de códigos através da noção de isometria, bem como uma breve apresentação dos códigos de Reed-Muller generalizados e seus parâmetros. No capítulo 3 são apresentados alguns resultados da teoria de Bases de Gröbner e a definição dos Códigos Cartesianos Afins, que são uma generalização dos códigos de Reed-Muller generalizados. Usamos ferramentas da teoria de bases de Gröbner para determinar a dimensão e distância mínima de Códigos Cartesianos Afins. Para finalizar nosso trabalho, no capítulo 4 determinamos o segundo peso de Hamming do Código de Reed-Muller generalizado na maioria dos casos.

Palavras-chave: Códigos de Reed-Muller generalizados, Distância Mínima, Segundo peso de Hamming, Códigos Cartesianos Afins.

ÁVILA, D. M. *The second Hamming weight of generalized Reed-Muller Code.* 2016. 53 p. M. Sc. Dissertation, Federal University of Uberlândia, Uberlândia-MG.

Abstract

In this work we present the determination of the second Hamming weight of generalized Reed-Muller codes in most cases (see Teorema 4.6). Our main reference is [13], although we have also used results from [3] and [5]. In the first chapter we describe finite fields e we show how they can be constructed. In chapter 2 we present the basics of coding theory. We define what are error correcting codes, the Hamming metric, the parameters of a code, the equivalence of codes through the concept of isometry, and we briefly present generalized Reed-Muller codes and their parameters. In chapter 3 we present some results from Gröbner bases theory and the defintion of Affine Cartesian codes, which generalize the generalized Reed-Muller codes. we use tools from Gröbner bases theory to determine the dimension and the minimum distance of Affine Cartesian codes. We finish our work in chapter 4, with the determination of the second Hamming weight for generalized Reed-Muller codes in most cases.

Key-words: Generalized Reed-Muller codes, minimum distance, second Hamming weight, affine cartesian codes.

Sumário

Introdução

Os códigos corretores de erros estão presentes em nosso cotidiano, sempre que fazemos uso de informações digitalizadas, tais como assistir a um programa de televisão, falar ao telefone, assistir um filme ou navegar pela internet.

Os parâmetros principais de um código são o comprimento, a dimensão e a distância mínima. No entanto, para analisar a performance de um código é necessário estudar os pesos de Hamming de ordem mais alta.

Em 1965, a nave espacial *Mariner 4* transmitiu 22 fotos em preto e branco de Marte, onde cada foto foi decomposta em 200×200 elementos de imagem, e a cada elemento de imagem foi atribuído um dos 64 tons de cinza previamente escolhidos. Esses 64 tons de cinza foram codificados como elementos de $\{0,1\}^6$(código fonte). Em 1972, a nave espacial *Mariner 9* transmitiu novas imagens de Marte. Desta vez cada imagem foi decomposta em 700×832 elementos, obtendo uma resolução maior. O código da fonte foi mantido, mas desta vez foi possível recodificar o código através de uma aplicação injetora φ de $\{0,1\}^6$ em $\{0,1\}^{32}$, de modo que o código de canal resultante fosse capaz de detectar e corrigir até sete erros. O sinal recebido era corrigido e decodificado utilizando-se a transformação φ^{-1}, achando-se o elementos de $\{0,1\}^6$ e, em seguida o tom de cinza a ele correspondente. Este código é um membro particular de uma família de códigos chamados de Códigos de Reed-Muller.

Os códigos de Reed-Muller são uma família de códigos corretores de erros usados em comunicação. Esse nome é devido a Irving S. Reed e David E. Muller. Muller descobriu os códigos, e Reed foi o primeiro a propôr a lógica de decodificação desses códigos.

Nesse trabalho apresentamos o cálculo do segundo peso de Hamming de códigos de Reed-Muller na maioria dos casos (v. Teorema 4.6). Nossa referência principal será [13], embora tenhamos utilizado também resultados de [3] e [5].

No primeiro capítulo descrevemos os corpos finitos e mostramos como podem ser construídos. No capítulo 2 apresentamos os conceitos básicos da teoria de códigos. Nele, definimos o que são os códigos corretores de erros, a métrica de Hamming, os parâmetros de um código, a equivalência de códigos através da noção de isometria, bem como uma breve apresentação dos códigos de Reed-Muller e seus parâmetros. No capítulo 3 são apresentados alguns resultados da teoria de Bases de Gröbner e a definição dos Códigos Cartesianos Afins, que são uma generalização dos códigos de Reed-Muller. Usamos ferramentas da teoria de bases de Gröbner para determinar a dimensão e distância mínima de Códigos Cartesianos Afins. Para finalizar nosso trabalho, no capítulo 4 determinamos o segundo peso de Hamming do Código de Reed-Muller generalizado na maioria dos casos.

Capítulo 1

Corpos Finitos

Iniciaremos este capítulo estudando as propriedades dos corpos finitos, culminando com a descrição completa desses corpos.

1.1 A característica de um corpo

Sejam A e B dois anéis. Uma função $f : A \longrightarrow B$ será chamada **homomorfismo** se, para todos os elementos a e b em A, vale que

(i) $f(a+b) = f(a) + f(b)$,

(ii) $f(a \cdot b) = f(a) \cdot f(b)$,

(iii) $f(1) = 1$.

Proposição 1.1 *Seja $f : A \longrightarrow B$ um homomorfismo entre anéis A e B e sejam $a, b \in A$. Temos que*

i) $f(0) = 0$.

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

iii) $f(a-b) = f(a) - f(b)$.

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

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

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

Demonstração:

(i) Note que $f(0) = f(0+0) = f(0)+f(0)$, logo, cancelando $f(0)$ nos extremos das igualdades acima, obtemos que $0 = f(0)$.

(ii) Note que $0 = f(0) = f(a + (-a)) = f(a) + f(-a)$. Somando $-f(a)$ aos extremos da igualdade acima, obtemos que $-f(a) = f(-a)$.

(iii) Segue imediatamente de (ii), notando que $a - b = a + (-b)$.

(iv) Se a é invertível, temos que $f(a)^{-1} = f(a^{-1})$, pois

$$1 = f(1) = f(a \cdot a^{-1}) = f(a) \cdot f(a^{-1}).$$

(v) Se f é um homomorfismo bijetor, segue que $f^{-1}(1) = 1$, pois $f(1) = 1$, por definição. Sejam $c, d \in B$, $a = f^{-1}(c)$ e $b = f^{-1}(d)$, temos que

$$f^{-1}(c+d) = f^{-1}(f(a)+f(b)) = f^{-1}(f(a+b)) = a+b = f^{-1}(c)+f^{-1}(d),\ e$$

$$f^{-1}(c\cdot d) = f^{-1}(f(a)\cdot f(b)) = f^{-1}(f(a\cdot b)) = a\cdot b = f^{-1}(c)\cdot f^{-1}(d).$$

(vi) Suponhamos que A e B sejam corpos. Se $f(a) = f(b)$, por (iii), segue que $f(a-b) = f(a) - f(b) = 0$. Se $a \neq b$, então $a - b$ seria invertível. Por (iv), $f(a-b)$ seria invertível, e portanto, não nulo; o que seria um absurdo. Consequentemente, $a = b$ e, portanto, f é injetora.

Para provar que $f(A)$ é um subcorpo de B, temos que mostrar apenas que, se $\alpha, \beta \in f(A)$ com $\beta \neq 0$, então $\alpha - \beta \in f(A)$ e $\dfrac{\alpha}{\beta} \in f(A)$. Suponhamos então que $\alpha = f(a)$ e $\beta = f(b)$, temos que

$$\alpha - \beta = f(a) - f(b) = f(a-b) \in f(A),\ e$$

$$\frac{\alpha}{\beta} = \frac{f(a)}{f(b)} = f(a\cdot b^{-1}) \in f(A).$$

□

Definição 1.2 *Um homomorfismo bijetor de corpos será chamado de* ***isomorfismo****. Dois corpos serão ditos* ***isomorfos*** *se existir um isomorfismo entre eles.*

Seja K um corpo finito com elemento unidade 1. Considere o conjunto

$$\Lambda_K = \{n \in \mathbb{N};\ n1 = 0\} \subset \mathbb{N} := \{0, 1, 2, \cdots\}.$$

Pelo fato de K ser finito, temos que existem dois inteiros $n_1 < n_2$ tais que $n_1 1 = n_2 1$. Logo, $(n_2 - n_1)1 = 0$ com $n_2 - n_1 > 0$ e, portanto, $\Lambda_K \setminus \{0\} \neq \emptyset$.

Definição 1.3 *Define-se a* ***característica*** *de um corpo finito K, como sendo o inteiro positivo*

$$car(K) = \min\{\Lambda_K \setminus \{0\}\} = \min\{n \in \mathbb{N} \setminus \{0\};\ n1 = 0\}.$$

Se um corpo F é um subcorpo de um corpo K, então $car(K) = car(F)$, pois $\Lambda_F = \Lambda_K$. Temos também que K é um espaço vetorial sobre F.

Proposição 1.4 *Seja K um corpo finito, então $car(K)$ é um número primo.*

Demonstração: Seja $m = car(K)$ e suponhamos que m não seja primo. Logo, $m = m_1 \cdot m_2$, onde m_1 e m_2 são inteiros maiores que 1 e menores do que m. Logo

$$0 = m1 = (m_1 \cdot m_2)1 = m_1(m_2 1) = (m_1 1)\cdot(m_2 1).$$

Como K é um domínio, temos que $m_1 1 = 0$ ou $m_2 1 = 0$, o que contradiz a minimalidade de m. □

Proposição 1.5 *Seja K um corpo finito com $car(K) = p$. Se para $m \in \mathbb{Z}$ e $a \in \mathbb{K}$ tem-se $ma = 0$, então m é um múltiplo de p ou $a = 0$.*

Demonstração: Suponhamos que $ma = 0$, logo, $(m1) \cdot a = 0$. E como K é um corpo, temos que $m1 = 0$ ou $a = 0$. Basta agora mostrar que, se $m1 = 0$, então m é um múltiplo de p. De fato, suponhamos que $m1 = 0$. Pelo algoritmo da divisão, temos que $m = \lambda p + r$, onde $0 \leq r < p$. Logo,

$$0 = m1 = (\lambda p + r)1 = \lambda(p1) + r1 = \lambda 0 + r1 = r1,$$

e como p é o menor inteiro positivo tal que $p1 = 0$, segue que $r = 0$. Portanto, m é múltiplo de p. □

Teorema 1.6 *Seja K um corpo finito com $car(K) = p$, onde p é um número primo. Então, K contém um subcorpo isomorfo a $\mathbb{Z}_p$ (que ainda denotaremos por $\mathbb{Z}_p$). Em particular, K tem p^n elementos para algum número natural n.*

Demonstração: Considere a aplicação

$$\begin{array}{rccc} \varphi: & \mathbb{Z}_p & \longrightarrow & K \\ & [n] & \longmapsto & n1 \end{array}$$

Primeiramente, é preciso mostrar que essa função está bem definida. De fato, se $[n] = [m]$, onde m e n são dois inteiros, então existe um inteiro λ tal que $n = m + \lambda p$. Logo,

$$n1 = (m + \lambda p)1 = m1 + (\lambda p)1 = m1 + \lambda(p1) = m1 + \lambda 0 = m1 + 0 = m1,$$

Agora, é imediato verificar que φ é um homomorfismo. Logo, pela Proposição 1.1(vi), temos que $\varphi(\mathbb{Z}_p)$ é um subcorpo de K, isomorfo a $\mathbb{Z}_p$.

Temos, portanto, que K é um espaço vetorial sobre $\mathbb{Z}_p$; e como K é finito, segue que tem dimensão finita sobre $\mathbb{Z}_p$. Seja $\alpha_1, \cdots, \alpha_n$ uma base de K sobre $\mathbb{Z}_p$. Então, todo elemento de K se escreve de modo único na forma

$$\lambda_1\alpha_1 + \cdots + \lambda_n\alpha_n,$$

com os $\lambda_i \in \mathbb{Z}_p$, $i = 1, \cdots, n$. Contando esses elementos, segue que $|K| = p^n$. □

1.2 Potências da Característica

As potências da característica de um corpo finito possuem propriedades importantes que serão aqui apresentadas.

Proposição 1.7 *Seja K um corpo finito de característica p e seja $q = p^r$, para algum inteiro positivo r. Se $a, b \in K$, temos que*

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

Demonstração: Pelo Binômio de Newton, temos que

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

Como p é primo temos que $p \mid \binom{p}{i}$, $i = 1, \cdots, p-1$ e logo

$$(a \pm b)^p = a^p \pm b^p.$$

Agora a prova segue por indução observando que

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

□

Observação 1.8 *Segue por indução, da proposição acima, que, se $a_1, \cdots, a_n$ são elementos de um corpo finito K, de característica p, e se q é uma potência de p, então*

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

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

$$P(X)^q = a_0^q + a_1^qX^q + \cdots + a_n^qX^{nq}.$$

Corolário 1.9 *Seja K um corpo finito de característica p. Se $q = p^r$ para algum inteiro positivo r, então a aplicação f_q é um isomorfismo de corpos, onde*

$$f_q : \begin{array}{ccc} K & \longrightarrow & K \\ x & \longmapsto & x^q \end{array}.$$

Demonstração: Temos claramente que

$$f_q(ab) = (ab)^q = a^qb^q = f_q(a)f_q(b),$$

e pela proposição acima,

$$f_q(a + b) = (a + b)^q = a^q + b^q = f_q(a) + f_q(b).$$

Como $f_q(1) = 1$, segue que f_q é um homomorfismo. Pela Proposição 1.1(vi), f_q é injetora e, como K é finito, segue que f_q é bijetora; logo, é um isomorfismo. □

Corolário 1.10 *Sejam F um corpo de característica $p > 0$ e q uma potência inteira de p. O conjunto $K = \{\alpha \in F;\ \alpha^q - \alpha = 0\}$ é um subcorpo de F.*

Demonstração: Temos apenas que mostrar que, se $\alpha, \beta \in K$, onde $\beta \neq 0$, então $\alpha - \beta$ e $\dfrac{\alpha}{\beta}$ estão em K, e isto segue imediatamente da Proposição 1.7. □

Proposição 1.11 *Sejam K um corpo finito de característica p e $P(X) \in K[X]$ Temos que $P'(X) = 0$ se, e somente se, existe um polinômio $Q(X) \in K[X]$ tal que $P(X) = Q(X)^p$.*

Demonstração: Se $P(X) = a_0 + a_1X + \cdots + a_iX^i + \cdots + a_nX^n$ é tal que $P'(X) = 0$, segue que $ia_i = 0$ para todo $i = 1, \cdots, n$. Consequentemente, pela Proposição 1.5, temos que ter i múltiplo de p sempre que $a_i \neq 0$, ou seja,

$$P(X) = a_0 + a_pX^p + a_{2p}X^{2p} + \cdots .$$

Escolhendo $b_i \in K$ tal que $b_i^p = a_{ip}$, o que é possível pelo Corolário 1.9 acima, para $q = p$, o resultado segue pondo $Q(X) = b_0 + b_1X + b_2X^2 + \cdots$.

A recíproca é imediata uma vez que

$$P'(X) = (Q(X)^p)' = pQ(X)^{p-1} \cdot Q'(X) = (p1) \cdot Q(X)^{p-1} \cdot Q'(X) = 0 \cdot Q(X)^{p-1} \cdot Q'(X) = 0.$$

□

Proposição 1.12 *Seja K um corpo finito de característica p e seja $q = p^r$, para algum inteiro positivo r. O polinômio $F(X) = X^q - X$ não possui fatores irredutíveis múltiplos em $K[X]$.*

Demonstração: Suponha que $F(X) = g(X)^2G(X)$, com grau de $g(X)$ positivo. Temos então que $F'(X) = 2g(X)G(X) + g(X)^2G'(X)$ e logo $g(X) \mid F'(X)$. Observe, no entanto, que $F'(X) = -1$, logo não existe um tal fator $g(X)$. □

Lema 1.13 *Seja K um corpo finito com q elementos. Para todo $\alpha \in K^*$, onde $K^* = K \setminus \{0\}$, temos que*

$$\alpha^{q-1} = 1.$$

Demonstração: Seja $\alpha \in K^*$ e considere a aplicação

$$\begin{array}{rccc} \varphi_\alpha : & K^* & \longrightarrow & K^* \\ & a & \longmapsto & \alpha a \end{array}$$

É imediato verificar que φ_α é bijetora. Se $K^* = \{a_1, \cdots, a_{q-1}\}$, temos então que

$$\{\alpha a_1, \cdots, \alpha a_{q-1}\} = \{a_1, \cdots, a_{q-1}\},$$

e portanto,

$$\alpha a_1 \cdots \alpha a_{q-1} = a_1 \cdots a_{q-1},$$

e consequentemente,

$$\alpha^{q-1} = 1.$$

□

Segue imediato do lema acima o seguinte resultado:

Corolário 1.14 *Seja K um corpo finito com q elementos. Para todo $\alpha \in K$ e para todo $i \in \mathbb{N}$, temos que $\alpha^{q^i} = \alpha$.*

Corolário 1.15 *Seja K um corpo finito de característica p com q elementos. Seja F uma extensão de K. Então os elementos de K são os elementos de F que são raízes de $X^q - X = 0$, enquanto que os elementos do subcorpo $\mathbb{Z}_p$ de F são as raízes do polinômio $X^p - X = 0$.*

Demonstração: Pelo Corolário 1.14, temos que os elementos de K são raízes do polinômio $X^q - X$. Mas esse polinômio, tendo grau q, tem no máximo q raízes, logo, as suas raízes são todos os elementos de K. A segunda asserção segue do mesmo modo, considerando $\mathbb{Z}_p$ no lugar de K. □

Seja K um corpo finito e seja $\alpha \in K^*$. Sabemos do Lema 1.13 que

$$\{n \in \mathbb{N} \setminus \{0\};\ \alpha^n = 1\} \neq \emptyset.$$

Definição 1.16 *A **ordem** de $\alpha \in K^*$ é o inteiro positivo*

$$ord(\alpha) = \min\{n \in \mathbb{N} \setminus \{0\};\ \alpha^n = 1\}.$$

Proposição 1.17 *Seja K um corpo finito com q elementos e seja $\alpha \in K^*$. Se para algum inteiro positivo m temos que $\alpha^m = 1$, então $ord(\alpha) \mid m$. Em particular, $ord(\alpha) \mid (q-1)$.*

Demonstração: Pelo algoritmo da divisão, $m = (ord(\alpha))s + r$, para alguns inteiros $s \geq 0$ e r, com $0 \leq r < ord(\alpha)$. Portanto,

$$1 = \alpha^m = (\alpha^{ord(\alpha)})^s \alpha^r = 1 \cdot \alpha^r,$$

o que , pela minimalidade de $ord(\alpha)$, implica que $r = 0$; logo, $ord(\alpha) \mid m$.

Para finalizar, pelo Lema 1.13, temos que $\alpha^{q-1} = 1$. Logo, pelo que acabamos de provar, $ord(\alpha) \mid (q-1)$. □

Proposição 1.18 *Seja K um corpo finito. Sejam α e β elementos do corpo K tais que $mdc(ord(\alpha), ord(\beta)) = 1$. Então $ord(\alpha\beta) = ord(\alpha)ord(\beta)$.*

Demonstração: Sejam $m = ord(\alpha)$ e $n = ord(\beta)$. Temos então que

$$(\alpha\beta)^{mn} = (\alpha^m)^n(\beta^n)^m = 1.$$

Por outro lado, se $(\alpha\beta)^t = 1$, então

$$1 = ((\alpha\beta)^t)^m = \alpha^{tm}\beta^{tm} = 1\beta^{tm} = \beta^{tm}, \ e$$

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

Logo, pela Proposição 1.17, temos que $n \mid tm$ e $m \mid tn$. Como $mdc(m, n) = 1$, segue que $m \mid t$ e $n \mid t$. Novamente usando o fato de que $mdc(m, n) = 1$, segue que $mn \mid t$, o que prova que $mn = \min\{t > 0;\ (\alpha\beta)^t = 1\}$; concluindo assim que $ord(\alpha\beta) = mn$. □

Proposição 1.19 *Seja K um corpo finito e sejam $\alpha \in K^*$ e $i \in \mathbb{N} \setminus \{0\}$. Suponhamos que $ord(\alpha) = m$, então*

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

Demonstração: Seja $t = ord(\alpha^i)$, logo, t é o menor inteiro positivo tal que

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

Ou seja, t é o menor inteiro positivo tal que $m \mid it$; dito de outra forma, it é o menor múltiplo simultaneamente de m e de i. Logo, $it = mmc(m, i)$, ou seja,

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

□

1.3 Polinômios Irredutíveis

O resultado central desta seção é o Teorema 1.23, que nos assegurará a existência de polinômios irredutíveis de qualquer grau com coeficientes num corpo finito arbitrário.

Proposição 1.20 *Seja K um corpo finito com q elementos e seja $f(X)$ um polinômio mônico irredutível em $K[X]$, de grau d. Considere o corpo $F = K[X]/(f(X))$. Temos que*

i) $1, [X], [X^2], \cdots, [X^{d-1}]$ *formam uma base de F sobre K.*

ii) $[X]^{q^d} = [X]$ *em F.*

iii) $f(X)$ *divide* $X^{q^d} - X$ *em $K[X]$.*

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

Demonstração:

(i) É uma consequência da divisão euclidiana em $K[X]$.

(ii) Note que F é um corpo finito com q^d elementos, logo, pelo Corolário 1.14 do Lema 1.13, temos que $[X]^{q^d} = [X]$.

(iii) Isso segue imediatamente de (ii).

(iv) Considere o polinômio

$$g(Y) = (Y - [X])(Y - [X]^q) \cdots (Y - [X]^{q^{d-1}}) \in F[Y].$$

Temos de (ii) que

$$g(Y^q) = (Y^q - [X])(Y^q - [X]^q) \cdots (Y^q - [X]^{q^{d-1}}) =$$

$$= (Y^q - [X]^{q^d})(Y^q - [X]^q) \cdots (Y^q - [X]^{q^{d-1}}) =$$

$$= \Big((Y - [X]^{q^{d-1}})(Y - [X]) \cdots (Y - [X]^{q^{d-2}})\Big)^q = (g(Y))^q.$$

Assim, $g(Y^q) = (g(Y))^q$ e portanto se $g(Y) = b_0 + b_1 Y + \cdots + b_{d-1}Y^{d-1} + Y^d$, temos que $b_i^q = b_i$ para todo $i = 0, \cdots, d-1$. Sendo assim, do Corolário 1.15 do Lema 1.13, segue que $g(Y) \in K[Y]$.

Como $f(Y), g(Y) \in K[Y]$ possuem uma raiz comum numa extensão F de K, segue que o seu *mdc* em $F[Y]$ é não constante e pertence a $K[Y]$. Como $f(Y)$ é irredutível e mônico, ele coincide com o *mdc* de $f(Y)$ e $g(Y)$, logo, $f(Y)$ divide $g(Y)$. Como $f(Y)$ e $g(Y)$ são polinômios mônicos de mesmo grau, eles devem ser iguais.

Agora sabemos de (iii) que $g(Y)(= f(Y))$ divide $Y^{q^d} - Y$, que, pela Proposição 1.12, não tem fatores múltiplos em nenhuma extensão F de K, logo as raízes $[X], [X]^q, \cdots, [X]^{q^{d-1}}$ de $g(Y)$ são duas a duas distintas, provando assim (iv).

□

Note que com a proposição acima fica provado que, se $[X] \in K[X]/(f(X))$ com $f(X) \in K[X]$ irredutível de grau d, então

$$d = \min\{j \in \mathbb{N} \setminus \{0\};\ [X]^{q^j} = [X]\}.$$

No que se segue vamos usar que $mdc(X^{q^n} - X, X^{q^m} - X) = X^{q^{mdc(n,m)}} - X$(veja [10, pág. 46]).

Proposição 1.21 *Seja K um corpo finito com q elementos e seja n um inteiro positivo. Em $K[X]$ temos que:*

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

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

Demonstração: Seja $f(X) \in K[X]$ um polinômio mônico irredutível de grau d. Como $X^{q^n} - X$ não possui fatores múltiplos (veja Proposição 1.12), basta provar a seguinte asserção:

$$f(X) \text{ divide } X^{q^n} - X \iff d \text{ divide } n.$$

Suponhamos inicialmente que $f(X) \mid (X^{q^n} - X)$, daí segue (Proposição 1.20(iii)), que $f(X)$ divide $mdc(X^{q^n} - X, X^{q^d} - X)$; logo, da observação acima, vem que $f(X) \mid (X^{q^e} - X)$, onde $e = mdc(n, d) \leq d$. Sendo assim, $[X]^{q^e} = [X]$, o que pela Proposição 1.20(iv) só é possível se $e \geq d$. Segue, então, que $d = e = mdc(n, d)$, e portanto, $d \mid n$.

Reciprocamente, se d divide n, segue novamente da observação acima que $(X^{q^d} - X) \mid (X^{q^n} - X)$. Como $f(X) \mid (X^{q^d} - X)$ (Proposição 1.20(iii)), segue que $f(X) \mid (X^{q^n} - X)$. □

Corolário 1.22 *Seja $I(n)$ o número de polinômios mônicos irredutíveis de grau n em $K[X]$, onde K é um corpo finito com q elementos. Temos que*

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

Demonstração: Compare os graus dos polinômios em ambos os lados da igualdade da proposição acima. □

Teorema 1.23 *Seja K um corpo finito qualquer. Para cada inteiro positivo n, existe pelo menos um polinômio irredutível de grau n em $K[X]$.*

Demonstração: Para $n = 1$ o resultado é óbvio, pois o polinômio X é irredutível.

Suponhamos agora $n > 1$. Sejam $1 = d_1 < \cdots < d_s < n$, com $s \geq 1$, os divisores de n. Pondo $q = |K|$ e usando duas vezes o corolário acima, temos que

$$q^n = \sum_{d|n} dI(d) = \sum_{i=1}^{s} d_i I(d_i) + nI(n) \leq \sum_{i=1}^{s} \left(\sum_{d|d_i} dI(d) \right) + nI(n) =$$

$$= \sum_{i=1}^{s} q^{d_i} + nI(n) < \sum_{i=0}^{d_s} q^i + nI(n) = \frac{q^{d_s+1} - 1}{q - 1} + nI(n) < q^{d_s+1} + nI(n).$$

Portanto,

$$nI(n) > q^n - q^{d_s+1}.$$

Como d_s divide n e $d_s < n$, temos que $n = \lambda d_s$ com $\lambda > 1$. Logo, $d_s = \frac{n}{\lambda} \leq \frac{n}{2}$ e $q^{d_s+1} \leq q^{\frac{n}{2}+1}$. Consequentemente,

$$nI(n) > q^n - q^{\frac{n}{2}+1} = q^n \left(1 - q^{-\frac{n}{2}+1}\right),$$

o que acarreta $I(n) > 0$. □

1.4 Classificação dos Corpos Finitos

Nesta seção finalmente classificaremos todos os corpos finitos. Mais precisamente, dado um número primo positivo p, daremos uma receita, pelo menos teórica, para construir todos os corpos finitos de característica p.

Teorema 1.24 *(Existência de Corpos Finitos) Para todos os números inteiros positivos p e n, com p primo, existe um corpo com p^n elementos.*

Demonstração: Para todo primo positivo p e todo inteiro positivo n, existe, pelo Teorema 1.23, um polinômio irredutível $f(X) \in \mathbb{Z}_p[X]$ de grau n; logo, o corpo $K = \mathbb{Z}_p[X]/(f(X))$ é um dos corpos procurados, pois tem p^n elementos. □

Teorema 1.25 *(Unicidade dos Corpos Finitos) Dois corpos finitos com o mesmo número de elementos são isomorfos.*

Demonstração: Seja L um corpo finito com p^n elementos, logo, a característica de L é p e ele contêm um corpo isomorfo a $\mathbb{Z}_p$ (veja Teorema 1.6). Logo, L é um espaço vetorial sobre $\mathbb{Z}_p$ de dimensão n.

Como L é um corpo finito com p^n elementos, temos, pelo Corolário 1.15 do Lema 1.13, que a equação $X^{p^n} - X$ tem todas as suas raízes em L e todos os elementos de L são as raízes desse polinômio.

Seja $f(X) \in \mathbb{Z}_p[X]$ um polinômio mônico irredutível de grau n, cuja existência está garantida pelo Teorema 1.23. Pela Proposição 1.20(iii), sabemos que $f(X)$ divide $X^{p^n} - X$ em $\mathbb{Z}_p[X]$, logo, existe β em L tal que $f(\beta) = 0$. Os elementos $1, \beta, \beta^2, \cdots, \beta^{n-1}$ de L são linearmente independentes sobre $\mathbb{Z}_p$, pois, caso contrário, existiria um polinômio não nulo $r(X) \in \mathbb{Z}_p[X]$ de grau menor do que o grau de $f(X)$ tal que $r(\beta) = 0$. Então teríamos que $mdc(f(X), r(X)) \neq 1$; e como $f(X)$ é irredutível, teríamos $f(X) \mid r(X)$, o que é impossível, pois $gr(r(X)) < n$. Logo, tais elementos formam uma base de L sobre $\mathbb{Z}_p$. Sabemos também, pela Proposição 1.20(i), que $1, [X], [X]^2, \cdots, [X]^{n-1}$ é um base de $\mathbb{Z}_p[X]/(f(X))$ sobre $\mathbb{Z}_p$.

Definindo

$$\begin{array}{rccc} \varphi: & \mathbb{Z}_p[X]/(f(X)) & \longrightarrow & L \\ & [a_0 + \cdots + a_{n-1}X^{n-1}] & \longmapsto & a_0 + \cdots + a_{n-1}\beta^{n-1} \end{array}$$

temos que φ está bem definida, pois, se

$$[a_0 + \cdots + a_{n-1}X^{n-1}] = [b_0 + \cdots + b_{n-1}X^{n-1}],$$

então, para algum $g(X) \in \mathbb{Z}_p[X]$, temos que

$$(a_0 + \cdots + a_{n-1}X^{n-1}) - (b_0 + \cdots + b_{n-1}X^{n-1}) = f(X)g(X).$$

Portanto,

$$(a_0 + \cdots + a_{n-1}\beta^{n-1}) - (b_0 + \cdots + b_{n-1}\beta^{n-1}) = f(\beta)g(\beta) = 0,$$

o que prova que

$$a_0 + \cdots + a_{n-1}\beta^{n-1} = b_0 + \cdots + b_{n-1}\beta^{n-1}.$$

Além disso, φ é claramente sobrejetora e é aditiva, isto é,

$$\varphi(u + v) = \varphi(u) + \varphi(v), \ \forall u, v \in \mathbb{Z}_p[X]/(f(x)).$$

Para mostrarmos que φ é um isomorfismo, basta provarmos que é multiplicativa, isto é,

$$\varphi(uv) = \varphi(u)\varphi(v), \ \forall u, v \in \mathbb{Z}_p[X]/(f(X)),$$

já que $\varphi([1]) = 1$ e todo homomorfismo de corpos é injetor. Sejam $u = [u(X)]$ e $v = [v(X)]$, em $\mathbb{Z}_p[X]/(f(X))$, onde $u(X)$ e $v(X)$ são polinômios em $\mathbb{Z}_p[X]$. Pelo algoritmo da divisão de polinômios, escrevemos

$$u(X)v(X) = f(X)q(X) + r(X),$$

onde $r(X)$ é zero ou é um polinômio de grau menor ou igual a $n - 1$. Logo, temos que

$$[u(X)v(X)] = [r(X)] \ e \ u(\beta)v(\beta) = r(\beta),$$

provando assim que

$$\varphi(uv) = r(\beta) = u(\beta)v(\beta) = \varphi(u)\varphi(v).$$

□

Se q é uma potência de um número primo p, denotaremos doravante por $\mathbb{F}_q$ o único corpo (a menos de isomorfismo) com q elementos.

1.5 Elementos Primitivos

Definição 1.26 *Um elemento α de um corpo finito $\mathbb{F}_q$ é chamado de* ***elemento primitivo*** *se*

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

ou seja, se $ord(\alpha) = q - 1$.

Explicitaremos, a seguir, uma propriedade fundamental dos corpos finitos que permite simplificar a operação de multiplicação, sendo, por isso, utilizada em computação.

Teorema 1.27 *Todo corpo finito possui elementos primitivos.*

Demonstração: Suponhamos que K tenha q elementos. Sabemos pelo Lema 1.13 que $x^{q-1} = 1$ para todo $x \in K^*$. Logo, todos os elementos de K^* tem ordem $\leq q - 1$.

Queremos provar que K^* possui um elementos de ordem $q - 1$. Seja $a \in K^*$ um elemento de ordem máxima m, queremos, portanto, mostrar que $m = q - 1$.

Vamos inicialmente provar que se $b \in K^*$, então $ord(b)$ divide $ord(a) = m$. Escrevamos $ord(b) = ds$, onde $d = mdc(ord(b), m)$. Segue que $mdc(s, m) = 1$. Queremos então provar que $s = 1$. De fato, se $s > 1$, teríamos pela Proposição 1.19, que $ord(b^d) = s > 1$ e, portanto, pela Proposição 1.18, que

$$ord(ab^d) = ms > m,$$

contradizendo a maximalidade de $m = ord(a)$.

Segue, então, que todo elemento de K^* satisfaz à equação

$$X^m - 1 = 0,$$

e portanto, $q - 1 = |K^*| \leq m$.

Como $m \leq q - 1$, pois todos os elementos de K^* tem ordem $\leq q - 1$, segue que

$$m = ord(a) = q - 1.$$

□

O Teorema acima significa que existe um elemento $\alpha \in \mathbb{F}_q^*$ tal que

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

É claro que $\alpha^{q-1} = 1$, e portanto, nesta representação, a operação de multiplicação fica extremamente simplificada. Mais precisamente,

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

onde $[i + j]$ representa o resto da divisão de $i + j$ por $q - 1$.

Em compensação, nesta representação a operação de adição se torna mais complicada e, para efetuá-la, usam-se certas tabelas chamadas de tabelas logarítmicas de Zech, que passamos a descrever.

Se $n \leq m$, temos que

$$\alpha^n + \alpha^m = \alpha^n(1 + \alpha^{m-n}).$$

Então, se soubéssemos para cada r determinar o inteiro $z(r)$ tal que $1 + \alpha^r = \alpha^{z(r)}$, teríamos que

$$\alpha^n + \alpha^m = \alpha^n \cdot \alpha^{z(m-n)}.$$

Portanto, para efetuar a adição em $\mathbb{F}_q$, escolhemos um elemento primitivo α de $\mathbb{F}_q^*$ e utilizamos uma tabela previamente preparada, chamada de tabela de Zech, que nos fornece os valores de $z(r)$ para $1 \leq r \leq q-2$.

Em [10], podemos encontrar alguns exemplos de como determinar a tabela de Zech para alguns corpos, utilizando do método de tentativa e erro para determinar um elemento primitivo para o corpo em questão.

Se K é um corpo com muitos elementos, pode ser difícil encontrar um elemento primitivo por tentativa e erro. Por tal razão, a seguir, descrevemos um algoritmo devido a Gauss, bastante eficiente para a determinação de elementos primitivos em um corpo K.

Seja K um corpo com q elementos. O algoritmo permitirá construir uma sequência de elementos $\alpha_1, \cdots, \alpha_n$ em K^* tais que, para todo $i = 1, \cdots, n-1$, tenha-se $ord(\alpha_i)$ divide $ord(\alpha_{i+1})$, e

$$ord(\alpha_1) < ord(\alpha_2) < \cdots < ord(\alpha_n) = q - 1.$$

O Algoritmo de Gauss

(1) Seja $i = 1$ e seja α_1 um elemento qualquer de K^*.
Ponha $t_1 = ord(\alpha_1)$.

(2) Se $t_i = q - 1$, pare. Temos que α_i é o elemento primitivo.

(3) Caso contrário, escolha um elemento β em K^* que não seja uma potência de α_i.
Seja $s = ord(\beta)$. Se $s = q - 1$, ponha $\alpha_{i+1} = \beta$ e pare.

(4) Caso contrário, determine t divisor de t_i e r divisor de s tais que $mdc(t, r) = 1$ e $tr = mmc(t_i, s)$ $(> t_i)$. Ponha $\alpha_{i+1} = \alpha_i^{\frac{t_i}{t}} \beta^{\frac{s}{r}}$; $t_{i+1} = tr$. Vá para (2).

O algoritmo fornece um elemento de ordem $q - 1$, desde que se justifique devidamente o passo (4). Tal justificativa pode ser encontrada em [10].

Capítulo 2

Códigos Corretores de Erros

Os códigos corretores de erros estão presentes em nosso cotidiano, sempre que fazemos uso de informações digitalizadas, tais como assistir a um programa de televisão, falar ao telefone, assistir um filme ou navegar pela internet.

Segundo [10], um código corretor de erros é, em essência, um modo organizado de acrescentar algum dado adicional a cada informação que se queira transmitir ou armazenar, que permita, ao recuperar a informação, detectar e corrigir erros.

2.1 Métrica de Hamming

O ponto de partida para a construção de um código corretor de erros é um conjunto finito A chamado de **alfabeto**. O número de elementos de A será denotado por q, ou seja, $\#A = q$

Definição 2.1 *Um código corretor de erros C é um subconjunto próprio qualquer de A^n, para algum número natural n. Os elementos de A^n são chamados de* ***palavras****.*

A fim de tornar precisa a noção intuitiva de proximidade entre palavras, apresentamos a seguir um modo de medir a distância entre palavras em A^n.

Definição 2.2 *Dados dois elementos $x = (x_1, x_2, \cdots, x_n)$ e $y = (y_1, y_2, \cdots, y_n)$ de A^n, chama-se* ***distância de Hamming*** *de x a y ao número de coordenadas em que estes elementos diferem; isto é:*

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

Observação 2.3 *A Distância de Hamming é uma métrica em A^n (uma verificação dessa afirmação não é difícil e pode ser encontrada em [10]). Por isso a distância de Hamming entre elementos de A^n é também chamada de* ***Métrica de Hamming****.*

Dados um elemento $a \in A^n$ e um número real $t \geq 0$, definimos o **disco** e a **esfera** de centro em a e raio t como sendo, respectivamente, os conjuntos

$$D(a, t) = \{x \in A^n \ ; \ d(x, a) \leq t\},$$
$$S(a, t) = \{x \in A^n \ ; \ d(x, a) = t\}.$$

Esses conjuntos são finitos e o próximo lema nos oferecerá as suas cardinalidades.

Lema 2.4 *Para todo $a \in A^n$ e todo número natural $r > 0$, temos que*

$$|D(a, r)| = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i.$$

Demonstração: Escolhendo i coordenadas dentre as n possíveis temos que o número de palavras que diferem de a, exatamente nessas i posições, é de $(q-1)^i$, logo

$$|S(a,i)| = \binom{n}{i}(q-1)^i.$$

Agora o resultado segue observando que $S(a,i) \cap S(a,j) = \emptyset$ se $i \neq j$, e que

$$\bigcup_{i=0}^{r} S(a,i) = D(a,r).$$

□

Note que a cardinalidade de $D(a,r)$ depende apenas de n, q e r.

Definição 2.5 *Seja C um código. A **distância mínima** de C é o número*

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

Dado um código C com distância mínima d, seja

$$b := \left[\frac{d-1}{2}\right]$$

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

Lema 2.6 *Seja C um código com distância mínima d. Se c e c' são palavras distintas de C, então*

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

Demonstração: De fato, se x pertence a $D(c,b) \cap D(c',b)$, temos $d(x,c) \leq b$ e $d(x,c') \leq b$, e, portanto

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

absurdo, pois $d(c,c') \geq d$. □

A importância da distância mínima d de um código será revelada a seguir.

Teorema 2.7 *Seja C um código com distância mínima d. Então C pode corrigir até $b = \left[\frac{d-1}{2}\right]$ erros e detectar até $d-1$ erros.*

Demonstração: Se na transmissão de uma palavra c do código são introduzidos t erros com $t \leq b$, resultando na palavra r, então $d(r,c) = t \leq b$; enquanto que, pelo Lema 2.6, a distância de r a qualquer outra palavra do código é maior do que b. Isso determina c univocamente a partir de r.

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

Note que, em virtude do Teorema 2.7, um código terá maior capacidade de correção de erros quanto maior for sua distância mínima. Portanto, é fundamental, para a Teoria dos Códigos, poder calcular d ou pelo menos determinar uma cota inferior para ele.

Definição 2.8 *Seja $C \subset A^n$ um código com distância mínima d e seja $b = \left[\frac{d-1}{2}\right]$. O código C será dito **perfeito** se*

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

Um código C sobre um alfabeto A possui três parâmetros fundamentais $[n, M, d]$, que são, respectivamente, o seu comprimento (o número n correspondente ao espaço ambiente A^n onde C se encontra), o seu número de elementos e a sua distância mínima.

Já vimos que a importância da distância mínima reside na sua relação com a capacidade de correção de erros do código. A dimensão de um código é definida como $\log_q(M)$, e a importância da dimensão de um código é que ela é uma medida da quantidade de informação que o código pode processar. A importância do comprimento n do código é que quanto mais longo o código é, mais energia deve-se gastar para transmitir cada palavra. O parâmetros relativos $\log_q(M)/n$ e d/n são os principais conceitos que aparecem na análise de um código, tendo também um papel importante quando se deseja comparar códigos distintos. O código ideal deve ter uma dimensão grande, uma distância mínima grande e um comprimento curto, mas esses requerimentos não podem ser atendidos todos ao mesmo tempo. De fato, uma relação básica entre esses parâmetros é a chamada **Desigualdade de Singleton** que afirma que $\log_q(M)+d \leq n+1$.

Mais informações sobre essa interdependência entre os parâmetros de um código pode ser encontrada em [10].

2.2 Equivalência de Códigos

Sempre que se define uma classe de objetos matemáticos, como por exemplo a classe dos códigos de comprimento n sobre um alfabeto A, define-se também a noção de equivalência entre esses objetos. A noção de equivalência de códigos repousa sobre o conceito de isometria que definiremos abaixo.

Definição 2.9 *Sejam A um alfabeto e n um número natural. Diremos que uma função $F : A^n \to A^n$ é uma **isometria** de A^n se ela preserva distâncias de Hamming. Em símbolos,*

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

Proposição 2.10 *Toda isometria de A^n é uma bijeção de A^n.*

Demonstração: Seja $F : A^n \to A^n$ uma isometria. Suponha que para $x, y \in A^n$ tenhamos $F(x) = F(y)$. Logo, $d(x, y) = d(F(x), F(y)) = 0$, o que implica que $x = y$. Assim, provamos que F é injetora, e como toda aplicação injetora de um conjunto finito nele próprio é sobrejetora, temos que F é uma bijeção. □

Proposição 2.11

(1) A função identidade de A^n é uma isometria.

(2) Se F é uma isometria de A^n, então F^{-1} é uma isometria de A^n.

(3) Se F e G são isometrias de A^n, então $F \circ G$ é uma isometria de A^n.

Demonstração:

(1) é imediato.

(2) Se F é uma isometria, pela Proposição 2.10, existe a função inversa F^{-1} de F. Como F é uma isometria, segue que

$$d(F^{-1}(x), F^{-1}(y)) = d(F(F^{-1}(x)), F(F^{-1}(y))) = d(x, y),$$

o que prova que F^{-1} é uma isometria.

(3) Sejam $x, y \in A^n$, usando o fato que F e G são isometrias, obtemos que

$$d(F(G(x)), F(G(y))) = d(G(x), G(y)) = d(x, y),$$

o que prova que $F \circ G$ é uma isometria.

□

Definição 2.12 *Dados dois códigos C e C' em A^n, diremos que C' é equivalente a C se existir uma isometria F de A^n tal que $F(C) = C'$.*

Segue da Proposição 2.11 que a equivalência de códigos é uma relação de equivalência. Além disso, decorre imediatamente da definição que dois códigos equivalentes têm os mesmos parâmetros.

Damos, a seguir, exemplos de duas famílias importantes de isometrias.

Exemplo 2.13 *Se $f : A \to A$ é uma bijeção, e i é um número inteiro tal que $1 \leq i \leq n$, a aplicação*

$$\begin{array}{rccc} T_f^i : & A^n & \longrightarrow & A^n \\ & (a_1, \cdots, a_n) & \longmapsto & (a_1, \cdots, f(a_i), \cdots, a_n) \end{array}$$

é uma isometria.

Exemplo 2.14 *Se π é uma bijeção do conjunto $\{1, \cdots, n\}$ nele próprio, também chamada de **permutação** de $\{1, \cdots, n\}$, a aplicação permutação de coordenadas*

$$\begin{array}{rccc} T_\pi : & A^n & \longrightarrow & A^n \\ & (a_1, \cdots, a_n) & \longmapsto & (a_{\pi(1)}, \cdots, a_{\pi(n)}) \end{array}$$

é uma isometria.

Na sequência, enunciaremos o teorema principal desta seção, o qual será consequência dos dois lemas seguintes.

Lema 2.15 *Dada uma isometria F de A^n com $n \geq 2$, e dados elementos $a_1, \cdots, a_{n-1} \in A$, existem $a'_1, \cdots, a'_{n-1} \in A$, uma bijeção $f_n : A \longrightarrow A$ e uma permutação σ de $\{1, \cdots, n\}$ tais que*

$$(T_\sigma \circ F)(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f_n(x)), \ \forall x \in A.$$

Demonstração: Se $q = 1$, então $A^n = \{(a, \cdots, a)\}$ e $F(a, \cdots, a) = (a, \cdots, a)$, portanto, o resultado segue trivialmente.

Considere, agora, $q \geq 2$. Sejam $a_n, b_n \in A$ tais que $a_n \neq b_n$ e ponhamos

$$u = (a_1, \cdots, a_{n-1}, a_n) \text{ e } v = (a_1, \cdots, a_{n-1}, b_n).$$

Temos, então, que

$$d(F(u), F(v)) = d(u, v) = 1.$$

Logo, $F(u)$ e $F(v)$ diferem apenas em uma componente. Escolhendo convenientemente a permutação σ de $\{1, \cdots, n\}$ - que depende em princípio de u e de v - podemos supor que

$$(T_\sigma \circ F)(u) = (a'_1, \cdots, a'_{n-1}, a'_n)$$

$$(T_\sigma \circ F)(v) = (a'_1, \cdots, a'_{n-1}, b'_n)$$

com $a'_n \neq b'_n$.

Se $q = 2$, o lema está provado, pois, nesse caso, a bijeção f_n procurada é definida por $a_n \longmapsto a'_n$ e $b_n \longmapsto b'_n$.

Se $q \geq 2$, ponhamos

$$w = (a_1, \cdots, a_{n-1}, x),$$

e como $T_\sigma \circ F$ é uma isometria (veja Proposição 2.11(iii)), temos, para $x \neq a_n$, que

$$d((T_\sigma \circ F)(w), (T_\sigma \circ F)(u)) = d(w, u) = 1.$$

Existe um único $y \in A$ tal que

$$(T_\sigma \circ F)(w) = (a'_1, \cdots, a'_{i-1}, y, a'_{i+1}, \cdots, a'_n)$$

com $y \neq a'_i$. Vamos agora mostrar que $i = n$, ou seja, que σ nã depende nem de u nem de v. De fato, se $x = b_n$, teríamos $w = v$ e, consequentemente, $(T_\sigma \circ F)(w) = (a'_1, \cdots, a_{n-1'}, b'_n)$ e $i = n$. Se $x \neq b_n$ e $i < n$, teríamos

$$1 = d(v, w) = d((T_\sigma \circ F)(v), (T_\sigma \circ F)(w)) = 2,$$

com a última igualdade valendo, pois $y \neq a'_i$ e $a'_n \neq b'_n$. Isso seria absurdo, logo $i = n$. Consequentemente

$$(T_\sigma \circ F)(w) = (a'_1 \cdots, a'_{n-1}, y),$$

e portanto, está bem definida uma função $f_n : A \longrightarrow A$ tal que

$$(T_\sigma \circ F)(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f_n(x)).$$

Como $(T_\sigma \circ F)$ é bijetora (veja Proposição 2.10), segue que f_n é injetora, e como A é finito, temos que f_n é bijetora. □

Lema 2.16 *Seja dada uma isometria G de A^n e sejam $a_1, \cdots, a_{n-1}, a'_1, \cdots, a'_{n-1}$ elementos fixos de A. Suponhamos que exista um bijeção $f : A \longrightarrow A$ tal que*

$$G(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f(x)), \ \forall x \in A.$$

Então, existe uma isometria H de A^{n-1} tal que

$$G(x_1, \cdots, x_{n-1}, x_n) = (H(x_1, \cdots, x_{n-1}), f(x_n)), \ \forall (x_1, \cdots, x_n) \in A^n.$$

Demonstração: Seja $(b_1, \cdots, b_{n-1}) \in A^{n-1}$ tal que

$$(b_1, \cdots, b_{n-1}) \neq (a_1, \cdots, a_{n-1}),$$

e seja $a_n \in A$. Ponhamos

$$u = (a_1, \cdots, a_n) \ e \ v = (b_1, \cdots, b_{n-1}, a_n).$$

Temos, por hipótese, que

$$G(u) = (a'_1, \cdots, a'_{n-1}, f(a_n)).$$

Agora escrevamos

$$G(v) = (c_1, \cdots, c_n).$$

Vamos, inicialmente, provar que $c_n = f(a_n)$. De fato, suponhamos por absurdo que $c_n \neq f(a_n)$. Como f é uma bijeção, existe $b_n \in A$ com $b_n \neq a_n$ tal que $c_n = f(b_n)$. Considere

$$w = (a_1, \cdots, a_{n-1}, b_n);$$

logo,

$$G(w) = (a'_1, \cdots, a'_{n-1}, f(b_n)).$$

Seja $r = d(u, v)$. Logo

$$d((a_1, \cdots, a_{n-1}), (b_1, \cdots, b_{n-1})) = d(u, v) = r. \tag{2.1}$$

Por outro lado,

$$d((a'_1, \cdots, a'_{n-1}), (c_1, \cdots, c_{n-1})) = d(G(u), G(v)) - 1 = d(u, v) - 1 = r - 1. \tag{2.2}$$

Como $a_n \neq b_n$, temos de (2.1) que

$$d(w, v) = d((a_1, \cdots, a_{n-1}), (b_1, \cdots, b_{n-1})) + 1 = r + 1. \tag{2.3}$$

Por outro lado, de (2.2), temos que

$$d(G(w), G(v)) = d((a'_1, \cdots, a'_{n-1}), (c_1, \cdots, c_{n-1})) = r - 1. \tag{2.4}$$

Mas como G é uma isometria, (2.3) e (2.4) geram uma contradição, consequentemente, $c_n = f(a_n)$.

Provamos então que, dado $(x_1, \cdots, x_n) \in A^n$ qualquer, existe $(y_1, \cdots, y_{n-1}) \in A^{n-1}$ tal que

$$G(x_1, \cdots, x_n) = (y_1, \cdots, y_{n-1}, f(x_n)),$$

logo, $y_1, \cdots, y_{n-1}$ são univocamente determinados por $x_1, \cdots, x_n$; e, para provar a existência da função H, é preciso mostrar que $y_1, \cdots, y_{n-1}$ dependem de $x_1, \cdots, x_{n-1}$ e não de x_n. Para isso, consideremos $z_n \in A$ tal que $z_n \neq x_n$ e suponhamos que

$$G(x_1, \cdots, x_{n-1}, z_n) = (y'_1, \cdots, y'_{n-1}, f(z_n)),$$

todavia

$$d(G(x_1, \cdots, x_{n-1}, x_n), G(x_1, \cdots, x_{n-1}, z_n)) =$$
$$d((x_1, \cdots, x_{n-1}, x_n), (x_1, \cdots, x_{n-1}, z_n)) = 1;$$

e como $f(x_n) \neq f(z_n)$, temos

$$y'_i = y_i, \ \forall i = 1, \cdots, n-1,$$

o que prova que está bem definida uma função $H : A^{n-1} \longrightarrow A^{n-1}$ tal que

$$G(x_1, \cdots, x_n) = (H(x_1, \cdots, x_{n-1}), f(x_n)).$$

Agora só falta provar que H é uma isometria. Sejam $(x_1, \cdots, x_{n-1})$ e $(x'_1, \cdots, x'_{n-1})$ em A^{n-1} e seja $x_n \in A$, logo, temos

$$d((x'_1, \cdots, x'_{n-1}), (x_1, \cdots, x_{n-1})) = d((x'_1, \cdots, x'_{n-1}, x_n), (x_1, \cdots, x_{n-1}, x_n)) =$$
$$= d(G(x'_1, \cdots, x'_{n-1}, x_n), G(x_1, \cdots, x_{n-1}, x_n)) =$$
$$= d((H(x'_1, \cdots, x'_{n-1}), f(x_n)), (H(x_1, \cdots, x_{n-1}), f(x_n))) =$$
$$= d(H(x'_1, \cdots, x'_{n-1}), H(x_1, \cdots, x_{n-1})),$$

provando assim que H é uma isometria. □

Teorema 2.17 *Seja $F : A^n \to A^n$ uma isometria, então existem uma permutação π de $\{1, \cdots, n\}$ e bijeções f_i de A, $i = 1, \cdots, n$, tais que*

$$F = T_\pi \circ T^1_{f_1} \circ \cdots \circ T^n_{f_n}.$$

Demonstração: A demonstração será feita por indução sobre n. Se $n = 1$, o resultado segue trivialmente. Suponhamos $n > 1$ e que o resultado vale para $n - 1$. Sejam $a_1, \cdots, a_{n-1} \in A$. Pelo Lema 2.15, existem $a'_1 \cdots, a'_{n-1} \in A$, uma bijeção $f_n : A \longrightarrow A$ e uma permutação σ de $\{1, \cdots, n\}$ tais que

$$(T_\sigma \circ F)(a_1, \cdots, a_{n-1}, x) = (a'_1, \cdots, a'_{n-1}, f_n(x)), \ \forall x \in A.$$

Pelo Lema 2.16, existe uma isometria H de A^{n-1} tal que

$$(T_\sigma \circ F)(x_1, \cdots, x_n) = (H(x_1, \cdots, x_{n-1}), f_n(x)). \tag{2.5}$$

Pela hipótese de indução, temos que existe uma permutação τ' de $1, \cdots, n-1$ e bijeções $f_1, \cdots, f_{n-1}$ de A tais que

$$H = (T_{\tau'})' \circ (T^1_{f_1})' \circ \cdots \circ (T^{n-1}_{f_{n-1}})', \tag{2.6}$$

onde

$$\begin{array}{rcl} (T_{\tau'})' : A^{n-1} & \longrightarrow & A^{n-1} \\ (x_1, \cdots, x_{n-1}) & \longmapsto & (x_{\tau'(1)}, \cdots, x_{\tau'(n-1)}) \end{array}$$

e, para $i = 1, \cdots, n-1$,

$$\begin{array}{rcl} (T^i_{f_i})' : A^{n-1} & \longrightarrow & A^{n-1} \\ (x_1, \cdots, x_{n-1}) & \longmapsto & (x_1, \cdots, f_i(x_i), \cdots, x_{n-1}). \end{array}$$

Definimos a permutação τ de $\{1, \cdots, n\}$ como segue:

$$\tau(i) = \begin{cases} \tau'(i) & se \quad 1 \leq i \leq n-1 \\ n & se \quad i = n \end{cases}$$

e ponhamos

$$\begin{array}{rcl} T_\tau : A^n & \longrightarrow & A^n \\ (x_1, \cdots, x_n) & \longmapsto & (x_{\tau(1)}, \cdots, x_{\tau(n)}) \end{array}$$

Para $i = 1, \cdots, n$, ponhamos

$$\begin{array}{rcl} T^i_{f_i} : A^n & \longrightarrow & A^n \\ (x_1, \cdots, x_n) & \longmapsto & (x_1, \cdots, f_i(x_i), \cdots, x_n). \end{array}$$

Segue de (2.5) e (2.6) que

$$T_\sigma \circ F = T_\tau \circ T^1_{f_1} \circ \cdots \circ T^n_{f_n}.$$

Usando o fato de que $T^{-1}_\sigma = T_{\sigma^{-1}}$ e que $T_\sigma \circ T_{\sigma'} = T_{\sigma \circ \sigma'}$, e pondo $\pi = \sigma^{-1} \circ \tau$, temos que

$$F = T_\pi \circ T^1_{f_1} \circ \cdots \circ T^n_{f_n},$$

o que prova o resultado. □

Corolário 2.18 *Sejam C e C' dois códigos em A^n. Temos que C e C' são equivalentes se, e somente se, existem uma permutação π de $\{1, \cdots .n\}$ e bijeções $f_1, \cdots, f_n$ de A tais que*

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

Daí decorre o resultado que enunciaremos a seguir, e que usualmente é apresentado em textos sobre códigos como definição de códigos equivalentes.

Dois códigos de comprimento n sobre um alfabeto A, cujos elementos serão chamados de letras, são equivalentes se, e somente se, um deles pode ser obtido do outro mediante uma sequência de operações do tipo:

1. Substituição das letras numa dada posição fixa em todas as palavras do código por meio de uma bijeção de A.

2. Permutação das posições das letras em todas as palavras do código, mediante uma permutação fixa de $\{1, 2, \cdots, n\}$.

2.3 Códigos Lineares

A classe de códigos mais utilizada na prática é a chamada classe dos códigos lineares.

Denotaremos por K um corpo finito com q elementos tomado como alfabeto. Temos, portanto, para cada número natural n, um K-espaço vetorial K^n de dimensão n.

Definição 2.19 *Um código $C \subset K^n$ será chamado de **código linear** se for um subespaço vetorial de K^n.*

Todo código linear é por definição um espaço vetorial de dimensão finita. Seja k a dimensão do código C e seja $v_1, v_2, \cdots, v_k$ uma de suas bases, portanto, todo elemento de C se escreve de modo único na forma

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k,$$

onde os $\lambda_i, i = 1 \cdots, k$, são elementos de K. Segue daí que

$$|C| = q^k.$$

Definição 2.20 *Dado $x = (x_1, \cdots, x_n) \in K^n$, define-se o **peso** de x como sendo o número inteiro*

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

Em outras palavras, temos que $\omega(x) = d(x, 0)$, onde d representa a métrica de Hamming.

Definição 2.21 *O **peso** de um código linear C é o inteiro*

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

Proposição 2.22 *Seja $C \subset K^n$ um código com distância mínima d. Temos que*

(i) $\forall x, y \in K^n,\ d(x, y) = \omega(x - y)$.

(ii) $d = \omega(C)$.

Demonstração: O item (i) segue imediatamente das definições de métrica da Hamming e da de peso de um código. O item (ii) decorre do fato que, para todo par de elementos x, y em C com $x \neq y$, tem-se $z = x - y \in C \setminus \{0\}$ e $d(x, y) = \omega(z)$. □

Em virtude da Proposição 2.22(ii), a distância mínima de um código linear C será também chamada de **peso do código** C. Em álgebra linear, não é incomum se descrever subespaços vetoriais C de um espaço vetorial K^n como imagem ou núcleo de uma transformação linear. Observemos como se obtém a representação de C como uma tal imagem. Escolha uma base $v_1, \cdots, v_k$ de C e considere a aplicação linear

$$\begin{array}{rccc} T: & K^k & \longrightarrow & K^n \\ & x = (x_1, \cdots, x_k) & \longmapsto & x_1 v_1 + \cdots + x_k v_k \end{array}.$$

Temos que T é uma transformação linear injetora, tal que a imagem de T é C, ou em símbolos,

$$Im(T) = C.$$

Portanto, dar um código $C \subset K^n$ de dimensão k é equivalente a dar uma transformação linear injetora

$$T : K^k \longrightarrow K^n$$

e definir $C = Im(T)$. Essa é a forma paramétrica do subespaço C, pois os elementos de C são parametrizados pelos elementos x de K^k através de T, o que torna fácil gerar todos os elementos de C. Note que nessa representação é, porém, difícil decidir se um dado elementos v de K^n pertence ou não a C, pois, para tal, é necessário resolver o sistema de n equações nas k incógnitas $x_1, \cdots, x_k$ abaixo

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k = v.$$

Essa resolução, em geral, representa um custo computacional muito elevado.

Vejamos agora, como descrever um código C como núcleo de uma transformação linear. Tome um subespaço C' de K^n complementar de C, isto é,

$$C \oplus C' = K^n,$$

e considere a aplicação linear

$$\begin{array}{rccc} H: & C \oplus C' & \longrightarrow & K^{n-k} \\ & u \oplus v & \longmapsto & v \end{array}$$

cujo núcleo é precisamente C. Computacionalmente é muito mais simples determinar se um certo elementos $v \in K^n$ pertence ou não a C; para isto, basta verificar se $H(v)$ é ou não o vetor nulo de K^{n-k}, o que tem custo bem pequeno.

Definição 2.23 *Seja K um corpo finito. Dois códigos lineares C e C' são* ***linearmente equivalentes*** *se existir uma isometria dada por uma transformação linear $T : K^n \longrightarrow K^n$ tal que $T(C) = C'$.*

Dos Exemplos 2.13 e 2.14 e do Teorema 2.17, segue que dois códigos lineares C e C' em K^n são linearmentes equivalentes se, e somente se, existem uma permutação π de $\{1, \cdots, n\}$ e elementos $c_1, \cdots, c_n$ de $K \setminus \{0\}$ tais que

$$C' = \{(c_1 x_{\pi(1)}, \cdots, c_n x_{\pi(n)});\ (x_1, \cdots, x_n) \in C\}.$$

Daí decorre o resultado a seguir, que usualmente é utilizado em textos sobre códigos como definição de códigos lineares equivalentes.

Dois códigos lineares são linearmente equivalentes se, e somente se, cada um deles pode ser obtido do outro mediante uma sequência de operações do tipo:

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

2. Permutação das posições das letras em todas as palavras do código, mediante uma permutação fixa de $\{1, 2, \cdots, n\}$.

2.4 Códigos de Reed-Muller

Os Códigos de Reed-Muller (ou RM) são uma das famílias mais antigas e melhor compreendidas de códigos. Os códigos de Reed-Muller foram originalmente definidos sobre o corpo binário [11] [12], e posteriormente a definição foi estendida para corpos com q elementos, sendo então conhecido como código de Reed-Muller generalizado.

Definição 2.24 *Seja $\mathbb{F}_q$ um corpo finito com q elementos e $n \geq 1$ inteiro. Seja d inteiro tal que $1 \leq d \leq n(q-1)$. O código de Reed-Muller generalizado de ordem d é o seguinte subespaço do espaço $\mathbb{F}_q^{(q^n)}$:*

$$RM_q(d, n) = \{(f(x))_{x \in \mathbb{F}_q^n} \mid f \in \mathbb{F}_q[X_1, \cdots, X_n] \text{ e } deg(f) \leq d\}.$$

O código $RM_q(d, n)$ tem os seguintes parâmetros:

1. comprimento $m = q^n$,

2. dimensão $M = \sum_{t=0}^{d} \sum_{j=0}^{n} (-1)^j \binom{n}{j} \binom{t-jq+n-1}{t-jq}$

3. distância mínima $W_1 = (q - \ell)q^{n-k-1}$, onde k e ℓ são respectivamente o quociente e o resto na divisão euclidiana de d por $(q-1)$, ou seja $d = k(q-1) + \ell$ e $0 \leq \ell < q - 1$.

No próximo capítulo, utilizando conceitos da teoria de bases de Gröbner, vamos definir uma classe de códigos que tem como subclasse os códigos de Reed-Muller generalizados.

Capítulo 3

Bases de Gröbner e a Pegada de um Ideal

3.1 Monômios e Ordens Monomiais

Definição 3.1 *Um monômio em $X_1, \cdots, X_n$ é um produto da forma $X_1^{\alpha_1} \cdot ... \cdot X_n^{\alpha_n}$, onde todos os expoentes são inteiros não negativos. Vamos utilizar a notação multi-índice para monômios. Escreveremos $X^\alpha = X_1^{\alpha_1} \cdot ... \cdot X_n^{\alpha_n}$, onde $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$. O grau desse monômio é a soma $|\alpha| := \alpha_1 + ... + \alpha_n$.*

Seja $\mathbb{K}$ um corpo, vamos denotar por $\mathcal{M}$ o conjunto de todos os monômios em $\mathbb{K}[X] := \mathbb{K}[X_1, ..., X_n]$.

Um polinômio f em $\mathbb{K}[X]$ é uma combinação linear de monômios com coeficientes em $\mathbb{K}$. Dessa forma podemos escrever

$$f = \sum_\alpha a_\alpha X^\alpha, \ a_\alpha \in \mathbb{K}.$$

que é uma soma de um número finito de n-uplas $\alpha = (\alpha_1, ..., \alpha_n)$. Chamamos a_α de coeficiente do monômio X^α. Se $a_\alpha \neq 0$, então $a_\alpha X^\alpha$ é um termo de f. O grau total de f, denotado por $deg(f)$, é o máximo $|\alpha|$ tal que $a_\alpha \neq 0$.

Definição 3.2

(i) Sejam $f_1, ..., f_s$ polinômios em $\mathbb{K}$. Definimos o ideal gerado por $f_1, ..., f_s$ como sendo o conjunto

$$(f_1, ..., f_s) = \left\{ \sum_{i=1}^{s} h_i f_i \ : \ h_1, ..., h_s \in \mathbb{K}[X] \right\}.$$

Observe que esse conjunto é de fato um ideal de $\mathbb{K}[X]$.

(ii) Chamamos um ideal I de $\mathbb{K}[X]$ de finitamente gerado se existem $f_1, ..., f_s \in \mathbb{K}[X]$ tais que $I = (f_1, ..., f_s)$, e dizemos que $\{f_1, ..., f_s\}$ é uma base para I.

Definição 3.3 *Uma ordem monomial $\preceq$ em $\mathcal{M} \subset \mathbb{K}[X]$ é qualquer relação em $\mathbb{N}^n$ satisfazendo:*

(i) $\preceq$ é uma ordem total em $\mathcal{M}$;

(ii) se $X^\alpha \preceq X^\beta$ em $\mathcal{M}$ e $X^\gamma \in \mathcal{M}$, então $X^\alpha \cdot X^\gamma \preceq X^\beta \cdot X^\gamma$;

(ii) todo subconjunto não vazio de $\mathcal{M}$ possui elemento mínimo em relação a $\preceq$.

Vejamos dois exemplos de ordens monomiais:

Exemplo 3.4

(i) A ordem lexicográfica (com $X_n \preceq ... \preceq X_1$) é definida por $X^\alpha \preceq_{lex} X^\beta$ se $\alpha = \beta$ ou se a primeira entrada diferente de zero da esquerda para a direita de $\beta - \alpha$ for positiva. Assim, nós temos por exemplo que $X_2^{1000} \preceq_{lex} X_1$ e $X_1^2 X_3^{2015} \preceq_{lex} X_1^2 X_2$.

(ii) A ordem lexicográfica graduada (com $X_n \preceq ... \preceq X_1$) é definida por $X^\alpha \preceq X^\beta$ se $\alpha = \beta$ ou $\sum_{i=1}^n \alpha_1 < \sum_{i=1}^n \beta_i$, ou se $\sum_{i=1}^n \alpha_1 = \sum_{i=1}^n \beta_i$ então $X^\alpha \preceq_{lex} X^\beta$, onde $\preceq_{lex}$ é a ordem lexicográfica, definida anteriormente.

Temos por exemplo que $X_1^3 X_2^2 \preceq X_1 X_2^2 X_3^3$, pois o grau do primeiro é menor que o grau do segundo, e $X_1 X_2 X_3^5 \preceq X_1 X_2^2 X_3^4$, pois embora eles possuam o mesmo grau, temos que $X_1 X_2 X_3^5 \preceq_{lex} X_1 X_2^2 X_3^4$.

Definição 3.5 *Seja $f = \sum_{i=1}^m a_i x^{\alpha_i} \in \mathbb{K}[X]$ um polinômio não nulo, onde $a_i \in \mathbb{K}$, $a_i \neq 0$ para todo $i = 1, \cdots, m$, e seja $\preceq$ uma ordem monomial definida em $\mathcal{M}$. Então:*

*i) O **Multigrau** de f (com respeito a $\preceq$) é dado por $mdeg(f) := max\{\alpha_i \in \mathbb{N}^n : i = 1, \cdots, m\}$;*

*ii) O **Monômio Líder** de f (com respeito a $\preceq$) é $lm(f) := X^{mdeg(f)}$;*

*iii) O **Coeficiente Líder** de f (com respeito a $\preceq$) é $lc(f) := a_{mdeg(f)}$;*

*iv) O **Termo Líder** de f (com respeito a $\preceq$) é $lt(f) := a_{mdeg(f)} X^{mdeg(f)}$.*

Assim, por exemplo, se $f(X_1, X_2, X_3) = 4X_1^3 X_2^4 + 5X_1 X_3^8 + 2 \in \mathbb{R}[X_1, X_2, X_3]$ e nós consideramos o conjunto dos monômios com a ordem lexicográfica, então temos que $lm(f) = X_1^3 X_2^4$ e $lt(f) = 4X_1^3 X_2^4$, por outro lado, se considerarmos a ordem lexicográfica graduada, teremos que $lm(f) = X_1 X_3^8$ e $lt(f) = 5X_1 X_3^8$.

Um importante procedimento na teoria de bases de Gröbner é a divisão de um polinômio por uma lista de polinômio não nulos.

Definição 3.6 *Dividir $f \in \mathbb{K}[X]$ por $\{g_1, \cdots, g_t\} \subset \mathbb{K}[X] \setminus \{0\}$, com respeito a uma ordem monomial $\preceq$, significa encontrar quocientes $q_1, \cdots, q_t$ e um resto r em $\mathbb{K}[X]$ tais que $f = q_1 g_1 + \cdots + q_t g_t + r$, e também $r = 0$ ou nenhum monômio que aparece em r é múltiplo de $lm(g_i)$, para todo $i \in \{1, \cdots, t\}$.*

Na literatura sobre Bases de Gröbner o leitor poderá encontrar uma descrição do algoritmo usado para determinar os quocientes e o resto, bem como uma prova de que o algoritmo termina, de fato, depois de um número finito de passos. Aqui nós apenas descrevemos o algoritmo e mostramos como ele funciona em dois exemplos.

A ideia básica do algoritmo é a mesma do caso de uma variável: queremos cancelar o termo líder de f (com respeito a $\preceq$) pela multiplicação de algum g_i por um apropriado monômio e subtrair. A novidade aqui é que às vezes o termo líder de um "polinômio intermediário" não é um múltiplo de nenhum $lm(g_1), \cdots, lm(g_t)$, então devemos movê-lo para o resto para continuar com a divisão.

Exemplo 3.7 *Primeiramente vamos dividir $f = XY^2 + 1$ por $g_1 = XY + 1$ e $g_2 = Y + 1$, usando a ordem lexicográfica com $Y \preceq X$. Listando os divisores de g_1 e g_2 na chave e os quocientes q_1 e q_2 abaixo da chave, temos o seguinte esquema:*

$$\begin{array}{r|l} XY^2+1 & XY+1 \\ & Y+1 \\ \hline & q_1: \\ & q_2: \end{array}$$

Os termos líderes são $lt(g_1) = XY$ *e* $lt(g_2) = Y$*, e ambos dividem o* $lt(f) = XY^2$*. Então dividindo* f *por* g_1*, temos que dividir* XY^2 *por* XY*, assim, basta multiplicar* g_1 *por* Y *e depois subtrair* g_1 *de* f *obtendo:*

$$\begin{array}{r|l} XY^2+1 & XY+1 \\ & Y+1 \\ \hline -Y+1 & q_1: Y \\ & q_2: \end{array}$$

ou seja

$$XY^2+1 = Y(XY+1) + 0(Y+1) + (-Y+1).$$

Agora repetimos o mesmo processo para $-Y+1$*. Desta fez devemos dividir por* g_2 *pois* $lt(g_1) = XY$ *não divide o* $lt(-Y+1) = -Y$*. Daí*

$$\begin{array}{r|l} XY^2+1 & XY+1 \\ & Y+1 \\ \hline -Y+1 & q_1: Y \\ 2 & q_2: -1 \end{array}$$

Notando que o $lt(g_1)$ *e o* $lt(g_2)$ *não dividem 2, o resto é igual a* $r = 2$ *e acabamos. Desse modo escrevemos* $f = XY^2 + 1$ *da forma:*

$$XY^2+1 = Y(XY+1) + (-1)(Y+1) + 2.$$

Exemplo 3.8 *Neste exemplo, encontraremos uma inesperada sutileza que pode ocorrer quando trabalhamos com polinômios de mais de uma variável. Vamos dividir* $f = X^2Y + XY^2 + Y$ *por* $g_1 = XY - 1$ *e* $g_2 = Y^2 - 1$*. Como no exemplo anterior, usaremos a ordenação lexicográfica, com* $Y \preceq X$*. As duas primeiras etapas do algoritmo seguem abaixo (notando que quando os dois termos lideres dividem f, usamos o primeiro):*

$$\begin{array}{r|l} X^2Y + XY^2 + Y & XY-1 \\ & Y^2-1 \\ \hline XY^2 + X + Y^2 & q_1: X+Y \\ X + Y^2 + Y & q_2: \end{array}$$

ou seja

$$X^2Y + XY^2 + Y = (X+Y)(XY-1) + 0(Y^2-1) + (X + Y^2 + Y).$$

Note que nenhum dos dois $lt(g_1) = XY$ *e* $lt(g_2) = Y^2$ *dividem o* $lt(X + Y^2 + Y) = X$*. Entretanto,* $X + Y^2 + Y$ *não é o resto, pois* $lt(g_2)$ *divide* Y^2*. Deste modo, se nós movermos* X *para o resto podemos continuar dividindo.*

Para implementar essa ideia, criaremos uma coluna para o resto r*, a esquerda dos dividendos, onde colocaremos os termos pertencentes ao resto.*

Além disso, chamaremos os polinômios debaixo do dividendo, de dividendo intermediário, e continuaremos dividindo até que ele se anule.

r		$X^2Y + XY^2 + Y$	$XY - 1$
			$Y^2 - 1$
		$XY^2 + X + Y^2$	$q_1 : X + Y$
X	$\leftarrow$	$X + Y^2 + Y$	$q_2 : 1$
		$Y^2 + Y$	
$Y + 1$	$\leftarrow$	$Y + 1$	
$r = X + Y + 1$		0	

Assim, o resto é $X + Y + 1$*, e obtemos*

$$X^2Y + XY^2 + Y^2 = (X + Y) \cdot (XY - 1) + 1 \cdot (Y^2 - 1) + (X + Y + 1).$$

Note que o resto é uma soma de monômios, na qual nenhum dos monômios é divisível pelos termos lideres de g_1 *e* g_2*.*

É importante observar que no algoritmo da divisão temos que, se o resto r *é diferente de zero, então o monômio líder de* r *é menor ou igual ao monômio líder de* f*.*

Além disso, olhando atentamente para o algoritmo observamos que estamos levando em conta a ordem em que os divisores $g_1, \cdots, g_t$ *são escritos e podemos perguntar se uma mudança nesta ordem irá produzir uma mudança nos quocientes e no resto. A resposta à esta pergunta é sim, e pode-se verificar que a aplicação do procedimento acima para dividir* $X^2Y + XY + 2 + Y^2$ *por* $\{Y^2 - 1, XY - 1\}$*(nesta ordem) nos dá que*

$$X^2Y + XY^2 + Y^2 = (X + 1) \cdot (Y^2 - 1) + X \cdot (XY - 1) + (2X + 1).$$

3.2 Bases de Gröbner

O conceito de bases de Gröbner apareceu pela primeira vez na tese de doutorado do matemático austríaco Bruno Buchberger, publicada em 1965. Seu orientador, Wolfgang Gröbner, propôs o seguinte problema para sua tese: dado um ideal $I \subset \mathbb{K}[X]$, encontrar uma base para $\mathbb{K}[X]/I$ como um $\mathbb{K}$-espaço vetorial.

Quando estamos trabalhando com um anel de polinômios em uma única variável a resposta é conhecida: I é gerado por um certo polinômio de grau d(no caso em que $I \neq 0$) e $\{1 + I, X + I, \cdots, X^{d-1} + I\}$ forma uma base para $\mathbb{K}[X]/I$.

Agora, quando estamos trabalhando com um anel de mais de uma variável a situação muda radicalmente. Pelo Teorema da Base de Hilbert, nós sabemos que I é gerado por um número finito de polinômios, mas I não é necessariamente um ideal principal; mais ainda o anel quociente $\mathbb{K}[X]/I$ pode ser um $\mathbb{K}$-espaço vetorial de dimensão infinita.

A ideia de Buchberger para solucionar o problema acima foi fixar uma ordem monomial em $\mathcal{M}$ e determinar um conjunto especial de geradores para I cuja propriedade principal é que as classes dos monômios que não são múltiplos de nenhum dos monômios líderes dos polinômios que estão nessa base especial, formam uma base para $\mathbb{K}[X]/I$ como $\mathbb{K}$-espaço vetorial. Em 1976 Buchberger decidiu chamar essa base especial para I de Base de Gröbner em virtude da influência das ideias de seu orientador em seu trabalho de tese.

Definição 3.9 *Seja* $I \subset \mathbb{K}[X]$ *um ideal não nulo e seja* $\mathcal{M}$ *com a ordem monomial* $\preceq$*. Um conjunto* $\{g_1, \cdots, g_s\} \subset I$ *é uma* ***Base de Gröbner*** *para* I *(com respeito a* $\preceq$*) se para todo* $f \in I$*,* $f \neq 0$*, temos que* $lm(f)$ *é um múltiplo de* $lm(g_i)$ *para algum* $i \in \{1, \cdots, s\}$*.*

Exemplo 3.10 *Seja* $I = (XY - 1, Y^2 - 1) \subset \mathbb{R}[X, Y]$ *e considere a ordem lexicográfica (com* $Y \preceq X$*) definida no conjunto dos monômios de* $\mathbb{R}[X, Y]$*. Então* $Y(XY - 1) - X(Y^2 - 1) = -Y + X \in I$ *e* $lm(X - Y) = X$ *não é um múltiplo de* $lm(XY - 1) = XY$ *ou* $lm(Y^2 - 1) = Y^2$*, assim* $\{XY - 1, Y^2 - 1\}$ *não é uma Base de Gröbner para* I*.*

Assumimos a partir de agora que $\mathcal{M}$ é dotado com alguma ordem monomial fixa e que $I \neq (0)$. O seguinte resultado mostra que uma Base de Gröbner para I é de fato uma base para I, no sentido de ser um conjunto de geradores, e que podemos utilizar isso para decidir se um dado polinômio está em I.

Lema 3.11 *Seja* $\{g_1, \cdots, g_s\} \subset I$ *uma Base de Gröbner para* I*, então* $f \in I$ *se, e somente se, o resto da divisão de* f *por* $\{g_1, \cdots, g_s\}$ *é zero. Consequentemente* $I = (g_1, \cdots, g_s)$*.*

Demonstração: A "volta"dessa afirmação é trivial. Por outro lado para $f \in I$ seja $f = \sum_{i=1}^{s} q_i g_i + s$ a divisão de f por $\{g_1, \cdots, g_s\}$. Então, como $r = f - \sum_{i=1}^{s} q_i g_i \in I$, devemos ter $r = 0$, caso contrário r seria um polinômio não nulo cujo monômio líder não é um múltiplo de $lm(g_i)$ para algum $i = 1, \cdots, s$, contradizendo o fato que $\{g_1, \cdots, g_s\}$ é uma base de Gröbner para I. Isso mostra que $I \subset (g_1, \cdots, g_s)$ e consequentemente $I = (g_1, \cdots, g_s)$. □

Uma importante propriedade das Bases de Gröbner é a seguinte.

Proposição 3.12 *Seja* $\{g_1, \cdots, g_s\} \subset I$ *uma Base de Gröbner para* I*. Na divisão de* $f \in \mathbb{K}[X]$ *por* $\{g_1, \cdots, g_s\}$ *o resto é sempre o mesmo, independentemente da ordem que escolhemos para* $g_1, \cdots, g_s$ *no algoritmo da divisão.*

Demonstração: Assuma que $f = q_1 g_1 + \cdots + q_s g_s + r = q_1' g_1 + \cdots + q_s' g_s + r'$, onde $q_i, q_i' \in \mathbb{K}[X]$ para todo $i = 1, \cdots, s$, $r, r' \in \mathbb{K}[X]$ e não aparecendo em r ou r' um múltiplo de $lm(g_i)$ para todo $i = 1, \cdots, s$. De $r - r' = \sum_{i=1}^{s} (q_i' - q_i) g_i \in I$ devemos ter $r - r' = 0$ caso contrário $r - r'$ seria um polinômio não nulo cujo monômio líder não é um múltiplo do $lm(g_i)$ para algum $i = 1, \cdots, s$, contradizendo o fato de que $\{g_1, \cdots, g_s\}$ é uma Base de Gröbner para I. □

Os resultados acima listam algumas propriedades agradáveis de Bases de Gröbner, mas até agora não está claro se todo ideal $I \subset \mathbb{K}[X]$ admite tal base. Esta parte é a principal contribuição de Buchberger em sua tese. Lá, ele apresenta um algoritmo, o qual parte de uma base finita dada para I, e acrescenta novos elementos, se necessário, até o ponto em que a base aumentada seja uma Base de Gröbner. A seguir, um conceito chave no Algoritmo de Buchberger.

Definição 3.13 *Sejam* $f, g \in \mathbb{K}[X] - \{0\}$*, com* $lt(f) = aX^{\alpha}$ *e* $lt(g) = bX^{\beta}$*, com* $\alpha = (\alpha_1, \cdots, \alpha_n), \beta = (\beta_1, \cdots, \beta_n) \in \mathbb{N}^n$*. Seja ainda* $\gamma_i = \max\{\alpha_i, \beta_i\}$*, para* $i = 1, \cdots, n$ *e* $\gamma = (\gamma_1, \cdots, \gamma_n) \in \mathbb{N}^n$*. O* ***S-polinômio*** *de* f *e* g *é definido por:*

$$S(f, g) = \frac{X^{\gamma - \alpha}}{a} f - \frac{X^{\gamma - \beta}}{b} g.$$

Observe que $lt((1/a)X^{\gamma - \alpha} f) = X^{\gamma} = lt((1/b)X^{\gamma - \beta} g)$. Buchberger provou que $\{g_1, \cdots, g_s\} \subset I$ é uma Base de Gröbner para I se, e somente se, o resto da divisão de $S(g_i, g_j)$ por $\{g_1, \cdots, g_s\}$ é zero para todos distintos $i, j = \{1, \cdots, s\}$. Ele também provou que o seguinte procedimento pode ser usado em um algoritmo que produz uma Base de Gröbner para $I = (g_1, \cdots, g_s)$ em um número finito de passos: assuma que para algum par de distintos inteiros $i, j \in \{1, \cdots, s\}$ o resto $R_{i,j}$ da divisão de $S(g_i, g_j)$ por $\{g_1, \cdots, g_s\}$ é diferente de zero. Defina $g_{s+1} := R_{i,j}$ e considere o conjunto $\{g_1, \cdots, g_s, g_{s+1}\}$. Claramente $I = (g_1, \cdots, g_s, g_{s+1})$ pois $g_{s+1} \in I$. Se para algum par

de inteiros $i, j \in \{1, \cdots, s+1\}$ o resto $R_{i,j}$ da divisão de $S(g_i, g_j)$ por $\{g_1, \cdots, g_{s+1}\}$ é diferente de zero então defina $g_{s+2} := R_{i,j}$ e considere o conjunto $\{g_1, \cdots, g_{s+1}, g_{s+2}\}$. Buchberger provou que depois de um número finito de passos, os restos nas divisões serão sempre iguais a zero e assim esse processo produzirá um conjunto $\{g_1, \cdots, g_t\}$ que é uma Base de Gröbner para I.

Exemplo 3.14 *Vimos no exemplo 3.10 que* $\{XY-1, Y^2-1\}$ *não é uma Base de Gröbner para* $I = (XY-1, Y^2-1) \subset \mathbb{R}[X,Y]$ *com respeito a ordem lexicográfica onde* $Y \preceq X$. *Vamos aplicar o Algoritmo de Buchberger a fim encontrar uma Base de Gröbner para* I. *Seja* $g_1 = XY - 1$ *e* $g_2 = Y^2 - 1$, *então* $S(g_1, g_2) = Yg_1 - Xg_2 = X - Y$ *e o resto da divisão de* $S(g_1, g_2)$ *por* $\{XY - 1, Y^2 - 1, X - Y\}$ *é zero. Pode-se também facilmente verificar que o resto da divisão de* $S(g_1, g_3) = Y^2 - 1$ *e* $S(g_2, g_3) = Y^3 - x$ *por* $\{XY - 1, Y^2 - 1, X - Y\}$ *é zero, então* $\{XY - 1, Y^2 - 1, X - Y\}$ *é uma Base de Gröbner para* I *(com respeito a* $\preceq$*).*

Introduziremos agora o conceito que resolveu o problema da tese de Buchberger.

Definição 3.15 *Seja* $I \subset \mathbb{K}[X]$ *um ideal. A* ***pegada*** *de* I *(com respeito a uma ordem monomial em* $\mathcal{M}$ *fixada) é o conjunto:*

$$\Delta(I) = \{M \in \mathcal{M} \ : \ M \text{ não é monômio líder de nenhum polinômio em } I\}.$$

A pegada de um ideal I tem uma correlação com uma Base de Gröbner para I (sendo ambos definidos em relação à mesma ordem monomial em $\mathcal{M}$) .

Proposição 3.16 *Seja* $I \subset \mathbb{K}[X]$ *um ideal e seja* $\{g_1, \cdots, g_s\}$ *uma Base de Gröbner para* I. *Então, um monômio* M *está em* $\Delta(I)$ *se, e somente se,* M *não é múltiplo de* $lm(g_i)$, *para todo* $i = 1, \cdots, s$.

Demonstração: A "ida"é trivial pela definição de $\Delta(I)$. Por outro lado, da definição de Base de Gröbner sabemos que se M não é múltiplo de $lm(g_i)$ para todo $i = 1, \cdots, s$ então M não é o monômio líder de nenhum polinômio em I. □

A prova acima é muito simples e utiliza a definição de $\Delta(I)$ em uma direção e a definição de base de Gröbner na outra. Isso sugere que os conceitos de Base de Gröbner e Pegada podem ser equivalentes, e de fato são. Tendo definido o que é uma Base Gröbner para um ideal I podemos definir a pegada de I usando a afirmação da proposição acima. Por outro lado, podemos começar com a definição 3.15 e então definir uma Base de Gröbner para I como sendo um conjunto $\{g_1, \cdots, g_s\} \subset I$ tal que o conjunto dos monômios que são múltiplos de $lm(g_i)$ para algum $i \in \{1, \cdots, s\}$ é exatamente $\mathcal{M} \setminus \Delta(I)$. Em seguida, pode-se provar que tal conjunto $\{g_1, \cdots, g_s\}$ de fato existe e satisfaz a condição da Definição 3.9.

No exemplo a seguir mostramos como usar o resultado acima para obter uma representação gráfica da pegada de um ideal.

Exemplo 3.17 *Seja* $I = (X^3 - X, Y^3 - Y, X^Y - Y) \subset \mathbb{R}[X,Y]$, *e tome* $\mathcal{M}$ *com ordem lexicográfica, onde* $Y \preccurlyeq X$. *Não é difícil de checar que* $\{X^3 - X, Y^3 - Y, X^Y - Y\}$ *é uma Base de Gröbner para* I. *Temos que* $lm(X^3 - X) = X^3$, $lm(Y^3 - Y) = Y^3$ *e* $lm(X^2Y - Y) = X^2Y$, *e aplicamos a proposição acima para determinar* $\Delta(I)$. *É fácil "ver"a pegada de* I *na figura abaixo, onde representamos o monômio* $X^\alpha Y^\beta$ *pelo par de inteiros não negativos* (α, β).
Observe que as bolinhas "fechadas"representam os monômios líderes da Base de Gröbner para I, *e as bolinhas "abertas"representam os monômios da pegada de* I.

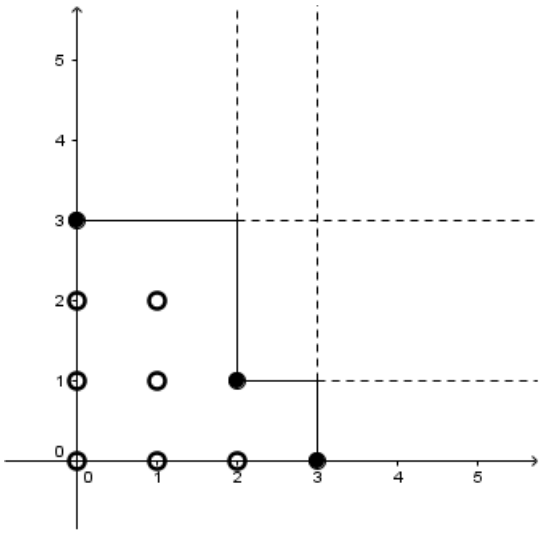

Figura 3.1: Representação gráfica da pegada de um ideal

De fato, os pontos $(3,0)$, $(0,3)$ *e* $(2,1)$ *correspondem aos monômios líderes da Base de Gröbner e assim fica fácil determinar os monômios que são múltiplos de cada um desses monômios líderes (assim determinando o conjunto dos monômios que são monômios líderes de polinômios em* I*). Desse conjunto e do resultado acima temos que* $\Delta(I) = \{1, X, X^2, Y, XY, Y^2, XY^2\}$.

Agora apresentamos a solução para o problema tese de Buchberger, que será muito útil na próxima seção.

Teorema 3.18 *Seja* $I \subset \mathbb{K}[X]$. *Então*

$$\mathcal{B} := \{M + I \mid M \in \Delta(I)\}$$

é uma base para $\mathbb{K}[X]/I$ *como um* $\mathbb{K}$*-espaço vetorial.*

Demonstração: Seja $\mathcal{G}$ uma Base de Bröbner para I com respeito à mesma ordem monomial usada para determinar $\Delta(I)$, e seja $f \in \mathbb{K}[X]$. Dividindo f por $\mathcal{G}$ temos que o resto é da forma $r = \sum_{i=1}^{t} a_i M_i$ onde $a_i \in \mathbb{K}[X]$ e $M_i \in \Delta(I)$ para todo $i = 1, \cdots, t$. Uma vez que $f + I = r + I$ temos que $\mathcal{B}$ gera $\mathbb{K}[X]/I$ como uma $\mathbb{K}$-espaço vetorial. Agora assuma que $\sum_{i=1}^{\ell} b_i(M_i + I) = 0 + I$ onde $b_i \in \mathbb{K}[X]$ para todo $i = 1, \cdots, \ell$. Então $\sum_{i=1}^{\ell} b_i M_i \in I$ e assim devemos ter $b_i = 0$ para todo $i = 1, \cdots, \ell$, caso contrário $\sum_{i=1}^{\ell} b_i M_i$ poderia ser um elemento não nulo de I cujo monômio líder não é um monômio líder de um polinômio em I. Isso mostra que $\mathcal{B}$ é um conjunto linearmente independente sobre $\mathbb{K}$. □

Exemplo 3.19 *Continuando com o que foi feito no exemplo 3.17, do resultado acima temos que* $\mathbb{R}[X,Y]/I$ *é um* $\mathbb{R}$*-espaço vetorial de dimensão* 7 *e* $\{1 + I, X + I, X^2 + I, Y + I, XY + I, Y^2 + I, XY^2 + I\}$ *é uma base para esse espaço vetorial.*

Seja $I \subset \mathbb{K}[X]$ um ideal e seja o conjunto $\{f_1, \cdots, f_t\}$ uma base para I. Denotaremos por $\Delta(lm(f_1), \cdots, lm(f_t))$ o conjunto:

$$\Delta(lm(f_1), \cdots, lm(f_t)) := \{M \in \mathcal{M} \mid M \text{ não é múltiplo de } lm(f_i), \forall i \in \{1, \cdots, t\}\}.$$

Observação 3.20 *Observe que* $\Delta(I) \subset \Delta(lm(f_1), \cdots, lm(f_t))$. *Na verdade, pela proposição 3.16 temos que* $\Delta(I) = \Delta(lm(f_1), \cdots, lm(f_t))$ *se, e somente se,* $\{f_1, \cdots, f_t\}$ *é uma Base de Gröbner para* I.

Definição 3.21 *Sejam $f, g \in \mathbb{K}[X] - \{0\}$. Se $mdeg(f) = \alpha$ e $mdeg(g) = \beta$, então seja $\gamma = (\gamma_1, \cdots, \gamma_n)$ onde $\gamma_i = \max(\alpha_i, \beta_i)$ para todo $i = 1, \cdots, n$. Chamamos X^γ de* ***mínimo múltiplo comum*** *de $lm(f)$ e $lm(g)$, e escrevemos $X^\gamma = mmc(lm(f), lm(g))$.*

Anunciamos a seguir um teorema importante que será necessário no próximo capítulo.

Teorema 3.22 *Dado um conjunto finito $G \subset \mathbb{K}[X]$ suponha que tenhamos $f, g \in G$ tais que*

$$mmc(lm(f), lm(g)) = lm(f) \cdot lm(g).$$

Isso significa que os monômios líderes de f e g são relativamente primos. Então o resto da divisão de $S(f, g)$ por G é zero.

Demonstração: Para simplificar, assumimos que f e g foram multiplicadas por constantes apropriadas para termos $lc(f) = lc(g) = 1$. Escreva $f = lm(f) + p$, $g = lm(g) + q$. Então, como $mmc(lm(f), lm(g)) = lm(f) \cdot lm(g)$, temos

$$\begin{array}{rcl} S(f,g) & = & lm(g) \cdot f - lm(f) \cdot g \\ & = & (g - q) \cdot f - (f - p) \cdot g \\ & = & g \cdot f - q \cdot f - f \cdot g + p \cdot g \\ & = & p \cdot g - q \cdot f \end{array} \tag{3.1}$$

Afirmamos que

$$mdeg(S(f,g)) = max(mdeg(p \cdot g), mdeg(q \cdot f)). \tag{3.2}$$

Note que (3.1) e (3.2) implicam que o resto da divisão de $S(f, g)$ por G é zero, desde que $f, g \in G$. Para provar (3.2) observe que no último polinômio de (3.1), os monômios líderes de $p \cdot g$ e $q \cdot f$ são distintos e, por isso, não é possível cancelá-los. Pois se os monômios líderes fossem os mesmos, teríamos

$$lm(p) \cdot lm(g) = lm(q) \cdot lm(f).$$

No entanto, isso é impossível se $lm(f)$ e $lm(g)$ são primos relativos: da última equação, $lm(g)$ teria que dividir $lm(q)$, o que é absurdo pois $lm(q) \preceq lm(g)$. □

3.3 Códigos Cartesianos Afins e seus Parâmetros

Começamos apresentando um conceito chave na geometria algébrica, que dá uma iteração entre álgebra e geometria.

Definição 3.23 *Seja $I \subset \mathbb{K}[X]$ um ideal. A* ***Variedade Afim*** *associada ao ideal I é o conjunto:*

$$V(I) = \{(a_1, \cdots, a_n) \in \mathbb{K}^n \mid f(a_1, \cdots, a_n) = 0, \forall f \in I\}.$$

É fácil ver que se $I = (g_1, \cdots, g_t)$ então $(a_1, \cdots, a_n) \in V(I)$ se, e somente se, $g_i(a_1, \cdots, g_n) = 0$, para todo $i = 1, \cdots, t$.

Dado $V = V(I)$ podemos nos perguntar sobre o conjunto de todos os polinômios que se anulam em V. É fácil ver que esse conjunto é um ideal de $\mathbb{K}[X]$ que contém I, e é chamado de **Ideal das Variedades de** V e denotado por $\mathcal{I}(V)$. Um famoso teorema de Hilbert afirma se $\mathbb{K}$ é algebricamente fechado então $\mathcal{I}(V(I)) = \sqrt{I}$, onde $\sqrt{I} := \{f \in \mathbb{K}[X] \mid f^m \in I \text{ para algum } m \in \mathbb{N}\}$ é o ideal chamado de **Radical de I**.

Uma variedade $V(I)$ pode ter infinitos pontos (tome por exemplo $I = (x - y^2) \subset \mathbb{R}[X, Y]$) ou um número finito de pontos (tome por exemplo $I = (X^2 - 1, Y^2 - 1) \subset \mathbb{R}(X, Y)$). Para provar uma importante relação entre variedades de I e a pegada de I quando $\Delta(I)$ é finito precisamos do seguinte resultado auxiliar.

Lema 3.24 *Seja $I \subset \mathbb{K}[X]$ um ideal e seja $P_1, \cdots, P_r$ pontos distintos de $V(I)$. Então, existem polinômios $p_1, \cdots, p_r \in \mathbb{K}[X]$ tais que $p_i(P_j) = \delta_{ij}$ para todo $i, j \in \{1, \cdots, r\}$.*

Demonstração: Seja $P_i = (a_{i1}, \cdots, a_{in}) \in \mathbb{K}^n$ onde $i = 1, \cdots, r$, vamos mostrar como obter p_1 satisfazendo o lema. Como todos os pontos são distintos, para $i \in \{2, \cdots, r\}$ existe $j_i \in \{1, \cdots, n\}$ tal que $a_{1j_i} \neq a_{ij_i}$. Seja $h_i = (X_{j_i} - a_{ij_i})/(a_{1j_i} - a_{ij_i})$, então $h_i(P_1) = 1$ e $h_i(P_i) = 0$ para todo $i = 2, \cdots, r$. Tomando $p_1 = \prod_{i=2}^{r} h_i$ temos $p_1(P_1 = 1)$ e $p_1(P_i) = 0$ para todo $i = 2, \cdots, r$. Da mesma forma obtemos $p_2, \cdots, p_r$ como no lema. □

Proposição 3.25 *Seja $I \subset \mathbb{K}[X]$ um ideal tal que $\Delta(I)$ é um conjunto finito. Então $V(I)$ também é um conjunto finito e $\#V(I) \leq \#\Delta(I)$.*

Demonstração: Sejam $P_1, \cdots, P_r$ elementos distintos de $V(I)$, vamos encontrar um conjunto linearmente independente em $\mathbb{K}[X]/I$ e que tenha r elementos. Isso provará a proposição pois como nós vimos $\#\Delta(I)$ é a dimensão de $\mathbb{K}[X]/I$ como $\mathbb{K}$-espaço vetorial. Pelo lema acima, sabemos que existem $p_1, \cdots, p_r$ em $\mathbb{K}[X]$ tais que $p_i(P_j) = \delta_{ij}$ para todo $i, j \in \{1, \cdots, r\}$. Assuma que $\sum_{i=1}^{r} a_i(p_i + i) = 0 + I$ onde $a_1, \cdots, a_r \in \mathbb{K}$, então $\sum_{i=1}^{r} a_i p_i \in I$ sempre que $\sum_{i=1}^{r} a_i p_i(P_j) = 0$, isto é, $a_j = 0$ para todo $j \in \{1, \cdots, r\}$. Assim $\{p_1 + I, \cdots, p_r + I\}$ é um conjunto linearmente independente em $\mathbb{K}[X]/I$, completando a prova. □

Na verdade pode-se provar um resultado mais refinado (veja [1], Thm. 8.32). Recorde que um ideal I é chamado de ideal radical se $I = \sqrt{I}$.

Teorema 3.26 *Seja $I \subset \mathbb{K}[X]$ um ideal tal que $\Delta(I)$ é um conjunto finito e seja L uma extensão algebricamente fechada de $\mathbb{K}$. Então $V_L(I) := \{(a_1, \cdots, a_n) \in L^n \mid f(a_1, \cdots, a_n) = 0, \forall f \in I\}$ é um conjunto finito e $\#V_L(I) \leq \#\Delta(I)$. Mais ainda, se $\mathbb{K}$ é um corpo perfeito e I é um ideal radical então $\#V_L(I) = \#\Delta(I)$.*

Em 1998, Fitzgerald e Lax propuseram a seguinte construção de códigos lineares. Seja $I = (g_1, \cdots, g_t) \subset \mathbb{F}_q[X]$ e seja $I_q = (g, \cdots, g_t, X_1^q - X_1, \cdots, X_n^q - X_n)$. Recorde que $\prod_{q \in \mathbb{F}_q}(X - a) = X^q - X$ então $V(I) = V(I_q)$. A partir de agora vamos sempre considerar a ordem lexicográfica graduada em $\mathcal{M} \subset \mathbb{F}_q[X]$. Temos que $\#(\Delta(I_q)) \leq \#(\Delta(lm(g_1), \cdots, lm(g_t), X_1^q, \cdots, X_n^q) \leq q^n$ então da proposição 3.25 temos que $\#(V(I_q)) \leq \#(\Delta(I_q))$. Seja $V(I_q) = \{P_1, \cdots, P_m\}$ e seja φ o homomorfismo

$$\begin{array}{rccc} \varphi: & \mathbb{F}_q[X]/I_q & \longrightarrow & \mathbb{F}_q^m \\ & f + I_q & \longmapsto & (f(P_1), \cdots, f(P_m)) \end{array} \tag{3.3}$$

Proposição 3.27 *O homomorfismo φ é um isomorfismo de $\mathbb{F}$-espaços vetoriais.*

Demonstração: É claro que φ é uma transformação linear. Como $X_i^q - X_i \in I_q$, para todo $i = 1, \cdots, n$ temos que I_q é um ideal radical (v. [1, Prop. 8.14]), e também para qualquer extensão L algebricamente fechada de F_q temos que $V_L(I_q) = V_{\mathbb{F}_q}(I_q)$, assim de 3.18 e 3.26 temos que $dim(\mathbb{F}_q[X]/I_q) = \#(\Delta(I_q)) = m$. Do lema 3.24 sabemos que existem polinômios $p_1, \cdots, p_m \in \mathbb{F}_q[X]$ tais que $p_i(P_j) = \delta_{ij}$ para todo $i, j \in \{1, \cdots, m\}$, assim $\varphi(p_i + I_q) = e_i$. Isso prova que φ é sobrejetivo e consequentemente um isomorfismo. □

O seguinte conceito foi introduzido por Fitzgerald e Lax em [9].

Definição 3.28 *Seja $L \subset \mathbb{F}_q[X]/I_q$ um F_q subespaço vetorial de $\mathbb{F}_q[X]/I_q$. A imagem $\varphi(L) := C(L)$ é chamada de* ***código de variedade afim*** *associado a L.*

Em [9] os autores provam que todo código $\mathbb{F}_q$-linear é igual a $C(L)$ para n, I e L convenientemente escolhidos.

Vamos definir a seguir um tipo de código que tem como caso particular os códigos de Reed-Muller.

Sejam $A_1, \cdots, A_n$ conjuntos não vazios de $\mathbb{F}_q$ e seja $X := A_1 \times \cdots \times A_n$. Seja $f_i = \prod_{c \in A_i}(X_i - c)$ para todo $i \in \{1, \cdots, n\}$ e seja $I := (f_1, \cdots, f_n)$, claramente $V(I) = X$. Como acima montamos $I_q = (f_1, \cdots, f_n, X_1^q - X, \cdots, X_n^q - X_n)$ e observe que neste caso $I_q = I$ por que f_i é um fator de $X_i^q - X_i$, para todo $i = 1, \cdots, n$.

Considere, para todos inteiros $d \geq 0$, o $\mathbb{F}_q$ subespaço vetorial de $\mathbb{F}_q[X]/I$ dado por

$$L_d := \{p + I \mid p = 0 \text{ ou } deg(p) \leq d\}$$

onde $deg(p)$ é o grau total de um polinômio $p \in \mathbb{F}_q[X]$.

Definição 3.29 *O* ***código cartesiano afim*** *$C(d)$ é a imagem $\varphi(L_d)$.*

Observe que quando tomamos $A_i = \mathbb{F}_q$ para todo $i = 1, \cdots, n$ obtemos os códigos de Reed-Muller generalizados.

Vamos determinar a dimensão e a distância mínima para esses códigos usando técnicas envolvendo a teoria apresentada até aqui. Seja $d_i := \#(A_i)$ para todo $i = 1, \cdots, n$, então $\#(V(I)) = d_1 \cdot \cdots \cdot d_n$ e este é o comprimento de $C(d)$, para todo $d \geq 0$.

Lema 3.30 *O conjunto $\{f_1, \cdots, f_n\}$ é uma Base de Gröbner para I.*

Demonstração: Claramente $lm(f_i) = X_i^{d_i}$ para todo $i = 1, \cdots, n$, de modo que

$$\Delta(I) \subset \{X_1^{\alpha_1} \cdot \cdots \cdot X_n^{\alpha_n} \mid 0 \leq \alpha_i < d_i, \ \forall i = 1, \cdots, n\}.$$

De $\#(V(I)) = d_1 \cdot \cdots \cdot d_n \leq \#(\Delta(I)) \leq d_1 \cdot \cdots \cdot d_n$ temos em particular que $\#(\Delta(I)) = d_1 \cdot \cdots \cdot d_n$. Isto mostra que $B := \{f_1, \cdots, f_n\}$ é uma Base de Gröbner para I, pois de outro modo a partir do algoritmo de Buchberger teríamos que adicionar a B é um polinômio cujo monômio líder não é múltiplo de $X_i^{d_i}$ para todo $i = 1, \cdots, n$, mas isso implicaria $\#(\Delta(I)) < d_1 \cdot \cdots \cdot d_n$, uma contradição. □

Lema 3.31 *O ideal de X é I.*

Demonstração: Claramente $I \subset \mathcal{I}(X)$ então $\Delta(\mathbb{I}(X)) \subset \Delta(I)$. Pela proposição 3.25 e o lema acima temos que $d_1 . \cdots . d_n = \#(V(\mathcal{I}(X))) \leq \#(\Delta(\mathcal{I}(X))) \leq \#(\Delta(I)) = d_1 . \cdots . d_n$ então $\#\Delta(\mathcal{I}(X)) = d_1 . \cdots . d_n$. Como $\{f_1, \cdots . f_n\} \subset \mathcal{I}(X)$, pelo lema anterior temos que $\{f_1, \cdots, f_n\}$ é uma base (de Gröbner) para $\mathcal{I}(X)$ e então $\mathcal{I}(X) = I$. □

Agora queremos calcular a dimensão de $C(d)$. Como φ é um isomorfismo e $C(d) = \varphi(L_d)$ temos que $dim(C(d)) = dim(L_d)$. Seja

$$\Delta(I)_{\leq d} := \{M \in \Delta(I) \mid deg(M) \leq d\}.$$

Proposição 3.32 *O conjunto* $\{M + I \mid M \in \Delta(I)_{\leq d}\}$ *é uma base de* L_d.

Demonstração: Do teorema 3.18 sabemos que $\{M + I \mid M \in \Delta(I)_{\leq d}\}$ é um conjunto linearmente independente, e claramente está contido em L_d. Seja $f \in \mathbb{F}_q[X]$, $f \neq 0$ tal que $deg(f) \leq d$. Seja r o resto da divisão de f por $\{f_1, \cdots, f_n\}$. Do algoritmo da divisão, o fato de $\{f_1, \cdots, f_n\}$ ser uma Base de Gröbner para I e da proposição 3.16 temos que r é uma combinação linear de monômios em $\Delta(I)_{\leq d}$, o que termina a prova. □

Como consequência do resultado acima nós temos o seguinte resultado.

Lema 3.33 *A dimensão de* $C(d)$ *é* $dim(C(d)) = \#(\Delta(I)_{\leq d})$, *em particular* $dim(C(d)) = d_1.\cdots.d_n$ *e* $d_{min}(C(d)) = 1$, *para todo* $d \geq \sum_{i=1}^{n}(d_i - 1)$.

Demonstração: A primeira afirmação é uma consequência da proposição acima e do fato de φ ser um isomorfismo. Para a segunda e terceira, observe que desde que $\{f_1, \cdots, f_n\}$ seja uma Base de Gröbner para I, temos:

$$\Delta(I) = \{X_1^{\alpha_1}.\cdots.X_n^{\alpha_n} \mid 0 \leq \alpha_i \leq d_i - 1,\ \forall i = 1, \cdots, n\}.$$

Assim, $\Delta(I)_{\leq d} = \Delta(I)$ sempre que $d \geq \sum_{i=1}^{n}(d_i - 1)$. O resultado segue do fato de $\#(\Delta(I)) = d_1.\cdots.d_n$ e do fato de $\varphi(L(d)) = \mathbb{F}_q^{d_1.\cdots.d_n}$. □

Teorema 3.34 *A dimensão de* $C(d)$ *para* $0 \leq d < \sum_{i=1}^{n}(d_i - 1)$ *é dada por:*

$$dim(C(d)) = \binom{n+d}{d} - \sum_{i=1}^{n}\binom{n+d-d_i}{d-d_i} + \cdots + (-1)^j \sum_{1\leq i_1<\cdots<i_j<n}\binom{n+d-d_{i_1}-\cdots-d_{i_j}}{d-d_{i_1}-\cdots-d_{i_j}} +$$

$$+\cdots+(-1)^n\binom{n+d-d_1-\cdots-d_n}{d-d_1-\cdots-d_n}.$$

onde $\binom{a}{b} = 0$ *se* $b < 0$.

Demonstração: De acordo com o resultado anterior a dimensão de $C(d)$ é igual a cardinalidade de $\Delta(I)_{\leq d}$, isto é, do número de monômios em $\Delta(I)$ da forma $X_1^{\alpha_1}.\cdots.X_n^{\alpha_n}$ com $\sum_{i=1}^{n}\alpha_i \leq d$. Seja

$$h(t) := (1 + t + \cdots + t^{d_1-1}).\cdots.(1 + t + \cdots + t^{d_n-1}),$$

é fácil ver que o coeficiente de t^e em $h(t)$ é igual ao número de monômios em $\Delta(I)$ que tem grau e, para todo $e \in \{0, \cdots, \sum_{i=1}^{n}(d_i - 1)\}$. Assim, uma forma de obter o que queremos é calculando os coeficientes de $t^0, t, \cdots, t^d$ e depois resumi-los. Uma maneira mais rápida é observar que existe uma bijeção entre os conjuntos $\Delta(I)_{\leq d}$ e

$$\square_d := \{X_0^{\alpha_0}.X_1^{\alpha_1}.\cdots.X_n^{\alpha_n} \in \mathbb{F}_q[X_0, X_1, \cdots, X_n] \mid \begin{array}{c} com\ \sum_{i=0}^{n}\alpha_i = d\ e \\ 0 \leq \alpha_i \leq d_i - 1,\ \forall i = 1, \cdots, n\} \end{array}$$

dado por $\beta : \Delta(I)_{\leq d} \longrightarrow \square_d$ onde $\beta(M) = X_0^d M(X_1/X_0, \cdots, X_n/X_0)$ e $\beta^{-1} : \square_d \longrightarrow \Delta(I)_{\leq d}$ dado por $\beta^{-1}(N) = N(1, X_1, \cdots, X_n)$. Agora considere

$$H(t) := (1 + t + t^2 + \cdots)(1 + t + \cdots + t^{d_1-1}).\cdots.(1 + t + \cdots + t^{d_n-1}),$$

então o coeficiente de t^d é a cardinalidade de $\square_d$. Para calcular este coeficiente notamos que podemos pensar em $H(t)$ como uma função real de uma variável t definida em uma vizinhança adequada de 0, $|T| < 1$. Como $1 + t + t^2 + \cdots = \frac{1}{(1-t)}$ então

$$H(t) = \frac{1}{1-t} \cdot \frac{1-t^{d_1}}{1-t} \cdot \cdots \cdot \frac{1-t^{d_n}}{1-t}.$$

Assim, $H(t) = (1/(1-t)^{n+1}) \prod_{i=1}^{n}(1-t^{d_i})$. Usando que $1/(1-t)^{n+1} = \sum_{j=0}^{\infty} \binom{n+j}{j} t^j$ temos que

$$H(t) = \left(\sum_{j=0}^{\infty} \binom{n+j}{j} t^j \right) \cdot$$

$$\cdot \left(1 - \sum_{i=1}^{n} t^{d_i} + \sum_{1 \le i_1 < i_2 \le n} t^{d_{i_1}+d_{i_2}} + \cdots + (-1)^j \sum_{1 \le i_j < \cdots < i_j \le n} t_{d_{i_1}+\cdots+d_{i_j}} + \cdots + (-1)^n t^{d_{i_1}+\cdots+d_{i_n}} \right).$$

A expressão para a $dim(C(d))$ na demonstração do teorema é o coeficiente de t^α em $H(t)$ calculado usando o produto acima. □

Para encontrar a distância mínima de $C(d)$, para $0 \le d < \sum_{i=1}^{n}(d_i - 1)$, precisamos do seguinte resultado auxiliar.

Lema 3.35 *Sejam os inteiros $0 < d_1 \le \cdots \le d_n$ e $d < \sum_{i=1}^{n}(d_i - 1)$. Seja $m(\alpha_1, \cdots, \alpha_n) := \prod_{i=1}^{n}(d_i - \alpha_i)$, onde $0 \le \alpha_i < d_i$ é um inteiro para todo $i = 1, \cdots, n$. Então*

$$\min\{m(\alpha_1, \cdots, \alpha_n) \mid \alpha_1 + \cdots + \alpha_n \le d\} = (d_{k+1} - \ell) \prod_{i=k+2}^{n} d_i$$

onde k e ℓ são unicamente definidas por $s = \sum_{i=1}^{k}(d_i - 1) + \ell$ com $0 \le \ell < d_{k+1} - 1$. Aqui, se $k + 1 = n$ nós entendemos que $\prod_{i=k+2}^{n} d_i$, e se $s < d_1 - 1$ então usamos $k = 0$ e $\ell = d$.

Demonstração: Observe que o mínimo deve ser alcançado quando $\sum_{i=1}^{n} \alpha_i = d$, e o lema afirma que é atingido na n-upla $(d_1 - 1, \cdots, d_k - 1, \ell, 0, \cdots, 0)$. Assim seja $\alpha = (\alpha_1, \cdots, \alpha_n)$ com $\sum_{i=1}^{n} \alpha_i = d$ e assuma que $\alpha_{i_1} < d_{i_1} - 1$ para algum $i_1 \in \{1, \cdots, k\}$. Se existir $i_2 \in \{k+1, \cdots, n\}$ tal que $\alpha_{i_2} > 0$ e $\alpha_{i_1} + \alpha_{i_2} \le d_{i_1} - 1$ então denotando por α' a n-upla obtida de α substituindo α_{i_1} por $\alpha_{i_1} + \alpha_{i_2}$ e α_{i_2} por 0 temos que:

$$m(\alpha) - m(\alpha') = \prod_{i=1}^{n}(d_i - \alpha_i) - \prod_{i=1; i \ne i_1, i_2}^{n} (d_i - \alpha_i)(d_{i_1} - (\alpha_{i_1} + \alpha_{i_2}))(d_{i_2} - 0) =$$

$$= \prod_{i=1; i \ne i_1, i_2}^{n} (d_i - \alpha_i) \left[(d_{i_1} - \alpha_{i_1})(d_{i_2} - \alpha_{i_2}) - (d_{i_1} - \alpha_{i_1} - \alpha_{i_2}) d_{i_2} \right] =$$

$$= \prod_{i=1; i \ne i_1, i_2}^{n} (d_i - \alpha_i) \left[d_{i_1} d_{i_2} - d_{i_1}\alpha_{i_2} - d_{i_2}\alpha_{i_1} + \alpha_{i_1}\alpha_{i_2} - d_{i_1} d_{i_2} + d_{i_1}\alpha_{i_1} + d_{i_2}\alpha_{i_2} \right] =$$

$$= (\alpha_{i_1}\alpha_{i_2} + (d_{i_2} - d_{i_1})\alpha_{i_2}) \cdot \prod_{1=1; i \ne i_1, i_2}^{n} (d_i - \alpha_i) \ge 0$$

então $m(\alpha) \geq m(\alpha')$. Se existir $i_2 \in \{k+1, \cdots, n\}$ tal que $\alpha_{i_2} > 0$ e $\alpha_{i_1} + \alpha_{i_2} > d_{i_1} - 1$ então denotando por α'' a n-upla obtida de α substituindo α_{i_1} por $d_{i_1} - 1$ e α_{i_2} por $\alpha_{i_2} - (d_{i_1} - 1 - \alpha_{i_1})$ temos que

$$m(\alpha) - m(\alpha'') = \prod_{i=1}^{n}(d_i - \alpha_i) - \prod_{i=1;i\neq i_1,i_2}^{n}(d_i - \alpha_i)(d_{i_1} - (d_{i_1} - 1))(d_{i_2} - (\alpha_{i_2}(d_{i_1} - 1 - \alpha_{i_1}))) =$$

$$= \prod_{i=1;i\neq i_1,i_2}^{n}(d_i - \alpha_i)\left[(d_{i_1} - \alpha_{i_1})(d_{i_2} - \alpha_{i_2}) - (d_{i_2} - \alpha_{i_2} + d_{i_1} - 1 - \alpha_{i_1})\right] =$$

$$= \prod_{i=1;i\neq i_1,i_2}^{n}(d_i - \alpha_i)\left[d_{i_1}d_{i_2} - d_{i_1}\alpha_{i_2} - d_{i_2}\alpha_{i_1} + \alpha_{i_1}\alpha_{i_2} - d_{i_2} + \alpha_{i_2} - d_{i_1} + 1 + \alpha_{i_1}\right] =$$

$$= \prod_{i=1;i\neq i_1,i_2}^{n}(d_i - \alpha_i)\left[d_{i_1}(d_{i_2} - 1 - \alpha_{i_2}) - (d_{i_2} - 1 - \alpha_{i_2}) - \alpha_{i_1}(d_{i_2} - 1 - \alpha_{i_2})\right] =$$

$$= (d_{i_1} - 1 - \alpha_{i_1})(d_{i_2} - 1 - \alpha_{i_2}) \cdot \prod_{1=1;i\neq i_1,i_2}^{n}(d_i - \alpha_i) \geq 0$$

então $m(\alpha) \geq m(\alpha'')$. Isso prova que, se m atinge seu mínimo em α então temos que $\alpha_i = d_i - 1$ para todo $i = 1, \cdots, k$ e também que $\alpha_{k+1} = \ell$. □

Vamos denotar por $\mu_1(d)$ o valor do mínimo determinado no Lema acima. O próximo resultado mostra que a distância mínima de $C(d)$ é exatamente $\mu_1(d)$.

Teorema 3.36 *Seja $0 \leq d < \sum_{i=1}^{n}(d_i - 1)$. Então a distância mínima de $C(d)$ é $(d_{k+1} - \ell)\prod_{i=k+2}^{n} d_i$ onde k e ℓ são unicamente definidos por $d = \sum_{i=1}^{k}(d_i - 1) + \ell$ com $0 \leq \ell \leq d_{k+1} - 1$. Como no resultado acima se $k + 1 = n$, entendemos que $\prod_{i=k+2}^{n} d_i = 1$, e se $s < d_1 - 1$, então definimos $k = 0$ e $\ell = d$.*

Demonstração: Seja $F \in L_d$ e seja $J_F := (F, f_1, \cdots, f_n)$. Então o número de zeros na palavra $\varphi(F + I)$ é igual a $\#(V(J_F))$ de modo que o peso é $\omega(\varphi(F + I)) = \prod_{i=1}^{n} d_i - \#(V(J_F))$. Da proposição 3.25 temos que $\#(V(J_F)) \leq \#(\Delta(J_F))$. Seja $M := X_1^{\alpha_1} \cdot \cdots \cdot X_n^{\alpha_n}$ o monômio líder de F, da observação 3.20 temos que $\Delta(J_F) \subset \Delta(M, X_1^{d_1} \cdots, X_n^{d_n})$ então $\#(\Delta(J_F)) \leq \prod_{i=1}^{n} d_i - \prod_{i=1}^{n}(d_i - \alpha_i)$. Assim, $\omega(\varphi(F + I)) \geq \prod_{i=1}^{n}(d_i - \alpha_i)$ e do lema anterior temos $\omega(\varphi(F+I)) \geq (d_{k+1} - \ell)\prod_{i=k+2}^{n} d_i$. Para ver que este limite é realmente atingido nós escrevemos $A_i = \{a_{i_1}, \cdots, a_{i_{d_i}}\}$ para $i = 1, \cdots, n$ e seja

$$G(X_1, \cdots, X_n) = \left(\prod_{i=1}^{k}\prod_{j=1}^{d_i - 1}(X_i - \alpha_{i_j})\right)\prod_{j=1}^{\ell}(X_{k+1} - a_{k+1_j}),$$

então $\deg(G) = d$, G tem $\prod_{i=1}^{n} d_i - (d_{k+1} - \ell)\prod_{i=k+2}^{n} d_i$ zeros em $A_1 \times \cdots \times A_n$ então $\omega(\varphi(G + I)) = (d_{k+1} - \ell)\prod_{i=k+2}^{n} d_i$. □

Observe que, se nos Teoremas 3.34 e 3.36 fazemos $d_1 = \cdots = d_n = q$ obtemos as fórmulas escritas no final do Capítulo 2.

Capítulo 4

O Segundo Peso Mínimo

4.1 Resultados Iniciais

Nesta seção queremos encontrar o segundo valor mínimo da função m, definida no Lema 3.35, e restrita às n-uplas cujas entradas somam no máximo d, no caso em que $d_1 = \cdots = d_n = q$ e para isso vamos precisar da seguinte definição.

Definição 4.1 *Dizemos que $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}^n$ é uma n-upla normalizada sempre que tivermos $\alpha_1 \geq \ldots \geq \alpha_n$.*

Seja d um inteiro não negativo e escreva $d = k(q-1) + \ell$, com $0 \leq \ell < q-1$ como no Lema 3.35.

Lema 4.2 *A n-upla normalizada α, com $\alpha_1 + \cdots + \alpha_n \leq d$ e $0 \leq \alpha_i < q$ para todo $i = 1, \cdots, n$, tal que $m(\alpha)$ é mínimo, isto é, $m(\alpha) = (q-\ell)q^{n-k-1}$ é da forma*

$$a = (q-1, \cdots, q-1, \ell, 0, \cdots, 0),$$

onde ℓ aparece na entrada $k+1$.

Demonstração: Começamos introduzindo algumas notações. Seja $i, j \in \{1, \cdots, n\}$ com $i \neq j$. Aplicar $M_1(i,j)$ na n-upla α obtendo a n-upla α' significa fazer $\alpha'_i = \alpha_i + \alpha_j$, $\alpha'_j = 0$ e $\alpha'_\ell = \alpha_\ell$ para todo $\ell \in \{1, \cdots, n\} \setminus \{i, j\}$, assim, para aplicar $M_1(i,j)$ em α devemos ter $\alpha_i + \alpha_j \leq q-1$, e assim temos que

$$m(\alpha) - m(\alpha') = \prod_{i=1}^{n}(q-\alpha_i) - \prod_{i=1}^{n}(q-\alpha'_i) =$$

$$= ((q-\alpha_i)(q-\alpha_j) - (q-(\alpha_i+\alpha_j))(q-0)) \prod_{\ell=1,\ell\neq i,j}^{n}(q-\alpha_\ell) =$$

$$= (q^2 - q\alpha_j - q\alpha_i + \alpha_i\alpha_j - q^2 + q\alpha_i + q\alpha_j) \prod_{\ell=1,\ell\neq i,j}^{n}(q-\alpha_\ell) =$$

$$= (\alpha_i\alpha_j) \prod_{\ell=1,\ell\neq i,j}^{n}(q-\alpha_\ell).$$

Da mesma forma, definimos que aplicar $M_2(i,j)$ em α obtendo α'' significa fazer $\alpha''_i = q-1$, $\alpha''_j = \alpha_j - (q-1-\alpha_i)$ e $\alpha''_\ell = \alpha_\ell$ para todo $\ell \in \{1,\cdots,n\} \setminus \{i,j\}$, assim, para aplicar $M_2(i,j)$ devemos ter $\alpha_i + \alpha_j \geq q-1$, e temos que

$$m(\alpha) - m(\alpha'') = \prod_{i=1}^{n}(q-\alpha_i) - \prod_{i=1}^{n}(q-\alpha''_i) =$$

$$= ((q-\alpha_i)(q-\alpha_j) - (q-(q-1))(q-(\alpha_j-(q-1-\alpha_i)))) \prod_{\ell=1,\ell\neq i,j}^{n}(q-\alpha_\ell) =$$

$$= (q^2 - q\alpha_j - q\alpha_i + \alpha_i\alpha_j - q + \alpha_j - q + 1 + \alpha_i) \prod_{\ell=1,\ell\neq i,j}^{n}(q-\alpha_\ell) =$$

$$= (q(q-\alpha_j-1) - \alpha_i(q-\alpha_j-1) - (q-\alpha_j-1)) \prod_{\ell=1,\ell\neq i,j}^{n}(q-\alpha_\ell) =$$

$$= (q-\alpha_i-1)(q-\alpha_j-1) \prod_{\ell=1,\ell\neq i,j}^{n}(q-\alpha_\ell).$$

Suponha que $\alpha = (\alpha_1,\cdots,\alpha_n)$ é uma n-upla normalizada tal que $\alpha_i = q-1$ para todo $i \in \{1,\cdots,k\}$ e $\alpha \neq a$. Isto significa que $\alpha_{k+1} < \ell$ e para algum $t > k+1$ temos $1 \leq \alpha_t \leq \ell$, também temos $\alpha_{k+1} > 0$ pois α é normalizado. Aplicando $M_1(k+1,t)$ em α obtemos

$$m(\alpha) - m(\alpha') = (\alpha_{k+1}\alpha_t) \prod_{\ell=1;\ell\neq k+1,t}^{n}(q-\alpha_\ell) > 0,$$

logo, temos $m(\alpha) > m(\alpha')$.

Suponha agora que $\alpha = (\alpha_1,\cdots,\alpha_n)$ é uma n-upla normalizada tal que $0 \leq \alpha_i < q-1$ para algum $i \in \{1,\cdots,k\}$. Então para algum $t \geq k+1$ temos $\alpha_t > 0$ (observe que $\alpha_i \neq 0$ pois α é normalizado). Se $\alpha_i + \alpha_t \leq q-1$ então aplicando $M_1(i,t)$ em α obtemos uma n-upla α' tal que $m(\alpha) > m(\alpha')$.

Se $\alpha_i + \alpha_t > q-1$ então aplicando $M_2(i,t)$ em α obtemos uma n-upla α'' tal que $m(\alpha) > m(\alpha'')$ (observe que como α é normalizado não podemos ter $\alpha_t = q-1$ pois $\alpha_i < q-1$).

□

Agora procuramos o segundo valor mínimo da função m restrita às n-uplas cujas entradas somam no máximo d. Quando dizemos que aplicamos $N(i,t)$ em uma n-upla α e obtemos a n-upla α' significa que estamos fazendo $\alpha'_i = \alpha_i - 1$, $\alpha'_t = \alpha_t + 1$ e $\alpha'_j = \alpha_j$ para todo $j \in \{1,\cdots,n\} \setminus \{i,t\}$, e aplicar $L(i)$ em uma n-upla α obtendo a n-upla α'' significa fazer $\alpha''_i = \alpha_i - 1$, e $\alpha''_j = \alpha_j$ para todo $j \in \{1,\cdots,n\} \setminus \{i\}$. No que se segue, queremos aplicar essas operações para a n-upla a nos casos onde obtemos uma n-upla com entradas diferentes de zero e menores que q, pois então podemos aplicar a função m e comparar o resultado com $m(a)$. Dividimos essa análise em cinco casos.

Caso I. Seja $i \in \{1,\cdots,k\}$ e $t \in \{k+2,\cdots,n\}$ (observe que esse caso acontece apenas quando $k \leq n-2$). Aplicando $N(i,t)$ em a obtemos uma n-upla a', e

$$m(a') - m(a) = \prod_{j=1}^{n}(q-\alpha'_i) - \prod_{j=1}^{n}(q-\alpha) =$$

$$= [(q-(q-2))(q-\ell)(q-1)q^{n-k-2}] - [(q-\ell)q^{n-k-1}] =$$

$$= (q-\ell)q^{n-k-2}[2(q-1)-q]$$

$$= (q-2)(q-\ell)q^{n-k-2} > 0.$$

Caso II. Seja agora $i \in \{1, \cdots, k\}$ e seja a' uma n-upla obtida aplicando $N(i, k+1)$ em a. Então

$$m(a') - m(a) = [(q-(q-2))(q-(\ell+1))q^{n-k-1}] - [(q-\ell)q^{n-k-1}] =$$

$$= q^{n-k-1}[2(q-\ell-1)-(q-\ell)] =$$

$$(q-\ell-2)q^{n-k-1}$$

e assim temos $m(a') - m(a) > 0$ sempre que $\ell < q-2$.

Caso III. Seja $t \in \{k+2, \cdots, n\}$, assuma $\ell > 0$ e seja a' a n-upla obtida aplicando $N(k+1, t)$ em a (observe que esse caso só acoontece se $k \leq n-2$). Assim temos que

$$m(a') - m(a) = [(q-(\ell-1))(q-1)q^{n-k-2}] - [(q-\ell)q^{n-k-1}] =$$

$$= q^{n-k-2}[(q-\ell+1)(q-1) - q(q-\ell)] =$$

$$= q^{n-k-2}[q^2 - q - \ell q + \ell + q - 1 - q^2 + \ell q] =$$

$$= (\ell-1)q^{n-k-2}$$

consequentemente $m(a') - m(a) > 0$ sempre que $\ell > 1$.

Caso IV. Assuma que $\ell > 0$ e seja a'' a n-upla obtida aplicando $L(k+1)$ em a. Assim temos que

$$m(a'') - m(a) = [(q-(\ell-1))q^{n-k-1}] - [(q-\ell)q^{n-k-1}] =$$

$$= q^{n-k-1}[q-\ell+1-q+\ell] =$$

$$= q^{n-k-1}.$$

Caso V. Seja $i \in \{1, \cdots, k\}$ e aplique $L(i)$ em a obtendo a''. Temos

$$m(a'') - m(a) = [(q-(q-2))(q-\ell)q^{n-k-1}] - [(q-\ell)q^{n-k-1}] =$$

$$= (q-\ell)q^{n-k-1}.$$

Mostramos a seguir que o segundo valor mínimo de m é alcançado em uma das n-uplas a' e a'' que aparecem nos casos estudados acima. Na realidade, podemos nos concentrar nas n-uplas que aparecem nos casos I-IV, pois nos casos I e V não temos nenhuma condição sobre ℓ e temos que

$$(q-\ell)q^{n-k-1} > (q-2)(q-\ell)q^{n-k-1}$$

de modo que a menor diferença aparece no caso I.

Lema 4.3 *O segundo valor mínimo de $m(\alpha)$, onde $\sum_{i=1}^{n} \alpha_i \leq d$ e $0 \leq \alpha_i < q$ para todo $i \in \{1, \cdots, n\}$, ocorre quando α é igual a uma das n-uplas a' ou a'' obtidas nos casos I-IV acima.*

Demonstração: Seja α uma n-upla normalizada com $\sum_{i=1}^{n}\alpha_i \leq d$, $0 \leq \alpha_i < q$ para todo $i \in \{1, \cdots, n\}$ e distinta de a, a' e a'' obtidos no casos I-IV acima, para todo $j \in \{1, \cdots, k\}$.

Suponha que existam i_1, i_2 tais que $0 < \alpha_{i_1} < q-1$, $0 < \alpha_{i_2} < q-1$ e $\alpha_{i_1} \leq \alpha_{i_2}$. Seja α' obtido de α tomando $\alpha'_{i_1} = \alpha_{i_1} - 1$, $\alpha'_{i_2} = \alpha_{i_2} + 1$ e $\alpha'_t = \alpha_t$ para todo $t \in \{1, \cdots, n\} \setminus \{i_1, i_2\}$. Como α é diferente de a' temos que α' é diferente de a. Temos

$$m(\alpha) - m(\alpha') = ((q-\alpha_{i_1})(q-\alpha_{i_2}) - (q-\alpha_{i_1}+1)(q-\alpha_{i_2}-1)) \cdot \prod_{s=1, s\neq i_1, i_2}^{n} (q-\alpha_s) =$$

$$= (\alpha_{i_2} - \alpha_{i_1} + 1) \prod_{s=1, s\neq i_1, i_2}^{n} (q-\alpha_s) > 0$$

de modo que $m(\alpha) > m(\alpha') > m(a)$, consequentemente $m(\alpha)$ não pode ser o segundo valor mínimo da função m.

Suponha agora que para algum índice i_1, temos $0 < \alpha_{i_1} < q-1$ e para todos os outros temos $\alpha_t = 0$ ou $\alpha_t = q-1$. Uma vez que α é normalizada e diferente de a, a' e a'' devemos ter $\sum_{i=1}^{n}\alpha_i \leq d-2$. Então existe uma n-upla α' com $0 \leq \alpha_i \leq \alpha'_i \leq q-1$ para todo $i \in \{1, \cdots, n\}$ e $\sum_{i=1}^{n}\alpha'_i = 1 + \sum_{i=1}^{n}\alpha_i \leq d-1$. Temos $m(\alpha) > m(\alpha') > m(a)$, então $m(\alpha)$ não pode ser o segundo valor mínimo da função m.

Finalmente, suponha que para todos os índices tenhamos $\alpha_t = 0$ ou $\alpha_t = q-1$. Então, como no parágrafo anterior, temos $\sum_{i=1}^{n}\alpha_i \leq d-2$ e assim o resultado segue. □

Determinaremos agora o segundo valor mínimo de m restrito às n-uplas α, onde $0 \leq \alpha_i < q$ para todo $i \in \{1, \cdots, n\}$ e $\sum_{i=1}^{n}\alpha_i \leq d$, e para fazer isso resumimos os resultados dos casos I-IV acima de uma maneira que será útil no que se segue. Recorde, do Lema 3.35, que o valor mínimo $\mu_1(d)$ para m quando restrito a tais n-uplas é $\mu_1(d) = (q-\ell)q^{n-k-1}$.

No caso I temos que existem n-uplas α tais que

$$m(\alpha) = (q-\ell)q^{n-k-1} + (q-2)(q-\ell)q^{n-k-2} =$$

$$= (q-\ell)q^{n-k-1} \cdot \left(1 + \frac{q-2}{q}\right) = \mu_1(d) \cdot \left(1 + \frac{q-2}{q}\right) > \mu_1(d)$$

No caso II temos que existem n-uplas β tais que

$$m(\beta) = (q-\ell)q^{n-k-1} + (q-\ell-2)q^{n-k-1} =$$

$$= (q-\ell)^{n-k-1} \cdot \left(1 + \frac{q-\ell-2}{q-\ell}\right) = \mu_1(d) \cdot \left(1 + \frac{q-\ell-2}{q-\ell}\right) > \mu_1(d)$$

desde que $\ell < q-2$.

No caso III temos que existem n-uplas γ tais que

$$m(\gamma) = (q-\ell)q^{n-k-1} + (\ell-1)q^{n-k-2} =$$

$$= (q-\ell)q^{n-k-1} \cdot \left(1 + \frac{\ell-1}{(q-1)q}\right) = \mu_1(d) \cdot \left(1 + \frac{\ell-1}{(q-\ell)q}\right) > \mu_1(d)$$

desde que $\ell > 1$.

No caso IV temos que existem n-uplas δ tais que

$$m(\delta) = (q-\ell)q^{n-k-1} + q^{n-k-1} =$$

$$= (q-\ell)q^{n-k-1} \cdot \left(1 + \frac{1}{q-\ell}\right) = \mu_1(d) \cdot \left(1 + \frac{1}{q-\ell}\right) > \mu_1(d)$$

desde que $\ell \neq 0$.

Seja $\mu_2(d)$ o segundo valor mínimo de m quando restrito às n-uplas α, onde $0 \leq \alpha_1 < q$ para todo $i \in \{1, \cdots, n\}$ e $\sum_{i=1}^{n} \alpha_i \leq d$.

Teorema 4.4 *O valor de $\mu_2(d)$ é como se segue:*

Caso $0 \leq k \leq n-2$.

a) Se $\ell \geq 2$, temos

$$\mu_2(d) = \mu_1(d) \cdot \left(1 + \frac{\ell - 1}{(q-\ell)q}\right) = q^{n-k-2}(q-1)(q-\ell+1).$$

b) Se $\ell = 1$ e $q \geq 3$ então

$$\mu_2(d) = \mu_1(d) \cdot \left(1 + \frac{1}{q-1}\right) = q^{n-k} \text{ se } q > 3.$$

$$\mu_2(d) = \mu_1(d) \cdot \left(1 + \frac{1}{3}\right) = 8 \cdot 3^{n-k-2} \text{ se } q = 3, \text{ e}$$

c) Se $\ell = 0$ então

$$\mu_2(d) = \mu_1(d) \cdot \left(1 + \frac{q-2}{q}\right) = 2 \cdot q^{n-k-1}(q-1).$$

Caso $k = n-1$.

d) Se $\ell = 0$ então

$$\mu_2(d) = \mu_1(d) \cdot \left(1 + \frac{q-2}{q}\right) = 2 \cdot (q-1).$$

e) Se $1 \leq \ell \leq n-2$ então

$$\mu_2(d) = \mu_1(d) \cdot \left(1 + \frac{1}{q-\ell}\right) = 2 \cdot (q-\ell-1).$$

Demonstração: Antes de começar, notamos que os valores de m encontrados nos casos I e III acima, são uma função crescente de q.
Suponha $0 \leq k \leq n-2$.
a) Se $1 < \ell < q-2$ então devemos considerar os valores encontrados em todos os casos I-IV e comparando-os chegamos que

$$\frac{\ell - 1}{(q-\ell)q} = \frac{1}{q-\ell}\left(1 - \frac{q-\ell+1}{q}\right) < \frac{1}{q-\ell} \leq$$

$$\frac{q-\ell-2}{q-\ell} = 1 - \frac{2}{q-\ell} \leq 1 - \frac{2}{q} = \frac{q-2}{q}.$$

Se $\ell = q-2$ então não temos que considerar o caso II. Como acima nós temos

$$\frac{\ell-1}{(q-\ell)q} = \frac{q-3}{2q} < \frac{1}{q-\ell} = \frac{1}{2}.$$

b) No caso em que $\ell = 1$ se $q > 3$ temos que considerar os casos I, II e IV e de

$$\frac{1}{q-\ell} \leq \frac{q-\ell-2}{q-\ell} = 1 - \frac{2}{q-\ell} \leq 1 - \frac{2}{q} = \frac{q-2}{q},$$

temos que a afirmação segue.

Agora, no caso em que $\ell = 1$ e $q = 3$ temos que considerar apenas os casos I e IV, e o valor mínimo acontece no caso I, pois nesse caso

$$\frac{q-2}{q} = \frac{1}{3} < \frac{1}{2} = \frac{1}{q-\ell},$$

e a afirmação segue.

c) Se $\ell = 0$ temos que comparar os valores nos casos I e II, e nesse caso temos que

$$\frac{q-\ell-2}{q-\ell} = \frac{q-2}{q}$$

e a afirmação segue.
No caso em que $k = n-1$ fazemos uma análise análoga à de acima, comparando os casos II e IV. □

Observe que para alguma n-upla α tal que $0 \leq \alpha_i < q$ para todo $i \in \{1, \cdots, n\}$ existe um elemento no código de Reed-Muller $RM_q(d,n)$ com peso $m(\alpha)$. De fato, para todo $i \in \{1, \cdots, n\}$ seja $B_i \subset \mathbb{F}_q$ um conjunto com α_i elementos e seja

$$f_\alpha(X) = \prod_{i=1}^{n} \prod_{a \in B_i} (X_i - a).$$

Então o peso da palavra $(f_\alpha(x))_{x \in \mathbb{F}_q^n}$ é $\omega(\varphi_d(f_\alpha)) = |\{P \in \mathbb{F}_q^n | f(P) \neq 0\}| = \prod_{i=1}^{n}(q - \alpha_i)$.

No que se segue denotaremos o segundo peso de Hamming de um código de Reed-Muller $RM_q(d,n)$ por $W_2(d)$. Escolhendo α tal que $m(\alpha) = \mu_2(d)$ concluímos, da observação acima, que $W_2(d) \leq \mu_2(d)$.

4.2 O Segundo Peso de Hamming de $RM_q(d,n)$

Apresentamos a seguir o teorema que nos fornece o resultado principal desse trabalho, mas antes enunciamos o seguinte lema, cuja demonstração pode ser encontrada no apêndice da referência [13].

Lema 4.5 *Sejam q, n e ℓ inteiros tais que $q \geq 3$, $n \geq 3$, $q \leq d \leq (n-1)(q-1)$. Denotamos por k e ℓ o quociente e resto da divisão de d por $q-1$, isto é, $d = k(q-1)+\ell$ com $0 \leq b < q-1$. Denotamos por V o conjunto das sequências finitas de inteiros $\alpha = (\alpha_1, \cdots, \alpha_n)$, de comprimento n tais que*

(1) para $i = 1, \cdots, n$ *temos* $0 \leq \alpha_i \leq q-1$;

(2) $\sum_{i=1}^{n} \alpha_i \leq K$ *onde* $K = d+1$ *se* $l = 0$ *e* $K = d+q-\ell$ *se* $\ell > 0$*;*

(3) se $\alpha_1 = \alpha_2 = \cdots = \alpha_k = q-1$*, então* $\alpha_{k+1} < \ell$.

Vamos definir $\gamma = \max\{\alpha_{k+1}, \ell\}$.
Então, temos que

$$\min_{\alpha \in V} \left\{ \prod_{i=1}^{n} (q-\alpha_i) - (q-\gamma) \prod_{i=k+2}^{n} (q-\alpha_i) \right\} = \mu$$

onde

$$\mu = \begin{cases} (q-2)q^{n-k-1} \text{ se } \ell = 0 \\ (q-1)q^{n-k-3} \text{ se } \ell = 1,\ k < n-2 \\ (q-2)q^{n-k-2} \text{ se } \ell = 1,\ k = n-2 \\ (b-1)q^{n-k-2} \text{ se } 2 \leq k < q-1 \end{cases}.$$

O resultado a seguir dá o valor exato para o segundo peso de Hamming de $RM_q(d,n)$ nos casos em que $\ell \neq 1$ e fornece limitações para tal peso no caso em que $\ell = 1$.

Teorema 4.6 *Para* $n \geq 3$, $q \geq 3$ *e* $q-1 < d \leq (n-1)(q-1)$ *temos os seguintes valores para o segundo peso de Hamming* $W_2(d)$ *do código de Reed-Muller generalizado* $RM_q(d,n)$*:*

(1) se $1 \leq k \leq n-1$ *e* $\ell = 0$ *então*

$$W_2(d) = \mu_2(d) = 2q^{n-k-1}(q-1);$$

(2) se $1 \leq k < n-1$ *e* $\ell = 1$ *então*
(a) se $k < n-2$ *então*

$$q^{n-k} - q^{n-k-1} + q^{n-k-2} - q^{n-k-3} \leq W_2(d) \leq q^{n-k} = \mu_2(d);$$

(b) se $k = n-2$ *então*

$$q^2 - 2 \leq W_2(d) \leq q^2 = \mu_2(d);$$

(3) se $1 \leq k < n-1$ *e* $2 \leq \ell < q-1$ *então*

$$W_2(d) = \mu_2(d) = q^{n-k-2}(q-1)(q-\ell+1).$$

Demonstração: Seja $F(X_1, X_2, \cdots, X_n)$ um polinômio de grau d. Como $\{X_1^q - X_1, \cdots, X_n^q - X_n\}$ é uma base de Gröbner para o ideal I gerado por esse conjunto (v. Lema 3.30) da Proposição 3.32 vem que podemos assumir que os monômios que aparecem em F estão em $\Delta(I)_{\leq d}$. Seja $lm(F) = X_1^{u_1} X_2^{u_2} \cdots X_n^{u_n}$ seu monômio líder. Suponhamos que as variáveis X_i são numeradas de tal maneira que $u_1 \geq u_2 \geq \cdots \geq u_n$. Considere o ideal

$$J = (F, X_1^q - X_1, \cdots, X_n^q - X_n).$$

Usando a pegada de J temos que

$$\#\Delta(J) \leq q^n - \prod_{i=1}^{n} (q-u_i) < \infty.$$

Como $\Delta(J)$ é um conjunto finito temos pela Proposição 3.25 que $V(J)$ também é um conjunto finito e além disso, temos que $\#V(J) \leq \#\Delta(J)$.

Seja φ o homomorfismo definido em (3.3) no capítulo anterior. Temos

$$\omega(\varphi(F+I_q)) = q^n - V(F),$$

e observe que $V(F) = V(F, X_1^q - X_1, \cdots, X_n^q - X_n) = V(J)$. Temos ainda que J é um ideal radical (v. [1, Lemma 8.13]) logo $\#V(J) = \#\Delta(J)$ (v. Teorema 3.26).

$$q^n - \#V(J) = q^n - \#\Delta(J) \geq q^n - \left(q^n - \prod_{i=1}^{n}(q-u_i)\right) = \prod_{i=1}^{n}(q-u_i) = m(u).$$

Pelo Lema 4.2, temos que se $u \neq (q-1, \cdots, q-1, \ell, 0, \cdots, 0)$ então $m(u) \geq \mu_2(d)$. Assim, para obter uma palavra u tal que $m(u) \leq \mu_2(d)$ temos que ter $u = (q-1, \cdots, q-1, \ell, 0, \cdots, 0)$ e também que $\{F, X_1^q - X_1, \cdots, X_n^q - X_n\}$ não seja uma Base de Gröbner, pois caso seja, a desigualdade acima é uma igualdade e temos $m(u) = \mu_1(d)$. Observe no entanto que o conjunto $\{X_1^q - X_1, \cdots, X_n^q - X_n\}$ é uma Base de Gröbner, pois $mdc(lm(X_i^q - X_i), lm(X_j^q - X_j)) = 1$ para todo $i, j \in \{1, \cdots, n\}$, $i \neq j$ (Ver Teorema 3.22). Nesse caso, para algum $i \in \{1, \cdots, k+1\}$(ou $i \in \{1, \cdots, k\}$ se $\ell = 0$) devemos ter que $S(F, X_i^q - X_i)$ dividido por $\{F, X_1^q - X_1, \cdots, X_n^q - X_n\}$ não deixa resto zero. Seja R esse resto, e seja $M = lm(R)$, onde $M = X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. Temos que

$$S(F, X_i^q - X_i) = X_i^{q-u_i} F - \left(\prod_{j=1, j\neq i}^{n} X_j^{u_i}\right)(X_i^q - X_i),$$

e do algoritmo da divisão temos que $gr(R) \leq gr(S(F, X_i^q - X_i))$. Logo, se $i \in \{1, \cdots, k\}$ temos $\sum_{j=1}^{n} \alpha_j \leq d+1$, e se $i = k+1$ (com $\ell > 0$) então $\sum_{j=1}^{n} \alpha_j \leq d+q-\ell$. Em ambos os casos temos que $0 \leq \alpha_j \leq q-1$, $\forall j = 1, \cdots, n$ e mais ainda, $lm(F) = X_1^{q-1} \cdots X_k^{q-1} X_{k+1}^{\ell}$ não divide M.

Agora, temos que

$$R \in (F, X_1^q - X_1, \cdots, X_n^q - X_n) = J$$

logo

$$J = (F, R, X_1^q - X_1, \cdots, X_n^q - X_n).$$

Já vimos que

$$\omega(\varphi(F+I_q)) = q^n - \#V(J) = q^n - \#\Delta(J)$$

e seja $A = \{X_1^{\beta_1} \cdots X_n^{\beta_n},\ \forall i = 1, \cdots, n\}$. Já vimos que

$$\Delta(J) \subset \{X_1^{\beta_1} \cdots X_n^{\beta_n} \mid \quad 0 \leq \beta_i \leq q-1\ ,\ \forall i = 1, \cdots, n\ ,\ X_1^{u_1} \cdots X_n^{u_n}\ e\ X_1^{\alpha_1} \cdots X_n^{\alpha_n} \text{ não dividem } X_1^{\beta_1} \cdots X_n^{\beta_n}\} =$$

$$= \{X_1^{\beta_1} \cdots X_n^{\beta_n} \mid 0 \leq \beta_i \leq q-1\ ,\ \forall i = 1, \cdots, n\} -$$
$$(\{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M\} \cup \{M \in A \mid X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\}) \cup$$
$$\{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M\ e\ X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\}.$$

Então

$$|\Delta(J)| \leq q^n - |\{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M\}| - |\{M \in A | X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\}| +$$
$$|\{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M\ e\ X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\}|,$$

logo

$$\omega(\varphi(F+I_q)) \geq |\{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M\}| +$$
$$|\{M \in A | X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\}| - |\{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M\ e\ X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\}|.$$

Sejam

$$A_1 := \{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M\};$$
$$A_2 := \{M \in A | X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\};$$
$$A_3 := \{M \in A \mid X_1^{u_1} \cdots X_n^{u_n} | M \text{ e } X_1^{\alpha_1} \cdots X_n^{\alpha_n} | M\},$$

então

$$|A_1| = (q-\ell)q^{n-k-1};$$
$$|A_2| = \prod_{i=1}^{n}(q-\alpha_i);$$
$$|A_3| = (q-\gamma)\prod_{j=k+2}^{n}(q-\alpha_j), \text{ onde } \gamma := \max\{\ell, \alpha_{k+1}\}.$$

Logo

$$\omega(\varphi(F)) \geq (q-\ell)q^{n-k-1} + \prod_{i=1}^{n}(q-\alpha_i) - (q-\gamma)\prod_{j=k+2}^{n}(q-\alpha_j).$$

Seja V como no enunciado do Lema 4.5. Tomando

$$\mu = \min_{\alpha \in V}\left\{\prod_{i=1}^{n}(q-\alpha_i) - (q-\gamma)\prod_{i=k+2}^{n}(q-\alpha_i)\right\}$$

temos pelo Lema 4.5 que o resultado segue. □

Observe que para $\ell = 1$ apresentamos apenas uma cota superior e inferior para o segundo peso de Hamming do código de Reed-Muller. Em sua tese de doutorado (v. [8]) Erickson prova que se nós sabemos o segundo peso de $RM_q(d,2)$, então nós sabemos o segundo peso para todos os códigos de Reed-Muller generalizados. A partir de uma conjectura sobre blocking sets, Erickson conjectura que o segundo peso de $RM_q(d,2)$ é $(q-d)q+d-1$. Essa conjectura foi provada por Bruen (v. [2]). Na referência [7] são apresentados os valores exatos do segundo peso de Hamming para esse caso.

Referências Bibliográficas

[1] T. Becker and V. Weispfenning, Gröbner Bases - A computational approach to commutative algebra, Berlin, Germany: Springer Verlag, 1998, 2nd. pr.

[2] A. Bruen, Blocking sets and low-weight codewords in the generalized Reed-Muller codes. In: Bruen A., Wehlau D., Society C.M. (eds.) Error-Correcting Codes, Finite Geometries, and Cryptography, Contemporary Mathematics, vol. 525, pp. 161-164. American Mathematical Society (2010).

[3] C. Carvalho, Gröbner bases methods in coding theory, Publications of CIMPA, Mexico, 2012.

[4] C. Carvalho, On the second Hamming weight of some Reed-Muller type codes, Finite Filds Appl. 24 (2013), 88-94.

[5] C. Carvalho e V.G.L. Neumann, On the next-to-minimal weight of affine cartesian codes, preprint.

[6] D. Cox, J. Little, D. O'Shea, Ideals, Varieties e Algorithms, 3rd ed., Springer, New York, 2007.

[7] L. Elodie, Second weight codewords of generalized Reed-Muller codes, Cryptogr. Commun. 5 (2013) 241-276.

[8] D. Erickson, Counting zeros of polynomials over finite fields. PhD Thesis, California Institute of Technology, Pasadena (1974)

[9] J. Fitzgerald and R.F. Lax, "Decoding affine variety codes using Gröbner bases,"Des. Codesand Cryptogr., vol. 13, pp. 147-158, 1998.

[10] A. Hefez, M.L.T. Villela, , Códigos Corretores de Erros, Série de Computação e Matemática, 2002.

[11] D. E. Muller. Application of boolean algebra to switching circuit design and to error detection. IRE Transactions on Electronic Computers, 3:6?12, 1954.

[12] Irving S. Reed. A class of multiple-error-correcting codes and the decoding scheme. Transactions of the IRE Professional Group on Information Theory, 4:38?49, 1954.

[13] R.Rolland, The second weight of generalized Reed-Muller codes in most cases, Cryptogr. Commun. 2 (2010) 19-40.