UNIVERSIDADE FEDERAL DE SERGIPE
PRÓ-REITORIA DE PÓS-GRADUÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL - PROFMAT

Sistemas de equações polinomiais e base de Gröbner

Fábio Fontes Vilanova

Orientador: Dr. Zaqueu Alves Ramos

São Cristóvão, 2015.

Fábio Fontes Vilanova

Sistemas de equações polinomiais e base de Gröbner

Dissertação apresentada ao Departamento de Matemática da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do título de Mestre em Matemática.

Orientador: Dr. Zaqueu Alves Ramos

São Cristóvão, 2015

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE

Vilanova, Fábio Fontes
V696 Sistemas de equações polinomiais e base de Grobner / Fábio Fontes Vilanova ; orientador Zaqueu Alves Ramos. – São Cristóvão, 2015.
149 f. : il.

Dissertação (mestrado em Matemática)– Universidade Federal de Sergipe, 2015.

1. Matemática. 2. Equações polinomiais. 3. Algoritmos. I. Ramos, Zaqueu Alves, orient. II. Título.

CDU 517.583

Dissertação submetida à aprovação pelo Programa de Pós-Graduação em Matemática da Universidade Federal de Sergipe, como parte dos requisitos para obtenção do grau de Mestre em Matemática.

Sistemas de Equações Polinomiais e Bases de Grobner

por

Fábio Fontes Vilanova

Aprovada pela Banca Examinadora:

Prof. Dr. Zaqueu Alves Ramos - UFS
Orientador

Prof. Dr. André Vinicius Santos Dória - UFS
Primeiro Examinador

Prof. Dra. Crislene Santos da Paixão - IFS
Segundo Examinador

São Cristóvão, 10 de Abril de 2015.

Agradecimentos

"Grandes coisas o Senhor tem feito por nós e por isso estamos alegres".
Salmos 126:3

Eu não poderia iniciar este texto de outra forma senão agradecendo primeiro Àquele que diariamente me capacita, estimula, guia, proteje e me ama incondicionalmente. O Senhor me fortalece, me ajuda e me sustenta com Sua destra fiel. Obrigado Pai!

À minha querida mãe Graça, mulher guerreira, sofrida, que sacrificou muito de sua vida em prol da minha educação. Saiba que te amo muito e agradeço a Deus pela "graça"que me deu de tê-la como mãe.

À minha amada esposa Cainã, presente precioso que Deus me confiou, agradeço profundamente pela generosidade de me amar, algo que, certamente, não mereço mas me completa e me deixa muito feliz. Você com certeza foi a pessoa que mais sofreu comigo durante esses mais de dois anos de curso, por isso te peço perdão e te agradeço pelo apoio, paciência e palavras de incentivo. *"Te adoro em tudo, tudo, tudo; te quero mais que tudo, tudo, tudo; te amar sem limites, viver uma grande história".*

A meu amigo, professor e orientador Zaqueu Alves Ramos, muito obrigado. *Greatest of All Time* ou simplesmente GOAT é a frase usada pelos americanos para se referir àqueles que são os melhores naquilo que fazem, como, por exemplo, Pelé no futebol, Michael Jordan no basquete, Tiger Woods no golfe e Roger Federer no tênis. Para mim essa frase se aplica perfeitamente a você. Agradeço o companheirismo e a confiança sempre depositada em mim. Foi uma honra ser orientado por você, GOAT.

Aos professores Almir, Humberto, Débora, Kalasas, Naldisson, Dalino Felizardo, Allyson, Evilson, Danilo Dias, Leandro, Lucas, Anderson e Romero, que estiveram conosco durante todo o curso, obrigado pelos ensinamentos, pela troca de experiências e pela oportunidade de conviver com profissonais tão qualificados. Cada um de vocês contribuiu de alguma forma para minha capacitação.

Aos professores Fábio dos Santos, André Vinícius e Bruno Luis, que assumiram a função de coordenadores locais do PROFMAT, obrigado pelo tempo despendido na função.

À SBM, responsável direta pela criação do PROFMAT, obrigado pela oportunidade dada. Saiba que hoje certamente sou um profissonal melhor do que era há dois anos e que, por isso, me sinto muito mais preparado para ajudar a melhorar a qualidade do ensino básico no país.

À CAPES, pelo apoio financeiro que nos foi concedido durante o curso.

Aos meus colegas de turma, que estiveram *in loco* comigo durante toda a jornada, agradeço pela oportunidade de conhecê-los. Certamente levarei comigo boas e preciosas lições que aprendi com cada um de vocês.

Resumo

O objetivo principal deste trabalho é, usando bases de Gröbner, apresentar um método algébrico capaz de determinar a solução, quando existir, de sistemas de equações polinomiais não lineares. Para tanto, necessitamos inicialmente apresentar alguns conceitos e teoremas ligados a anéis de polinômios com várias indeterminadas e de ideais monomiais, dentre os quais destacamos o algoritmo extendido da divisão, o teorema da Base de Hilbert e o algoritmo de Buchberger. Além disso, usando noções básicas da Teoria de eliminação e extensão, apresentamos uma solução algébrica para o problema da coloração de mapas usando três cores, bem como uma solução geral para o puzzle Sudoku.

Palavras chave: Sistemas de equações polinomiais, Algoritmo extendido da divisão, Ideais Monomiais, Bases de Hilbert, Algoritmo de Buchberger, Base de Gröbner, coloração de mapas, Sudoku.

Abstract

The main objective of this dissertation is to present an algebraic method capable of determining a solution, if any, of a non linear polynomial equation systems using Gröbner basis. In order to accomplish that, we first present some concepts and theorems linked to polynomial rings with several undetermined and monomial ideals where we highlight the division extended algorithm, the Hilbert Basis and the Buchberger's algorithm. Beyond that, using basics of Elimination and Extension Theorems, we present an algebraic solution to the map coloring that use 3 colors as well as a general solution to the Sudoku puzzle.

Keywords: Polynomial equation systems, Division extended algorithm, monomial ideals, Hilbert Base, Buchberger's algorithm, Gröbner basis, map coloring, Sudoku.

Introdução

"Uma vida sem desafios não vale a pena ser vivida"
Sócrates

Uma equação é algo tão comum em nossa vivência matemática que muitas vezes não nos damos conta de quão fascinante ela é. Descobrir o "valor desconhecido"em uma determinada igualdade intriga o homem desde os tempos remotos. Entre os hindus, por exemplo, um passatempo muito popular e que rendia ao vencedor grande status eram as competições públicas de resolução de problemas matemáticos, nas quais um competidor propunha problemas para o outro resolver. A matemática naquela época era bastante exclusivista. Não havia sinais, nem variáveis, apenas algumas poucas "mentes privilegiadas"eram capazes de resolver problemas, muitas vezes usando complexos artifícios e/ou trabalhosas construções geométricas. O que fascinava esses homens era o desafio.

Para os egípcios, babilônicos e hindus, resolver uma simples equação do tipo $ax + b = c$ era um desafio. Já para gregos e, principalmente, árabes, esse era um problema de simples solução.

É interessante ver como a busca por algoritmos capazes de resolver tipos específicos de equações sempre instigou o homem. Entre os gregos era um desafio determinar a solução de igualdades do tipo $ax^2+bx+c = 0$, eles resolviam algumas dessas equações por meio de construções geométricas com régua e compasso. Já entre os árabes, um matemático que viveu no século IX chamado Al Khowarisma fez grandes avanços que ajudaram o matemático Bhascara Akaria a desenvolver, no século XII, um algoritmo capaz de resolver qualquer equação desse tipo, que conhecemos hoje como fórmula de Bhaskara.

O primeiro matemático a desenvolver um método para resolver equações do terceiro grau foi o francês Scipione del Ferro em meados do século XVI. Ele descobriu uma solução algébrica para equações do tipo $x^3 + px = q$. Mas foi Nicoli Fontana,

mais conhecido como Tartaglia, o primeiro a apresentar um método geral para resolução de uma cúbica completa da forma $ax^3 + bx^2 + cx + d = 0$. Coube a Lodovico Ferrari, matemático bolonhês, apresentar uma solução geral para equações de quarto grau.

Entretanto o desafio de determinar um método para resolver equações de grau maior que 4 ficou em aberto até que Évariste Galois, um polêmico porém brilhante matemático francês, mostrou que, para o caso de equações de grau maior ou igual a cinco, não é possível apresentar uma fórmula resolvente utilizando um número finito de adições, multiplicações e extração de raízes. Ao chegar a essa conclusão, Galois não só resolveu um antigo desafio, como acabou criando uma nova ramificação da álgebra abstrata: a teoria dos grupos.

Aprender a resolver uma equação é uma habilidade cujo desenvolvimento é incentivado desde as séries iniciais do ensino básico, quando um professor propõe ao aluno o desafio de descobrir o termo desconhecido em

$$10 + \square = 15 \quad \text{ou} \quad 8 \times \square = 40$$

Durante o ensino básico, o processo de resolução de sistemas de equações polinomiais lineares é abordado por meio da regra de Cramer e do método de eliminação de Gauss, mais conhecido como escalonamento. Desse modo, uma pergunta natural seria: e se alguma dessas equações fosse não linear? Como esse sistema poderia ser resolvido? É curioso perceber que em nenhum momento um método para resolução de sistemas polinomiais desse tipo seja sequer citado, mesmo que estes surjam naturalmente quando se estuda a interseção de figuras geométricas como elipses, hipérboles, etc.

Responder a pergunta acima é um desafio que inspira muitos matemáticos. Na segunda metade do século XX, dois matemáticos em particular, Wolfgang Gröbner e Bruno Buchberger, professor e discípulo, usando um conjunto especial gerador de um ideal I, denominado *base de Gröbner*, em um anel de polinômios, desenvolveram um algoritmo capaz de, entre outras coisas, determinar se um sistema de equações polinomiais tem solução, caso possua, como estima-las e se estas forem em número finito, como obtê-las. A principio, a comunidade científica não deu a devida importância a esse trabalho mas, a partir dos anos 80, talvez por conta da massificação/popularização dos microcomputadores, pesquisadores passaram a investigar mais profundamente essa nova teoria, o que fez surgir uma ampla variedade de

aplicações desta. Neste trabalho pretendemos usar *bases de Gröbner* para, em um primeiro momento, definir se um polinômio f pertence ou não a um ideal polinomial I e, em um segundo momento, apresentar um método algébrico para resolver um sistema de equações polinomiais não necessariamente lineares.

No capítulo 1 deste trabalho apresentaremos algumas propriedades algébricas dos anéis de polinômios úteis para o desenvolvimento do texto. Neste capítulo damos uma definição formal para anéis de polinômios, apresentamos uma prova para o algoritmo da divisão de polinômios em uma variável e discutimos sobre as diferenças entre polinômio e função polinomial, assim como sob que condições estas coincidem.

Iniciaremos o capítulo 2 dando uma definição formal para um sistema de equações polinomiais em um dado anel de polinômios e, em seguida, apresentaremos algumas proposições relacionadas ao conjunto solução do sistema. Além disso, daremos uma abordagem histórica e mostraremos como modelar, usando um sistema de equações polinomiais, dois conhecidos problemas: o da *coloração de mapas usando apenas 3 cores* e o *puzzle Sudoku*.

No capítulo 3, ponto central deste trabalho, discutiremos sobre as *Bases de Gröbner* e suas aplicações. Antes disso, porém, apresentaremos o conceito de ordenação monomial, ideais monomiais, bem como alguns teoremas fundamentais para o desenvolvimento do texto, entre os quais destacamos: o algoritmo extendido da divisão, o Lema de Dickson, o teorema da Base de Hilbert e o critério de Buchberger. Finalizando o capítulo, usando *base de Gröbner*, apresentaremos uma solução para o problema da coloração de mapas com apenas três cores, assim como uma solução para uma versão mais fraca do Sudoku, o Shidoku.

Sumário

Capítulo 1

Generalidades sobre anéis de polinômios

O objetivo deste trabalho é estudar conjuntos soluções de sistemas de equações polinomiais em várias variáveis. Assim, para uma melhor compreensão do referido assunto, faz-se necessário o conhecimento de algumas particularidades dos anéis de polinômios. Com isso em mente, elaboramos o presente capítulo, cujo o intuito é apresentar propriedades algébricas dos anéis de polinômios úteis para o desenvolvimento do texto. Observamos que as demonstrações de alguns resultados desse capítulo serão omitidas, pois este não é o foco principal do trabalho. Focaremos somente nas provas de proposições que serão generalizadas mais adiante ou que sejam menos elementares. Assumiremos como conhecido noções básicas de teoria de anéis comutativos como as definições de anel, subanel, domínio, corpo, ideal, homomorfismo de anéis, etc.

1.1 Sobre a definição de anel de polinômios

Segundo alguns autores (ver por exemplo [1, Capítulo 3]), um polinômio com coeficientes em um anel $\mathbf{A}$ é uma expressão da forma

$$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$$

onde $a_0, a_1, a_2, \ldots, a_n$ são elementos de $\mathbf{A}$ chamados coeficientes do polinômio. Dessa forma, polinômios são completamente determinados por seus coeficientes, ou seja, dois polinômios são iguais se seus coeficientes correspondentes são iguais.

Apesar da definição acima prestar-se útil na prática, entende-se que a frase "*uma*

expressão da forma"em uma definição matemática não atende aos padrões modernos de rigor. De acordo com tais padrões, sempre que possível, devemos construir novos conceitos a partir de algum outro previamente conhecido. A definição abaixo dá um tratamento mais formal aos polinômios.

Definição 1.1.1. Seja $\mathbf{A}$ um anel. *Um polinômio em uma variável com coeficientes sobre* $\mathbf{A}$ é uma sequência $(a_0, a_1, \ldots, a_n, \ldots)$, onde $a_i \in \mathbf{A}$ para todo i e $a_i \neq 0$ apenas para um número finito de índices.

Denotemos o conjunto de todos os polinômios com coeficientes sobre $\mathbf{A}$ por $\mathcal{A}$. Definimos uma adição e uma multiplicação em $\mathcal{A}$, respectivamente, por:

$$\begin{array}{rccc} + : & \mathcal{A} \times \mathcal{A} & \longrightarrow & \mathcal{A} \\ & (a_0, a_1, \ldots), (b_0, b_1, \ldots) & \longmapsto & (a_0 + b_0, a_1 + b_1, \ldots) \\ \cdot : & \mathcal{A} \times \mathcal{A} & \longrightarrow & \mathcal{A} \\ & (a_0, a_1, \ldots), (b_0, b_1, \ldots) & \longmapsto & (c_0, c_1, \ldots) \end{array}$$

onde $c_i = \sum_{k=0}^{i} a_k b_{i-k}$.

É de fácil verificação que $\mathcal{A}$ com as operações de adição e multiplicação acima definidas é um anel onde:

(i) o elemento neutro da adição é $(0, 0, 0, \ldots)$.

(ii) o elemento neutro da multiplicação é $(1, 0, 0, \ldots)$.

(iii) o inverso de $(a_0, a_1, \ldots, a_n, \ldots)$ com respeito a adição é $(-a_0, -a_1, \ldots, -a_n, \ldots)$.

Se $(a_0, a_1, \ldots)$ é um elemento de $\mathcal{A}$, então $(a_0, a_1, \ldots)^n$ representa o elemento

$$\underbrace{(a_0, a_1, \ldots) \cdot (a_0, a_1, \ldots) \cdot \ldots \cdot (a_0, a_1, \ldots)}_{n \text{ vezes}}$$

Note ainda que:

- $(0, \ldots, 0, \underbrace{a_n}_{\text{lugar } n+1}, 0, 0, \ldots) = (a_n, 0, 0, \ldots) \cdot (0, \ldots, 0, \underbrace{1}_{\text{lugar } n+1}, 0, 0, \ldots)$
- $(0, \ldots, 0, \underbrace{1}_{\text{lugar } n+1}, 0, 0, \ldots) = (0, 1, 0, 0, \ldots)^n$

Desse modo, um polinômio $(a_0, a_1, \ldots, a_n, 0, 0, \ldots)$ pode ser escrito como a soma abaixo

$$(a_0, 0, 0, \ldots) + (a_1, 0, 0, \ldots) \cdot (0, 1, 0, 0, \ldots) + \ldots + (a_n, 0, 0, \ldots) \cdot (0, 1, 0, \ldots, 0, \ldots)^n$$

Se usarmos o símbolo X para representar o elemento $(0, 1, 0, \ldots)$ e a_i para representar $(a_i, 0, 0, \ldots)$, concluímos que o elemento $(a_0, a_1, \ldots, a_n, 0, 0, \ldots)$ é igual a soma $a_0 + a_1X + \ldots + a_nX^n$, ou seja,

$$\mathcal{A} = \left\{ \sum_{i=0}^{n} a_i X^i ; n \in \mathbb{N} \text{ e } a_i \in \mathbf{A} \right\}.$$

Denotamos o anel $(\mathcal{A}, +, \cdot)$ por $\mathbf{A}[X]$ e o denominamos *anel de polinômios em uma variável com coeficientes sobre* $\mathbf{A}$.

Definição 1.1.2. Seja $\mathbf{A}$ um anel e seja $f(X) = a_0 + a_1X + \ldots + a_nX^n \in \mathbf{A}[X]$ com $a_n \neq 0$. Definimos o *grau de* $f(X)$, ou simplesmente gr(f(X)), por esse número natural n. Além disso, o coeficiente a_n é chamado *coeficiente líder de* $f(X)$. Observe ainda que não definimos a noção de grau para o polinômio nulo.

Convém observar que:

(i) Se $\mathbf{A}$ é um domínio, então $\text{grau}(f(X) \cdot g(X)) = \text{grau}\, f(X) + \text{grau}\, g(X)$, para todo $f(X), g(X) \in \mathbf{A}[X] \setminus 0$.

(ii) $\mathbf{A}[X]$ é um domínio se, e somente se, $\mathbf{A}$ é um domínio.

Podemos ainda definir um *anel de polinômios em n variáveis* $X_1, \ldots, X_n$ indutivamente por

$$\mathbf{A}[X_1, \ldots, X_n] = (\mathbf{A}[X_1, \ldots, X_{n-1}])[X_n].$$

1.2 Divisão Euclidiana em anéis de polinômios

No estudo da aritmética de $\mathbb{Z}$ aprendemos que uma ferramenta muito útil é a divisão euclidiana em $\mathbb{Z}$ (por exemplo, no cálculo do máximo divisor comum de dois números inteiros fazemos uso da divisão euclidiana de maneira sistemática). No contexto dos anéis de polinômios temos a seguinte versão de divisão euclidiana.

Teorema 1.2.1. *Seja* $\mathbf{A}[X]$ *um anel de polinômios em uma variável com coeficientes sobre um anel* $\mathbf{A}$. *Sejam* $f(X), g(X) \in \mathbf{A}[X]$, *com* $g(X)$ *não nulo e tendo coeficiente líder invertível em* $\mathbf{A}$. *Então,*

(i) *Existem* $t(X), r(X) \in \mathbf{A}[X]$ *tais que* $f(X) = g(X) \cdot t(X) + r(X)$ *com* $\mathrm{gr(r(X))} < \mathrm{gr(g(X))}$ *ou* $r(X) = 0$.

(ii) *Tais polinômios* $t(X)$ *e* $r(X)$ *são unicamente determinados.*

Prova. (i) Provaremos inicialmente a existência. Notemos que os casos em que $\mathrm{gr}f(X) < \mathrm{gr}(g(X))$ são triviais pois, fazendo $t(X) = 0$ e $r(X) = f(X)$ tem-se o desejado. Suponhamos então $f(X) \neq 0$ e $\mathrm{gr}(f(X)) = n \geq \mathrm{gr}(g(X)) = m$. Digamos que

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0$$

$$g(X) = b_m X^m + b^{m-1} X^{m-1} + \ldots + b_1 X + b_0$$

Definamos

$$r_1(X) = f(X) - a_n b_m^{-1} X^{n-m} \cdot g(X).$$

Note que $\mathrm{gr}(f(X)) > \mathrm{gr}(r_1(X))$. Se $r_1(X) = 0$ ou $\mathrm{gr}(r_1(X)) < \mathrm{gr}(g(X))$ então a prova está encerrada. Caso contrário, defina

$$r_2(X) = r_1(X) - a_{1n_1} b_m^{-1} X^{n_1 - m} \cdot g(X),$$

onde a_{1n_1} é o coeficiente líder de $r_1(X)$ e $n_1 = \mathrm{gr}(r_1(X))$. Note que $\mathrm{gr}(f(X)) > \mathrm{gr}(r_1(X)) > \mathrm{gr}(r_2(X))$. Se $r_2(X) = 0$ ou $\mathrm{gr}(r_2(X)) < \mathrm{gr}(g(X))$, então

$$f(X) = (a_n b_m^{-1} X^{n-m} + a_{1n_r} b_m^{-1} X^{n_1 - n}) g(X) + r_2(X)$$

e a prova está encerrada. Caso contrário, construímos $r_3(X)$ de forma análoga a $r_2(X)$. Procedendo de maneira iterada construimos uma sequência desses r_i. Note que esse processo não pode continuar indefinidamente, ou seja, em algum momento $r_i(X) = 0$ ou $\mathrm{gr}(r_i(X)) < \mathrm{gr}(g(X))$, pois, caso contrário, construiríamos uma sequência infinita decrescente

$$\mathrm{gr}(f(X)) > \mathrm{gr}(r_1(X)) > \mathrm{gr}(r_2(X)) \ldots$$

de inteiros positivos. Digamos, então que ℓ seja o índice tal que $r_\ell(X) = 0$ ou $\text{gr}(r_\ell(X)) < \text{gr}(g(X))$. Somando as igualdades abaixo

$$\begin{array}{rcl} r_1(X) & = & f(X) - a_n b_m^{-1} X^{n-m} \cdot g(X) \\ r_2(X) & = & r_1(X) - a_{1n_1} b_m^{-1} X^{n_1-m} \cdot g(X) \\ \ldots\ldots & \ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ r_\ell(X) & = & r_{\ell-1}(X) - a_{\ell-1 n_{\ell-1}} b_m^{-1} X^{n_{\ell-1}-m} \cdot g(X) \end{array}$$

obtemos

$$f(X) = t(X) \cdot g(X) + r(X)$$

com

$$t(X) = a_n b_m^{-1} X^{n-m} + a_{1n_r} b_m^{-1} X^{n_1-n} + \ldots + a_{\ell-1 n_{\ell-1}} b_m^{-1} X^{n_{\ell-1}-m}$$

e

$$r(X) = r_\ell(X)$$

satisfazendo as propriedades desejadas.

(ii) Suponha que $t(X)$ e $r(X)$ não sejam únicos. Assim, existem $t_1(X), r_1(X) \in A[X]$, com $r_1(X) = 0$ ou $\text{gr}(r_1(X)) < \text{gr}(g(X))$ tais que

$$f(X) = g(X) \cdot t(X) + r(X) = g(X) \cdot t_1(X) + r_1(X).$$

Assim,

$$g(X)\,(t_1(X) - t(X)) = r(X) - r_1(X)$$

$$g(X)\,(t_1(X) - t(X)) \neq 0 \text{ pois } r(X) - r_1(X) \neq 0.$$

Desse modo, concluímos que

$$\text{gr}\,(g(X)(t_1(X) - t(X)) = \text{gr}(g(X)) + \text{gr}(t_1(X) - t(X)) \geq \text{gr}(g(X))$$

Ou seja,

$$\text{gr}\,(r(X) - r_1(X)) \geq \text{gr}(g(X))$$

o que é uma contradição. □

Os polinômios $f(X)$, $g(X)$, $t(X)$ e $r(X)$ do teorema acima são chamados respec-

tivamente de *dividendo*, *divisor*, *quociente* e *resto* da divisão euclidiana.

Observação 1.2.2. A prova do Teorema 1.2.1 nos fornece um algoritmo para calcular de forma explícita o quociente e o resto. De fato, não é difícil perceber que esse algoritmo nada mais é que o algoritmo que utilizamos desde o ensino básico para efetuar a divisão entre dois polinômios.

A condição na proposição 1.2.1 do coeficiente líder é automática se supusermos que o anel $\mathbf{A}$ é um corpo. Assim, temos a seguinte proposição.

Proposição 1.2.3. *Seja $\mathbf{A}[X]$ um anel de polinômios em uma variável com coeficientes sobre um corpo $\mathbf{A}$. Sejam $f(X), g(X) \in \mathbf{A}[X]$ com $g(X)$ não nulo. Então,*

(i) *Existem $t(X), r(X) \in \mathbf{A}[X]$ tais que $f(X) = g(X) \cdot t(X) + r(X)$ com* $\mathrm{gr(r(X))} < \mathrm{gr(g(X))}$ *ou $r(X) = 0$.*

(ii) *Tais polinômios $t(X)$ e $r(X)$ são unicamente determinados.*

Para entendermos um pouco mais sobre a utilidade da divisão euclidiana, consideremos previamente algumas definições.

Definição 1.2.4. Dado um polinômio $f(X_1, \ldots, X_n) = \sum_{\mathbf{v} \in \mathbb{N}^n} a_{\mathbf{v}} X_1^{v_1} \cdots X_n^{v_n} \in \mathbf{A}[X_1, \ldots, X_n]$ e um elemento $(\alpha_1, \ldots, \alpha_n) \in \mathbf{A}^n$, definimos a *avaliação* de $f(X_1, \ldots, X_n)$ em $(\alpha_1, \ldots, \alpha_n)$, denotada $f(\alpha_1, \ldots, \alpha_n)$, pela seguinte igualdade

$$f(\alpha_1, \ldots, \alpha_n) = \sum_{\mathbf{v} \in \mathbb{N}^n} a_{\mathbf{v}} \alpha_1^{v_1} \cdots \alpha_n^{v_n}.$$

É de fácil verificação que para cada $f(X_1, \ldots, X_n), g(X_1, \ldots, X_n) \in \mathbf{A}[X_1, \ldots, X_n]$ e $(\alpha_1, \ldots, \alpha_n) \in \mathbf{A}^n$ tem-se

$$(f + g)(\alpha_1, \ldots, \alpha_n) = f(\alpha_1, \ldots, \alpha_n) + g(\alpha_1, \ldots, \alpha_n)$$

e

$$(f \cdot g)(\alpha_1, \ldots, \alpha_n) = f(\alpha_1, \ldots, \alpha_n) \cdot g(\alpha_1, \ldots, \alpha_n),$$

ou seja, em linguagem mais elaborada estas igualdades nos dizem que para cada $(\alpha_1, \ldots, \alpha_n) \in \mathbf{A}^n$ a aplicação de $\mathbf{A}[X_1, \ldots, X_n]$ em $\mathbf{A}$ que envia $f(X_1, \ldots, X_n)$ em $f(\alpha_1, \ldots, \alpha_n)$ é um homomorfismo de anéis.

Dizemos que $(\alpha_1, \ldots, \alpha_n) \in \mathbf{A}^n$ é *raiz* (ou *zero*) de $f(X_1, \ldots, X_n)$ se

$$f(\alpha_1, \ldots, \alpha_n) = 0.$$

No caso de uma única variável, temos o seguinte critério para decidir quando um elemento de $\mathbf{A}$ é raiz de um polinômio

Proposição 1.2.5. *Seja* $\mathbf{A}$ *um anel e* $f(X) \in \mathbf{A}[X]$ *um polinômio não constante. Um elemento* $\alpha \in \mathbf{A}$ *é raiz de* $f(X)$ *se, e somente se, o resto da divisão de* $f(X)$ *por* $X - \alpha$ *é zero.*

Prova. ($\Rightarrow$) Se α é raiz de $f(X)$ então, por definição $f(\alpha) = 0$. Além disso, pela divisão euclidiana de $f(X)$ por $X - \alpha$, existem $q(X), r(X) \in A[X]$ tais que

$$f(X) = q(X)(X - \alpha) + r(X)$$

onde $r(X) = 0$ ou $\text{gr}(r(X)) < \text{gr}(X - \alpha) = 1$. Assim, $r(X) = \beta$ com $\beta \in A$. Desse modo,

$$f(X) = q(X).(X - \alpha) + \beta$$

Avaliando $f(X)$ em α, temos

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

ou seja, $f(X) = q(X).(X - \alpha)$ donde concluímos que $(X - \alpha)$ divide $f(X)$.

($\Leftarrow$) Se $(X - \alpha)$ divide $f(X)$ então existe $q(X) \in \mathbf{A}[X]$ tal que

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

Portanto,

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

□

A proposição acima generaliza-se para n variáveis da seguinte maneira

Corolário 1.2.6. *Seja* $\mathbf{A}$ *um anel e* $f(X_1, \ldots, X_n) \in \mathbf{A}[X_1, \ldots, X_n]$ *um polinômio não constante. Um elemento* $(\alpha_1, \ldots, \alpha_n) \in \mathbf{A}^n$ *é raiz de* $f(X_1, \ldots, X_n)$ *se, e somente*

se, existem $g_1, \ldots, g_n \in \mathbf{A}[X_1, \ldots, X_n]$ *tais que*

$$f(X_1, \ldots, X_n) = (X_1 - \alpha_1)g_1 + \ldots + (X_n - \alpha_n)g_n$$

Prova. Se $f(X_1, \ldots, X_n) = (X_1 - \alpha_1)g_1 + \ldots + (X_n - \alpha_n)g_n$ então

$$f(\alpha_1, \ldots, \alpha_n) = (\alpha_1 - \alpha_1)g_1(\alpha_1, \ldots, \alpha_n) + \ldots + (\alpha_n - \alpha_n)g_n(\alpha_1, \ldots, \alpha_n) = 0$$

o que mostra que $(\alpha_1, \ldots, \alpha_n)$ é raiz de $f(X_1, \ldots, X_n)$.

Para provar a recíproca aplicaremos indução sobre o número de variáveis. Para $n = 1$ o resultado segue da Proposição 1.2.5. Agora suponhamos o resultado válido para $n - 1$ ($n > 1$). Fazendo a divisão euclidiana de $f(X_1, \ldots, X_n)$ por $X_n - \alpha_n$ obtemos a igualdade

$$f(X_1, \ldots, X_n) = (X_n - \alpha_n)g_n + r$$

onde r um polinômio constante na variável X_n. Desse modo, $r \in A[X_1, \ldots, X_{n-1}]$. Além disso,

$$r(\alpha_1, \ldots, \alpha_{n-1}) = f(\alpha_1, \ldots, \alpha_n) - (\alpha_n - \alpha_n)g(\alpha_1, \ldots, \alpha_n) = 0.$$

Logo, por hipótese de indução, devemos ter

$$r = (X_1 - \alpha_1)g_1 + \ldots + (X_{n-1} - \alpha_{n-1})g_{n-1}$$

com $g_1, \ldots, g_{n-1} \in A[X_1, \ldots, X_{n-1}] \subset A[X_1, \ldots, X_n]$. Portanto, $f(X_1, \ldots, X_n) = (X_1 - \alpha_1)g_1 + \ldots + (X_n - \alpha_n)g_n$ como desejávamos. □

Uma consequência da caracterização contida na Proposição 1.2.5 acima é o seguinte resultado:

Proposição 1.2.7. *Seja* $\mathbf{A}$ *um domínio e* $f(X) \in \mathbf{A}[X]$ *um polinômio não constante de grau* n*. Então* $f(X)$ *contém no máximo* n *raízes.*

Prova. Seja $f(X) = a_0 + a_1X + \ldots + a_nX^n \in \mathbf{A}[X]$ um polinômio de grau n. Em particular $a_n \neq 0$. Provaremos essa proposição por indução sobre $n = \mathrm{gr}(f(X))$.

Se $n = 0$, então $f(X) = a_0 \neq 0$ para todo $x \in \mathbf{A}$. logo $f(X)$ não tem raízes em $\mathbf{A}$ e, portanto, a proposição é válida.

Suponhamos que a proposição seja válida para todo polinômio $q(X) \in A[X]$ tal que $\text{gr}(q(X)) = n$. Considere $f(X) \in \mathbf{A}[X]$ tal que $\text{gr}(f(X)) = n + 1$.

Se $f(X)$ não possui raízes em $\mathbf{A}$ não há o que demonstrar. Suponhamos então que $f(X)$ tenha uma raiz $\alpha \in \mathbf{A}$, ou seja, $f(\alpha) = 0$. Pela Proposição 1.2.5 segue que $(X - \alpha)$ divide $f(X)$ em $\mathbf{A}[X]$; logo, existe $q(X) \in \mathbf{A}[X]$ tal que

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

Note que $\text{gr}(q(X)) = n$ assim, por hipótese de indução, $q(X)$ tem no máximo n raízes em A. Note ainda que

$$\begin{aligned} \beta \text{ é raiz de } f(X) &\iff 0 = f(\beta) = q(\beta)(\beta - \alpha) \\ &\iff q(\beta) = 0 \text{ ou } \beta - \alpha = 0 \\ &\iff \beta \text{ é raiz de } \ q(X) \text{ ou } \beta = \alpha. \end{aligned}$$

Logo, $f(X)$ tem no máximo $n + 1$ raízes em $\mathbf{A}$, o que conclui a demonstração. □

Observação 1.2.8. Notemos que os resultados contidos nas três últimas proposições acima dizem respeito à discussão das raízes de um polinômio particular. Para deduzir todas essas afirmações foi imprescindível a utilização do algorítmo da divisão. O que iremos discutir mais adiante é uma versão mais geral de algorítmo da divisão que permita realizar uma discussão mais ampla sobre raízes de polinômios.

1.3 Distinção entre polinômios e funções polinomiais.

Dado um polinômio $f(X_1, \ldots, X_n) \in \mathbf{A}[X_1, \ldots, X_n]$ definimos a seguinte função $\widetilde{f}$ com domínio em $\mathbf{A}^n$ e contradomínio em $\mathbf{A}$

$$\begin{array}{rccc} \widetilde{f}: & \mathbf{A}^n & \to & \mathbf{A} \\ & (\alpha_1, \ldots, \alpha_n) & \mapsto & f(\alpha_1, \ldots, \alpha_n) \end{array}$$

Esta função $\widetilde{f}$ é denominada de *função polinomial* (induzida pelo polinômio f).

Notemos que desde o ensino fundamental somos apresentados às funções polinomiais, sendo que neste contexto básico a ênfase ao estudo dessas funções é dada para aquelas em que o número de variáveis é 1 e o grau é pequeno, tipicamente 1 ou 2.

Outra particularidade do ensino das funções polinomiais em nível elementar é o anel envolvido, que neste caso é $\mathbb{R}$ ou $\mathbb{C}$.

Como é claro pela definição, as funções polinomiais surgem de polinômios e, portanto, são objetos formalmente distintos. Inclusive, a correspondência $f \mapsto \widetilde{f}$ não é necessariamente bijetora como nos mostra o seguinte exemplo

Exemplo 1.3.1. Considere $\mathbf{A} = \mathbb{Z}_p$, com p primo. Os polinômios $p(X) = 0$ e $q(X) = X^p - X$ são distintos. Contudo, como $\alpha^p - \alpha = 0$ para cada $\alpha \in \mathbb{Z}_p$, (vide pequeno teorema de Fermat em [6, capítulo 7]) temos que p e q induzem as mesmas funções polinomiais.

Apesar da distinção formal entre as noções de polinômio e função polinomial, na maioria dos livros de ensino médio tal distinção não é mencionada. De fato, o que veremos na proposição a seguir é que tal postura adotada pelos livros didáticos não é absurda e, em um certo sentido, é correta.

Proposição 1.3.2. *Seja* $\mathbf{A}$ *um domínio infinito e* $f(X) \in \mathbf{A}[X]$. *Então* $f(X)$ *é o polinômio nulo se, e somente se, a função polinomial correspondente é identicamente nula. Em particular, dois polinômios em* $\mathbf{A}[X]$ *são iguais se, e somente se, as respectivas funções polinomiais são iguais.*

Prova. ($\Rightarrow$) Se $f(X)$ é o polinômio nulo, então $f(\alpha) = 0$ para todo $\alpha \in \mathbf{A}$, ou seja, a correspondente função polinomial é $f(X) = 0$.

($\Leftarrow$) Suponha que o polinômio $f(X)$ seja não nulo. Pela Proposição 1.2.7 esse polinômio possui uma quantidade finita de raízes. Se a função polinomial correspondente é identicamente nula, então $f(\alpha) = 0$ para todo $\alpha \in \mathbf{A}$. Como $\mathbf{A}$ é infinito, então $f(X)$ tem infinitas raízes, o que contraria a hipótese. Logo $f(X)$ é um polinômio nulo.

Para a segunda parte da proposição, sejam $f(X)$ e $g(X)$ dois polinômios em $\mathbf{A}[X]$ tais que $f(\alpha) = g(\alpha)$ para todo $\alpha \in \mathbf{A}$. Considere o polinômio $f(X) - g(X)$. Notemos que todo $\alpha \in \mathbf{A}$ é raiz desse polinômio. Assim, pela primeira parte da proposição, o polinômio $f(X) - g(X)$ deve ser nulo. Logo, $f(X) = g(X)$. A recíproca é trivial. □

1.4 Teorema da base de Hilbert

Seja $\mathbf{A}$ um anel. Um subconjunto I de $\mathbf{A}$ é chamado um *ideal* de $\mathbf{A}$ se satisfaz as seguintes condições:

(i) $0 \in I$.

(ii) Se $a, b \in I$ então $a - b \in I$.

(iii) Se $a \in \mathbf{A}$ e $b \in I$, então $a \cdot b \in I$.

Exemplo 1.4.1. Em um anel $\mathbf{A}$, temos que $I = \{0\}$ e $I = \mathbf{A}$ são exemplos óbvios de ideais. Tais ideais são chamados *ideais triviais* de $\mathbf{A}$.

Exemplo 1.4.2. Dado um anel $\mathbf{A}$ e um subconjunto S de $\mathbf{A}$, consideremos o seguinte conjunto

$$I = \{a_1x_1 + \ldots + a_nx_n \mid a_i \in \mathbf{A},\ x_i \in S,\ n \in \mathbb{N}\}.$$

É de fácil verificação que I assim definido é um ideal de $\mathbf{A}$. Dizemos neste caso que I é um ideal *gerado* por S ou que S é um conjunto gerador de I. Também é comum dizermos que S é uma *base* para o ideal I. Doravante, utilizaremos a notação $I = \langle S \rangle$ para denotar que I é gerado por S.

Dentre todos os ideais, os mais manejáveis são aqueles que possuem um conjunto finito de geradores. Tais ideais são chamados de *ideais finitamente gerados*.

Definição 1.4.3. Um anel $\mathbf{A}$ é dito *anel Noetheriano* se todo ideal de $\mathbf{A}$ é finitamente gerado.

Teorema 1.4.4. *Seja* $\mathbf{A}$ *um anel. As seguintes afirmações são equivalentes:*

(i) $\mathbf{A}$ *é anel Noetheriano.*

(ii) *Para toda cadeia ascendente de ideais de* $\mathbf{A}$, $I_1 \subset I_2 \subset \ldots \subset I_r \subset \ldots$ *existe* $n_0 \in \mathbb{N}$ *tal que* $I_{n_0} = I_n$ *para cada* $n \geq n_0$.

(iii) *Toda família* $\mathfrak{F}$ *não vazia de ideais de* $\mathbf{A}$ *possui elemento máximo* (*com respeito a ordem de inclusão*).

Teorema 1.4.5. (Teorema da base de Hilbert) *Se* $\mathbf{A}$ *é anel Noetheriano então* $\mathbf{A}[X]$ *também é anel Noetheriano.*

Prova. Mostraremos que se I é um ideal de $\mathbf{A}[X]$ então ele é finitamente gerado. Note que se $I = \{0\}$ então não há o que provar. Assim, considere $I \neq \{0\}$ e defina f_1 como sendo um polinômio de I que possua grau mínimo. Se $I = \langle f_1 \rangle$, então a prova está encerrada. Caso contrário, escolhemos $f_2 \in I - \langle f_1 \rangle$ com grau mínimo e

se $I = \langle f_1, f_2 \rangle$ encerramos a prova. Continuando dessa maneira de forma indutiva, definamos para cada i,

$$a_i := \text{ coeficiente líder de } f_i.$$

Temos assim uma cadeia ascendente de ideais $\langle a_1 \rangle \subset \langle a_1, a_2 \rangle \subset \ldots$ em $\mathbf{A}$. Como $\mathbf{A}$ é Noetheriano, existe índice m tal que esta cadeia estabiliza em $\langle a_1, \ldots, a_m \rangle$. Afirmamos que $I = \langle f_1, \ldots, f_m \rangle$. Caso contrário, o processo de escolha acima seleciona um elemento $f_{m+1} \in I - \langle f_1, \ldots, f_m \rangle$. Note que podemos escrever $a_{m+1} = u_1 a_1 + \ldots + u_m a_m$. Como o grau de f_{m+1} é pelo menos $\max\{\text{gr}(f_i) \mid 1 \leq i \leq m\}$, então podemos definir o seguinte polinômio tendo o mesmo grau e coeficiente líder de f_{m+1}

$$g = \sum_{j=1}^{m} u_j f_j X^{\text{gr}(f_{m+1}) - \text{gr}(f_j)} \in \langle f_1, \ldots, f_m \rangle$$

Note que $f_{m+1} - g \in I - \langle f_1, \ldots, f_m \rangle$ e $\text{gr}(f_{m+1} - g) < \text{gr}(f_{m+1})$, o que contraria a escolha de f_{m+1}. □

Corolário 1.4.6. *Se* $\mathbf{A}$ *é anel Noetheriano, então* $\mathbf{A}[X_1, \ldots, X_n]$ *também é anel Noetheriano.*

Prova. Em virtude da igualdade $A[X_1, \ldots, X_n] = (A[X_1, \ldots, X_{n-1}])[X_n]$, o resultado segue utilizando indução e o Teorema da Base de Hilbert □

Corolário 1.4.7. *Se* k *é um corpo, então* $k[X_1, \ldots, X_n]$ *é um anel Noetheriano.*

Prova. Ora, os únicos ideais de um corpo k são os triviais que são finitamente gerados (de fato, $\{0\} = \langle 0 \rangle$ e $k = \langle 1 \rangle$). Assim, k é Noetheriano. Logo, pelo corolário acima temos o desejado. □

Capítulo 2

Sistemas de equações polinomiais

Neste capítulo apresentamos o principal objeto deste trabalho, ou seja, os sistemas de equações polinomiais. Destacamos a relação existente entre os conjuntos soluções de tais sistemas e os ideais de um anel de polinômios apropriado. Esta relação será a ponte para as bases de Grobner, como veremos no capítulo 3. Colocaremos também, em relevo, a importância dos sistemas de equações polinomiais para a matemática e para problemas do mundo real.

2.1 Terminologia

Seja k um corpo (que para fixação de ideias podemos pensar $k = \mathbb{R}$ ou $k = \mathbb{C}$) e n um inteiro positivo. Diante desses dados, podemos pensar nos seguintes universos: $k[X_1, \ldots, X_n]$ (de natureza algébrica) e $k^n = \underbrace{k \times \ldots \times k}_{n\ vezes}$ (de natureza geométrica).

Primeiro estabelecemos uma correspondência de $k[X_1, \ldots, X_n]$ para k^n através da seguinte definição.

Definição 2.1.1. Seja S um subconjunto de $k[X_1, \ldots, X_n]$. O *conjunto dos zeros* de S, denotado $Z(S)$, é

$$Z(S) = \{(\alpha_1, \ldots, \alpha_n) \in k^n \mid f(\alpha_1, \ldots, \alpha_n) = 0 \text{ para cada } f(X_1, \ldots, X_n) \in S\}.$$

Se S é finito, digamos $S = \{f_1, \ldots, f_m\}$, então $Z(S)$ é o conjunto solução do

sistema de equações polinomiais

$$\begin{cases} f_1(x_1, \ldots, x_n) &= 0 \\ f_2(x_1, \ldots, x_n) &= 0 \\ \vdots & \vdots \\ f_m(x_1, \ldots, x_n) &= 0 \end{cases} \tag{2.1}$$

Exemplo 2.1.2. Para $k = \mathbb{R}$, $n = 2$, os seguintes exemplos nos são familiares desde o ensino básico:

(a) Para $S = \{2X + 3Y - 1\}$, $Z(S)$ é a reta determinada pelos pontos $(-1, 3)$ e $(2, -1)$.

(b) Para $S = \{X^2 + Y^2 - 1\}$, $Z(S)$ é a circunferência unitária centrada na origem.

(c) Para $S = \{4X^2 + 9Y^2 - 36\}$, $Z(S)$ é a elipse com focos nos pontos $(-\sqrt{5}, 0)$ e $(\sqrt{5}, 0)$.

A seguir mostramos como os conjuntos de zeros comportam-se diante da relação de inclusão

Proposição 2.1.3. *Sejam S_1 e S_2 subconjunto de $k[X_1, \ldots, X_n]$ tais que $S_1 \subset S_2$. Então, $Z(S_1) \supset Z(S_2)$.*

Prova. Seja p um ponto de $Z(S_2)$. Como, $S_1 \subset S_2$ e p é anulado por todo polinômio de S_2, então, por mais forte razão p é anulado por todo polinômio de S_1. Logo, p também pertence a $Z(S_1)$. Portanto, temos a inclusão desejada. □

A proposição a seguir nos mostra que para qualquer $S \subset k[X_1, \ldots, X_n]$, o conjunto $Z(S)$ pode ser visto como o conjunto solução de um sistema de equações polinomiais tal como em (2.1)

Proposição 2.1.4. *Seja $S \subset k[X_1, \ldots, X_n]$ e $I = (S)$. Então, $Z(S) = Z(I)$. Em particular, existem $f_1, \ldots, f_n \in k[X_1, \ldots, X_n]$ tais que $Z(S) = Z(f_1, \ldots, f_r)$.*

Prova. Como $S \subset I$, então pela Proposição 2.1.3 temos $Z(I) \subset Z(S)$. Agora, suponhamos $(\alpha_1, \ldots, \alpha_n)$ um ponto de $Z(S)$. Um elemento arbitrário $f(X_1, \ldots, X_n)$ de I se escreve na forma

$$f(X_1, \ldots, X_n) = h_1(X_1, \ldots, X_n)g_1(X_1, \ldots, X_n) + \ldots + h_m(X_1, \ldots, X_n)g_m(X_1, \ldots, X_n)$$

onde $h_1, \ldots, h_m \in k[X_1, \ldots, X_n]$ e $g_1(X_1, \ldots, X_n), \ldots g_m(X_1, \ldots, X_n) \in S$. Assim,

$$\begin{aligned} f(\alpha_1, \ldots, \alpha_n) &= h_1(\alpha_1, \ldots, \alpha_n)g_1(\alpha_1, \ldots, \alpha_n) + \ldots + h_m(\alpha_1, \ldots, \alpha_n)g_m(\alpha_1, \ldots, \alpha_n) \\ &= h_1(\alpha_1, \ldots, \alpha_n) \cdot 0 + \ldots + h_m(\alpha_1, \ldots, \alpha_n) \cdot 0 = 0. \end{aligned}$$

Portanto, $(\alpha_1, \ldots, \alpha_n)$ é zero de qualquer polinômio de $Z(I)$. Logo, a igualdade $Z(S) = Z(I)$ segue.

Pelo teorema da base de Hilbert temos a existência de $f_1, \ldots, f_r$ tais que $I = (f_1, \ldots, f_r)$. Usando a primeira parte da proposição conluimos que $Z(S) = Z(f_1, \ldots, f_r)$.

□

Diante desse resultado, podemos agora restringir a discussão a sistemas de equações como em (2.1). A pergunta natural é:

Questão 2.1.5. *Como determinar o conjunto solução de um sistema como em* (2.1)?

Esta questão, por sua vez, nos remete a um outro questionamento

Questão 2.1.6. *O que significa determinar o conjunto solução de um sistema de equações polinomiais?*

No caso em que o conjunto solução é finito, o termo "determinar" fica melhor comprendido. De fato, em tal caso determinar significa listar todas as soluções. Entretanto, quando o conjunto solução é infinito o termo fica meio vago. Nessa situação, determinar pode significar fornecer propriedades qualitativas que individualizem um dado conjunto solução de sistema de equações polinomiais, ou fornecer uma descrição paramétrica para o conjunto solução. Neste trabalho, estamos mais interessados no segundo significado.

Notemos que quando os polinômios f_i em (2.1) são de grau 1, sabemos por métodos da álgebra linear como parametrizar as soluções. Nesse caso particular as parametrizações são obtidas efetuando-se operações sistemáticas com as equações e obtendo sistemas equivalentes cada vez mais simples. De fato, essa é uma filosofia que de certa forma sobrevive ao caso geral como veremos no Capítulo 3.

Agora apresentamos uma correspondência que vai no sentido contrário da que definimos em 2.1.1.

Definição 2.1.7. Seja X um subconjunto de k^n. Definimos o conjunto $I(X)$ por

$$I(X) = \{f(X_1, \ldots, X_n) \in k[X_1, \ldots, X_n] \mid f(\alpha_1, \ldots, \alpha_n) = 0 \text{ para cada } (\alpha_1, \ldots, \alpha_n) \in X\}.$$

É de fácil verificação que o conjunto $I(X)$ é um ideal de $k[X_1, \ldots, X_n]$.

Exemplo 2.1.8. Para $X = \{(\alpha_1, \ldots, \alpha_n)\} \subset k^n$ temos que o ideal $I(X)$ é igual a $(X_1 - \alpha_1, \ldots, X_n - \alpha_n)$. A inclusão $(X_1 - \alpha_1, \ldots, X_n - \alpha_n) \subset K[X_1, \ldots, X_n]$ é trivial. Esta afirmação é consequência imediata do Corolário 1.2.6.

Temos, assim, as seguintes correspondências

$$\{\text{subconjunto de } k[\mathbf{X}]\} \xrightarrow{Z} \{\text{subconjunto de } k^n\}$$

$$\{\text{subconjunto de } k^n\} \xrightarrow{I} \{\text{subconjunto de } k[\mathbf{X}]\}.$$

Temos as seguintes propriedades

Proposição 2.1.9. *Se V_1, V_2 são subconjuntos de k^n tais que $V_1 \subset V_2$ então $I(V_2) \subset I(V_1)$.*

Prova. Seja f um polinômio de $I(V_2)$. Então, cada elemento de V_2 é zero de f. Como $V_1 \subset V_2$, então, por mais forte razão, cada elemento de V_1 é zero de f. Logo, $f \in I(V_1)$. Portanto, temos a inclusão desejada. □

2.2 Problemas modelados por sistemas de equações polinomiais

Estudar conjuntos soluções de sistemas de equações polinomiais é algo interessante tanto do ponto de vista teórico quanto de suas aplicações. Teoricamente eles são utilizados, por exemplo, para realizar estudos aproximados de objetos mais complexos. Por outro lado, também temos diversas situações reais que são diretamente descritas por conjuntos soluções de equações polinomiais. Nessa seção apresentaremos algumas destas situações.

2.2.1 Conjunto de configurações de um braço mecânico

Apresentamos um modelo simples de braço mecânico composto por duas hastes de comprimentos 1 e 2 como nos mostra a Figura 2.1

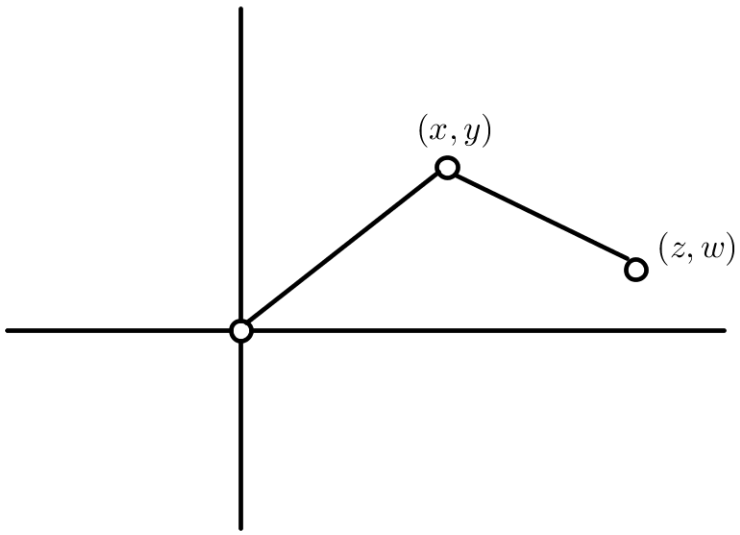

Figura 2.1:

A configuração do braço é completamente determinada pelas coordenadas (x, y) e (z, w) como ilustrado na figura. Assim, se quisermos determinar todas as configurações possíveis para esse braço mecânico, devemos determinar o conjunto solução do seguinte sistema de equações polinomiais

$$\begin{cases} X^2 + Y^2 & = & 4 \\ (X - Z)^2 + (Y - W)^2 & = & 1 \end{cases}$$

2.2.2 O problema de coloração

A história envolvendo problemas sobre coloração de mapas começou em meados do século XIX, quando Francis Guthrie tentava colorir os distritos do mapa da Inglaterra de modo que dois distritos vizinhos tivessem cores distintas. Após analisar cuidadosamente o problema, Francis Guthrie conjecturou que qualquer mapa poderia ser colorido com apenas quatro cores, porém não conseguiu desenvolver uma prova formal para o problema.

Apesar de diversas tentativas terem sido feitas por renomados matemáticos, a demonstração para a conjectura proposta por Francis ficou em aberto por cerca de 120 anos, até que em 1976, com a ajuda de um IBM 360, Kenneth Appel e Wolfgang Haken apresentaram uma prova do que é conhecido hoje como *teorema das quatro cores*. Entretanto, o fato de Appel e Haken terem usado computadores de alta velocidade por

mais de mil horas para concluir a demonstração do teorema não agradou a comunidade científica, que julgava impossível verificar os cálculos apresentados.

Desse modo, muitos matemáticos continuaram a buscar uma prova mais simples para o teorema das quatro cores. Em 1994, Paul D. Seymour apresentou uma prova que reduzia a quantidade de cálculos a níveis bastante aceitáveis, entretanto não conseguiu dispensar o uso do computador. Uma prova que não necessite do auxílio de computadores é uma questão que até hoje continua em aberto.

Como já foi citado acima, é sempre possível colorir um mapa com quatro cores. Em algumas situações específicas é possível usar menos cores. Por exemplo, se o mapa não tiver regiões de tríplice frontera, ou seja, uma região que tem fronteira com outras duas, podemos colori-lo com apenas duas cores. Entretanto, sob que circunstâncias podemos colorir um mapa com apenas três cores? Seria possível, por exemplo, usar apenas três cores para colorir o mapa da região nordeste do Brasil? Nessa seção apresentamos como estas questões podem ser traduzidas por meio de um sistema de equações polinomiais. No capítulo seguinte, depois de apresentada a ferramenta da base de Grobner, mostraremos a solução final do problema.

Como queremos usar três cores para colorir o mapa, com o intuito de "algebrizar"o problema, vamos representar cada cor por uma raiz cúbica da unidade, ou seja, cada cor será uma raiz do polinômio $f = X^3 - 1$ e cada região por uma incógnita X_i que pode assumir apenas um desses três valores/cores possíveis. Então temos que $X_i^3 - 1 = 0$ para todo $i = 1, \ldots, n$, onde n representa o número de regiões do mapa.

Como duas regiões vizinhas X_i e X_j são raízes da unidade, temos

$$X_i^3 - 1 = 0 \quad \text{e} \quad X_j^3 - 1 = 0,$$

ou seja, $X_i^3 = X_j^3$. Logo,

$$X_i^3 - X_j^3 = 0 \Rightarrow (X_i - X_j)(X_i^2 + X_iX_j + X_j^2) = 0.$$

Como $X_i \neq X_j$ (pois são regiões vizinhas), temos

$$X_i^2 + X_iX_j + X_j^2 = 0.$$

Desse modo, o problema da coloração de mapas usando três cores se resume a obter as soluções do seguinte sistema de equações polinomiais:

$$\begin{cases} X_i^3 - 1 = 0, & \text{para} \quad 1 \leq i \leq n \\ X_i^2 + X_i X_j + X_j^2 = 0, & \text{sempre que} \quad X_i \text{ e } X_j \text{ forem regiões vizinhas.} \end{cases} \tag{2.2}$$

Observemos agora o mapa da região nordeste do Brasil:

Figura 2.2: Região nordeste do Brasil

A cada um dos nove estados que compõem esta região vamos atribuir uma variável de $\mathbb{C}[X_1, \ldots, X_9]$. Definamos então, sem perda de generalidade, que

$X_1 =$ Sergipe	$X_4 =$ Paraíba	$X_7 =$ Piauí
$X_2 =$ Alagoas	$X_5 =$ Rio Grande do Norte	$X_8 =$ Maranhão
$X_3 =$ Pernambuco	$X_6 =$ Ceará	$X_9 =$ Bahia

Assim, temos a seguinte distribuição no mapa:

Figura 2.3: Estados do nordeste representados por variáveis X_i com $i = 1, \ldots, 9$.

Determinar se é possível colorir o mapa acima com apenas três cores é o mesmo que determinar as soluções, se existirem, do seguinte sistema de equações polinomiais:

$$
\begin{array}{lll}
\bullet X_1^3 - 1 = 0 & \bullet X_1^2 + X_1X_2 + X_2^2 = 0 & \bullet X_4^2 + X_4X_6 + X_6^2 = 0 \\
\bullet X_2^3 - 1 = 0 & \bullet X_1^2 + X_1X_9 + X_9^2 = 0 & \bullet X_5^2 + X_5X_6 + X_6^2 = 0 \\
\bullet X_3^3 - 1 = 0 & \bullet X_2^2 + X_2X_3 + X_3^2 = 0 & \bullet X_6^2 + X_6X_7 + X_7^2 = 0 \\
\bullet X_4^3 - 1 = 0 & \bullet X_2^2 + X_2X_9 + X_9^2 = 0 & \bullet X_7^2 + X_7X_8 + X_8^2 = 0 \\
\bullet X_5^3 - 1 = 0 & \bullet X_3^2 + X_3X_4 + X_4^2 = 0 & \bullet X_7^2 + X_7X_9 + X_9^2 = 0 \\
\bullet X_6^3 - 1 = 0 & \bullet X_3^2 + X_3X_9 + X_9^2 = 0 & \\
\bullet X_7^3 - 1 = 0 & \bullet X_3^2 + X_3X_6 + X_6^2 = 0 & \\
\bullet X_8^3 - 1 = 0 & \bullet X_3^2 + X_3X_7 + X_7^2 = 0 & \\
\bullet X_9^3 - 1 = 0 & \bullet X_4^2 + X_4X_5 + X_5^2 = 0 &
\end{array}
\tag{2.3}
$$

Na última seção do capítulo 3 mostraremos como esse sistema pode ser resolvido à luz das bases de Grobner.

2.2.3 O Puzzle Sudoku

O Sudoku é um *puzzle* (ou quebra cabeça) que se tornou muito popular no Japão, país onde jogos numéricos são bem mais populares que jogos do tipo *palavras cruzadas* e *caça palavras*, no ano de 1986, quando a Nikoli, revista de raciocínio lógico oriental, que descobriu o jogo em 1984, resolveu levá-lo para aquele país. Apesar de muitos acreditarem que o sudoku é de origem nipônica, ele foi projetado pelo arquiteto e designer norte-americano Howard Garns em 1979, que batizou o puzzle por *Number Place*, nome que é usado até hoje nos Estados Unidos.

A partir de 2004, quando o jornal britânico *The Times* passou a publicar o Sudoku em suas edições diárias, o jogo finalmente se tornou mundialmente popular.

O jogo é composto por uma grade 9×9 dividida em nove sub-grades 3×3 onde algumas cédulas já contém números entre 1 e 9. O objetivo do jogo é preeencher as cédulas vazias, com um número em cada cédula, de modo que em cada linha, coluna ou sub-grade, os números de 1 a 9 apareçam apenas uma vez. Segue abaixo um exemplo e sua respectiva solução:

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Figura 2.4: Exemplo

5	3	4	6	7	8	9	1	2
6	7	2	1	9	5	3	4	8
1	9	8	3	4	2	5	6	7
8	5	9	7	6	1	4	2	3
4	2	6	8	5	3	7	9	1
7	1	3	9	2	4	8	5	6
9	6	1	5	3	7	2	8	4
2	8	7	4	1	9	6	3	5
3	4	5	2	8	6	1	7	9

Figura 2.5: Solução

Embora o objetivo seja resolver o jogo usando basicamente o raciocínio lógico, é perfeitamente possível obter sua solução associando-o a um sistema de equações polinomiais. Para tanto, vamos transformar o problema de resolver um sudoku em um problema de coloração de mapas de modo que:

(i) cada cédula será uma região do mapa a ser pintada.

(ii) duas regiões serão consideradas vizinhas (ou adjacentes) se estiverem na mesma linha, coluna ou sub-grade.

(iii) duas regiões vizinhas não poderão ser pintadas de uma mesma cor.

Desse modo, a solução desse problema se dará de maneira bastante semelhante ao problema de coloração do mapa da região nordeste discutido anteriormente. Nesse

caso, queremos colorir o mapa (grade 9×9) usando 9 cores (números de 1 a 9) sem que regiões vizinhas sejam pintadas com a mesma cor. Assim, associando cada cor a uma raiz nona da unidade, e cada região a uma incógnita que só pode assumir um desses nove valores/cores possíveis, temos 81 equações do tipo $X_i^9 - 1 = 0$ com $i \in \{1, \ldots, 81\}$ compondo o sistema. Além disso, sendo X_i e X_j duas regiões distintas do mapa, temos que:

$$\begin{aligned} X_i^9 - X_j^9 = 0 &\Leftrightarrow \left(X_i^3\right)^3 - \left(X_j^3\right)^3 = 0 \\ &\Leftrightarrow (X_i^3 - X_j^3)(X_i^6 + X_i^3 X_j^3 + X_j^6) = 0 \\ &\Leftrightarrow (X_i - X_j)(X_i^2 + X_i X_j + X_j^2)(X_i^6 + X_i^3 X_j^3 + X_j^6) = 0 \\ &\Leftrightarrow X_i - X_j = 0 \text{ ou } (X_i^2 + X_i X_j + X_j^2)(X_i^6 + X_i^3 X_j^3 + X_j^6) = 0 \end{aligned}$$

Quando X_i e X_j são regiões vizinhas, $X_i - X_j \neq 0$, então

$$(X_i^2 + X_i X_j + X_j^2)(X_i^6 + X_i^3 X_j^3 + X_j^6) = 0 \tag{2.4}$$

sempre que X_i e X_j tem fronteira em comum. Note que cada região do mapa/tabuleiro faz fronteira com com 20 outras regiões, como mostra a seguinte figura

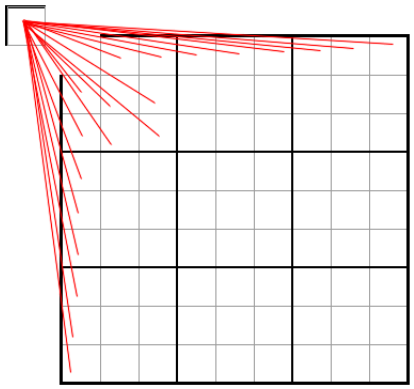

Figura 2.6: Regiões de fronteira

Como o mapa tem 81 regiões e cada uma delas faz fronteira com outras 20, o sistema terá $\frac{81 \cdot 20}{2} = 810$ equações do tipo 2.4. Adicionando a esse número as 81 equações do tipo $X_i^9 - 1 = 0$, temos um total de $810 + 81 = 891$ equações.

Embora o processo de resolução do Sudoku usando bases de Gröbner seja perfeitamente possível por tudo que discutimos acima, o grande número de equações que

compõem o sistema torna o processo demasiadamente longo, mesmo com o auxílio de softwares. Assim, decidimos descrever o processo de resolução de uma versão mais fraca do Sudoku conhecida como *Shidoku*.

O Shidoku consiste numa grade 4×4 dividida em 4 subgrades 2×2 onde algumas cédulas já contém números entre 1 e 4. O objetivo do jogo é preeencher as cédulas vazias, com um número em cada cédula, de modo que em cada linha, coluna ou sub-grade, os números de 1 a 4 apareçam apenas uma vez.

Vamos então determinar a solução do Shidoku abaixo:

			4
4		2	
	3		1
1			

Figura 2.7: Shidoku

Note que esse é um problema de coloração de mapas usando 4 cores. Vamos então, inicialmente, representar cada cédula (região) por X_i com $i \in \{1, \ldots, 16\}$ como mostra a seguinte figura:

X_1	X_2	X_3	X_4
X_5	X_6	X_7	X_8
X_9	X_{10}	X_{11}	X_{12}
X_{13}	X_{14}	X_{15}	X_{16}

Figura 2.8: Cada região é representada por X_i com $i \in \{1, \ldots, 16\}$

Como queremos colorir o mapa com 4 cores, vamos associar cada cor $(1, 2, 3, 4)$ a uma raiz quarta da unidade $(1, i, -1, -i)$ tal que

$$1 \leftrightarrow 1, \quad 2 \leftrightarrow i \quad 3 \leftrightarrow -1, \quad 4 \leftrightarrow -i.$$

Desse modo, temos 16 equações do tipo $X_i^4 - 1 = 0$ com $i \in \{1, \ldots, 16\}$. Além disso, dadas duas regiões X_i e X_j quaisquer, temos

$$
\begin{aligned}
X_i^4 - X_j^4 = 0 &\Leftrightarrow \left(X_i^2\right)^2 - \left(X_j^2\right)^2 = 0 \\
&\Leftrightarrow (X_i^2 - X_j^2)(X_i^2 + X_j^2) = 0 \\
&\Leftrightarrow (X_i - X_j)(X_i + X_j)(X_i^2 + X_j^2) = 0 \\
&\Leftrightarrow (X_i - X_j)(X_i^3 + X_i X_j^2 + X_i^2 X_j + X_j^3) = 0 \\
&\Leftrightarrow X_i - X_j = 0 \text{ ou } (X_i^3 + X_i X_j^2 + X_i^2 X_j + X_j^3) = 0
\end{aligned}
$$

Assim, sempre que X_i e X_j são regiões vizinhas temos que

$$X_i^3 + X_i X_j^2 + X_i^2 X_j + X_j^3 = 0. \tag{2.5}$$

Como cada região é vizinha de outras 7 regiões, teremos um total de $\frac{7\times 16}{2} = 56$ equações do tipo (2.5), ou seja, o sistema de equações polinomiais associado ao problema será formado por $16 + 56 = 72$ equações, a saber

$$
\begin{array}{lll}
\bullet X_1^4 - 1 = 0 & \bullet X_2^3 + X_2 X_4^2 + X_2^2 X_4 + X_4^3 = 0 & \bullet X_7^3 + X_7 X_{11}^2 + X_7^2 X_{11} + X_{11}^3 = 0 \\
\bullet X_2^4 - 1 = 0 & \bullet X_2^3 + X_2 X_5^2 + X_2^2 X_5 + X_5^3 = 0 & \bullet X_7^3 + X_7 X_{15}^2 + X_7^2 X_{15} + X_{15}^3 = 0 \\
\bullet X_3^4 - 1 = 0 & \bullet X_2^3 + X_2 X_6^2 + X_2^2 X_6 + X_6^3 = 0 & \bullet X_8^3 + X_8 X_{12}^2 + X_8^2 X_{12} + X_{12}^3 = 0 \\
\bullet X_4 + i = 0 & \bullet X_2^3 + X_2 X_{10}^2 + X_2^2 X_{10} + X_{10}^3 = 0 & \bullet X_8^3 + X_8 X_{16}^2 + X_8^2 X_{16} + X_{16}^3 = 0 \\
\bullet X_5 + i = 0 & \bullet X_2^3 + X_2 X_{14}^2 + X_2^2 X_{14} + X_{14}^3 = 0 & \bullet X_9^3 + X_9 X_{10}^2 + X_9^2 X_{10} + X_{10}^3 = 0 \\
\bullet X_6^4 - 1 = 0 & \bullet X_3^3 + X_3 X_4^2 + X_3^2 X_4 + X_4^3 = 0 & \bullet X_9^3 + X_9 X_{11}^2 + X_9^2 X_{11} + X_{11}^3 = 0 \\
\bullet X_7 - i = 0 & \bullet X_3^3 + X_3 X_7^2 + X_3^2 X_7 + X_7^3 = 0 & \bullet X_9^3 + X_9 X_{12}^2 + X_9^2 X_{12} + X_{12}^3 = 0 \\
\bullet X_8^4 - 1 & \bullet X_3^3 + X_3 X_8^2 + X_3^2 X_8 + X_8^3 = 0 & \bullet X_9^3 + X_9 X_{13}^2 + X_9^2 X_{13} + X_{13}^3 = 0 \\
\bullet X_9^4 - 1 = 0 & \bullet X_3^3 + X_3 X_{11}^2 + X_3^2 X_{11} + X_{11}^3 = 0 & \bullet X_9^3 + X_9 X_{14}^2 + X_9^2 X_{14} + X_{14}^3 = 0 \\
\bullet X_{10} + 1 = 0 & \bullet X_3^3 + X_3 X_{15}^2 + X_3^2 X_{15} + X_{15}^3 = 0 & \bullet X_{10}^3 + X_{10} X_{11}^2 + X_{10}^2 X_{11} + X_{11}^3 = 0 \\
\bullet X_{11}^4 - 1 = 0 & \bullet X_4^3 + X_4 X_7^2 + X_4^2 X_7 + X_7^3 = 0 & \bullet X_{10}^3 + X_{10} X_{12}^2 + X_{10}^2 X_{12} + X_{12}^3 = 0 \\
\bullet X_{12} - 1 = 0 & \bullet X_4^3 + X_4 X_8^2 + X_4^2 X_8 + X_8^3 = 0 & \bullet X_{10}^3 + X_{10} X_{13}^2 + X_{10}^2 X_{13} + X_{13}^3 = 0 \\
\bullet X_{13} - 1 = 0 & \bullet X_4^3 + X_4 X_{12}^2 + X_4^2 X_{12} + X_{12}^3 = 0 & \bullet X_{10}^3 + X_{10} X_{14}^2 + X_{10}^2 X_{14} + X_{14}^3 = 0 \\
\bullet X_{14}^4 - 1 = 0 & \bullet X_4^3 + X_4 X_{16}^2 + X_4^2 X_{16} + X_{16}^3 = 0 & \bullet X_{11}^3 + X_{11} X_{12}^2 + X_{11}^2 X_{12} + X_{12}^3 = 0 \\
\bullet X_{15}^4 - 1 = 0 & \bullet X_5^3 + X_5 X_6^2 + X_5^2 X_6 + X_6^3 = 0 & \bullet X_{11}^3 + X_{11} X_{15}^2 + X_{11}^2 X_{15} + X_{15}^3 = 0 \\
\bullet X_{16}^4 - 1 = 0 & \bullet X_5^3 + X_5 X_7^2 + X_5^2 X_7 + X_7^3 = 0 & \bullet X_{11}^3 + X_{11} X_{16}^2 + X_{11}^2 X_{16} + X_{16}^3 = 0 \\
\bullet X_1^3 + X_1 X_2^2 + X_1^2 X_2 + X_2^3 = 0 & \bullet X_5^3 + X_5 X_8^2 + X_5^2 X_8 + X_8^3 = 0 & \bullet X_{12}^3 + X_{12} X_{15}^2 + X_{12}^2 X_{15} + X_{15}^3 = 0 \\
\bullet X_1^3 + X_1 X_3^2 + X_1^2 X_3 + X_3^3 = 0 & \bullet X_5^3 + X_5 X_9^2 + X_5^2 X_9 + X_9^3 = 0 & \bullet X_{12}^3 + X_{12} X_{16}^2 + X_{12}^2 X_{16} + X_{16}^3 = 0 \\
\bullet X_1^3 + X_1 X_4^2 + X_1^2 X_4 + X_4^3 = 0 & \bullet X_5^3 + X_5 X_{13}^2 + X_5^2 X_{13} + X_{13}^3 = 0 & \bullet X_{13}^3 + X_{13} X_{14}^2 + X_{13}^2 X_{14} + X_{14}^3 = 0 \\
\bullet X_1^3 + X_1 X_5^2 + X_1^2 X_5 + X_5^3 = 0 & \bullet X_6^3 + X_6 X_7^2 + X_6^2 X_7 + X_7^3 = 0 & \bullet X_{13}^3 + X_{13} X_{15}^2 + X_{13}^2 X_{15} + X_{15}^3 = 0 \\
\bullet X_1^3 + X_1 X_6^2 + X_1^2 X_6 + X_6^3 = 0 & \bullet X_6^3 + X_6 X_8^2 + X_6^2 X_8 + X_8^3 = 0 & \bullet X_{13}^3 + X_{13} X_{16}^2 + X_{13}^2 X_{16} + X_{16}^3 = 0 \\
\bullet X_1^3 + X_1 X_9^2 + X_1^2 X_9 + X_9^3 = 0 & \bullet X_6^3 + X_6 X_{10}^2 + X_6^2 X_{10} + X_{10}^3 = 0 & \bullet X_{14}^3 + X_{14} X_{15}^2 + X_{14}^2 X_{15} + X_{15}^3 = 0 \\
\bullet X_1^3 + X_1 X_{13}^2 + X_1^2 X_{13} + X_{13}^3 = 0 & \bullet X_6^3 + X_6 X_{14}^2 + X_6^2 X_{14} + X_{14}^3 = 0 & \bullet X_{14}^3 + X_{14} X_{16}^2 + X_{14}^2 X_{16} + X_{16}^3 = 0 \\
\bullet X_2^3 + X_2 X_3^2 + X_2^2 X_3 + X_3^3 = 0 & \bullet X_7^3 + X_7 X_8^2 + X_7^2 X_8 + X_8^3 = 0 & \bullet X_{15}^3 + X_{15} X_{16}^2 + X_{15}^2 X_{16} + X_{16}^3 = 0
\end{array}
\tag{2.6}
$$

Desse modo, para determinar a solução do problema do shidoku devemos resolver

esse sistema de equações polinomiais. Assim como no caso do problema de coloração do mapa da região nordeste, apresentaremos a solução desse sistema de equações polinomiais na última seção do capítulo 3 utilizando as bases de Grobner.

Capítulo 3

Base de Gröbner

Neste capítulo, estudaremos a teoria das bases de Gröbner para anéis de polinômios, desenvolvida por Bruno Buchberger em 1965, que foi assim batizada em homenagem a seu orientador Wolfgang Gröbner. Nosso objetivo é usar as bases de Gröbner para resolver sistemas de equações polinomiais de qualquer grau e com qualquer número de variáveis no corpo dos Complexos.

3.1 Ordem monomial

Definição 3.1.1. Uma relação R sobre um conjunto S é uma *relação de ordem* se

(i) $(\forall x)(x \in S \Rightarrow xRx)$ (R é reflexiva)

(ii) $(\forall x, y \in S)(xRy \text{ e } yRx \Rightarrow x = y)$ (R é anti-simétrica)

(iii) $(\forall x, y, z \in S)(xRy \text{ e } yRz \Rightarrow xRz)$ (R é transitiva)

Vemos que a relação dada por $\alpha \leq \beta$ sobre $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ ou $\mathbb{R}$, assim como a relação $\alpha|\beta$ sobre $\mathbb{N}$ são relações de ordem. Entretanto, a relação $\alpha|\beta$ sobre $\mathbb{Z}, \mathbb{Q}$ ou $\mathbb{R}$ não é de ordem pois $1|-1$ e $-1|1$, porém $-1 \neq 1$ o que mostra que a relação não é anti-simétrica.

Definição 3.1.2. Uma relação de ordem R sobre um conjunto S é dita *relação de ordem total* se para qualquer $x, y \in S$, xRy ou yRx.

Note que a relação de ordem $\alpha \leq \beta$ é total sobre $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ ou $\mathbb{R}$ pois dados α, β quaisquer nesses conjuntos temos que $\alpha < \beta$ ou $\beta < \alpha$ ou $\alpha = \beta$. Já a relação de ordem $\alpha|\beta$ não é total sobre $\mathbb{N}$.

Definição 3.1.3. Polinômios em $\mathbb{C}[X_1, \ldots, X_n]$ da forma $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$, com $(\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+^n$, são chamados de *monômios*.

Notação: Para cada $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+^n$ representaremos simplificadamente o monômio $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ por $\mathbf{X}^{\boldsymbol{\alpha}}$.

Quando trabalhamos com polinômios em uma variável, existe uma relação de ordem trivial para os monômios: a relação que os organiza no polinômio em ordem decrescente (ou crescente) de grau. No caso de polinômios em várias variáveis não existe, à princípio, uma ordenação trivial. Em geral é possível ordenar os monômios de um polinômio em várias variáveis de diversos modos dependendo, entre outros fatores, de qual variável é "mais significativa".

A seguir definiremos uma *ordenação monomial* e apresentaremos três tipos de ordenação que podem ser usadas para ordenar monômios em um polinômio com várias variáveis.

Definição 3.1.4. Uma *ordenação monomial sobre* $\mathbb{C}[X_1, ..., X_n]$ é uma relação $>$ sobre $\mathbb{Z}_+^n$, ou equivalentemente, uma relação $>$ sobre o conjunto dos monômios $\mathbf{X}^{\boldsymbol{\alpha}} = X_1^{\alpha_1} ... X_n^{\alpha_n}$ com $\boldsymbol{\alpha} \in \mathbb{Z}_+^n$, satisfazendo as seguintes condições:

(i) $>$ é uma relação de ordem total sobre $\mathbb{Z}_+^n$

(ii) Se $\alpha > \beta$ e $\gamma \in \mathbb{Z}_+^n$, então $\alpha + \gamma > \beta + \gamma$.

(iii) $>$ é uma *boa ordenação* em $\mathbb{Z}_+^n$, ou seja, todo subconjunto não-vazio de $\mathbb{Z}_+^n$ tem um elemento mínimo sobre $>$.

O lema abaixo é uma reformulação da condição (iii) na definição acima.

Lema 3.1.5. *Uma relação de ordem $>$ sobre $\mathbb{Z}_+^n$ é uma boa ordenação se, e somente se, toda sequência estritamente decrescente em $\mathbb{Z}_+^n$*

$$\alpha(1) > \alpha(2) > \alpha(3) > ...$$

é finita.

Prova. Usando a contrapositiva, queremos mostrar que existe uma sequência infinita estritamente decrescente em $\mathbb{Z}_+^n$ se, e somente se, $>$ não é uma boa ordenação em $\mathbb{Z}_+^n$.

($\Leftarrow$) Se $>$ não é uma boa ordenação em $\mathbb{Z}_+^n$, então algum subconjunto não vazio $S \subset \mathbb{Z}_+^n$ não tem um menor elemento. Considere $\alpha(1) \in S$. Note que $\alpha(1)$ não é

o menor elemento de S. Assim, existe $\alpha(2) \in S$ tal que $\alpha(1) > \alpha(2)$ e $\alpha(2)$ não é o menor elemento de S. Continuando esse processo, concluímos que

$$\alpha(1) > \alpha(2) > \alpha(3) > ...$$

é uma sequência infinita estritamente decrescente.

($\Rightarrow$) Se $S = \{\alpha(1), \alpha(2), \alpha(3), ...\} \subset \mathbb{Z}_+^n$ é uma sequência infinita estritamente decrescente, então S não tem um elemento mínimo, o que mostra que $>$ não é uma boa ordenação. □

Observação 3.1.6. Notemos que a ordem usual dos elementos de $\mathbb{Z}_+$,

$$... > m+1 > m > ... > 3 > 2 > 1 > 0$$

satisfaz as três condições da Definição 3.1.4. Assim, o grau de ordenação de monômios em $\mathbb{C}[X]$ é uma ordenação monomial.

Nas subseções a seguir apresentaremos três tipos de ordenação monomial sobre n-uplas.

3.1.1 Ordenação Lexicográfica

Definição 3.1.7. Sejam $\alpha = (\alpha_1, ..., \alpha_n)$ e $\beta = (\beta_1, ..., \beta_n) \in \mathbb{Z}_+^n$. Dizemos que α *é uma ordenação lexicográfica em* β*, ou simplesmente* $\alpha >_{lex} \beta$, se no vetor diferença $\alpha - \beta \in \mathbb{Z}_+^n$ a coordenada não-nula mais à esquerda é positiva.

Os exemplos a seguir ilustram bem a definição acima:

(a) $\alpha = (3,4,2) >_{lex} \beta = (1,5,2)$ pois $\alpha - \beta = (2,-1,0)$.

(b) $\alpha = (1,2,5) >_{lex} \beta = (1,2,3)$ pois $\alpha - \beta = (0,0,2)$

Proposição 3.1.8. *A ordenação lexicográfica sobre* $\mathbb{Z}_+^n$ *é uma ordenação monomial.*

Prova. (i) A justificativa para $>_{lex}$ ser uma ordenação total segue da definição e do fato de que a ordem numérica sobre $\mathbb{Z}_+$ ser uma ordenação total.

(ii) Se $\alpha >_{lex} \beta$, então $\alpha_k - \beta_k > 0$, onde k é a posição da primeira coordenada não nula de $\alpha - \beta$. Note que, dado $\gamma \in \mathbb{Z}_+^n$, $(\alpha + \gamma) - (\beta + \gamma) = \alpha - \beta$. Assim, a primeira coordenada não nula é novamente $\alpha_k - \beta_k > 0$.

(iii) Suponha por absurdo que $>_{lex}$ não é uma boa ordenação. Assim, pelo Lema 3.1.5, existe uma sequência infinita estritamente decrescente

$$\alpha(1) >_{lex} \alpha(2) >_{lex} > \alpha(3) >_{lex} ...$$

de elementos de $\mathbb{Z}_+^n$. Considere a primeira coordenada de cada um dos vetores $\alpha(i) \in \mathbb{Z}_+^n$. Pela definição de ordem lexicográfica, estas coordenadas formam uma sequência não crescente de números inteiros não negativos, portanto devem se estabilizar. Assim, existe um k tal que a primeira coordenada de todos os $\alpha(i)$ com $i \geq k$ é igual.

A partir de $\alpha(k)$, as segundas coordenadas determinam a ordem lexicográfica e, pelo mesmo motivo apresentado acima, formam uma sequência não crescente de números inteiros não negativos e, portanto, devem também se estabilizar. Continuando esse processo vemos que para algum ℓ, $\alpha(\ell), \alpha(\ell+1), ...$ são todos iguais, o que é um absurdo pois $\alpha(\ell) >_{lex} \alpha(\ell+1)$. □

Convém observar que uma ordenação lexicográfica está associada ao modo como as variáveis estão ordenadas. Assim, para evitar dúbias interpretações, quando usarmos a ordenação lexicográfica, iremos considerar $X_1 > ... > X_n$.

Desse modo, numa ordenação lexicográfica, uma variável domina qualquer monômio que contem variáveis menores. Assim, se $X > Y > Z$ então $X^2 >_{lex} XY^5Z^3$.

Observação 3.1.9. O termo *ordem lexicográfica* é usado para determinar o modo como as palavras aparecem no dicionário. Observe que, de certo modo, é isso que estamos fazendo com os monômios. Se, por exemplo, escrevêssemos os monômios X^2Y^3Z e $X^2Y^2Z^3$ como as "palavras" $XXYYYZ$ e $XXYYZZZ$, então a "palavra" $XXYYYZ$ apareceria antes no dicionário, o que coincide com a ordem lexicográfica dos monômios em questão.

3.1.2 Ordenação lexicográfica graduada

Definição 3.1.10. Sejam $\alpha, \beta \in \mathbb{Z}_+^n$. Diremos que $\alpha >_{glex} \beta$ se

$$|\alpha| = \sum_{i=1}^{n} \alpha(i) > |\beta| = \sum_{i=1}^{n} \beta(i) \text{ ou quando } |\alpha| = |\beta| \text{ e } \alpha >_{lex} \beta.$$

Assim, a ordem *glex* ordena primeiro pelo grau total e usa a ordem *lex* quando há

igualdade no grau total dos monômios. Abaixo seguem alguns exemplos para facilitar o entendimento:

(a) $\alpha = (1,3,2) >_{glex} \beta = (2,1,2)$ pois $|\alpha| = 6 > 5 = |\beta|$

(b) $\alpha = (1,3,4) >_{glex} \beta = (1,1,6)$ pois $|\alpha| = 8 = |\beta|$ e $(1,3,4) >_{lex} (1,1,6)$

(c) $X^2Y^3Z^2 >_{glex} X^3YZ^2$

3.1.3 Ordenação Lexicográfica Graduada Reversa

Definição 3.1.11. Sejam $\alpha, \beta \in \mathbb{Z}_+^n$. Diremos que $\alpha >_{grevlex} \beta$ se

$$|\alpha| = \sum_{i=1}^{n} \alpha(i) > |\beta| = \sum_{i-1}^{n} \beta(i) \text{ ou quando } |\alpha| = |\beta|$$

e em $\alpha - \beta$, a coordenada não nula mais à direita é negativa.

Alguns exemplos:

(a) $\alpha = (1,3,2) >_{grevlex} \beta = (2,1,2)$ pois $|\alpha| = 6 > 5 = |\beta|$

(b) $\alpha = (3,3,4) >_{grevlex} \beta = (1,5,4)$ pois $|\alpha| = 10 = |\beta|$ e $\alpha - \beta = (2,-2,0)$

Agora vamos ver como aplicar uma ordenação monomial a polinômios em geral. Para isso, considere $f = \sum_{\alpha} a_\alpha \mathbf{X}^{\boldsymbol{\alpha}}$ em $\mathbb{C}[X_1, ..., X_n]$. Se selecionarmos uma ordenação monomial $>$, então é possível ordenar os monômios de f de modo não ambíguo com respeito a $>$. Por exemplo, seja $f = XY^3Z + XY^3Z^2 + X^2Z^3 \in \mathbb{C}[X,Y,Z]$. Então, colocando os termos de f em ordem decrescente,

(i) Considerando a ordenação lex, temos:

$$f = X^2Z^3 + XY^3Z^2 + XY^3Z$$

(ii) Considerando a ordenação $grlex$, temos:

$$f = XY^3Z^2 + X^2Z^3 + XY^3Z$$

(iii) Considerando a ordenação $grevlex$, temos:

$$f = XY^3Z^2 + XY^3Z + X^2Z^3$$

Definição 3.1.12. Seja $f = \sum_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} \mathbf{X}^{\boldsymbol{\alpha}}$ um polinômio não nulo em $\mathbb{C}[X_1, ..., X_n]$ e seja $>$ uma ordenação monomial.

(i) O multigrau de f é:

$$\text{multigrau}(f) := \max\{\boldsymbol{\alpha} \in \mathbb{Z}_+^n; a_{\boldsymbol{\alpha}} \neq 0\}$$

onde o máximo é escolhido com respeito a $>$.

(ii) O *coeficiente líder* de f é:

$$CL(f) := a_{\text{multigrau}(f)} \in \mathbb{C}$$

(iii) O *monômio líder* de f é:

$$ML(f) := \mathbf{X}^{\text{multigrau}(f)}$$

(iv) o *termo líder* de f é:

$$TL(f) := CL(f) \cdot ML(f)$$

Exemplo 3.1.13. Seja $f = 2X^2Y^7Z - 3X^5YZ^4 + XYZ^3 - XY^4 + 4X^6Z^3 \in \mathbb{C}[X, Y, Z]$.

- Considerando a ordem $>_{lex}$ temos:

$$\text{multigrau}(f) = (6, 0, 3), \quad CL(f) = 4, \quad ML(f) = X^6Z^3 \quad \text{e} \quad TL(f) = 4X^6Z^3$$

- Considerando a ordem $>_{glex}$ temos:

$$\text{multigrau}(f) = (5, 1, 4), \quad CL(f) = -3, \quad ML(f) = X^5YZ^4 \quad \text{e} \quad TL(f) = -3X^5YZ^4.$$

- Considerando a ordem $>_{grevlex}$ temos:

$$\text{multigrau}(f) = (2, 7, 1) \quad CL(f) = 2 \quad ML(f) = X^2Y^7Z \quad \text{e} \quad TL(f) = 2X^2Y^7Z.$$

No lema a seguir listamos algumas propriedades básicas da função multigrau.

Lema 3.1.14. *Sejam $f, g \in \mathbb{C}[X_1, \ldots, X_n]$ polinômios não-nulos. Então:*

(i) $\text{multigrau}(f \cdot g) = \text{multigrau}(f) + \text{multigrau}(g)$

(ii) *Se* $f+g \neq 0$, *então* $\text{multigrau}(f+g) \leq \max\{\text{multigrau}(f), \text{multigrau}(g)\}$. *Além disso, se* $\text{multigrau}(f) \neq \text{multigrau}(g)$, *então*

$$\text{multigrau}(f+g) = \max\{\text{multigrau}(f), \text{multigrau}(g)\}.$$

Prova. (i) Sejam $f = \sum_{i=1}^{n} a_i \mathbf{X}^{\boldsymbol{\alpha}(i)}$ e $g = \sum_{j=1}^{m} b_j \mathbf{X}^{\boldsymbol{\beta}(j)}$ polinômios em $\mathbb{C}[X_1, \ldots, X_n]$. Suponha sem perda de generalidade que $\text{multigrau}(f) = \boldsymbol{\alpha}(1)$ e $\text{multigrau}(g) = \boldsymbol{\beta}(1)$. Assim,

$$\begin{aligned} f \cdot g &= \sum_{i=1}^{n} a_i \mathbf{X}^{\boldsymbol{\alpha}(i)} \cdot \sum_{j=1}^{m} b_j \mathbf{X}^{\boldsymbol{\beta}(j)} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \mathbf{X}^{\boldsymbol{\alpha}(i)+\boldsymbol{\beta}(j)} \\ &= \sum_{i=1}^{n} a_i b_1 \mathbf{X}^{\boldsymbol{\alpha}(i)+\boldsymbol{\beta}(1)} + \sum_{i=1}^{n} \sum_{j=2}^{m} a_i b_j \mathbf{X}^{\boldsymbol{\alpha}(i)+\boldsymbol{\beta}(j)} \\ &= a_1 b_1 \mathbf{X}^{\boldsymbol{\alpha}(1)+\boldsymbol{\beta}(1)} + \sum_{i=2}^{n} a_i b_1 \mathbf{X}^{\boldsymbol{\alpha}(i)+\boldsymbol{\beta}(1)} + \sum_{i=1}^{n} \sum_{j=2}^{m} a_i b_j \mathbf{X}^{\boldsymbol{\alpha}(i)+\boldsymbol{\beta}(j)}. \end{aligned}$$

Como $\boldsymbol{\alpha}(1) + \boldsymbol{\beta}(1) > \boldsymbol{\alpha}(i) + \boldsymbol{\beta}(j)$ para todo $1 \leq i \leq n, 1 \leq j \leq m$ e $a_1 \cdot b_1 \neq 0$ (pois $\mathbb{C}$ é um corpo), então

$$\text{multigrau}(f \cdot g) \quad = \boldsymbol{\alpha}(1) + \boldsymbol{\beta}(1) = \text{multigrau}(f) + \text{multigrau}(g).$$

(ii) Se f e g são polinômios tais que $f + g \neq 0$, então o $\text{multigrau}(f + g)$ está definido. Suponhamos a princípio que $\text{multigrau}(f) \neq \text{multigrau}(g)$. Nesse caso, é fácil ver que

$$\text{multigrau}(f+g) = \max\{\text{multigrau}(f), \text{multigrau}(g)\}.$$

Entretanto, se $\text{multigrau}(f) = \text{multigrau}(g)$ temos que analisar duas possibilidades:

(1) $TL(f) = -TL(g)$.

Neste caso os termos líderes se anulam e, por isso,

$$\text{multigrau}(f+g) < \text{multigrau}(f) = \text{multigrau}(g),$$

ou seja,

$$\text{multigrau}(f+g) < \max\{\text{multigrau}(f), \text{multigrau}(g)\}.$$

(2) $TL(f) \neq -TL(g)$.

Neste caso, os termos líderes não se anulam e, por isso,

$$\text{multigrau}(f+g) = \max\{\text{multigrau}(f), \text{multigrau}(g)\}.$$

□

3.2 Algoritmo da Divisão em $\mathbb{C}[X_1, \ldots, X_n]$

Após estabelecer critérios através dos quais conseguimos ordenar os monômios que compõem um polinômio em $\mathbb{C}[X_1, \ldots, X_n]$, a pergunta que naturalmente surge é: seria possível estender o algoritmo da divisão para polinômios de uma variável de modo a utilizá-lo em problemas com polinômios de várias variáveis? Veremos em seguida que a resposta para essa pergunta é **sim**.

Em geral, nosso objetivo é dividir $f \in \mathbb{C}[X_1, \ldots, X_n]$ por $f_1, \ldots, f_s \in \mathbb{C}[X_1, \ldots, X_n]$, ou seja, queremos expressar f na forma

$$f = a_1 f_1 + \ldots + a_n f_s + r$$

onde os "quocientes" $a_1, \ldots, a_s$ e o "restos" r pertencem a $\mathbb{C}[X_1, \ldots, X_n]$.

Basicamente vamos proceder com n do mesmo modo que com uma variável: tentaremos cancelar o termo líder de f (fixada uma ordem monomial) subtraindo-a pelo produto de alguma f_i por um monômio convenientemente escolhido. Antes de formalizar o algoritmo vamos considerar alguns exemplos.

Exemplo 3.2.1. Dividir $f = XY^2 + 1$ por $f_1 = XY + 1$ e $f_2 = Y + 1$, usando a ordenação lexicográfica com $X > Y$. Colocando os divisores f_1 e f_2 na chave e os quocientes a_1 e a_2 abaixo dela, temos o seguinte esquema:

$$\begin{array}{rl|l} XY^2+1 & XY+1 & \\ & Y+1 & \\ \hline & a_1: & \\ & a_2: & \end{array}$$

Note que os termos líderes $TL(f_1) = XY$ e $TL(f_2) = Y$ dividem o $TL(f) = XY^2$. Vamos então dividir primeiro f por f_1. Assim, para iniciar o processo, inicialmente dividimos XY^2 por XY cujo resultado é Y, e subtraímos f do produto de Y por f_1, obtendo:

$$\begin{array}{rl|l} XY^2+1 & XY+1 & \\ & Y+1 & \\ \hline -Y+1 & a_1: Y & \\ & a_2: & \end{array}$$

Em seguida repetiremos o mesmo processo para $-Y + 1$, porém dessa vez dividiremos por f_2 pois $TL(f_1) = XY$ não divide o $TL(-Y + 1) = -Y$. Assim,

$$\begin{array}{rl|l} XY^2+1 & XY+1 & \\ & Y+1 & \\ \hline -Y+1 & a_1: Y & \\ 2 & a_2: -1 & \end{array}$$

Como os termos líderes de f_1 e f_2 não dividem 2, o processo de divisão acaba de modo que $r = 2$. Assim, temos

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

No próximo exemplo veremos uma inesperada sutileza que pode ocorrer quando trabalhamos com polinômios com mais de uma variável.

Exemplo 3.2.2. Dividir $f = X^2Y + XY^2 + Y$ por $f_1 = XY - 1$ e $f_2 = Y^2 - 1$ considerando a ordenação lexicográfica com $X > Y$.

As etapas iniciais do processo de divisão seguem do mesmo modo que no exemplo anterior, lembrando que quando os dois termos líderes dividem f usamos o primeiro:

$$\begin{array}{rl|l} X^2Y+XY^2+Y^2 & XY-1 & \\ & Y^2-1 & \\ \hline XY^2+X+Y^2 & a_1: X+Y & \\ X+Y^2+Y & a_2: & \end{array}$$

Observe que os termos líderes de f_1 e f_2 não dividem o $TL(X + Y^2 + Y) = X$. Porém $X + Y^2 + Y$ não é o resto pois $TL(f_2)$ divide Y^2. Desse modo, "movendo" X para o resto podemos continuar a divisão.

A fim de deixar o algoritmo organizado, vamos criar uma coluna para o resto r, à direita dos divisores, onde colocaremos os termos que o compõem.

Além disso, denominaremos cada polinômio abaixo do dividendo por *dividendo intermediário*, e continuaremos o processo até que ele se anule.

$$\begin{array}{rl|lcc} X^2Y + XY^2 + Y^2 & XY - 1 & & & \text{resto} \\ & Y^2 - 1 & & & \hline \\ XY^2 + X + Y^2 & a_1 : X + Y & & & \\ X + Y^2 + Y & a_2 : 1 & & \to & X \\ Y^2 + Y & & & & \\ Y + 1 & & & \to & Y + 1 \\ 0 & & & & r = X + Y + 1 \end{array}$$

Assim, o resto é $X + Y + 1$ e

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

Convém observar que o resto dessa divisão é uma soma de monômios não divisíveis pelos termos líderes de f_1 e f_2.

Finalmente, vamos enunciar o algoritmo da divisão em $\mathbb{C}[X_1, \ldots, X_n]$.

Teorema 3.2.3. (Algoritmo da divisão em $\mathbb{C}[X_1, \ldots, X_n]$) *Fixe uma ordem monomial sobre $\mathbb{Z}_+^n$ e seja $F = (f_1, \ldots, f_s)$ uma s-upla ordenada de polinômios em $\mathbb{C}[X_1, \ldots, X_n]$. Então todo $f \in \mathbb{C}[X_1, \ldots, X_n]$ pode ser escrito como*

$$f = a_1 f_1 + \ldots + a_s f_s + r$$

onde $a_i \in \mathbb{C}[X_1, \ldots, X_n]$ e $r = 0$ ou r é uma combinação linear de monômios com coeficientes em $\mathbb{C}$, de modo que nenhum deles é divisível por algum termo líder de $f_1, \ldots, f_s$. Chamaremos r de um resto de f numa divisão por F. Além disso, se $f_i \neq 0$, então

$$\text{multigrau}(f) \geq \text{multigrau}(a_i \cdot f_i)$$

Prova. Podemos ter as seguintes situações:

(A_1) ou existe $TL(f_{i_1})$ que divide $TL(f)$

(B_1) ou não existe $TL(f_{i_1})$ que divida $TL(f)$.

Definamos

$$p_1 = \begin{cases} f - (TL(f)/TL(f_{i_1}))f_{i_1} & \text{caso aconteça } (A_1) \\ f - TL(f) & \text{caso aconteça } (B_1) \end{cases}$$

Em todo caso, podemos escrever

$$f = (\sum a_i f_i) + p_1 + r_1$$

onde $r_1 = 0$ ou $r_1 = TL(f)$

Se $TL(f_i)$ não divide nenhum termo de p_1 para qualquer $1 \leq i \leq s$ ou $p_1 = 0$ então o teorema fica provado, basta fazer $r = p_1 + r_1$. Caso contrário, temos as seguintes situações:

(A_2) ou existe $TL(f_{i_2})$ que divide $TL(p_1)$

(B_2) ou não existe $TL(f_{i_2})$ que divida $TL(p_1)$.

Definamos

$$p_2 = \begin{cases} p_1 - (TL(p_1)/TL(f_{i_2}))f_{i_2} & \text{caso aconteça } (A_2) \\ p_1 - TL(p_1) & \text{caso aconteça } (B_2) \end{cases}$$

Em todo caso, podemos escrever

$$f = (\sum a_i f_i) + p_1 + r_1 = (\sum a_i f_i) + p_2 + r_1 + r_2$$

onde $r_2 = 0$ ou $r_2 = TL(p_1)$.

Se $TL(f_i)$ não divide nenhum termo de p_2 para qualquer $1 \leq i \leq s$ então o teorema fica provado, basta fazer $r = p_2 + r_1 + r_2$. Caso contrário, de maneira análoga construimos um p_3 e assim sucessivamente. Afirmamos que esse processo de construção dos p_i não pode continuar idefinidamente. Para verifcar esta afirmação notemos que, em virtude do Lema 3.1.14 e da forma que os p_i são definidos, temos

$$\text{multigrau}(p_{i-1}) > \text{multigrau}(p_i)$$

caso $p_i \neq 0$ Mas, pelo Lema 3.1.5, a sequência

$$\text{multigrau}(p_1) > \text{multigrau}(p_2) > \dots$$

é finita. Assim, devemos chegar a uma certa etapa j em que $p_j = 0$. Consequentemente, o processo de construção dos p_i deve encerrar nessa etapa e temos o teorema demonstrado. □

Notemos que o algoritmo da divisão no caso do anel de polinômios em uma única variável $\mathbb{C}[X]$ decide completamente se um polinômio pertence ou não a um determinado ideal. De fato, neste caso temos que uma condição necessária e suficiente para um polinômio f pertença a um ideal $I = \langle g \rangle$ é que o resto r da divisão de f por g seja zero. No caso de um anel de polinômios em várias variáveis $\mathbb{C}[X_1, \dots, X_n]$, é fato óbvio que se o resto da divisão de um polinômio f por polinômios $f_1, \dots, f_s$ é zero então $f \in \langle f_1, \dots, f_s \rangle$, ou seja, o resto ser zero é condição suficiente para f pertencer ao ideal $\langle f_1, \dots, f_s \rangle$. Todavia, o resto ser zero não é condição necessaria para f pertencer ao ideal como nos revela o seguinte exemplo

Exemplo 3.2.4. Sejam $f_1 = XY+1$ e $f_2 = Y^2-1 \in \mathbb{C}[X, Y]$. Considere a ordenação lexicográfica. Dividindo $f = XY^2 - X$ por f_1, f_2 temos que

$$\begin{array}{rl|l} XY^2 - X & XY + 1 & \\ & Y^2 - 1 & \\ \hline -X - Y & a_1 : Y & \\ -X - Y & a_2 : 0 & \end{array}$$

ou seja,

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

Note que este resultado, a princípio, não garante que $f \in \langle f_1, f_2 \rangle$.

Entretanto, dividindo f por f_2, f_1 temos:

$$\begin{array}{rl|l} XY^2 - X & Y^2 - 1 & \\ & XY + 1 & \\ \hline 0 & a_1 : X & \\ 0 & a_2 : 0 & \end{array}$$

donde obtemos

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

o que mostra que $f \in \langle f_1, f_2 \rangle$

Assim, do ponto de vista do problema de pertinência de um elemento em um ideal, podemos concluir que o algoritmo da divisão em $\mathbb{C}[X_1, \ldots, X_n]$ é uma generalização imperfeita do algoritmo da divisão em $\mathbb{C}[X]$. Levando em consideração que quando lidamos com uma coleção de polinômios $f_1, \ldots, f_s$ em $\mathbb{C}[X_1, \ldots, X_n]$ é desejável determinarmos o ideal gerado por eles, uma pergunta que naturalmente surge é: qual seria um bom conjunto gerador de I? Para tal conjunto, queremos que o resto r seja unicamente determinado e que $r = 0$ seja condição necessária e suficiente para que f pertença ao ideal. Mais a frente veremos que as *Bases de Gröbner* são justamente esse bom conjunto gerador de I.

3.3 Ideais Monomiais

Nesta seção estudaremos, de modo cuidadoso, as propriedades dos chamados *ideais monomiais*. Inicialmente, definiremos ideais monomiais em $\mathbb{C}[X_1, \ldots, X_n]$.

Definição 3.3.1. Um ideal $I \subset \mathbb{C}[X_1, \ldots, X_n]$ é um *ideal monomial* se existe um subconjunto $A \subset \mathbb{Z}_+^n$ de modo que I consiste de todos os polinômios que são somas finitas da forma $\sum_{\boldsymbol{\alpha} \in A} h_{\boldsymbol{\alpha}} \mathbf{X}^{\boldsymbol{\alpha}}$, onde $h_{\boldsymbol{\alpha}} \in \mathbb{C}[X_1, \ldots, X_n]$, $\alpha \in A$. Nesse caso, vamos escrever $I = \langle \mathbf{X}^{\alpha} : \boldsymbol{\alpha} \in A \rangle$.

O lema abaixo nos dá uma caracterização alternativa para saber quando um dado monômio pertence a um ideal monomial

Lema 3.3.2. *Seja $I = \langle \mathbf{X}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in A \rangle$ um ideal monomial. Um monômio $\mathbf{X}^{\beta}$ pertence a I se, e somente se, $\mathbf{X}^{\beta}$ é divisível por $\mathbf{X}^{\boldsymbol{\alpha}}$ para algum $\boldsymbol{\alpha} \in A$.*

Prova. Se $\mathbf{X}^{\boldsymbol{\beta}}$ é múltiplo de $\mathbf{X}^{\boldsymbol{\alpha}}$, então, pela definição de ideal, concluímos que $\mathbf{X}^{\boldsymbol{\beta}} \in I$. Reciprocamente, se $\mathbf{X}^{\boldsymbol{\beta}} \in I$, então $\mathbf{X}^{\boldsymbol{\beta}} = \sum_{i=1}^{s} h_i \mathbf{X}^{\boldsymbol{\alpha}_i}$, onde $h_i \in \mathbb{C}[X_1, \ldots, X_n]$ e $\boldsymbol{\alpha}_i \in A$. Escrevendo cada h_i como uma combinação linear de monômios, distribuindo os produtos e agrupando os termos semelhantes, temos que

$$\mathbf{X}^{\beta} = \sum_{k=1}^{t} a_{\boldsymbol{\gamma}_k} \mathbf{X}^{\boldsymbol{\gamma}_k}$$

onde $a_{\boldsymbol{\gamma}_k} \in \mathbb{C}$.

Note que cada parcela do lado direito dessa igualdade é divisível por algum $\mathbf{X}^{\boldsymbol{\alpha}_i}$ com $\alpha_i \in A$. Além disso, se o lado esquerdo da igualdade é um monômio, então o lado direito também é. Assim, $\mathbf{X}^{\boldsymbol{\beta}} = \sum_{k=1}^{t} a_{\boldsymbol{\gamma}_k} \mathbf{X}^{\boldsymbol{\gamma}_k}$ se, e somente se, $t = 1$ e $a_{\boldsymbol{\gamma}_1} = 1$, ou seja,

$$\mathbf{X}^{\boldsymbol{\beta}} = \mathbf{X}^{\boldsymbol{\gamma}_1}.$$

Como $\mathbf{X}^{\boldsymbol{\alpha}_i}$ divide $\mathbf{X}^{\boldsymbol{\gamma}_1}$ então $\mathbf{X}^{\boldsymbol{\alpha}_i}$ divide $\mathbf{X}^{\boldsymbol{\beta}}$, o que conclui a demonstração. □

Observe que $\mathbf{X}^{\boldsymbol{\beta}}$ é divisível por $\mathbf{X}^{\boldsymbol{\alpha}}$ quando $\mathbf{X}^{\boldsymbol{\beta}} = \mathbf{X}^{\boldsymbol{\alpha}} \cdot \mathbf{X}^{\boldsymbol{\gamma}}$ para algum $\boldsymbol{\gamma} \in \mathbb{Z}_+^n$, ou seja, $\boldsymbol{\beta} = \boldsymbol{\alpha} + \boldsymbol{\gamma}$. Assim, o conjunto

$$\boldsymbol{\alpha} + \mathbb{Z}_+^n = \{\boldsymbol{\alpha} + \boldsymbol{\gamma} : \boldsymbol{\gamma} \in \mathbb{Z}_+^n\}$$

é formado pelos expoentes de todos os monômios divisíveis por $\mathbf{X}^{\alpha}$.

Se, por exemplo, $I = \langle X^4Y^2, X^3Y^4, X^2Y^5 \rangle$, então os expoentes dos monômios em I formam o conjunto

$$\left((4,2) + \mathbb{Z}_+^2\right) \cup \left((3,4) + \mathbb{Z}_+^2\right) \cup \left((2,5) + \mathbb{Z}_+^2\right).$$

O lema abaixo mostra se é possível decidir se um polinômio f pertence ou não a um ideal monomial verificando os monômios de f.

Lema 3.3.3. *Seja I um ideal monomial e seja $f \in \mathbb{C}[X_1, \ldots, X_n]$. Então são equivalentes:*

(i) *$f \in I$.*

(ii) *Todo termo de f pertence a I.*

(iii) *f é uma combinação linear de monômios em I.*

Prova. As implicações (iii) ⇒ (ii) ⇒ (i) são triviais. Assim, basta mostrar que (i)⇒(iii). Se $f \in I$ então, por definição de ideal monomial,

$$f = \sum_{i=1}^{k} h_k \mathbf{X}^{\boldsymbol{\alpha}_k}$$

onde $h_k \in \mathbb{C}[X_1, \ldots, X_n]$ e $\boldsymbol{\alpha}_k \in A$.

Se expandirmos cada h_k como combinação linear de monômios, efetuarmos os produtos e agruparmos os termos semelhantes, é fácil ver que toda parcela do lado

direito da igualdade é divisível por algum $\mathbf{X}^{\alpha_k}$. Assim, pelo Lema 3.3.2 , cada parcela de f pertence a I. Portanto f é uma combinação linear de monômios de I. □

Corolário 3.3.4. *Dois ideais monomiais são iguais se, e somente se, contêm os mesmos monômios.*

Finalmente podemos enunciar o importante lema abaixo.

Lema 3.3.5. (Lema de Dickson). *Um ideal monomial* $I = \langle \mathbf{X}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in A \rangle \subset \mathbb{C}[X_1, \ldots, X_n]$ *pode ser escrito na forma*

$$I = \langle \mathbf{X}^{\boldsymbol{\alpha}(1)}, \ldots, \mathbf{X}^{\boldsymbol{\alpha}(s)} \rangle$$

onde $\boldsymbol{\alpha}(1), \ldots, \boldsymbol{\alpha}(s) \in A$. *Em particular,* I *tem uma base finita.*

Prova. Vamos provar esse lema por indução sobre o número n de variáveis.

Se $n = 1$, então I é um ideal em $\mathbb{C}[X_1]$ gerado por monômios do tipo $X_1^{\alpha_i}$, onde $\alpha_i \in A \subset \mathbb{Z}_+$. Seja β o menor elemento de A. Como $\beta \leq \alpha_i$ para todo i, temos que X_1^{β} divide todos os outros geradores $X_1^{\alpha_i}$, ou seja, $I = \langle X_1^{\beta} \rangle$, donde concluímos que o Lema de Dickson vale para $n = 1$.

Suponhamos agora que $n > 1$ e que o lema seja válido para $n-1$. Representemos as variáveis como $X_1, \ldots, X_{n-1}, Y$. Assim, os monômios em $\mathbb{C}[X_1, \ldots, X_{n-1}, Y]$ podem ser escritos como $\mathbf{X}^{\boldsymbol{\alpha}} \cdot Y^m$, onde $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{n-1}) \in \mathbb{Z}_+^{n-1}$ e $m \in \mathbb{Z}_+$.

Suponhamos que $I \subset \mathbb{C}[X_1, \ldots, X_{n-1}, Y]$ seja um ideal monomial. Seja J um ideal em $\mathbb{C}[X_1, \ldots, X_{n-1}]$ gerado por monômios $\mathbf{X}^{\alpha}$ tais que $\mathbf{X}^{\alpha} \cdot Y^m \in I$ para algum $m \geq 0$. Como J, por construção, é um ideal monomial temos, por hipótese indutiva, que J possui um número finito $\mathbf{X}^{\alpha}$ de geradores, isto é

$$J = \langle \mathbf{X}^{\boldsymbol{\alpha}(1)}, \ldots, \mathbf{X}^{\boldsymbol{\alpha}(s)} \rangle.$$

Note que $\mathbf{X}^{\boldsymbol{\alpha}(i)} \cdot Y^{m_i} \in I$ para todo $1 \leq i \leq s$ e para algum $m_i \geq 0$. Seja m o maior dos m_i. Assim, para cada k entre 0 e $m-1$, consideremos o ideal $J_k \subset \mathbb{C}[X_1, \ldots, X_{n-1}]$ gerado pelos monômios $\mathbf{X}^{\beta}$ tais que $\mathbf{X}^{\boldsymbol{\beta}(i)} \cdot Y^k \in I$. Por hipótese indutiva J_k tem um número finito de monômios geradores, digamos $J_k = \langle \mathbf{X}^{\boldsymbol{\alpha}_k(1)}, \ldots, \mathbf{X}^{\boldsymbol{\alpha}_k(s_k)} \rangle$.

Podemos então afirmar que I é gerado pelos monômios da seguinte lista:

$$\begin{array}{llll} \text{de} & J & : & \mathbf{X}^{\boldsymbol{\alpha}(1)} \cdot Y^m, \ldots, \mathbf{X}^{\boldsymbol{\alpha}(s)} \cdot Y^m \\ \text{de} & J_0 & : & \mathbf{X}^{\boldsymbol{\alpha}_0(1)}, \ldots, \mathbf{X}^{\boldsymbol{\alpha}_0(s_0)} \\ \text{de} & J_1 & : & \mathbf{X}^{\boldsymbol{\alpha}_1(1)} Y, \ldots, \mathbf{X}^{\boldsymbol{\alpha}_1(s_1)} Y \\ & & \vdots & \\ \text{de} & J_{m-1} & : & \mathbf{X}^{\boldsymbol{\alpha}_{m-1}(1)} Y^{m-1}, \ldots, \mathbf{X}^{\boldsymbol{\alpha}_{m-1}(s_{m-1})} Y^{m-1} \end{array}$$

Note primeiramente que todo monômio em I é divisível por algum monômio da lista. Para justificar esse fato basta considerarmos um monômio $\mathbf{X}^{\boldsymbol{\alpha}} Y^p \in I$. Assim,

(i) Se $p \geq m$, então $\mathbf{X}^{\alpha} Y^p$ é divisível por algum $\mathbf{X}^{\boldsymbol{\alpha}(i)} Y^m$ pela construção de J.

(ii) Se $p \leq m-1$, então $\mathbf{X}^{\alpha} Y^p$ é divisível por algum $\mathbf{X}^{\boldsymbol{\alpha}_p(j)} Y^p$ pela construção de J_p

Assim, pelo Lema 3.3.2, os monômios acima geram o ideal I. Pelo Corolário 3.3.4, isso implica que os ideais são iguais, o que prova nossa afirmação.

Para completar a demonstração desse importante lema, precisamos ainda mostrar que um conjunto finito de geradores pode ser escolhido a partir de um dado conjunto de geradores (possivelmente infinito) do ideal.

Seja $I = \langle \mathbf{X}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in A \rangle \subset \mathbb{C}[X_1, \ldots, X_n]$ um ideal monomial. Queremos mostrar que I é gerado por um número finito de monômios $\mathbf{X}^{\boldsymbol{\beta}(i)} \in I$. Pelo que argumentamos anteriormente, sabemos que I tem um número finito de geradores, ou seja,

$$I = \langle \mathbf{X}^{\boldsymbol{\beta}(1)}, \ldots, \mathbf{X}^{\boldsymbol{\beta}(s)} \rangle$$

para certos monômios $\mathbf{X}^{\boldsymbol{\beta}(i)} \in I$. Como $\mathbf{X}^{\boldsymbol{\beta}(i)} \in I = \langle \mathbf{X}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in A \rangle$, temos que, pelo Lema 3.3.2, cada $\mathbf{X}^{\boldsymbol{\beta}(i)}$ é divisível por $\mathbf{X}^{\boldsymbol{\alpha}(i)}$ para algum $\boldsymbol{\alpha}(i) \in A$ e, portanto

$$\mathbf{X}^{\boldsymbol{\beta}(i)} = h_i \mathbf{X}^{\boldsymbol{\alpha}(i)}$$

onde $h_i \in \mathbb{C}[X_1, \ldots, X_n]$ para $i = 1, 2, \ldots, s$.

Assim, dado $f \in I$, podemos escrever

$$f = \sum_{i-1}^{s} a_i \mathbf{X}^{\boldsymbol{\beta}(i)} = \sum_{i=1}^{s} (a_i h_i) \mathbf{X}^{\boldsymbol{\alpha}(i)}$$

onde $a_i h_i \in \mathbb{C}[X_1, \ldots, X_n]$ e, portanto $f \in \langle \mathbf{X}^{\boldsymbol{\alpha}(1)}, \ldots, \mathbf{X}^{\boldsymbol{\alpha}(s)} \rangle$.

Logo, $I = \langle \mathbf{X}^{\boldsymbol{\alpha}(1)}, \ldots, \mathbf{X}^{\boldsymbol{\alpha}(s)} \rangle$, o que completa a demonstração. □

3.4 Base de Gröbner

Nessa seção veremos como solucionar definitivamente o problema da descrição de um ideal, ou seja, seremos capazes de reconhecer quando um polinômios f de $\mathbb{C}[X_1, \ldots, X_n]$ pertence ou não a um ideal I desse mesmo anel. Nosso estudo nos guiará a bases ideais com "boas"propriedades relativas ao algoritmo da divisão em $\mathbb{C}[X_1, \ldots, X_n]$. A ideia chave é que, uma vez escolhida uma ordem monomial, todo polinômio f em $\mathbb{C}[X_1, \ldots, X_n]$ terá um único termo líder $TL(f)$. Assim, para qualquer ideal I, podemos associá-lo a um ideal de termos líderes, como sugere a definição abaixo.

Definição 3.4.1. Seja $I \subset \mathbb{C}[X_1, \ldots, X_n]$ um ideal não nulo.

(i) Denotamos por $TL(I)$ o conjunto formado pelos termos líderes dos elementos de I, isto é

$$TL(I) = \{c\mathbf{X}^{\boldsymbol{\alpha}} | \text{ existe } f \in I \text{ com } TL(f) = c\mathbf{X}^{\boldsymbol{\alpha}}\}.$$

(ii) Denotamos por $\langle TL(I) \rangle$ o ideal gerado pelos elementos de $TL(I)$

Proposição 3.4.2. *Seja $I \subset \mathbb{C}[X_1, \ldots, X_n]$ um ideal.*

(i) *$\langle TL(I) \rangle$ é um ideal monomial.*

(ii) *Existem $g_1, \ldots g_s \in I$ tais que $\langle TL(I) \rangle = \langle TL(g_1), \ldots, TL(g_s) \rangle$.*

Prova. (i) Os monômios líderes $ML(g)$ de elementos de $g \in I - \{0\}$ geram o ideal monomial $\langle ML(g) : g \in I - \{0\} \rangle$. Note que $TL(g)$ e $ML(g)$ se diferem apenas por uma constante não-nula, logo

$$\langle ML(g) : g \in I - \{0\} \rangle = \langle TL(g) : g \in I - \{0\} \rangle = \langle TL(I) \rangle$$

ou seja, $\langle TL(I) \rangle$ é um ideal monomial.

(ii) Observe que $\langle TL(I)\rangle$ é gerado por monômios $ML(g)$ com $g \in I - \{0\}$. Assim, pelo Lema de Dickson, temos

$$\langle TL(I)\rangle = \langle ML(g_1), \ldots, ML(g_s)\rangle$$

para um número finito de $g_1, \ldots, g_s \in I$. Como $ML(g_i)$ e $TL(g_i)$ se diferem apenas por uma constante não-nula, concluímos que

$$\langle TL(I)\rangle = \langle TL(g_1), \ldots, TL(g_s)\rangle$$

o que completa a demonstração. □

Como pudemos ver na seção 3.2 deste capítulo, os termos líderes de um polinômio têm grande importância no algoritmo da divisão. Isso nos mostra uma sutil, porém importante, ponto a cerca do $\langle TL(I)\rangle$. Mostra que dado um conjunto gerador finito para I, por exemplo $I = \langle f_1, \ldots, f_s\rangle$, então $\langle TL(f_1), \ldots, TL(f_s)\rangle$ e $\langle TL(I)\rangle$ podem ser ideais diferentes. Como $TL(f_i) \in TL(I) \subset \langle TL(I)\rangle$ temos que $\langle TL(f_1), \ldots, TL(f_s)\rangle \subset \langle TL(I)\rangle$. Entretanto, $\langle TL(I)\rangle$ pode ser estritamente maior. Para observar esse fato, consideremos o exemplo a seguir.

Exemplo 3.4.3. Seja $I = \langle f_1, f_2\rangle$, tal que $f_1 = X^3 - 2XY$ e $f_2 = X^2Y - 2Y^2 + X$. Usando a ordenação lexicográfica graduada em $\mathbb{C}[X, Y]$ com $X > Y$ temos

$$X \cdot (X^2Y - 2Y^2 + X) - Y \cdot (X^3 - 2XY) = X^2$$

donde concluímos que $X^2 \in I$. Assim, $X^2 = TL(X^2) \in \langle TL(I)\rangle$. Porém X^2 não é divisível por $TL(f_1) = X^3$ nem por $TL(f_2) = X^2Y$; logo, pelo Lema 3.3.2, temos $X^2 \notin \langle TL(f_1), TL(f_2)\rangle$.

Finalmente, chegamos à definição do que é uma base de Gröbner

Definição 3.4.4. Seja I um ideal de $\mathbb{C}[X_1, \ldots, X_n]$. Um conjunto $\{g_1, \ldots, g_n\}$ de geradores de I é chamado de base de *Gröbner* de I se a inclusão $\langle TL(f_1), \ldots, TL(f_s)\rangle \subset \langle TL(I)\rangle$ for uma igualdade.

Teorema 3.4.5. *Fixe uma ordem monomial. Então todo ideal $I \subset \mathbb{C}[X_1, \ldots, X_n]$ não nulo tem uma base de Gröbner.*

Prova. Se $I = \{0\}$ não há o que demonstrar. Suponhamos então que I contém algum polinômio não nulo. Assim, pela Proposição 3.4.2, existem $g_1, \ldots, g_s \in I$ tais

que

$$\langle TL(I)\rangle = \langle TL(g_1), \ldots, TL(g_s)\rangle$$

Vamos provar que $I = \langle g_1, \ldots, g_s\rangle$. Notemos inicialmente que $\langle g_1, \ldots, g_s\rangle \subset I$ pois cada $g_i \in I$. Reciprocamente, se f é um polinômio qualquer tal que $f \in I$ então, se dividirmos f por $g_1, \ldots, g_s$ chegaremos a uma expressão do tipo

$$f = a_1 g_1 + \ldots + a_s g_s + r$$

onde r não é divisível por nenhum dos termos líderes de $g_1, \ldots, g_s$. Reescrevendo a expressão acima de modo conveniente temos

$$r = f - a_1 g_1 + \ldots + a_s g_s.$$

Se $r \neq 0$, então $TL(r) \in \langle TL(I)\rangle = \langle TL(g_1), \ldots, TL(g_s)\rangle$. Então, pelo Lema 3.3.2, temos que $TL(r)$ é divisível por algum $TL(g_i)$, o que é uma contradição. Logo $r = 0$ e, portanto

$$f = a_1 g_1 + \ldots + a_s g_s \in \langle TL(I)\rangle = \langle TL(g_1), \ldots, TL(g_s)\rangle$$

o que mostra a inclusão $I \subset \langle g_1, \ldots, g_s\rangle$. □

3.5 Propriedades das Bases de Gröbner

Na seção anterior, vimos que todo ideal em $\mathbb{C}[X_1, \ldots, X_n]$ possui uma base de Gröbner. Veremos agora algumas propriedades das bases de Gröbner que nos permitem decidir se um polinômio pertence ou não a um ideal. Mais precisamente, veremos que o resto obtido no algoritmo da divisão é único quando dividimos por uma base de Gröbner.

Proposição 3.5.1. *Seja $G = \{g_1, \ldots, g_s\}$ uma base de Gröbner para um ideal $I \subset \mathbb{C}[X_1, \ldots, X_n]$ e seja $f \in \mathbb{C}[X_1, \ldots, X_n]$. Então, existe um único $r \in \mathbb{C}[X_1, \ldots, X_n]$ com as seguintes propriedades:*

(i) *Nenhum termo de r é divisível por algum dos $TL(g_1), \ldots, TL(g_s)$.*

(ii) *Existe $g \in I$ tal que $f = g + r$.*

Em particular, r é o resto da divisão de f por g, não importando como os elementos de G estão ordenados quando usamos o algoritmo da divisão.

Prova. (*Existência*) Sabemos, pelo algoritmo da divisão que

$$f = a_1 g_1 + \ldots + a_s g_s + r$$

tal que r, pela construção do algoritmo da divisão, satisfaz i). Como $g_1, \ldots, g_s \in I$, então $a_1 g_1 + \ldots + a_s g_s = g \in I$. Assim, $f = g + r$ com $g \in I$, o que prova a existência de r.
(*Unicidade*) Suponhamos que existam g_1, g_2, r_1, r_2 tais que $f = g_1 + r_1 = g_2 + r_2$ que satisfaçam (i) e (ii). Assim, $r_2 - r_1 = g_1 - g_2 \in I$. Se $r_1 \neq r_2$, então $TL(r_2 - r_1) \in \langle TL(I) \rangle = \langle TL(g_1), \ldots, TL(g_s) \rangle$. Desse modo, pelo Lema 3.3.2, concluímos que $TL(r_2 - r_1)$ é divisível por algum $TL(g_i)$ o que é uma contradição pois nenhum termo líder de r_1 ou r_2 é divisível por $TL(g_1), \ldots, TL(g_s)$. Logo, $r_1 - r_2 = 0$ e, portanto, $r_1 = r_2$. □

Note que a propriedade acima garante que quando dividimos um polinômio por uma base de Gröbner o resto é único, independente da ordem dos geradores. Entretanto os quocientes a_i produzidos pelo algoritmo da divisão em $f = a_1 g_1 + \ldots + a_s g_s + r$ provavelmente mudam se listarmos os geradores em uma ordem diferente.

Notação: Se $G = \{g_1, \ldots, g_s\}$ é uma base de Gröbner de um ideal I então denotamos o resto da divisão de um polinômio f por G pelo símbolo $\overline{f}^G$.

Segue abaixo um critério para determinar se um polinômio pertence ou não a um ideal.

Corolário 3.5.2. *Seja $G = \{g_1, \ldots, g_s\}$ uma base de Gröbner para o ideal $I \subset \mathbb{C}[X_1, \ldots, X_n]$ e seja $f \in \mathbb{C}[X_1, \ldots, X_n]$. Então $f \in I$ se, e somente se, o resto da divisão de f por G é zero.*

Prova. Se $r = 0$ então $f = a_1 g_1 + \ldots + a_s g_s \in I$ pois $g_1, \ldots, g_s \in I$. Reciprocamente, se $f \in I$ então a igualdade $f = f + 0$ satisfaz as duas condições da proposição anterior, ou seja, zero é o resto da divisão de f por G. □

Definição 3.5.3. Sejam $f, g \in \mathbb{C}[X_1, \ldots, X_n]$ polinômios não nulos.

i) Se multigrau$(f) = \boldsymbol{\alpha}$ e multigrau$(g) = \boldsymbol{\beta}$, consideramos $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$, onde $\lambda_i = \max(\alpha_i, \beta_i)$ para todo $i = 1, \ldots, n$. Chamamos $\mathbf{X}^{\boldsymbol{\lambda}}$ o *mínimo múltiplo comum* de $ML(f)$ e $ML(g)$ e escrevemos

$$\mathbf{X}^{\boldsymbol{\lambda}} = \text{MMC}(ML(f), ML(g)).$$

ii) O *S-polinômio* de f e g é a combinação

$$S(f,g) = \frac{\mathbf{X}^{\boldsymbol{\lambda}}}{TL(f)} \cdot f - \frac{\mathbf{X}^{\boldsymbol{\lambda}}}{TL(g)} \cdot g.$$

Vejamos um exemplo para ajudar a fixar esse conceito.

Exemplo 3.5.4. Sejam $f = X^2Y^2 + X^3 + XY^2 + 2X$ e $g = X^4Y^2 + X^5 + 2XY$ polinômios em $\mathbb{C}[X, Y]$. Considere a ordenação lexicográfica graduada com $X > Y$. Assim, pela definição acima, $\boldsymbol{\lambda} = (4, 2)$ e

$$\begin{aligned} S(f,g) &= \frac{X^4Y^2}{X^2Y^2} \cdot (X^2Y^2 + X^3 + XY^2 + 2X) - \frac{X^4Y^2}{X^4Y^2} \cdot (X^4Y^2 + X^5 + 2XY) \\ &= X^2 \cdot (X^2Y^2 + X^3 + XY^2 + 2X) - (X^4Y^2 + X^5 + 2XY) \\ &= X^3Y^2 + 2X^3 - 2XY. \end{aligned}$$

Observe que um *S-polinômio* $S(f, g)$ é construído de modo que haja cancelamento de termos líderes. O lema abaixo mostra que todo cancelamento de termos líderes entre polinômios com mesmo multigrau resulta desse tipo de cancelamento.

Lema 3.5.5. *Considere a soma* $\sum_{i=1}^{s} c_i f_i$, *onde* $c_i \in \mathbb{C}$ *e* multigrau$(f_i) = \boldsymbol{\delta} \in \mathbb{Z}_+^n$ *para todo* $i = 1, \ldots, s$. *Se* multigrau $\left(\sum_{i=1}^{s} c_i f_i \right) < \boldsymbol{\delta}$, *então* $\sum_{i=1}^{s} c_i f_i$ *é uma combinação linear com coeficientes em* $\mathbb{C}$ *dos S-polinômios* $S(f_j, f_k)$, *para* $1 \leq j, k \leq s$. *Além disso,* multigrau $(S(f_j, f_k)) < \boldsymbol{\delta}$.

Prova. Seja $d_i = CL(f_i)$,, de modo que $c_i d_i$ é o coeficiente líder de $c_i f_i$. Como $c_i f_i$ tem multigrau $= \boldsymbol{\delta}$ e a soma tem multigrau estritamente menor, segue que $\sum_{i=1}^{s} c_i d_i = 0$.

Considere $p_i = \frac{f_i}{d_i}$. Note que $CL(p_i) = 1$. Considere ainda a soma

$$\begin{aligned}\sum_{i=1}^{s} c_i f_i &= \sum_{i=1}^{s} c_i d_i p_i \\ &= c_1 d_1 (p_1 - p_2) + (c_1 d_1 + c_2 d_2)(p_2 - p_3) + \ldots + \\ &\quad (c_1 d_1 + \ldots + c_{s-1} d_{s-1})(p_{s-1} - p_s) + (c_1 d_1 + \ldots + c_s d_s) p_s.\end{aligned}$$

Por construção, $TL(f_i) = d_i \mathbf{X}^{\boldsymbol{\delta}}$. Logo, o MMC$(ML(f_i), ML(f_k)) = \mathbf{X}^{\boldsymbol{\delta}}$ e portanto

$$\begin{aligned}S(f_j, f_k) &= \frac{\mathbf{X}^{\boldsymbol{\delta}}}{TL(f_j)} f_j - \frac{\mathbf{X}^{\boldsymbol{\delta}}}{TL(f_k)} f_k \\ &= \frac{\mathbf{X}^{\boldsymbol{\delta}}}{d_j \mathbf{X}^{\boldsymbol{\delta}}} f_j - \frac{\mathbf{X}^{\boldsymbol{\delta}}}{d_k \mathbf{X}^{\boldsymbol{\delta}}} f_k \\ &= p_j - p_k \qquad (3.1)\end{aligned}$$

Usando esta equação e $\sum_{i=1}^{s} c_i d_i = 0$, a soma telescópica acima torna-se

$$\sum_{i=1}^{s} = c_1 d_1 S(f_1, f_2) + (c_1 d_1 + c_2 d_2) S(f_2, f_3) + \ldots + (c_1 d_1 + \ldots + c_{s-1} d_{s-1}) S(f_{s-1}, f_s)$$

que é a soma da forma desejada. Como p_j e p_k têm multigrau $\boldsymbol{\delta}$ e coeficiente líder 1, então multigrau$(p_j - p_k) < \delta$. Pela equação (3.1), o mesmo vale para $S(f_j, f_k)$ e isso conclui a demonstração. □

Usando S-polinômios e o Lema 3.5.5, podemos finalmente provar o *Critério de Buchberger* que determina quando uma base de um ideal é uma base de Gröbner.

Teorema 3.5.6. (*Critério de Buchberger*) *Seja I um ideal polinomial. Então a base $G = \{g_1, \ldots, g_s\}$ de I é uma base de Gröbner de I se, e somente se, para todo $i \neq j$, o resto da divisão de $S(g_i, g_j)$ por G (listados em qualquer ordem) é zero.*

Prova. ($\Rightarrow$) Se G é uma base de Gröbner então, como $S(g_1, g_j) \in I$, o resto da divisão por G é zero (vide Corolário 3.5.2).

($\Leftarrow$) Seja $f \in I$ um polinômio não nulo. Nosso objetivo é mostrar que se todos os S-polinômios têm resto zero quando divididos por G, então $TL(f) \in \langle TL(g_1), \ldots, TL(g_t) \rangle$.

Dado $f \in I = \langle g_1, \ldots, g_t \rangle$, existem polinômios $h_i \in \mathbb{C}[X_1, \ldots, X_n]$ tais que

$$f = \sum_{i=1}^{t} h_i g_i. \tag{3.2}$$

Pelo Lema 3.1.14, temos

$$\text{multigrau}(f) \leq \max(\text{multigrau}(h_i g_i)). \tag{3.3}$$

Se a igualdade não ocorrer, então deve ocorrer o cancelamento dos termos líderes de (3.2). Como, pelo Lema 3.5.5, todo cancelamento de termos líderes se dá por S-polinômios, então podemos reescrever isto em termos deste. Desse modo, já que por hipótese os S-polinômios têm resto zero, podemos substitui-los por expressões que envolvam menos cancelamentos. Assim, vamos obter uma expressão para f com menos cancelamentos dos termos líderes. Continuando esse processo, obteremos em alguma etapa uma expressão para f tal que

$$\text{multigrau}(f) = \max(\text{multigrau}(h_i g_i))$$

para algum i, ou seja, $TL(f)$ é divisível por $TL(g_i)$. Daí concluiremos que $TL(f) \in \langle TL(g_1), \ldots, TL(g_t) \rangle$ e isso completará nossa prova.

Seja (3.2) uma expressão para f e $m(i) = \text{multigrau}(h_i g_i)$. Seja ainda $\boldsymbol{\delta} = \max\{m(1), \ldots, m(t)\}$. Temos então que

$$\text{multigrau}(f) \leq \boldsymbol{\delta}.$$

Considere agora todos os possíveis modos de escrever f na forma (3.2). Para cada possibilidade, temos possivelmente um $\boldsymbol{\delta}$ diferente. Porém, como toda ordem monomial é uma boa ordenação, podemos escolher uma expressão para f tal que δ é mínimo.

Devemos mostrar que para este $\boldsymbol{\delta}$ mínimo escolhido, o $\text{multigrau}(f) = \boldsymbol{\delta}$, pois assim, vale a igualdade (3.3) e, consequentemente, $TL(f) \in \langle TL(g_1), \ldots, TL(g_s) \rangle$.

Suponhamos por contradição que $\text{multigrau}(f) < \boldsymbol{\delta}$ e escrevamos f conveniente-

mente de modo a isolar os termos de multigrau $\boldsymbol{\delta}$

$$\begin{aligned} f &= \sum_{m(i)=\boldsymbol{\delta}} h_i g_i + \sum_{m(i)<\boldsymbol{\delta}} h_i g_i \\ &= \sum_{m(i)=\boldsymbol{\delta}} (TL(h_i) + h_i - TL(h_i)) g_i + \sum_{m(i)<\boldsymbol{\delta}} h_i g_i \\ &= \sum_{m(i)=\boldsymbol{\delta}} TL(h_i) g_i + \sum_{m(i)=\boldsymbol{\delta}} (h_i - TL(h_i)) g_i + \sum_{m(i)<\boldsymbol{\delta}} h_i g_i. \end{aligned} \tag{3.4}$$

Note que todos os polinômios que aparecem na segunda e terceira somas da terceira linha têm multigrau $< \boldsymbol{\delta}$. Desse modo, a hipótese de que multigrau$(f) < \boldsymbol{\delta}$ implica que a primeira soma tem multigrau menor que $\boldsymbol{\delta}$, ou seja,

$$\text{multigrau} \left(\sum_{m(i)=\boldsymbol{\delta}} TL(h_i) g_i \right) < \boldsymbol{\delta}.$$

Seja $TL(h_i) = c_i \mathbf{X}^{\boldsymbol{\alpha}(i)}$. Então

$$\sum_{m(i)=\boldsymbol{\delta}} TL(h_i) g_i = \sum_{m(i)=\boldsymbol{\delta}} c_i \mathbf{X}^{\boldsymbol{\alpha}(i)} g_i$$

satisfaz as hipóteses do Lema 3.5.5 com $f_i = \mathbf{X}^{\boldsymbol{\alpha}(i)} g_i$, e portanto este cancelamento pode ser escrito como uma combinação linear de S-polinômios $S(\mathbf{X}^{\boldsymbol{\alpha}(j)} g_j, \mathbf{X}^{\boldsymbol{\alpha}(k)} g_k)$. Contudo,

$$\begin{aligned} S(\mathbf{X}^{\boldsymbol{\alpha}(j)} g_j, \mathbf{X}^{\boldsymbol{\alpha}(k)} g_k) &= \frac{\mathbf{X}^{\boldsymbol{\delta}}}{\mathbf{X}^{\boldsymbol{\alpha}(j)} TL(g_j)} \mathbf{X}^{\boldsymbol{\alpha}(j)} g_j - \frac{\mathbf{X}^{\boldsymbol{\delta}}}{\mathbf{X}^{\boldsymbol{\alpha}(k)} TL(g_k)} \mathbf{X}^{\boldsymbol{\alpha}(k)} g_k \\ &= \frac{\mathbf{X}^{\boldsymbol{\delta}}}{TL(g_j)} g_j - \frac{\mathbf{X}^{\boldsymbol{\delta}}}{TL(g_k)} g_k \\ &= \frac{\mathbf{X}^{\boldsymbol{\delta} - \boldsymbol{\gamma}_{jk} + \boldsymbol{\gamma}_{jk}}}{TL(g_j)} g_j - \frac{\mathbf{X}^{\boldsymbol{\delta} - \boldsymbol{\gamma}_{jk} + \boldsymbol{\gamma}_{jk}}}{TL(g_k)} g_k \\ &= \mathbf{X}^{\boldsymbol{\delta} - \boldsymbol{\gamma}_{jk}} S(g_j, g_k) \end{aligned}$$

onde $\mathbf{X}^{\boldsymbol{\gamma}_{jk}} = MMC(ML(g_j), ML(g_k))$. Logo, existem constantes $c_{jk} \in \mathbb{C}$ tais que

$$\sum_{m(i)=\delta} TL(h_i) g_i = \sum_{jk} c_{jk} \mathbf{X}^{\boldsymbol{\delta} - \boldsymbol{\gamma}_{jk}} S(g_j, g_k). \tag{3.5}$$

O próximo passo é usar nossa hipótese de que o resto de $S(g_j, g_k)$ na divisão por $g_1, \ldots, g_t$ é zero. Usando o algoritmo da divisão, isso significa que cada S-polinômio pode ser escrito na forma

$$S(g_j, g_k) = \sum_{i=1}^{t} a_{ijk} g_i, \tag{3.6}$$

onde $a_{ijk} \in \mathbb{C}[X_1, \ldots, X_n]$.

Sabemos ainda, pelo algoritmo da divisão, que

$$\text{multigrau}(a_{ijk} g_i) \leq \text{multigrau}(S(g_j, g_k)) \tag{3.7}$$

para todo i, j, k.

Intuitivamente, isso mostra que quando o resto é zero, podemos encontrar uma expressão para $S(g_j, g_k)$ em termos de G onde nem todos os termos líderes se cancelam.

Para explorar isso, multipliquemos $S(g_j, g_k)$ por $\mathbf{X}^{\boldsymbol{\delta}-\boldsymbol{\gamma}_{jk}}$ para obter

$$\mathbf{X}^{\boldsymbol{\delta}-\boldsymbol{\gamma}_{jk}} S(g_j, g_k) = \sum_{i=1}^{s} b_{ijk} g_i,$$

onde $b_{ijk} = \mathbf{X}^{\boldsymbol{\delta}-\boldsymbol{\gamma}_{jk}} a_{ijk}$. Então, pelo Lema 3.5.5, temos

$$\text{multigrau}(b_{ijk} g_i) \leq \text{multigrau}(\mathbf{X}^{\boldsymbol{\delta}-\boldsymbol{\gamma}_{jk}} S(g_j, g_k)) < \boldsymbol{\delta} \tag{3.8}$$

Substituindo a expressão acima em (3.5) obtemos

$$\begin{aligned} \sum_{m(i)=\boldsymbol{\delta}} TL(h_i) g_i &= \sum_{jk} c_{jk} \mathbf{X}^{\boldsymbol{\delta}-\boldsymbol{\gamma}_{jk}} S(g_j, g_k) \\ &= \sum_{jk} c_{jk} \left(\sum_{i} b_{ijk} g_i \right) \\ &= \sum_{i} \widetilde{h}_i g_i \end{aligned}$$

donde concluímos que $\text{multigrau}(\widetilde{h}_i g_i) < \boldsymbol{\delta}$, para todo i.

Finalmente, substituindo $\sum_{m(i)=\delta} TL(h_i) g_i = \sum_i \widetilde{h}_i g_i$ em (3.4), obtemos uma expressão para f como combinação dos polinômios $g_i's$, onde todos têm multigrau menos que $\boldsymbol{\delta}$, o que contradiz a minimalidade de $\boldsymbol{\delta}$ e conclui a prova do teorema. □

O teorema acima é conhecido como "*critério dos S-pares de Buchberger*"e é um dos resultados chave na teoria das Bases de Gröbner pois fornece um método simples para identificar quando uma base dada é uma base de Gröbner. Mais à frente veremos que esse teorema nos levará a um algoritmo para encontrar as bases de Gröbner de um determinado ideal. Para fixar melhor esse teorema, vejamos o exemplo abaixo.

Exemplo 3.5.7. Seja o ideal $I = \langle -Y + X^2, -Z + X^3 \rangle$. Vamos mostrar que $G = \{-Y + X^2, -Z + X^3\}$ é uma base de Gröbner considerando a ordem *lex* com $Y > Z > X$. Considere $f = -Y + X^2$ e $g = -Z + X^3$. Note que multigrau$(f) = (1, 0, 0)$ e multigrau$(g) = (0, 1, 0)$. Assim,

$$\mathbf{X^{\lambda}} = \text{MMC}(ML(f), ML(g)) = YZ.$$

Logo, o S-polinômio $S(f, g)$ é dado por

$$\begin{aligned} S(f,g) &= \frac{\mathbf{X^{\lambda}}}{TL(f)} f - \frac{\mathbf{X^{\lambda}}}{TL(g)} g \\ &= \frac{YZ}{-Y}(-Y + X^2) - \frac{YZ}{-Z}(-Z + X^3) \\ &= YX^3 - XZ^2 \end{aligned}$$

Dividindo $S(f, g)$ por f e g, temos

$$YX^3 - XZ^2 = -X^3(-Y + X^2) + X^2(-Z + X^3).$$

Então, pelo Teorema 3.5.6, concluímos que G é uma base de Gröbner para I.

3.6 Algoritmo de Buchberger

Já sabemos, pelo Teorema 3.4.5, que todo ideal não nulo $I \subset \mathbb{C}[X_1, \ldots, X_n]$ tem uma base de Gröbner. Entretanto este teorema não nos mostrou como produzir a referida base. Nesta seção apresentaremos o *Algoritmo de Buchberger*, capaz de computar uma base de Gröbner de um ideal $I \subset \mathbb{C}[X_1, \ldots, X_n]$.

A ideia chave desse algoritmo é tentar expandir o conjunto original de geradores a uma base de Gröbner, adicionando os restos não nulos de $S(f_i, f_j)$ na divisão por F, onde F é o conjunto de geradores em um determinado momento do processo.

Porém, como podemos garantir que o processo termina em algum momento?

Observe que, em cada etapa do processo, se o resto de $S(f_i, f_j)$ na divisão por F é diferente de zero, ou seja, se $\overline{S(f_i, f_j)}^F \neq 0$, faremos $f_k = \overline{S(f_i, f_j)}^F$ e o acrescentaremos a F obtendo um novo conjunto F' de geradores. Note ainda que, em cada etapa, o conjunto antigo de geradores (digamos, F_k) está contido no novo conjunto de geradores (F_{k+1}). Dessa inclusão temos que $\langle \mathrm{TL}(F_k) \rangle \subset \langle \mathrm{TL}(F_{k+1}) \rangle$. Assim, se o processo de construção desses F_{k_s} fosse infinito, teríamos uma cadeia ascendente de ideais do tipo

$$\langle \mathrm{TL}(F_1) \rangle \subset \langle \mathrm{TL}(F_2) \rangle \subset \ldots \langle \mathrm{TL}(F_k) \rangle \subset \langle \mathrm{TL}(F_{k+1}) \rangle \subset \ldots$$

que cresceria indefinidamente, o que é um absurdo visto que $\mathbb{C}[X_1, \ldots, X_n]$ é um anel noetheriano.

Para fixar melhor essa ideia, vejamos o seguinte exemplo.

Exemplo 3.6.1. Seja $I = \langle f_1, f_2 \rangle = \langle X^2Y - 1, XY^2 - X \rangle \in \mathbb{C}[X, Y]$ e considere a ordem lexicográfica com $X > Y$. Temos multigrau$(f_1) = (2, 1)$ e multigrau$(f_2) = (1, 2)$. Assim,

$$S(f_1, f_2) = \frac{X^2Y^2}{X^2Y}(X^2Y - 1) - \frac{X^2Y^2}{XY^2}(XY^2 - X) = X^2 - Y.$$

Notemos que, pelo Teorema 3.5.6, $\{f_1, f_2\}$ não é uma base de Gröbner pois, considerando $F = (f_1, f_2)$, temos que $\overline{S(f_1, f_2)}^F = X^2 - Y \neq 0$. Porém se adicionarmos $f_3 = X^2 - Y$ ao conjunto original de geradores, $\overline{S(f_1, f_2)}^F = 0$. Entretanto,

$$S(f_1, f_3) = \frac{X^2Y}{X^2Y}(X^2Y - 1) - \frac{X^2Y}{X^2}(X^2 - Y) = Y^2 - 1$$

ou seja, $\overline{S(f_1, f_3)}^F = Y^2 - 1 \neq 0$. Então, acrescentemos também $f_4 = Y^2 - 1$ à base de geradores de I, o que faz com que $\overline{S(f_1, f_3)}^F = 0$. Note ainda que

$$S(f_1, f_4) = \frac{X^2Y^2}{X^2Y}(X^2Y - 1) - \frac{X^2Y^2}{Y^2}(Y^2 - 1) = X^2 - Y = f_3,$$

ou seja,

$$\overline{S(f_1, f_4)}^F = 0.$$

Porém,

$$S(f_2, f_3) = \frac{X^2Y^2}{XY^2}(XY^2 - X) - \frac{X^2Y^2}{X^2}(X^2 - Y) = -X^2 + Y^3$$

de modo que $\overline{S(f_2, f_3)}^F = -X^2 + Y^3 \neq 0$. Assim, devemos adicionar $f_5 = -X^2 + Y^3$ à base de geradores, o que faz com que $\overline{S(f_2, f_3)}^F = 0$. Temos ainda que,

$$S(f_1, f_5) = Y^4 - 1 = (Y^2 + 1)(Y^2 - 1) = (Y^2 + 1) \cdot f_4 \Rightarrow \overline{S(f_1, f_5)}^F = 0,$$

$$S(f_2, f_4) = 0 \Rightarrow \overline{S(f_2, f_4)}^F = 0,$$

$$S(f_2, f_5) = -X^2 + Y^5 \Rightarrow \overline{S(f_2, f_5)}^F = 0,$$

$$S(f_3, f_4) = X^2 - Y^3 = -f_5 \Rightarrow \overline{S(f_3, f_4)}^F = 0,$$

$$S(f_3, f_5) = -Y^4 - Y^2 = Y^2(Y^2 - 1) = Y^2 \cdot f_4 \Rightarrow \overline{S(f_3, f_5)}^F = 0,$$

e

$$S(f_4, f_5) = -X^2 + Y^5 \Rightarrow \overline{S(f_4, f_5)}^F = 0.$$

Assim, pelo Teorema 3.5.6,

$$\{f_1, f_2, f_3, f_4, f_5\} = \{X^2Y - 1, XY^2 - X, X^2 - Y, Y^2 - 1, -X^2 + Y^3\}$$

é uma base de Gröbner.

Em geral, bases de Gröbner construídas através do algoritmo de Buchberger são maiores que o necessário. O lema abaixo nos mostra como podemos, eventualmente, eliminar geradores "desnecessários".

Lema 3.6.2. *Seja G uma base de Gröbner do ideal I. Seja $P \in G$ um polinômio tal que $TL(P) \in \langle TL(G - \{P\})\rangle$. Então $G - \{P\}$ também é uma base de Gröbner para I.*

Prova. Já sabemos que $\langle TL(G)\rangle = \langle TL(I)\rangle$. Se $TL(P) \in \langle TL(G - \{P\})\rangle$, então $\langle TL(G - \{P\})\rangle = \langle TL(G)\rangle$. Logo, por definição, $G - \{P\}$ também é uma base de Gröbner para I. □

Definição 3.6.3. Uma *Base de Gröbner Minimal* para um ideal polinomial I é uma base de Gröbner G de I tal que:

(i) $CL(P) = 1 \ \ \forall P \in G$.

(ii) $\forall P \in G, TL(P) \notin \langle TL(G - \{P\})\rangle$.

Observe que, usando o algoritmo de Buchberger e o Lema 3.6.2, podemos construir uma base de Gröbner minimal para um determinado ideal não nulo. Consideremos, por exemplo, a base de Gröbner obtida no exemplo 3.6.1:

$$\begin{aligned} f_1 &= X^2Y - 1 \\ f_2 &= XY^2 - X \\ f_3 &= X^2 - Y \\ f_4 &= Y^2 - 1 \\ f_5 &= -X^2 + Y^3 \end{aligned}$$

Notemos que $TL(f_1) = X^2Y = Y \cdot TL(f_3)$. Então, pelo Lema 3.6.2, podemos retirar f_1. Do mesmo modo podemos retirar f_2 e f_5 pois $TL(f_2) = XY^2 = X \cdot TL(f_4)$ e $TL(f_5) = -X^2 = (-1) \cdot TL(f_3)$. Consequentemente,

$$\widetilde{f_3} = X^2 - Y, \quad \text{e} \quad \widetilde{f_4} = Y^2 - 1$$

formam uma base de Gröbner minimal de I

Definição 3.6.4. Uma *base de Gröbner reduzida* para um ideal I é uma base de Gröbner G para I tal que:

(i) $CL(P) = 1$ para todo $P \in G$.

(ii) Para todo $P \in G$, nenhum monômio de P pertence a $\langle TL(G - \{P\})\rangle$.

3.7 Teoria de eliminação

Na seção anterior apresentamos um algoritmo para encontrar uma base de Gröbner de um ideal polinomial I. Nessa seção vamos mostrar como podemos usar bases de Gröbner para resolver sistemas de equações polinomiais em várias variáveis. A ideia básica dessa técnica é transformar um dado sistema de equações com várias variáveis em outro sistema equivalente (com as mesmas soluções) de modo que este seja composto por equações com menos variáveis. Assim, podemos considerar essa técnica como uma generalização do *método de Gauss* para resolução de sistemas de equações lineares.

Para entender como o processo de eliminação funciona, consideremos o seguinte exemplo.

Exemplo 3.7.1. Resolver, em $\mathbb{C}^2$, o sistema de equações

$$\begin{cases} X^2 + 2Y^2 = 3 \\ X^2 + XY + Y^2 = 3 \end{cases} \tag{3.9}$$

Inicialmente consideremos o ideal

$$I = \langle X^2 + 2Y^2 - 3, X^2 + XY + Y^2 - 3 \rangle.$$

Note que a solução do sistema acima é simplesmente o conjunto de pontos de $Z(I)$. Pela Proposição 2.1.4, sabemos que podemos determinar $Z(I)$ usando qualquer base de I (uma base de Gröbner por exemplo). Considerando $F = \{f_1 = X^2 + 2Y^2 - 3, f_2 = X^2 + XY + Y^2 - 3\}$ e usando o algoritmo de Buchberger, vamos determinar uma base de Gröbner para I. Para isso, fixemos a ordenação lexicográfica com $X > Y$. Observe que o multigrau(f_1) = multigrau$(f_2) = (2, 0)$. Então

$$S(f_1, f_2) = \frac{X^2}{X^2}(X^2 + 2Y^2 - 3) - \frac{X^2}{X^2}(X^2 + XY + Y^2 - 3) = -XY + Y^2$$

e, consequentemente, $\overline{S(f_1, f_2)}^F = -XY + Y^2 \neq 0$. Adicionemos então o polinômio$f_3 = -XY + Y^2$ a F. Continuando a usar o algoritmo de Buchberger, temos que

$$S(f_1, f_3) = \frac{X^2Y}{X^2}(X^2 + 2Y^2 - 3) - \frac{X^2Y}{-XY}(-XY + Y^2) = XY^2 + 2Y^3 - 3Y$$

portanto $\overline{S(f_1, f_3)}^F = 3Y^3 - 3Y \neq 0$, ou seja, devemos adicionar $f_4 = 3Y^3 - 3Y$ a F também. Efetuando os cálculos necessários, obtemos

$$\overline{S(f_1, f_4)}^F = \overline{S(f_2, f_3)}^F = \overline{S(f_2, f_4)}^F = \overline{S(f_3, f_4)}^F = 0$$

o que mostra que

$$\begin{aligned} f_1 &= X^2 + 2Y^2 - 3 \\ f_2 &= X^2 + XY + Y^2 - 3 \\ f_3 &= -XY + Y^2 \\ f_4 &= 3Y^3 - 3Y \end{aligned}$$

é uma base de Gröbner para I.

Desse modo, segue que o sistema (3.9) e

$$\begin{cases} f_1 &= 0 \\ f_2 &= 0 \\ f_3 &= 0 \\ f_4 &= 0 \end{cases}$$

são equivalentes, ou seja, têm o mesmo conjunto solução. De $f_4 = 0$ temos

$$3Y^3 - 3Y = 0 \Rightarrow Y(Y^2 - 1) = 0$$

donde concluímos que os possíveis valores de $Y = -1$ ou $Y = 0$ ou $Y = 1$. Substituindo esses valores nas outras equações do novo sistema obtemos facilmente todos os pares (x, y) que são solução do mesmo, a saber

$$(-\sqrt{3}, 0), (\sqrt{3}, 0), (1, 1), (-1, -1).$$

Observe então que a ideia básica da *Teoria da Eliminação* consiste de duas etapas:

Etapa Eliminação: Nessa etapa buscamos encontrar um gerador f_i onde figure apenas uma variável (no exemplo anterior, $f_4 = 3Y^3 - 3Y$).

Etapa Extensão: Uma vez encontradas as raízes da equação mais simples $f_i = 0$, extendemos estas soluções para a solução do sistema original.

Caso o leitor deseje ter uma visão mais aprofundada sobre o tema sugerimos ver por exemplo [2, Capítulo 3].

3.8 Aplicações

Nesta seção mostraremos como certos sistemas de equações polinomiais apresentados no captítulo 2 podem ser resolvidos mediante o uso das bases de Grobner.

3.8.1 A solução do problema de coloração do mapa da região nordeste

Recordemos que para resolver o problema da coloração do mapa da região nordeste devemos determinar a solução do seguinte sistema de equações polinomiais

$$
\begin{array}{lll}
\bullet X_1^3 - 1 = 0 & \bullet X_1^2 + X_1X_2 + X_2^2 = 0 & \bullet X_4^2 + X_4X_6 + X_6^2 = 0 \\
\bullet X_2^3 - 1 = 0 & \bullet X_1^2 + X_1X_9 + X_9^2 = 0 & \bullet X_5^2 + X_5X_6 + X_6^2 = 0 \\
\bullet X_3^3 - 1 = 0 & \bullet X_2^2 + X_2X_3 + X_3^2 = 0 & \bullet X_6^2 + X_6X_7 + X_7^2 = 0 \\
\bullet X_4^3 - 1 = 0 & \bullet X_2^2 + X_2X_9 + X_9^2 = 0 & \bullet X_7^2 + X_7X_8 + X_8^2 = 0 \\
\bullet X_5^3 - 1 = 0 & \bullet X_3^2 + X_3X_4 + X_4^2 = 0 & \bullet X_7^2 + X_7X_9 + X_9^2 = 0 \\
\bullet X_6^3 - 1 = 0 & \bullet X_3^2 + X_3X_9 + X_9^2 = 0 & \\
\bullet X_7^3 - 1 = 0 & \bullet X_3^2 + X_3X_6 + X_6^2 = 0 & \\
\bullet X_8^3 - 1 = 0 & \bullet X_3^2 + X_3X_7 + X_7^2 = 0 & \\
\bullet X_9^3 - 1 = 0 & \bullet X_4^2 + X_4X_5 + X_5^2 = 0 &
\end{array}
\tag{3.10}
$$

Para isso, iremos determinar inicialmente uma base de Gröbner G para o ideal I cujos geradores são os polinômios que compõem o sistema de equações acima. Devido ao número elevado de operações, utilizaremos o software *Macaulay* 2 para determinar tal base, como mostra a figura abaixo.

```
Macaulay2, version 1.6
with packages: ConwayPolynomials, Elimination, IntegralClosure,
LLLBases,
               PrimaryDecomposition, ReesAlgebra, TangentCone

i1 : R=QQ[x_1..x_9,MonomialOrder=>Lex]

o1 = R

o1 : PolynomialRing
------------------------------------------------------------------------
------

i25 :
i=p1+p2+p3+p4+p5+p6+p7+p8+p9+p10+p11+p12+p13+p14+p15+p16+p17+p18+p19
+p20+p21+p22+p23

o25 : Ideal of R

i26 : g=gb i

o26 = GroebnerBasis[status: done; S-pairs encountered up to degree
6]

o26 : GroebnerBasis

i27 : I=gens g

o27 = | x_9^3-1 x_8^3-1 x_7x_8-x_7x_9+x_8^2-x_9^2 x_7^2+x_7x_9+x_9^2
x_6-x_9 x_5+x_7+x_9 x_4-x_7 x_3+x_7+x_9 x_2-x_7 x_1+x_7+x_9 |
```

Assim, considerando a ordenação lexicográfica com $X_1 > X_2 > \ldots > X_9$, uma base de Gröbner para o ideal I é

$$
\begin{aligned}
G \;=\; & \{X_6 - X_9, X_5 + X_7 + X_9, X_4 - X_7, X_3 + X_7 + X_9, X_2 - X_7, X_1 + X_7 + X_9, \\
& X_7X_8 + X_8^2 - X_7X_9 - X_9^2, X_7^2 + X_7X_9 + X_9^2, X_9^3 - 1, X_8^3 - 1\}
\end{aligned}
$$

Desse modo, o sistema (3.10) tem o mesmo conjunto solução do seguinte sistema:

$$\begin{cases} X_8^3 - 1 = 0 \\ X_9^3 - 1 = 0 \\ X_6 - X_9 = 0 \\ X_5 + X_7 + X_9 = 0 \\ X_4 - X_7 = 0 \\ X_3 + X_7 + X_9 = 0 \\ X_2 - X_7 = 0 \\ X_1 + X_7 + X_9 = 0 \\ X_7X_8 + X_8^2 - X_7X_9 - X_9^2 = 0 \\ X_7^2 + X_7X_9 + X_9^2 = 0 \end{cases} \tag{3.11}$$

Denotando $1, \xi$ e ξ^2 como as raízes cúbicas da unidade, é fácil provar que as únicas soluções da equação $r_1 + r_2 + r_3 = 0$ com $r_i \in \{1, \xi, \xi^2\}$ são aquelas onde r_1, r_2 e r_3 assumem valores distintos. Como cada variável do sistema de equações polinomiais (3.11) é uma raiz cúbica da unidade e cada uma destas raízes está associada a uma cor que irá colorir o mapa, vamos atribuir

$$1 = \text{verde} \quad \xi = \text{Amarelo} \quad \xi^2 = \text{Azul}$$

Analisando agora as equações que compõem o sistema (3.11), observamos que:

(i) Como na segunda equação figura apenas a variável X_9, podemos atribuir a esta, sem perda de generalidade, qualquer raiz cúbica da unidade. Consideremos então $X_9 = 1$.

(ii) Da terceira equação, concluímos que $X_6 = X_9$, ou seja, $X_6 = 1$.

(iii) Na quarta equação, temos que $X_5 + X_7 + X_9 = 0$ apenas se cada variável assumir valores distintos. Como $X_9 = 1$, então vamos considerar $X_5 = \xi$ e $X_7 = \xi^2$.

(iv) Na quinta equação temos $X_4 = X_7$. Logo, $X_4 = \xi^2$.

(v) Na sexta equação, o fato de $X_3 + X_7 + X_9 = 0$ implica em dizer que X_3, X_7 e X_9 são raízes cúbicas distintas da unidade. Como $X_7 = \xi^2$ e $X_9 = 1$, concluímos que $X_3 = \xi$.

(vi) Da sétima equação temos $X_2 = X_7$. Desse modo, $X_2 = \xi^2$.

(vii) Da equação número 8, temos $X_1 + X_7 + X_9 = 0$, ou seja, $X_1 = \xi$.

(ix) Da nona equação, temos

$$\begin{aligned} X_7X_8 + X_8^2 - X_7X_9 - X_9^2 = 0 \quad &\Leftrightarrow \quad X_7(X_8 - X_9) + (X_8 + X_9)(X_8 - X_9) = 0 \\ &\Leftrightarrow \quad (X_8 - X_9)(X_7 + X_8 + X_9) = 0 \\ &\Leftrightarrow \quad X_8 = X_9 \quad \text{ou} \quad X_7 + X_8 + X_9 = 0 \\ &\Leftrightarrow \quad X_8 = 1 \quad \text{ou} \quad X_8 = \xi \end{aligned}$$

Note que apesar de não termos feito nenhuma observação com relação às equações 1 e 10, os resultados obtidos através das outras equações do sistema satisfazem perfeitamente estas igualdades.

Assim, temos duas possíveis soluções para o problema proposto:

$$\begin{cases} X_1 = X_3 = X_5 = \xi \text{ (Amarelo)} \\ X_2 = X_4 = X_7 = \xi^2 \text{ (Azul)} \\ X_6 = X_8 = X_9 = 1 \text{ (Verde)} \end{cases} \quad \text{ou} \quad \begin{cases} X_1 = X_3 = X_5 = X_8 = \xi \text{ (Amarelo)} \\ X_2 = X_4 = X_7 = \xi^2 \text{ (Azul)} \\ X_6 = X_9 = 1 \text{ (Verde)} \end{cases}$$

Desse modo, o problema da coloração do mapa da região nordeste do Brasil usando três cores apresenta as seguintes soluções:

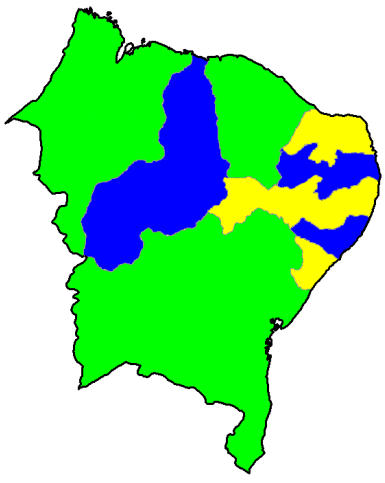

Figura 3.1: opção 1 - quando $X_8 = X_9$

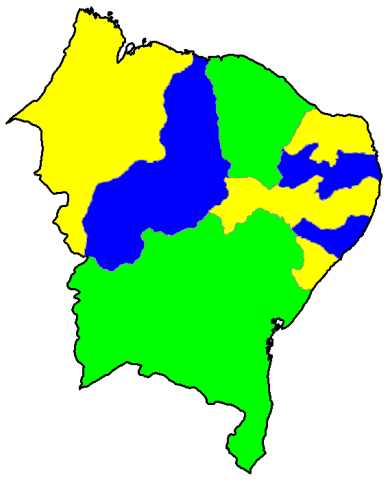

Figura 3.2: opção 2 - quando $X_8 \neq X_9$

3.8.2 A solução do Shidoku

Pelo que discutimos na subseção 2.2.3, para resolvermos o problema do Shidoku devemos determinar a solução do seguinte sistema de equações polinomiais

$$
\begin{array}{lll}
\bullet X_1^4 - 1 = 0 & \bullet X_2^3 + X_2X_4^2 + X_2^2X_4 + X_4^3 = 0 & \bullet X_7^3 + X_7X_{11}^2 + X_7^2X_{11} + X_{11}^3 = 0 \\
\bullet X_2^4 - 1 = 0 & \bullet X_2^3 + X_2X_5^2 + X_2^2X_5 + X_5^3 = 0 & \bullet X_7^3 + X_7X_{15}^2 + X_7^2X_{15} + X_{15}^3 = 0 \\
\bullet X_3^4 - 1 = 0 & \bullet X_2^3 + X_2X_6^2 + X_2^2X_6 + X_6^3 = 0 & \bullet X_8^3 + X_8X_{12}^2 + X_8^2X_{12} + X_{12}^3 = 0 \\
\bullet X_4 + i = 0 & \bullet X_2^3 + X_2X_{10}^2 + X_2^2X_{10} + X_{10}^3 = 0 & \bullet X_8^3 + X_8X_{16}^2 + X_8^2X_{16} + X_{16}^3 = 0 \\
\bullet X_5 + i = 0 & \bullet X_2^3 + X_2X_{14}^2 + X_2^2X_{14} + X_{14}^3 = 0 & \bullet X_9^3 + X_9X_{10}^2 + X_9^2X_{10} + X_{10}^3 = 0 \\
\bullet X_6^4 - 1 = 0 & \bullet X_3^3 + X_3X_4^2 + X_3^2X_4 + X_4^3 = 0 & \bullet X_9^3 + X_9X_{11}^2 + X_9^2X_{11} + X_{11}^3 = 0 \\
\bullet X_7 - i = 0 & \bullet X_3^3 + X_3X_7^2 + X_3^2X_7 + X_7^3 = 0 & \bullet X_9^3 + X_9X_{12}^2 + X_9^2X_{12} + X_{12}^3 = 0 \\
\bullet X_8^4 - 1 & \bullet X_3^3 + X_3X_8^2 + X_3^2X_8 + X_8^3 = 0 & \bullet X_9^3 + X_9X_{13}^2 + X_9^2X_{13} + X_{13}^3 = 0 \\
\bullet X_9^4 - 1 = 0 & \bullet X_3^3 + X_3X_{11}^2 + X_3^2X_{11} + X_{11}^3 = 0 & \bullet X_9^3 + X_9X_{14}^2 + X_9^2X_{14} + X_{14}^3 = 0 \\
\bullet X_{10} + 1 = 0 & \bullet X_3^3 + X_3X_{15}^2 + X_3^2X_{15} + X_{15}^3 = 0 & \bullet X_{10}^3 + X_{10}X_{11}^2 + X_{10}^2X_{11} + X_{11}^3 = 0 \\
\bullet X_{11}^4 - 1 = 0 & \bullet X_4^3 + X_4X_7^2 + X_4^2X_7 + X_7^3 = 0 & \bullet X_{10}^3 + X_{10}X_{12}^2 + X_{10}^2X_{12} + X_{12}^3 = 0 \\
\bullet X_{12} - 1 = 0 & \bullet X_4^3 + X_4X_8^2 + X_4^2X_8 + X_8^3 = 0 & \bullet X_{10}^3 + X_{10}X_{13}^2 + X_{10}^2X_{13} + X_{13}^3 = 0 \\
\bullet X_{13} - 1 = 0 & \bullet X_4^3 + X_4X_{12}^2 + X_4^2X_{12} + X_{12}^3 = 0 & \bullet X_{10}^3 + X_{10}X_{14}^2 + X_{10}^2X_{14} + X_{14}^3 = 0 \\
\bullet X_{14}^4 - 1 = 0 & \bullet X_4^3 + X_4X_{16}^2 + X_4^2X_{16} + X_{16}^3 = 0 & \bullet X_{11}^3 + X_{11}X_{12}^2 + X_{11}^2X_{12} + X_{12}^3 = 0 \\
\bullet X_{15}^4 - 1 = 0 & \bullet X_5^3 + X_5X_6^2 + X_5^2X_6 + X_6^3 = 0 & \bullet X_{11}^3 + X_{11}X_{15}^2 + X_{11}^2X_{15} + X_{15}^3 = 0 \\
\bullet X_{16}^4 - 1 = 0 & \bullet X_5^3 + X_5X_7^2 + X_5^2X_7 + X_7^3 = 0 & \bullet X_{11}^3 + X_{11}X_{16}^2 + X_{11}^2X_{16} + X_{16}^3 = 0 \\
\bullet X_1^3 + X_1X_2^2 + X_1^2X_2 + X_2^3 = 0 & \bullet X_5^3 + X_5X_8^2 + X_5^2X_8 + X_8^3 = 0 & \bullet X_{12}^3 + X_{12}X_{15}^2 + X_{12}^2X_{15} + X_{15}^3 = 0 \\
\bullet X_1^3 + X_1X_3^2 + X_1^2X_3 + X_3^3 = 0 & \bullet X_5^3 + X_5X_9^2 + X_5^2X_9 + X_9^3 = 0 & \bullet X_{12}^3 + X_{12}X_{16}^2 + X_{12}^2X_{16} + X_{16}^3 = 0 \\
\bullet X_1^3 + X_1X_4^2 + X_1^2X_4 + X_4^3 = 0 & \bullet X_5^3 + X_5X_{13}^2 + X_5^2X_{13} + X_{13}^3 = 0 & \bullet X_{13}^3 + X_{13}X_{14}^2 + X_{13}^2X_{14} + X_{14}^3 = 0 \\
\bullet X_1^3 + X_1X_5^2 + X_1^2X_5 + X_5^3 = 0 & \bullet X_6^3 + X_6X_7^2 + X_6^2X_7 + X_7^3 = 0 & \bullet X_{13}^3 + X_{13}X_{15}^2 + X_{13}^2X_{15} + X_{15}^3 = 0 \\
\bullet X_1^3 + X_1X_6^2 + X_1^2X_6 + X_6^3 = 0 & \bullet X_6^3 + X_6X_8^2 + X_6^2X_8 + X_8^3 = 0 & \bullet X_{13}^3 + X_{13}X_{16}^2 + X_{13}^2X_{16} + X_{16}^3 = 0 \\
\bullet X_1^3 + X_1X_9^2 + X_1^2X_9 + X_9^3 = 0 & \bullet X_6^3 + X_6X_{10}^2 + X_6^2X_{10} + X_{10}^3 = 0 & \bullet X_{14}^3 + X_{14}X_{15}^2 + X_{14}^2X_{15} + X_{15}^3 = 0 \\
\bullet X_1^3 + X_1X_{13}^2 + X_1^2X_{13} + X_{13}^3 = 0 & \bullet X_6^3 + X_6X_{14}^2 + X_6^2X_{14} + X_{14}^3 = 0 & \bullet X_{14}^3 + X_{14}X_{16}^2 + X_{14}^2X_{16} + X_{16}^3 = 0 \\
\bullet X_2^3 + X_2X_3^2 + X_2^2X_3 + X_3^3 = 0 & \bullet X_7^3 + X_7X_8^2 + X_7^2X_8 + X_8^3 = 0 & \bullet X_{15}^3 + X_{15}X_{16}^2 + X_{15}^2X_{16} + X_{16}^3 = 0
\end{array}
\tag{3.12}
$$

Assim, vamos então determinar uma base de Gröbner G para o ideal I cujos geradores são os polinômios que compõem o sistema de equações acima. Mais uma vez, devido ao número elevado de operações, utilizaremos o software *Macaulay* 2 para determinar tal base.

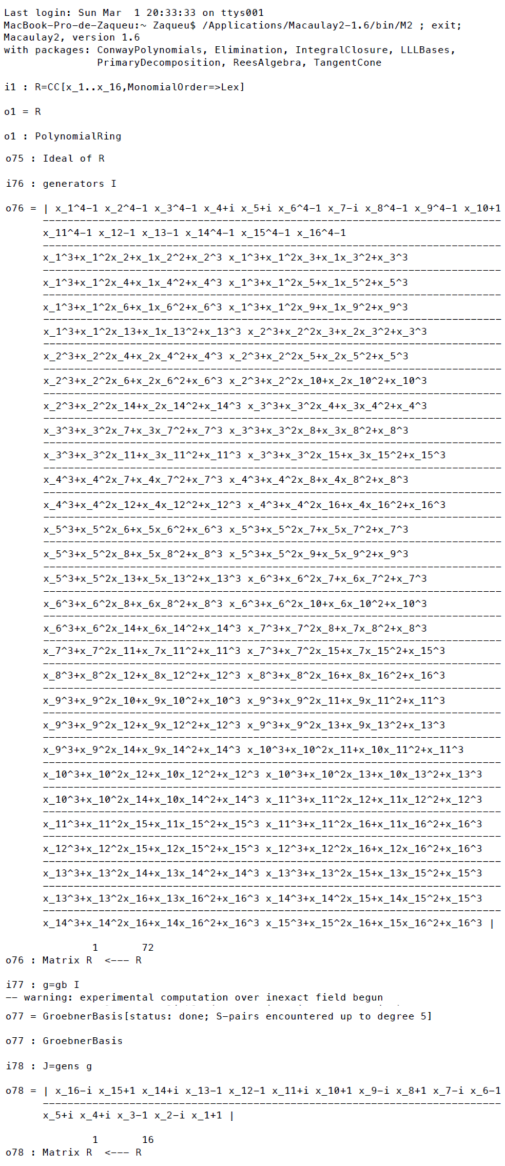

Figura 3.3: Cálculo da base de Gröbner usando o Macaulay 2

Desse modo, considerando a ordenação lexicográfica com $X_1 > X_2 > \ldots > X_{16}$, uma base de Gröbner para o ideal I é

$$\begin{aligned} G &= \{X_1+1, X_2-1, X_3-1, X_4+i, X_5+i, X_6-1, X_7-i, X_8+1, \\ &\quad X_9-i, X_{10}+1, X_{11}+i, X_{12}-1, X_{13}-1, X_{14}+i, X_{15}+1, X_{16}-i\} \end{aligned}$$

ou seja, o sistema 2.6 tem o mesmo conjunto solução do seguinte sistema:

$$\begin{cases} X_1 + 1 = 0 \\ X_2 - i = 0 \\ X_3 - 1 = 0 \\ X_4 + i = 0 \\ X_5 + i = 0 \\ X_6 - 1 = 0 \\ X_7 - i = 0 \\ X_8 + 1 = 0 \\ X_9 - i = 0 \\ X_{10} + 1 = 0 \\ X_{11} + i = 0 \\ X_{12} - 1 = 0 \\ X_{13} - 1 = 0 \\ X_{14} + i = 0 \\ X_{15} + 1 = 0 \\ X_{16} - i = 0 \end{cases} \tag{3.13}$$

Logo, concluímos que:

$$\begin{array}{llll} X_1 = -1 \textbf{ (cor 3)} & X_2 = i \textbf{ (cor 2)} & X_3 = 1 \textbf{ (cor 1)} & X_4 = -i \textbf{ (cor 4)} \\ X_5 = -i \textbf{ (cor 4)} & X_6 = 1 \textbf{ (cor 1)} & X_7 = i \textbf{ (cor 2)} & X_8 = -1 \textbf{ (cor 3)} \\ X_9 = i \textbf{ (cor 2)} & X_{10} = -1 \textbf{ (cor 3)} & X_{11} = -i \textbf{ (cor 4)} & X_{12} = 1 \textbf{ (cor 1)} \\ X_{13} = 1 \textbf{ (cor 1)} & X_{14} = -i \textbf{ (cor 4)} & X_{15} = -1 \textbf{ (cor 3)} & X_{16} = i \textbf{ (cor 2)} \end{array} \tag{3.14}$$

ou seja, a solução do Shidoku representado na figura 2.7 é

3	2	1	4
4	1	2	3
2	3	4	1
1	4	3	2

Figura 3.4: Solução do puzzle Shidoku

Note que, usando base de Gröbner, conseguimos transformar um sistema de 72 equações polinomiais, muitas delas não lineares e com mais de uma incógnita, em um

sistema equivalente (mesmo conjunto solução) formado por apenas 16 equações, todas elas lineares e com apenas uma incógnita. É claro que isso nem sempre ocorrerá, mas é fato que, em geral, o novo sistema é mais "simples"que o original.

Conclusão

Um elemento crucial no processo de construção do conhecimento é sem dúvida o questionamento. Sem este ingrediente nos tornamos meros receptores de informação, incapazes de exercer a crítica e de fornecer nossa própria contribuição ao desenvolvimento do conhecimento. No tocante ao papel do professor, acreditamos que o mesmo deve constantemente levantar indagações e incitar seus alunos a fazê-las.

Nessa perspectiva, nosso objetivo com este trabalho foi evidenciar como perguntas simples, que podem ocorrer de forma natural no âmbito do ensino básico, podem levar a um universo de conhecimento fascinante tanto do ponto de vista teórico quanto prático.

De fato, como se pode antever pelo próprio título do trabalho, esta dissertação foi essencialmente norteada pela seguinte questão:

(∗) *Como resolver um sistema de equações polinomiais em várias variáveis?*

Como é bem sabido, aprendemos a responder (∗) ainda no ensino básico em situações bastante particulares, como por exemplo, para um sistema constituído por uma única equação em uma única variável e com grau no máximo 2, ou para um sistema constituído unicamente por equações lineares. É no mínimo intrigante que, diante da amplitude de (∗), só consigamos respondê-la no ensino básico apenas para uma quantidade ínfima de situações. Dessa maneira, uma outra pergunta que surge é:

(∗∗) *O que impossibilita responder* (∗) *de maneira mais ampla no ensino básico?*

Para responder (∗∗) de forma honesta, faz-se necessário a pesquisa. O professor sem conhecimento de causa - e sem interesse de tê-lo - responderia de forma simplista e genérica, dizendo, por exemplo, que não se ensina porque é difícil. Todavia, o professor comprometido com seu ofício e com a verdade deve procurar conhecer o real

motivo estudando e capacitando-se. Acreditamos que a importância deste trabalho para o ensino básico esteja nesse sentido, pois ele revela, de maneira detalhada, quais são as dificuldades técnicas de se responder $(*)$ de forma mais ampla.

Como ficou claro pelo que foi apresentado no trabalho, os métodos utilizados para resolver $(*)$ envolvem conhecimento e maturidade matemática que talvez inviabilize o uso do mesmo no ensino básico. Entretanto, em nossa opinião, um bom professor de matemática deve ter três características fundamentais:

1. ter paixão pelo que faz - algo fundamental em qualquer profissão;

2. enxergar matemática sob um ponto de vista mais amplo ao que ensina - qualidade que só pode ser obtida através de um processo contínuo de capacitação;

3. ser hábil em instigar a curiosidade do aluno - fundamental para desenvolver o interesse deste pela matemática.

É interessante perceber que há uma relação recíproca entre professor capacitado e aluno curioso. Um professor capacitado gera alunos curiosos e alunos curiosos exigem professores capacitados.

Entendemos que, para que o ensino da matemática a nível básico se torne mais eficiente, é necessário que o professor se sinta mais bem preparado para desempenhar sua função. É importante que ele se sinta seguro para enfrentar o desafio de responder perguntas difíceis que eventualmente surgem de alunos curiosos, como: "professor, e se o sistema tiver equações não lineares?". Por isso, é fundamental que haja um investimento constante no processo de capacitação do professor e que seja dada a ele a oportunidade de desenvolver o interesse pela pesquisa acadêmica.

Apesar do uso de base de Gröbner como ferramenta de resolução de sistemas de equações polinomiais não se encaixar de modo satisfatório na grade curricular do ensino básico, entendemos que, principalmente com o auxílio de softwares como Macaulay, Maple ou até mesmo o WolfranAlpha, esta é sim uma alternativa que pode ser usada pelo professor para mostrar que existem modos de se obter a solução de determinados sistemas de equações não lineares. Mostrar ao aluno que a matemática não se restringe ao conteúdo normalmente abordado em sala de aula pode, entre outras coisas, despertar o interesse dele pelo universo acadêmico, além de levá-lo a perceber que a matemática é uma ciência que está em constante desenvolvimento; cada vez que se resolve um problema, automaticamente outro é criado. Esse processo é contínuo e, por isso mesmo, fascinante.

Finalmente, mostrar como uma ferramenta abstrata como a base de Gröbner pode ajudar a resolver problemas reais, como o problema da coloração de mapas, que faz parte da rotina escolar desde as séries iniciais e o problema de um puzzle tão conhecido como o Sudoku, nos ajuda a perceber a estreita relação que existe entre o abstrato e real; nos ajuda a perceber que a álgebra abstrata só existe por conta de desafios reais que precisam ser vencidos, ou seja, ela só existe por causa do mundo real e o mundo real só se desenvolve por causa da álgebra abstrata.

Desse modo, concluímos nosso trabalho agradecendo a Al Khowarisma, Bhascara Akaria, Scipione del Ferro, Nicoli Fontana, Girolamo Cardano, Lodovico Ferrari, Évariste Galois, Niels Henrik Abel, Wolfgang Gröbner, Bruno Buchberger, entre tantos outros, por terem aceitado o desafio de dedicar boa parte de suas vidas a desenvolver ferramentas e teorias que a princípio se mostravam tão abstratas e sem aplicação, mas que hoje nos ajudam a resolver tantos problemas no mundo real.

Referências Bibliográficas

[1] CLARK, A.; Elements of Abstract Algebra. Toronto: General Publishing Company, 1984. 205 p.

[2] COX, D.;LITTLE, J.;O'SHEA, D.; Ideals, Varieties, and Algorithms. New York: Springer, 2007. 3a edição. 539 p.

[3] DOMINGUES, H.H.;IEZZI, G.; Álgebra Moderna. São Paulo: Atual, 1982. 3a edição. 263 p.

[4] GARCIA, A.;LEQUAIN, Y.; Elementos de Álgebra. Rio de Janeiro: Associação IMPA, 2003. 326 p.

[5] GRAYSON, D. R.; STILLMAN, M.; *Macaulay 2*, a software system for research in algebraic geometry. Available at, http://www.math.uiuc.edu/Macaulay2/

[6] HEFEZ, A., Aritmética. Rio de Janeiro: SBM, 2013. Coleção Profmat. 1a edição. 338 p.

[7] Arnold, E.; Lucas, S.;Taalman, L. Gröbner Basis Representations of Sudoku. Virginia: Harrisonburg, 2009

[8] Escudeiro, Marcelo. Coloração de mapas e Bases de Grobner. Disponível em http://klein.sbm.org.br/wp-content/uploads/2012/11/coloracao-grobner1.pdf