



Universidade Federal da Paraíba  
Centro de Ciências Exatas e da Natureza  
Departamento de Matemática  
Mestrado Profissional em Matemática  
em Rede Nacional PROFMAT



# O método de circulantes, as fórmulas de Cardano e o teorema de Fermat para $n = 3^{\dagger}$

por

**Rômulo de Oliveira Lins Vieira de Melo**

sob orientação do

**Prof. Dr. Antônio de Andrade e Silva**

Dissertação apresentada ao Corpo Docente do Mestrado Profissional em Matemática em Rede Nacional PROFMAT-CCEN-UFPB, como requisito parcial para obtenção do título de Mestre em Matemática.

Agosto / 2017  
João Pessoa - PB

---

<sup>†</sup>O presente trabalho foi realizado com apoio da CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

Catálogo na publicação  
Seção de Catalogação e Classificação

|         |                                                                                                                                                                                                                                                                                                                                                                                                                   |
|---------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| M528m   | Melo, Rômulo de Oliveira Lins Vieira de.<br>O método de circulantes, as fórmulas de Cardano e o teorema de Fermat para $n = 3$<br>Rômulo de Oliveira Lins Vieira de Melo. –João Pessoa, 2017.<br>73 f. : il. -<br>Orientador: Dr. Antônio de Andrade e Silva.<br>Dissertação (Mestrado) - UFPB/CCEN/PROFMAT.<br>1. Matemática. 2. Matrizes circulantes. 3. Equações polinomiais. 4. Teorema de Fermat. I. Título. |
| UFPB/BC | CDU - 51(043)                                                                                                                                                                                                                                                                                                                                                                                                     |

# O método de circulantes, as fórmulas de Cardano e o teorema de Fermat para $n = 3$

por

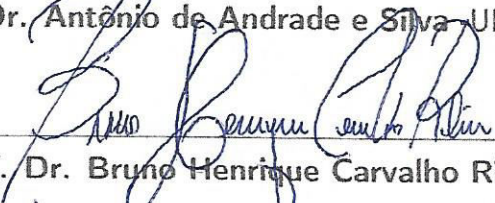
**Rômulo de Oliveira Lins Vieira de Melo**

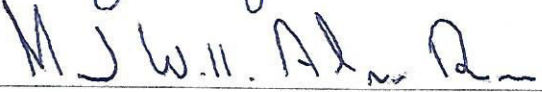
Dissertação apresentada ao Corpo Docente do Mestrado Profissional em Matemática em Rede Nacional PROFMAT CCEN-UEPB, como requisito parcial para obtenção do título de Mestre em Matemática.

Área de Concentração: Matemática

Aprovada por:

  
Prof. Dr. Antônio de Andrade e Silva - UEPB (Orientador)

  
Prof. Dr. Bruno Henrique Carvalho Ribeiro - UEPB

  
Prof. Dr. Manoel Wallace Alves Ramos - IFPB

Agosto / 2017

# Agradecimentos

Ao Supremo Arquiteto do Universo, por me proporcionar o dom da vida e a capacidade de raciocinar.

Aos meus pais, Samuel (in memorian) e Vitória, pela educação que me foi dada.

A minha esposa Carol pelo incentivo depositado em mim.

Ao Cátedra Antônio de Andrade e Silva pela orientação, dedicação e conhecimento a mim dedicados, além de me apresentar esse fascinante universo das circulantes.

Ao amigo Pedro Jerônimo Simões de Oliveira Júnior, que deu o pontapé inicial a esse trabalho e que abraçou esta causa, dedicando seu precioso tempo.

Ao Mestre Oriel de Carvalho Diniz, Tio Léo, meu primeiro mestre que me proporcionou o fascínio pelo aprender e ensinar.

A Liara Soares por me ajudar a exorcizar meus demônios.

Aos amigos da turma do PROFMAT, pelo companheirismo nessa caminhada ao longo dos últimos dois anos.

Aos valorosos professores do PROFMAT pelo conhecimento transmitido.

A todos que direta ou indiretamente colaboraram para a culminância desse trabalho.

A CAPES pelo incentivo financeiro proporcionado.

# Dedicatória

*Dedico este trabalho à minha mãe  
Vitória, por me fazer acreditar que a  
educação é um dos bens mais preciosos  
na formação de um cidadão, e ao meu  
filho Pedro, fonte de inspiração diária de  
minha busca em tornar-me um ser mais  
justo, fraterno e igualitário.*

# Resumo

No presente trabalho, princípios e teoremas associados aos números inteiros são retomados, bem como problemas de autovalores e autovetores, sendo ressaltada a matriz Hermitiana. Em seguida é dada ênfase às Matrizes Circulantes, através das quais verifica-se a associação a dois polinômios bem definidos: o representante e o característico. Posteriormente realiza-se um breve relato acerca da história das equações polinomiais, destacando-se as Fórmulas de Cardano-Tartaglia associadas às mesmas. Logo após é feita uma unificação no processo de resolução das equações polinomiais de graus menores do que o igual a 4, por meio das matrizes circulantes. O trabalho é finalizado, sendo provado o Teorema de Fermat para  $n = 3$ , recorrendo-se às Fórmulas de Cardano-Tartaglia.

**Palavras Chaves:** Matrizes Circulantes, Equações Polinomiais, Fórmulas de Cardano e Teorema de Fermat.

# Abstract

In this present work, principles and theorems associated to integers are returned, as well as eigenvalues and eigenvectors problems, highlighting a Hermitian matrix. Then it is emphasized to the Circulating Matrices, through which it is found the association to two well-defined polynomials: the representative and the characteristic. Later a brief account about the history of polynomial equations is made, drafting the Cardano-Tartaglia Formulas associated to them. Afterwards a unification is made in the resolution process of the polynomial equations of smaller degrees than the equal to 4, by means of the circulating matrices. The work is completed by proving a Fermat theorem for  $n = 3$ , using the Cardano-Tartaglia Formulas.

**Keywords:** Circulating Matrices, Polynomial Equations, Cardano Formulas and Fermat Theorem.

# Sumário

|                                                      |           |
|------------------------------------------------------|-----------|
| <b>Introdução</b>                                    | <b>xi</b> |
| <b>1 Resultados Básicos</b>                          | <b>1</b>  |
| 1.1 Princípios . . . . .                             | 1         |
| 1.2 Divisibilidade . . . . .                         | 3         |
| 1.3 Fatoração Única . . . . .                        | 5         |
| 1.4 Congruências . . . . .                           | 7         |
| <b>2 Problemas de Autovalores e Autovetores</b>      | <b>9</b>  |
| 2.1 Notações fundamentais . . . . .                  | 9         |
| 2.2 Problemas de autovalores e autovetores . . . . . | 13        |
| <b>3 Matrizes Circulantes</b>                        | <b>22</b> |
| 3.1 Matrizes de permutações . . . . .                | 22        |
| 3.2 Circulantes . . . . .                            | 26        |
| <b>4 Equações Polinomiais</b>                        | <b>33</b> |
| 4.1 História . . . . .                               | 33        |
| 4.2 Aplicação do Método Circulante . . . . .         | 41        |
| <b>5 Teorema de Fermat</b>                           | <b>55</b> |
| 5.1 Teorema de Fermat . . . . .                      | 55        |
| <b>Considerações Finais</b>                          | <b>61</b> |
| <b>Referências Bibliográficas</b>                    | <b>62</b> |

# Notações

## Notações Gerais

- $\mathbb{N}$  é o conjunto dos números naturais
- $\mathbb{Z}$  é o conjunto dos números inteiros
- $\mathbb{Z}_+$  é o conjunto dos números inteiros positivos
- $\mathbb{Z}^* = \mathbb{Z} - \{0\}$  é o conjunto dos números inteiros exceto o número 0
- $\mathbb{Q}$  é o conjunto dos números racionais
- $\mathbb{R}$  é o conjunto dos números reais
- $\mathbb{C}$  é o conjunto dos números complexos
- $\mathbb{Z}[i]$  é o anel dos inteiros Gaussianos
- $\mathcal{U}(\mathbb{Z}[i])$  é o grupo das unidades de  $\mathbb{Z}[i]$
- $R$  é um anel comutativo com identidade
- $M_2(R)$  é o anel das matrizes  $2 \times 2$  sobre  $R$
- $A$  representa uma matriz
- $N(A)$  é a norma de  $A$
- $A^t$  representa a matriz transposta de  $A$
- $\det(A)$  é o determinante de  $A$
- $\text{Tr}(A)$  é o traço de  $A$

- 
- $|$  é a operação divide
  - $\nmid$  é a operação não divide
  - $\equiv$  relação de congruência
  - mdc é o máximo divisor comum

# Introdução

Há algo fascinante sobre os procedimentos para a resolução de equações polinomiais com grau menor do que ou igual a 4. Por um lado, enquanto soluções gerais (usando radicais) são impossíveis para equações polinomiais de grau maior do que ou igual a cinco, elas têm sido encontradas para quadráticas, cúbicas e quárticas. Por outro lado, as soluções canônicas para a cúbica e a quártica são complicadas, e os métodos parecem ad hoc. Como é que alguém pode lembrar-se delas? Assim, é razoável pensar em obter um método simples (de fácil memorização) e unificado para todas as equações de graus até quatro.

As abordagens para a unificação têm sido algo tão longo quanto as próprias soluções. Cardano, em 1545, publicou soluções tanto para a equação cúbica quanto para equação quártica, atribuindo o primeiro mérito para Tartaglia e o segundo para Ferrari. Depois de várias tentativas subsequentes e sem êxito em resolver equações de grau maior do que ou igual a cinco, Lagrange, em 1770, apresentou uma análise detalhada para explicar por que os métodos de resolverem equações cúbicas e quárticas são bem sucedidos, utilizando-se de transformações lineares. A partir daquele momento até o presente, esforços têm sido apresentados para clarear as soluções de equações cúbicas e quárticas. Como uma aplicação provaremos o teorema de Fermat para  $n = 3$ .

No primeiro capítulo faremos uma breve revisão sobre números inteiros, os quais são prerequisites que alicerçam a teoria sobre a qual necessitamos no texto.

No segundo capítulo faremos uma breve revisão sobre polinômios, autovalores e autovetores, os quais são prerequisites que alicerçam a teoria sobre a qual necessitamos no texto.

No terceiro capítulo discorreremos sobre as matrizes circulantes e suas propriedades, dentre as quais poderemos destacar que cada matriz circulante possui dois polinômios naturalmente associados a ela, a saber, o seu representante e o seu polinômio característico. As “circulantes” foram introduzidas pela primeira vez em 1846 por Catalan.

No quarto capítulo começaremos com um pouco da história a respeito das equações de grau menor que ou igual a 4. Além disso, mostraremos como resolver equações polinomi-

---

ais usando o método circulante, chegando às fórmulas de Cardano-Tartaglia e concomitantemente com uma discussão sobre critérios para classificarmos as raízes das equações quadráticas, cúbicas e quárticas.

No quinto capítulo discutiremos a Conjectura de Fermat para  $n = 1, 2, 3$ , usando somente resultados elementares.

# Capítulo 1

## Resultados Básicos

Neste capítulo apresentaremos as terminologias e notações de determinados princípios, bem como, sobre divisibilidade, fatoração e congruência, que serão usadas em toda a dissertação. O leitor interessado em mais detalhes, assim como nas provas de certos lemas, corolários e teoremas, pode consultar [6, 10].

### 1.1 Princípios

Um dos axiomas fundamentais na teoria dos números inteiros e que será usado implicitamente muitas vezes, é o seguinte:

**Axioma 1.1 (Princípio da Boa Ordenação - PBO)** *Qualquer subconjunto não vazio dos inteiros positivos  $\mathbb{Z}_+$  possui um menor elemento.*

**Exemplo 1.2 (Princípio de Arquimedes)** *Sejam  $a, b \in \mathbb{Z}$ , com  $b > 0$ . Mostre que existe  $n \in \mathbb{N}$  tal que  $na \geq b$ .*

**Solução.** Suponhamos, por absurdo, que  $na < b$ , para todo  $n \in \mathbb{N}$ . Seja

$$S = \{b - na \in \mathbb{Z}_+ : n \in \mathbb{N}\},$$

então  $S \neq \emptyset$ , pois  $b - a \in S$ . Assim, pelo PBO, existe um  $s_0 = b - n_0a \in S$  tal que  $s_0 \leq x$ , para todo  $x \in S$ . Como  $b - (n_0 + 1)a \in S$ , pois  $S$  contém todos os inteiros desta forma, temos que

$$b - (n_0 + 1)a = s_0 - a < s_0,$$

o que é uma contradição. ■

**Proposição 1.3** Se  $x, y \in \mathbb{N}$  com  $x < y$ , então  $x + 1 \leq y$ .

**Prova.** Se  $x < y$ , então  $y - x > 0$ . Assim, basta provar que o conjunto

$$S = \{x \in \mathbb{N} : 0 < x < 1\}$$

é vazio. ■

**Exemplo 1.4 (Princípio do Máximo)** Qualquer subconjunto não vazio limitado superiormente de  $\mathbb{Z}$  possui um maior elemento.

**Solução.** Seja  $S$  um subconjunto não vazio de  $\mathbb{Z}$  limitado superiormente, isto é, existe um  $b \in \mathbb{Z}$  tal que  $s \leq b$ , para todo  $s \in S$ . Então o conjunto

$$T = \{n \in \mathbb{N} : s < n, \forall s \in S\}$$

é não vazio, pois se  $s \in S$ , então  $s \leq b < b + 1$ . Assim,  $b + 1 \in T$ . Logo, pelo PBO, existe  $t_0 \in T$  tal que  $t_0 \leq t$ , para todo  $t \in T$ . Note que existe um  $s_0 \in S$  tal que  $t_0 - 1 \leq s_0$ , ou seja,  $t_0 - 1 = s_0 \in S$ , pois  $s_0 < t_0$ . Por outro lado, como  $s < t_0$ , para todo  $s \in S$ , temos que  $s \leq t_0 - 1 = s_0$ . Portanto,  $s_0$  é o maior elemento de  $S$ .

**Teorema 1.5 (Princípio de Indução Finita)** Seja  $S$  um subconjunto de  $\mathbb{N}$  que goze das seguintes propriedades:

1.  $1 \in S$  (base de indução).
2. Se  $n \in S$ , então  $n + 1 \in S$  (PIF).

Então  $S = \mathbb{N}$ .

**Prova.** Basta provar que o conjunto

$$T = \{k \in \mathbb{N} : k \notin S\}$$

é vazio. ■

**Teorema 1.6 (Princípio de Indução Completo)** Seja  $S$  um subconjunto de  $\mathbb{N}$  que goze das seguintes propriedades:

1.  $1 \in S$  (base de indução).
2. Para cada  $n \in \mathbb{N}$ , se  $\{1, 2, \dots, n\} \subseteq S$ , então  $n + 1 \in S$ . (PIF)

Então  $S = \mathbb{N}$ .

**Prova.** Suponhamos, por absurdo, que  $S \neq \mathbb{N}$ . Então o conjunto

$$T = \mathbb{N} - S = \{n \in \mathbb{N} : n \notin S\} \neq \emptyset.$$

Assim, pelo PBO, existe um  $t_0 \in T$  tal que  $t_0 \leq t$ , para todo  $t \in T$ . Como  $1 \in S$  temos que  $t_0 \neq 1$  e  $t_0 > 1$ . Logo,  $0 < t_0 - 1 < t_0$ . Pela escolha de  $t_0$  temos que  $t_0 - 1 \notin T$  ou

$$1 \leq k \leq t_0 - 1, \forall k \in S, \text{ ou ainda, } \{1, 2, \dots, t_0 - 1\} \subseteq S.$$

Portanto, propriedade (2),

$$t_0 = (t_0 - 1) + 1 \in S \Rightarrow t_0 \in S \cap T = \emptyset,$$

o que é uma contradição. Consequentemente,  $S = \mathbb{N}$ . ■

## 1.2 Divisibilidade

Nesta seção provaremos um algoritmo, conhecido desde o ensino fundamental, que é o processo ordinário de dividir um inteiro positivo  $a$  por um inteiro positivo  $b$ , o qual fornece um quociente  $q$  e um resto  $r$ .

Sejam  $a, b \in \mathbb{Z}$ , com  $b \neq 0$ . Diremos que  $b$  divide  $a$  ou  $a$  é um múltiplo de  $b$  se existir um  $c \in \mathbb{Z}$  tal que

$$a = bc$$

e denotaremos por  $b \mid a$ . Caso contrário, diremos que  $b$  não divide  $a$  e denotaremos por  $b \nmid a$ . Diremos que  $a \in \mathbb{Z}$  é um número par se  $2 \mid a$  e ímpar se  $2 \nmid a$ .

**Teorema 1.7** *Sejam  $a, b, c, d \in \mathbb{Z}^*$ . Então as seguintes condições são satisfeitas:*

1. Se  $a \mid b$  e  $a \mid c$ , então  $a \mid (bx + cy)$ , para todos  $x, y \in \mathbb{Z}$ .
2.  $a \mid b \Rightarrow a \mid bc$ .
3.  $a \mid b$  e  $b \mid c \Rightarrow a \mid c$ .
4.  $a \mid b$ ,  $a > 0$  e  $b > 0 \Rightarrow a \leq b$ .
5.  $a \mid b$  e  $b \mid a \Rightarrow b = \pm a$ .
6. Seja  $c = ax + by$ , para alguns  $x, y \in \mathbb{Z}$ . Se  $d \mid b$ , mas  $d \nmid c$ , então  $d \nmid a$ .

**Prova.** Vamos provar apenas os itens (1) e (5): (1) Se  $a \mid b$  e  $a \mid c$ , então existem  $r, s \in \mathbb{Z}$  tais que  $b = ra$  e  $c = sa$ . Assim,

$$bx + cy = rax + say = (rx + sy)a.$$

Portanto,  $a \mid (bx + cy)$ , para todos  $x, y \in \mathbb{Z}$ .

(5) Se  $a \mid b$ , então existe um  $r \in \mathbb{Z}$  tal que  $b = ra$ . Assim,  $|b| = |r||a|$ , isto é,  $|a| \mid |b|$ . Logo, pelo item (4),  $|a| \leq |b|$ . De modo análogo,  $b \mid a$  implica que  $|b| \leq |a|$ . Portanto,  $|a| = |b|$  ou  $b = \pm a$ . ■

**Teorema 1.8 (Algoritmo da Divisão - AD)** *Sejam  $a, b \in \mathbb{Z}$ , com  $b > 0$ . Então existem únicos  $q, r \in \mathbb{Z}$  tais que*

$$a = qb + r, \text{ com } 0 \leq r < b.$$

**Prova.** Note que  $a < 0$  implica que  $-a > 0$ . O caso  $a = 0$ , existem  $q = r = 0$ . Assim, basta considerar o caso  $a > 0$ . O conjunto

$$S = \{n \in \mathbb{N} : nb > a\}$$

é, pelo Princípio de Arquimedes, não vazio. Logo, pelo PBO, existe um  $q + 1 \in S$  tal que  $qb \leq a < (q + 1)b$ . Pondo  $r = a - qb$ , obtemos  $r \geq 0$  e  $r < b$ . Portanto,

$$a = qb + r, \text{ com } 0 \leq r < b.$$

Para prova a unicidade, suponhamos que

$$a = q_1b + r_1, \text{ com } 0 \leq r_1 < b.$$

Como  $r \leq r_1$  ou  $r_1 \leq r$ , digamos  $r \leq r_1$ , temos que

$$0 \leq r_1 - r < b \text{ e } r_1 - r = (q - q_1)b.$$

Assim,  $b \mid (r_1 - r)$ . Se  $r < r_1$ , então, pelo item (4) do Teorema 1.7,  $b \leq r_1 - r$ , o que é uma contradição. Portanto,  $r = r_1$ . Consequentemente,  $(q - q_1)b = 0$  ou  $q = q_1$ . ■

Observe que o Algoritmo da Divisão é equivalente a:

$$\frac{a}{b} = q + \frac{r}{b}, \text{ com } 0 \leq \frac{r}{b} < 1.$$

Neste caso,  $q = \max\{n \in \mathbb{Z} : n \leq \frac{a}{b}\}$ .

Sejam  $a, b \in \mathbb{Z}$ , com  $a \neq 0$  ou  $b \neq 0$ . O *máximo divisor comum* de  $a$  e  $b$ , em símbolos  $\text{mdc}(a, b)$ , é um  $d \in \mathbb{Z}$  tal que

1.  $d > 0$ .
2.  $d \mid a$  e  $d \mid b$ .
3. Se  $c \mid a$  e  $c \mid b$ , então  $c \mid d$ .

Diremos que  $a$  e  $b$  são *relativamente primos* ou *primos entre si* quando  $\text{mdc}(a, b) = 1$ .

**Teorema 1.9** *Sejam  $a, b \in \mathbb{Z}$ , com  $a \neq 0$  ou  $b \neq 0$ . Então  $\text{mdc}(a, b)$  existe e é único.*

**Prova.** O conjunto

$$S = \{xa + by : x, y \in \mathbb{Z}\}$$

é não vazio, pois  $\pm a = (\pm 1) \cdot a + 0 \cdot b \in S$  e  $\pm b = 0 \cdot a + (\pm 1) \cdot b \in S$ . Assim, pelo PBO, existe um  $d \in S$  tal que  $d > 0$  e  $d \leq s$ , para todo  $s \in S$ . Como  $d \in S$  temos que existem  $x_0, y_0 \in \mathbb{Z}$  tais que

$$d = ax_0 + by_0.$$

Sendo assim, é fácil verificar que  $d = \text{mdc}(a, b)$ , por exemplo, pelo AD, existem únicos  $q, r \in \mathbb{Z}$  tais que

$$a = qd + r, \text{ com } 0 \leq r < d.$$

Logo,

$$r = a - qd = a(1 - qx_0) + b(-qy_0) \in S.$$

Portanto, pela minimalidade de  $d$ ,  $r = 0$  e  $d \mid a$ . ■

**Corolário 1.10** *Sejam  $a, b \in \mathbb{Z}$ , com  $a \neq 0$  ou  $b \neq 0$ . Se  $d = \text{mdc}(a, b)$ , então existem  $x, y \in \mathbb{Z}$  tais que*

$$d = ax + by.$$

**Lema 1.11 (Lema de Euclides)** *Sejam  $a, b, c \in \mathbb{Z}^*$ . Se  $c \mid ab$  e  $\text{mdc}(a, c) = 1$ , então  $c \mid b$ .*

## 1.3 Fatoração Única

Seja  $p \in \mathbb{Z}$ . Diremos que  $p$  é um *primo* se as seguintes condições são satisfeitas:

1.  $p > 1$ .
2. Se  $a \mid p$ , então  $a = \pm 1$  ou  $a = \pm p$ .

### 1.3. FATORAÇÃO ÚNICA

---

**Lema 1.12** *Sejam  $a, b \in \mathbb{Z}$  e  $p$  um primo. Se  $p \mid ab$ , então  $p \mid a$  ou  $p \mid b$ .*

**Prova.** Se  $p \nmid a$ , então  $\text{mdc}(a, p) = 1$ . Portanto, pelo Lema de Euclides,  $p \mid b$ .

**Lema 1.13** *Seja  $n \in \mathbb{Z}$ , com  $n > 1$ . Então  $n$  é um produto de primos.*

**Prova.** O conjunto

$$S = \{n \in \mathbb{Z} : n > 1 \text{ e } n \text{ não é um produto de primos}\}$$

é vazio. Caso contrário, pelo PBO, existe um  $d \in S$  tal que  $d > 1$  e  $d \leq s$ , para todo  $s \in S$ . Assim, pela condição (2), existe um  $a \in \mathbb{Z}$  tal que  $a \mid d$ , mas  $a \neq \pm 1$  e  $a \neq \pm d$ . Suponhamos que  $a > 0$ . Então existe  $b \in \mathbb{Z}$  tal que  $d = ab$ . Logo, pelo item (4) do Teorema 1.7,  $1 < a, b < d$ . Neste caso,  $a, b \notin S$ . Portanto,

$$a = p_1 \cdots p_m \text{ e } b = q_1 \cdots q_n$$

são produto de primos. Consequentemente,

$$d = p_1 \cdots p_m q_1 \cdots q_n$$

é um produto de primos, o que é uma contradição. ■

**Teorema 1.14 (Teorema Fundamental da Aritmética - TFA)** *Qualquer  $a \in \mathbb{Z} - \{-1, 0, 1\}$  pode ser escrito de modo único sob a forma*

$$a = up_1^{r_1} p_2^{r_2} \cdots p_n^{r_n},$$

com  $u = \pm 1, p_1 < p_2 < \cdots < p_n$  números primos e  $r_i \in \mathbb{Z}_+$ .

**Prova.** Note que  $a < -1$  implica que  $-a > 1$ . Então basta provar o caso  $a > 1$ . Já vimos, pelo Lema 1.13, que qualquer  $a > 1$ , pode ser fatorado em fatores primos. Para provar a unicidade. Consideremos o conjunto

$$S = \{a \in \mathbb{Z} : a > 1 \text{ e sua fatoração em fatores primos não é única}\}.$$

Assim, pelo PBO, existe um  $b \in S$  tal que  $b > 1$  e  $b \leq s$ , para todo  $s \in S$ . Neste caso,

$$b = p_1 \cdots p_m = q_1 \cdots q_n.$$

Mas, pelo Lema 1.12,  $p \mid q_j$ , para algum  $j = 1, \dots, n$ . Reordenando, se necessário, podemos supor que  $j = 1$  e  $p_1 = q_1$ . Logo,

$$c = p_2 \cdots p_m = q_2 \cdots q_n.$$

Como  $c \mid b$  temos, pelo item (4) do Teorema 1.7, que  $c < b$  e, pela minimalidade de  $b$ ,  $m - 1 = n - 1$ . Portanto, reordenando, se necessário,  $p_2 = q_2, \dots, p_n = q_n$ . Consequentemente,  $m = n$  e  $p_1 = q_1, p_2 = q_2, \dots, p_n = q_n$ . O resultado segue agrupando os fatores primos iguais. ■

**Lema 1.15** *Sejam  $a, b, n \in \mathbb{N}$  tais que  $\text{mdc}(a, b) = 1$ . Se  $ab = c^n$ , então existem  $u, v \in \mathbb{N}$  tais que  $a = u^n$  e  $b = v^n$ .*

**Prova.** Note, pelo TFA, que existe um primo  $p \in \mathbb{N}$  tal que  $p \mid a$ . Então o conjunto

$$S = \{m \in \mathbb{N} : p^m \mid a\}$$

é não vazio. Assim, pelo Princípio do Máximo, existe um  $r \in \mathbb{N}$  tal que  $p^r \mid a$  e  $p^{r+1} \nmid a$ . Como  $\text{mdc}(a, b) = 1$  temos que  $p^r \nmid b$ . Logo,  $p^r \mid ab$  e  $p^{r+1} \nmid ab$ . Pondo

$$c = p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m}$$

temos, pela hipótese, que  $r = nr_i$ , para algum  $i = 1, \dots, m$ , ou seja,  $n$  divide  $r$ . Portanto, existe um  $u \in \mathbb{N}$  tal que  $a = u^n$ . De modo análogo, existe um  $v \in \mathbb{N}$  tal que  $b = v^n$ . ■

## 1.4 Congruências

Sejam  $a, b \in \mathbb{Z}$  e  $n \in \mathbb{N}$ . Diremos que  $a$  e  $b$  são *congruentes módulo  $n$*  se  $n$  divide  $a - b$  e denotaremos por  $a \equiv b \pmod{n}$ . Caso contrário, diremos que  $a$  não é congruente a  $b$  módulo  $n$  e denotaremos por  $a \not\equiv b \pmod{n}$ .

**Teorema 1.16** *Sejam  $a, b, c, d, x \in \mathbb{Z}$  e  $n \in \mathbb{N}$ . Então as seguintes condições são satisfeitas:*

1.  $a \equiv a \pmod{n}$ .
2. Se  $a \equiv b \pmod{n}$ , então  $b \equiv a \pmod{n}$ .
3. Se  $a \equiv b \pmod{n}$  e  $b \equiv c \pmod{n}$ , então  $a \equiv c \pmod{n}$ .
4. Se  $a \equiv b \pmod{n}$  e  $c \equiv d \pmod{n}$ , então

$$a + c \equiv b + d \pmod{n} \text{ e } ac \equiv bd \pmod{n}.$$

5. Se  $a \equiv b \pmod{n}$ , então  $ax \equiv bx \pmod{n}$ .

6. Se  $a \equiv b \pmod{n}$  e  $a \equiv c \pmod{n}$ , então

$$ax \equiv c \pmod{n} \Leftrightarrow bx \equiv d \pmod{n}.$$

7. Se  $a \equiv b \pmod{n}$ , então  $a^k \equiv b^k \pmod{n}$ , para todo  $k \in \mathbb{N}$ .

Dado  $a \in \mathbb{Z}$  e  $n \in \mathbb{N}$ . Então existem únicos  $q, r \in \mathbb{Z}$  tais que  $a = qn + r$ , com  $0 \leq r < n$ . Logo,

$$a \equiv r \pmod{n}, \text{ onde } r \in \{0, 1, \dots, n-1\}.$$

**Lema 1.17** Seja  $a \in \mathbb{Z}$ . Então:

1.  $a \equiv r \pmod{2}$ , onde  $r \in \{0, 1\}$ .

2.  $a^2 \equiv r \pmod{4}$ , onde  $r \in \{0, 1\}$ .

3.  $a^2 \equiv r \pmod{8}$ , onde  $r \in \{0, 1, 4\}$ . Em particular, se  $a$  é ímpar, então

$$a^2 \equiv 1 \pmod{8}.$$

4.  $a \equiv r \pmod{3}$ , onde  $r \in \{-1, 0, 1\}$ .

5.  $a^3 \equiv a \pmod{3}$  se, e somente se,  $a^3 \equiv a \pmod{9}$ .

**Teorema 1.18** Se a equação Diofantina  $f(x, y, z) = 0$  possui uma solução  $a, b, c \in \mathbb{Z}$ , então

$$f(a, b, c) \equiv 0 \pmod{n},$$

para todo  $n \in \mathbb{N}$ .

**Prova.** Seja  $a, b, c \in \mathbb{Z}$  uma solução da equação  $f(x, y, z) = 0$ . Como  $n \mid 0$ , para todo  $n \in \mathbb{N}$ , temos que

$$f(a, b, c) \equiv 0 \pmod{n},$$

para todo  $n \in \mathbb{N}$ . ■

É muito importante ressaltar que o Teorema 1.18 nos fornece um **método** para decidir se uma dada equação Diofantina possui solução ou não: se existir um  $n \in \mathbb{N}$  tal que a equação

$$f(x, y, z) \equiv 0 \pmod{n},$$

não possui solução em  $\mathbb{Z}$ , então a equação Diofantina

$$f(x, y, z) = 0,$$

não possui solução em  $\mathbb{Z}$ .

## Capítulo 2

# Problemas de Autovalores e Autovetores

Neste capítulo apresentaremos as principais definições e resultados sobre matrizes, autovalores e autovetores que serão usadas em toda dissertação. O leitor interessado em mais detalhes, assim como nas provas de certos lemas, proposições, corolários e teoremas, pode consultar [3, 4, 11].

### 2.1 Notações fundamentais

Nesta seção introduziremos algumas definições, notações e terminologias. A notação

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}, \text{ com } i^2 = -1\}$$

representa o conjunto dos números complexos. Neste caso,

$$\begin{aligned}(a + bi) + (c + di) &= (a + c) + (b + d)i \\ (a + bi) \cdot (c + di) &= (ac - bd) + (bc + ad)i\end{aligned}$$

e  $\bar{z} = a - bi$  é o *conjugado complexo* de  $z = a + bi$ . O número real

$$|z|^2 = z \cdot \bar{z} = a^2 + b^2 \text{ ou } |z| = \sqrt{a^2 + b^2},$$

chama-se *norma* ou *módulo* de  $z$ . Assim, se  $z \neq 0$ , então

$$\frac{1}{z} = \frac{1}{|z|^2} \bar{z}.$$

As notações  $\Re(z) = \frac{1}{2}(z + \bar{z})$  e  $\Im(z) = \frac{1}{2}(z - \bar{z})$  representam a *parte real* e a *parte imaginária* de  $z$ . Note que

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

## 2.1. NOTAÇÕES FUNDAMENTAIS

---

Em tudo que segue, salvo menção explícita em contrário,  $F$  representa um subcorpo do corpo dos números complexos  $\mathbb{C}$ , em que seus elementos serão chamados de *escalares*.

Seja  $F[x]$  conjunto de todas as somas formais

$$f(x) = a_m x^m + \cdots + a_1 x + a_0 \in F[x].$$

A expressão

$$f(x) = a_m x^m + \cdots + a_1 x + a_0 \text{ ou } f(x) = a_0 + a_1 x + \cdots + a_m x^m$$

chama-se *polinômio* sobre  $F$ . Quando  $a_m \neq 0$ , diremos que  $f(x)$  possui *grau*  $m$  e será denotado por  $\partial(f) = m$ . Um escalar  $\lambda \in \mathbb{C}$  é uma *raiz* de  $f(x)$  se  $f(\lambda) = 0$ . Neste caso, é bem conhecido que  $x - \lambda$  divide  $f(x)$ , ou seja, existe  $g(x) \in F[x]$  tal que

$$f(x) = (x - \lambda)g(x).$$

Assim, indutivamente, obtemos

$$f(x) = a(x - \lambda_1) \cdots (x - \lambda_m),$$

onde  $a \in F$  e  $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ , são as raízes de  $f(x)$ . Note que se  $\beta \in \mathbb{C} - \mathbb{R}$  é uma raiz complexa de  $f(x) \in \mathbb{R}[x]$ , então

$$a_0 + a_1 \beta + \cdots + a_m \beta^m = 0.$$

Tomando a conjugação complexa desta equação, teremos

$$a_0 + a_1 \bar{\beta} + \cdots + a_m \bar{\beta}^m = 0.$$

Logo,  $\bar{\beta}$  também é uma raiz de  $f(x)$ , ou seja, as raízes complexas de  $f(x)$  ocorrem aos pares.

**Teorema 2.1 (Teorema Fundamental da Álgebra)** *Qualquer polinômio possui uma raiz em  $\mathbb{C}$ .*

**Proposição 2.2** *Seja*

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x].$$

*Então existem  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ , não necessariamente distintos, tais que:*

$$1. \ f(\lambda_i) = 0, \ i = 1, \dots, n.$$

2.  $f(x) = (x - \lambda_1) \cdots (x - \lambda_n)$ .

3.  $f(x)$  possui todas as raízes puramente imaginárias se, e somente se,  $\frac{f(ix)}{i^n}$  possui todas as raízes reais.

4. As relações de Girard:

$$f(x) = x^n - (\lambda_1 + \cdots + \lambda_n)x^{n-1} + \cdots + (-1)^n \lambda_1 \cdots \lambda_n.$$

Além disso, se  $f(\alpha) = 0$ , então  $\alpha = \lambda_i$ , para algum  $i = 1, \dots, n$ .

O item (4) da Proposição 2.2 produz uma conexão entre os coeficientes do polinômio  $f(x)$  e suas raízes.

A equação  $x^n - 1 = 0$  é de grande importância em diversos ramos da matemática e suas raízes chamam-se *raízes  $n$ -ésimas da unidade*. É comum escrever essas raízes como

$$\omega = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right) = e^{\frac{2\pi i}{n}} \in \mathbb{C},$$

sendo esta expressão conhecida como Fórmula de Euler. Neste caso, as raízes da equação  $x^n - 1 = 0$  são dadas pelas potências  $n$ -ésimas de  $\omega$  e, para cada  $k \in \mathbb{N}$ , com  $k = 0, 1, \dots, n-1$ , tem-se

$$\omega^k = e^{\frac{2\pi i k}{n}}.$$

Em particular,  $\omega^n = 1$ . Portanto,  $1, \omega, \omega^2, \dots, \omega^{n-1}$  são todas as raízes  $n$ -ésimas da unidade distintas, pois  $\omega^k \neq 1, k = 1, \dots, n-1$ .

**Lema 2.3** *Seja  $\omega \in \mathbb{C}$  uma raiz  $n$ -ésima da unidade. Então:*

1.  $\bar{\omega}^k = \omega^{-k} = \omega^{n-k}$  e  $\omega^{n+k} = \omega^k$ , para todo  $k \in \mathbb{N}$ .
2.  $1 + \omega^k + \omega^{2k} + \cdots + \omega^{k(n-1)} = 0$ , para todo  $k \in \mathbb{N}$ .

Denotaremos por

$$F^n = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in F\}$$

o conjunto de todas as  $n$ -uplas sobre  $F$  e por  $F^{m \times n}$  o conjunto de todas as matrizes de ordem  $mn$  sobre  $F$ . Em particular,  $F^{1 \times n}$  é o conjunto de todas as *matrizes linha* sobre  $F$  ou *vetores linhas* sobre  $F$ :

$$\mathbf{L}_i = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$$

e  $F^{m \times 1}$  é o conjunto de todas as *matrizes colunas* sobre  $F$  ou *vetores colunas* sobre  $F$ :

$$\mathbf{C}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Observe que todos esses conjuntos munidos com as operações de adição e multiplicação por escalar usuais de matrizes são espaços vetoriais sobre  $F$ . Além disso,  $F^{m \times n}$  munido com a multiplicação de matrizes é uma *álgebra linear* com identidade e não comutativa. Note que temos as identificações:  $F^n \longleftrightarrow F^{1 \times n}$  e  $F^m \longleftrightarrow F^{m \times 1}$ , por exemplo, a função  $T : F^n \rightarrow F^{1 \times n}$  definida como

$$T(a_1, \dots, a_n) = \mathbf{L} = (a_1 \ \cdots \ a_n)$$

é bijetora. Portanto, salvo menção explícita em contrário, não faremos distinção entre o vetor linha  $\mathbf{x} \in F^n$  e o vetor coluna  $\mathbf{X} \in F^n$ . A notação  $\mathbf{A}^* = \bar{\mathbf{A}}^t$  significa a *transposta conjugada* da matriz  $\mathbf{A} = (a_{ij})$ , em que  $\bar{\mathbf{A}}$  é a matriz conjugada de  $\mathbf{A}$ , isto é,  $\bar{\mathbf{A}} = (\bar{a}_{ij})$ .

Sejam  $\mathbf{X}, \mathbf{Y} \in F^n$ . Definimos o *produto escalar* (Hermitiano) sobre  $F^n$  como

$$\langle \mathbf{X}, \mathbf{Y} \rangle = x_1 \bar{y}_1 + \cdots + x_n \bar{y}_n = \mathbf{Y}^* \mathbf{X}.$$

Quando  $F \subseteq \mathbb{R}$ , obtemos o produto escalar usual

$$\langle \mathbf{X}, \mathbf{Y} \rangle = x_1 y_1 + \cdots + x_n y_n = \mathbf{Y}^t \mathbf{X}.$$

**Proposição 2.4** *Sejam  $\mathbf{A} \in F^{m \times n}$ ,  $\mathbf{X} \in F^n$  e  $\mathbf{Y} \in F^m$ . Então*

$$\langle \mathbf{A}\mathbf{X}, \mathbf{Y} \rangle = \langle \mathbf{X}, \mathbf{A}^* \mathbf{Y} \rangle.$$

Em tudo que segue o espaço vetorial  $F^n$  está munido com o produto escalar definido acima. Seja  $\mathbf{X} \in F^n$ . A *norma* de  $\mathbf{X}$  é definida como

$$\|\mathbf{X}\| = \sqrt{\langle \mathbf{X}, \mathbf{X} \rangle}.$$

Sejam  $\mathbf{X}, \mathbf{Y} \in F^n$ . Diremos  $\mathbf{X}$  e  $\mathbf{Y}$  são *ortogonais* se

$$\langle \mathbf{X}, \mathbf{Y} \rangle = 0.$$

É bem conhecido, via Processo de Ortogonalização de Gram-Schmidt, que a partir de qualquer base de  $F^n$  podemos obter uma base ortonormal de  $F^n$ , ou seja, uma base

$$\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$$

tal que

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j, \end{cases}$$

em que o símbolo  $\delta_{ij}$  chama-se *delta de Kronecker*.

## 2.2 Problemas de autovalores e autovetores

O problema de autovalores e autovetores é basicamente o seguinte: dado uma matriz  $A \in F^{n \times n}$ , devemos encontrar um vetor  $X \in F^n$ , com  $X \neq O$ , e um escalar  $\lambda \in F$  tal que

$$AX = \lambda X \Leftrightarrow (\lambda I - A)X = O. \quad (2.1)$$

O escalar  $\lambda$  chama-se *autovalor* de  $A$  e o vetor  $X \neq O$  chama-se *autovetor* de  $A$  associado a  $\lambda$ . Assim, para determinarmos um autovetor de  $A$ , primeiro devemos encontrar um escalar  $\lambda$  para o qual a matriz  $\lambda I - A$  seja singular e em seguida resolver o sistema homogêneo (2.1). Portanto, os autovalores são escalares  $\lambda$  para os quais

$$p(\lambda) = \det(\lambda I - A) = 0.$$

O polinômio  $p(x)$  chama-se *polinômio característico* de  $A$  e equação  $p(x) = 0$  chama-se *equação característica* de  $A$ .

**Proposição 2.5** *Sejam  $A \in F^{n \times n}$  e  $\lambda \in \mathbb{C}$ . Então as seguintes condições são equivalentes:*

1.  $\lambda I - A$  é singular;
2.  $\det(\lambda I - A) = 0$ ;
3. Existe um  $X$ , com  $X \neq O$ , tal que  $AX = \lambda X$ ;
4.  $\lambda$  é um autovalor de  $A$ .

É pertinente ressaltar que se  $\lambda$  é real, então existe um vetor real  $X$ . Por outro lado, se  $\lambda$  é complexo, o que pode ocorrer mesmo que a matriz  $A$  seja real, então o vetor  $X$  pode ser complexo.

**Exemplo 2.6** *Seja  $T = (t_{ij}) \in F^{n \times n}$  uma matriz triangular superior. Então os elementos da diagonal  $t_{11}, \dots, t_{nn}$  são os autovalores de  $T$ .*

**Solução.** Devemos encontrar o polinômio  $p(x) = \det(xI - T)$ . Para isto, vamos usar indução sobre  $n$ . Se  $n = 1$  ou  $n = 2$ , nada há para ser provado. Suponhamos que o resultado seja válido para todo  $k$ , com  $1 \leq k \leq n - 1$  e  $n > 2$ . Assim, pela Fórmula de Laplace em relação à primeira coluna de  $T$ ,

$$p(x) = (x - t_{11}) \det(T_{11}) = (x - t_{11})(x - t_{22}) \cdots (x - t_{nn}),$$

pois  $T_{11}$  é uma matriz triangular superior de ordem  $n - 1$ . Portanto,  $t_{11}, \dots, t_{nn}$  são os autovalores de  $T$ . ■

**Lema 2.7** Seja  $A \in F^{n \times n}$ .

1.  $A$  e  $P^{-1}AP$  possuem o mesmo polinômio característico, para toda matriz invertível  $P \in F^{n \times n}$ .
2. Se  $AX = \lambda X$ , com  $X \neq O$ , e  $P$  é uma matriz invertível cuja  $j$ -ésima coluna é  $X$ , então  $\lambda E_j$  é a  $j$ -ésima coluna de  $P^{-1}AP$ .

**Prova.** Vamos provar apenas o item (2). Suponhamos que

$$P = (C_1 \ \cdots \ X \ \cdots \ C_n), \text{ com } C_j = X.$$

Então

$$\begin{aligned} I &= P^{-1}P = P^{-1} (C_1 \ \cdots \ X \ \cdots \ C_n) \\ &= (P^{-1}C_1 \ \cdots \ P^{-1}X \ \cdots \ P^{-1}C_n). \end{aligned}$$

Assim,  $P^{-1}X = E_j$  é a  $j$ -ésima coluna de  $I$ . Logo,

$$\begin{aligned} P^{-1}AP &= P^{-1} (AC_1 \ \cdots \ AX \ \cdots \ AC_n) \\ &= P^{-1} (AC_1 \ \cdots \ \lambda X \ \cdots \ AC_n) \\ &= (P^{-1}AC_1 \ \cdots \ \lambda P^{-1}X \ \cdots \ P^{-1}AC_n) \\ &= (P^{-1}AC_1 \ \cdots \ \lambda E_j \ \cdots \ P^{-1}AC_n). \end{aligned}$$

Portanto,  $\lambda E_j$  é a  $j$ -ésima coluna de  $P^{-1}AP$ . ■

Seja  $A \in F^{n \times n}$  uma matriz. Diremos que  $A$  é uma *matriz diagonalizável* se existir uma matriz invertível  $P \in F^{n \times n}$  tal que

$$P^{-1}AP = D,$$

em que

$$D = (\lambda_1 E_1 \ \cdots \ \lambda_n E_n)$$

é uma matriz diagonal e  $\lambda_i$  os autovalores de  $A$ .

**Teorema 2.8** Sejam  $A \in F^{n \times n}$  e  $\lambda_1, \dots, \lambda_n$  autovalores de  $A$ , com autovetores associados  $X_1, \dots, X_n$ . Então  $A$  é diagonalizável se, e somente se,  $\{X_1, \dots, X_n\}$  é uma base de  $F^n$ .

**Proposição 2.9** Sejam  $A \in F^{n \times n}$  e  $AX = \lambda X$ , onde  $X \in F^n$ , com  $X \neq O$ .

1.  $\lambda + k$ ,  $k\lambda$  e  $\lambda^m$  são autovalores de  $A + kI$ ,  $kA$  e  $A^m$ , respectivamente.

2. A matriz

$$q(\mathbf{A}) = c_0 \mathbf{I} + c_1 \mathbf{A} + \cdots + c_n \mathbf{A}^n \in F^{n \times n}$$

possui autovalor

$$q(\lambda) = c_0 + c_1 \lambda + \cdots + c_n \lambda^n$$

e  $\mathbf{X}$  é o autovetor de  $q(\mathbf{A})$  associado ao autovalor  $q(\lambda)$ .

3. Se  $\mathbf{A}$  é diagonalizável, então  $q(\mathbf{A})$  também o é.

**Prova.** (1) Note que

$$\mathbf{A}^2 \mathbf{X} = \mathbf{A}(\mathbf{A} \mathbf{X}) = \lambda(\mathbf{A} \mathbf{X}) = \lambda^2 \mathbf{X}.$$

Assim, indutivamente,  $\mathbf{A}^m \mathbf{X} = \lambda^m \mathbf{X}$ , para todo  $m \in \mathbb{N}$ , ou seja,  $\lambda^m$  é o autovalor de  $\mathbf{A}^m$ , para todo  $m \in \mathbb{N}$ .

(2) Como

$$\begin{aligned} q(\mathbf{A}) \mathbf{X} &= (c_0 \mathbf{I} + c_1 \mathbf{A} + \cdots + c_n \mathbf{A}^n) \mathbf{X} \\ &= c_0 \mathbf{X} + c_1 \mathbf{A} \mathbf{X} + \cdots + c_n \mathbf{A}^n \mathbf{X} \\ &= c_0 \mathbf{X} + c_1 \lambda \mathbf{X} + \cdots + c_n \lambda^n \mathbf{X} \\ &= (c_0 + c_1 \lambda + \cdots + c_n \lambda^n) \mathbf{X} \\ &= q(\lambda) \mathbf{X} \end{aligned}$$

temos que  $q(\lambda)$  é o autovalor de  $q(\mathbf{A})$  associado ao autovetor  $\mathbf{X}$ .

(3) Suponhamos que  $\mathbf{A}$  seja uma matriz diagonalizável. Então existe uma matriz invertível  $\mathbf{P} \in F^{n \times n}$  tal que

$$\mathbf{P}^{-1} \mathbf{A} \mathbf{P} = \mathbf{D}.$$

Assim, indutivamente,

$$\mathbf{P}^{-1} \mathbf{A}^m \mathbf{P} = \mathbf{D}^m.$$

para todo  $m \in \mathbb{N}$ . Logo,

$$\mathbf{P}^{-1} q(\mathbf{A}) \mathbf{P} = q(\mathbf{D}) = \begin{pmatrix} q(\lambda_1) \mathbf{E}_1 & \cdots & q(\lambda_n) \mathbf{E}_n \end{pmatrix}.$$

Portanto,  $q(\mathbf{A})$  é uma matriz diagonalizável. ■

É importante ressaltar que a recíproca dos itens (2) e (3) da Proposição 2.9 é falsa. Por exemplo, consideremos a matriz

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in F^{2 \times 2}$$

e o polinômio  $q(x) = x^2$ . Assim,  $q(\mathbf{A}) = \mathbf{O}$ . Portanto, qualquer  $\mathbf{X} \in F^2$  é um autovetor de  $q(\mathbf{A})$ . Mas,  $\mathbf{X} = (1, 0)^t \in F^2$  é o único autovetor de  $\mathbf{A}$  associado ao autovalor 0.

**Teorema 2.10** Sejam  $\mathbf{A} \in F^{n \times n}$  e

$$p(x) = \det(x\mathbf{I} - \mathbf{A}) = (x - \lambda_1) \cdots (x - \lambda_n) \in F[x]$$

o polinômio característico de  $\mathbf{A}$ . Se

$$q(x) = c_0 + c_1x + \cdots + c_nx^n \text{ e } q(\mathbf{A}) = c_0\mathbf{I} + c_1\mathbf{A} + \cdots + c_n\mathbf{A}^n,$$

então

$$g(x) = (x - q(\lambda_1)) \cdots (x - q(\lambda_n)) \in \mathbb{C}[x]$$

é o polinômio característico de  $q(\mathbf{A})$ . Em particular, se  $\mu$  é um autovalor de  $q(\mathbf{A})$ , então  $\mu = q(\lambda)$ , para algum autovalor  $\lambda$  de  $\mathbf{A}$ .

**Prova.** Está além dos objetivos desta dissertação. ■

Seja  $\mathbf{A} \in F^{n \times n}$  uma matriz. Diremos que  $\mathbf{A}$  é uma *matriz unitária* se

$$\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A} = \mathbf{I}, \text{ com } \mathbf{A}^* = \mathbf{A}^{-1}.$$

Quando  $F \subseteq \mathbb{R}$ , diremos que  $\mathbf{A}$  é uma *matriz ortogonal*. Por exemplo, a matriz

$$\mathbf{U} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$$

é unitária, mas não é ortogonal. Observe que se  $\mathbf{U}$  e  $\mathbf{V}$  são matrizes unitárias, então

$$\mathbf{U}\mathbf{V}, \mathbf{U}^t, \mathbf{U}^* \text{ e } \bar{\mathbf{U}}$$

também o são. Além disso, se  $\mathbf{U}$  é uma matriz unitária e

$$\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$$

é a base de  $\mathbb{C}^n$  formada pelas colunas de  $\mathbf{U}$ , então

$$\begin{aligned} \mathbf{I} = \mathbf{U}^*\mathbf{U} &= \mathbf{U}^* \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_j & \cdots & \mathbf{u}_n \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{U}^*\mathbf{u}_1 & \cdots & \mathbf{U}^*\mathbf{u}_j & \cdots & \mathbf{U}^*\mathbf{u}_n \end{pmatrix} \\ &= (\langle \mathbf{u}_i, \mathbf{u}_j \rangle). \end{aligned}$$

Assim,  $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$ . Portanto,  $\mathcal{B}$  é uma base ortonormal de  $\mathbb{C}^n$ .

**Teorema 2.11** Seja  $\mathbf{U} \in F^{n \times n}$ .

1.  $U$  é unitária se, e somente se,

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \langle U\mathbf{X}, U\mathbf{Y} \rangle, \quad \forall \mathbf{X}, \mathbf{Y} \in F^n.$$

2. Se  $\lambda$  é um autovalor de  $U$ , então  $|\lambda| = 1$  e  $|\det(U)| = 1$ .

**Prova.** Vamos provar apenas o item (1). Suponhamos que  $U$  seja unitária. Então

$$\langle U\mathbf{X}, U\mathbf{Y} \rangle = \langle \mathbf{X}, U^*U\mathbf{Y} \rangle = \langle \mathbf{X}, \mathbf{Y} \rangle.$$

Reciprocamente, pondo  $\mathbf{X} = \mathbf{E}_i$  e  $\mathbf{Y} = \mathbf{E}_j$ , obtemos

$$\delta_{ij} = \langle \mathbf{E}_i, \mathbf{E}_j \rangle = \langle \mathbf{E}_i, U^*U\mathbf{E}_j \rangle.$$

Já vimos que  $U^*U\mathbf{E}_j$  é a  $j$ -ésima coluna do produto  $U^*U$ . Portanto,

$$U^*U = (\delta_{ij}) = \mathbf{I},$$

é o resultado desejado. ■

**Teorema 2.12 (Teorema de Schur)** *Seja  $\mathbf{A} \in F^{n \times n}$ . Então existe uma matriz unitária  $U \in F^{n \times n}$  e uma matriz triangular superior  $\mathbf{T} \in F^{n \times n}$  tal que  $U^*\mathbf{A}U = \mathbf{T}$ .*

**Prova.** Vamos usar indução sobre  $n$ . Se  $n = 1$  ou  $n = 2$ , nada há para ser provado. Suponhamos que o resultado seja válido para todo  $k$ , com  $1 \leq k \leq n - 1$  e  $n > 2$ . Seja  $\lambda \in \mathbb{C}$  um autovalor de  $\mathbf{A}$  e  $\mathbf{X} \in \mathbb{C}^n$  um autovetor associado, com  $\|\mathbf{X}\| = 1$ . Então estendendo o conjunto  $\{\mathbf{X}\}$  para uma base de  $\mathbb{C}^n$  e via o Processo de Gram-Schmidt, obtemos uma base ortonormal  $\mathcal{B} = \{\mathbf{X}, \mathbf{Y}_2, \dots, \mathbf{Y}_n\}$  de  $\mathbb{C}^n$ . Pondo

$$\mathbf{P} = \begin{pmatrix} \mathbf{X} & \mathbf{Y}_2 & \cdots & \mathbf{Y}_n \end{pmatrix},$$

é fácil verificar que  $\mathbf{P}^*\mathbf{P} = \mathbf{I}$  e

$$\mathbf{P}^*\mathbf{A}\mathbf{P} = \left( \begin{array}{c|c} \lambda & \mathbf{C} \\ \hline \mathbf{O} & \mathbf{B} \end{array} \right),$$

onde  $\mathbf{B} \in F^{(n-1) \times (n-1)}$ . Assim, existe uma matriz unitária  $\mathbf{V}$  e uma matriz triangular superior  $\mathbf{T}_1$  tal que  $\mathbf{V}^*\mathbf{A}\mathbf{V} = \mathbf{T}_1$ . Consideremos

$$\mathbf{Q} = \left( \begin{array}{c|c} 1 & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{V} \end{array} \right).$$

Então  $Q^*Q = I$  e  $U = PQ$  é unitária. Logo,

$$U^*AU = \left( \begin{array}{c|c} \lambda & CV^* \\ \hline O & T_1 \end{array} \right) = T,$$

que é o resultado desejado. ■

Seja  $A \in F^{n \times n}$  uma matriz. Diremos que  $A$  é uma *matriz Hermitiana* se  $A^* = A$  e é uma *matriz anti-Hermitiana* se  $A^* = -A$ . Quando  $F \subseteq \mathbb{R}$ , diremos que  $A$  é uma *matriz simétrica e antissimétrica*, respectivamente. Por exemplo, se

$$A = \begin{pmatrix} 2 & i \\ i & -1 \end{pmatrix}, B = \begin{pmatrix} 2 & i \\ -i & -1 \end{pmatrix}, C = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \in \mathbb{C}^{2 \times 2},$$

então  $A^t = A$ , mas  $A^* \neq A$ ;  $B^* = B$ , mas  $B^t \neq B$  e  $C^t = C^* = C$ . Note que se  $A$  é uma matriz anti-Hermitiana, então  $iA$  é uma matriz Hermitiana. Além disso, os elementos da diagonal de qualquer matriz Hermitiana  $A = (a_{ij})$  são reais, pois  $\bar{a}_{ii} = a_{ii}$ . Portanto, os elementos da diagonal de qualquer matriz anti-Hermitiana são imaginários puros.

**Teorema 2.13** *Seja  $A \in F^{n \times n}$  uma matriz Hermitiana. Então os autovalores de  $A$  são reais e existe uma matriz unitária  $U \in F^{n \times n}$  tal que  $U^*AU = D$ . Além disso, existe uma base ortonormal de autovetores de  $\mathbb{C}^n$ .*

**Prova.** Pelo Teorema de Schur, existe uma matriz unitária  $U \in F^{n \times n}$  tal que  $U^*AU = T$ , com  $T$  uma matriz triangular superior. Então

$$T^* = (U^*AU)^* = U^*A^*U = U^*AU = T.$$

Assim,  $T = (t_{ij})$  é uma matriz Hermitiana e  $t_{ij} = 0$ , quando  $i \neq j$ . Portanto,  $T$  é diagonal e real. ■

Seja  $A \in F^{n \times n}$  uma matriz. Diremos que  $A$  é uma *matriz normal* se

$$AA^* = A^*A.$$

Neste caso, qualquer matriz Hermitiana, anti-Hermitiana ou unitária é normal. No entanto, a matriz

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$$

é normal, mas não é Hermitiana, anti-Hermitiana e nem unitária. Observe que se  $\mathbf{A}$  é uma matriz normal e  $\mathbf{U}$  é uma matriz unitária, então as matrizes

$$k\mathbf{A}, \mathbf{A}^*, \mathbf{U}^*\mathbf{A}\mathbf{U} \text{ e } \mathbf{A}^m$$

são normais.

**Lema 2.14** *Seja  $\mathbf{T} \in F^{n \times n}$  uma matriz triangular superior e normal. Então  $\mathbf{T}$  é uma matriz diagonal.*

**Prova.** Sejam

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{pmatrix} \text{ e } \mathbf{T} = \begin{pmatrix} \bar{t}_{11} & 0 & \cdots & 0 \\ \bar{t}_{12} & \bar{t}_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{t}_{1n} & \bar{t}_{2n} & \cdots & \bar{t}_{nn} \end{pmatrix}.$$

Então  $\mathbf{T}\mathbf{T}^* = \mathbf{T}^*\mathbf{T}$  implica que

$$\sum_{k=1}^n t_{ik} \bar{t}_{jk} = \sum_{k=1}^n \bar{t}_{ki} t_{kj}, \quad i, j = 1, \dots, n.$$

Lembre-se que  $t_{rs} = 0$ , quando  $r > s$ . Assim, para  $i = j = 1$ , obtemos

$$|t_{11}|^2 = |t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2.$$

Logo,  $t_{1s} = 0$ , quando  $s > 1$ . Para  $i = j = 2$ , teremos

$$|t_{22}|^2 = |t_{22}|^2 + |t_{23}|^2 + \cdots + |t_{2n}|^2.$$

Donde,  $t_{2s} = 0$ , quando  $s > 2$ . Continuando deste modo, concluímos que  $t_{rs} = 0$ , quando  $r < s$ . Portanto,  $\mathbf{T}$  é uma matriz diagonal. ■

**Teorema 2.15** *Seja  $\mathbf{A} \in F^{n \times n}$ . Então  $\mathbf{A}$  é normal se, e somente se, existir uma matriz unitária  $\mathbf{U} \in F^{n \times n}$  tal que  $\mathbf{U}^*\mathbf{A}\mathbf{U}$  é diagonal.*

**Prova.** Suponhamos que  $\mathbf{A}$  seja normal. Então  $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$ . Por outro lado, pelo Teorema de Schur, existe uma matriz unitária  $\mathbf{U}$  tal que  $\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{T}$ , com  $\mathbf{T}$  uma matriz triangular superior. Assim,

$$\mathbf{T}\mathbf{T}^* = (\mathbf{U}^*\mathbf{A}\mathbf{U})(\mathbf{U}^*\mathbf{A}\mathbf{U})^* = \mathbf{U}^*\mathbf{A}\mathbf{A}^*\mathbf{U} = \mathbf{U}^*\mathbf{A}^*\mathbf{A}\mathbf{U} = \mathbf{T}^*\mathbf{T}.$$

Logo,  $T$  é normal. Portanto, pelo Lema 2.14,  $T$  é diagonal.

Reciprocamente, suponhamos que exista uma matriz unitária  $U$  tal que  $U^*AU = D$  seja diagonal. Então

$$DD^* = (U^*AU)(U^*AU)^* = U^*AA^*U$$

e

$$D^*D = (U^*AU)^*(U^*AU) = U^*A^*AU.$$

Como  $DD^* = D^*D$  temos que  $AA^* = A^*A$ . Portanto,  $A$  é uma matriz normal. ■

**Corolário 2.16** *Seja  $A \in F^{n \times n}$ . Se  $A$  é normal, então  $q(A)$  é normal, para todo  $q(x) \in F[x]$ .*

**Lema 2.17** *Seja  $A \in F^{n \times n}$  uma matriz normal.*

1.  *$A$  é Hermitiana se, e somente se, todos os seus autovalores são reais.*
2.  *$A$  é anti-Hermitiana se, e somente se, todos os seus autovalores são imaginários puros.*
3.  *$A$  é unitária se, e somente se, todos os seus autovalores são de módulo unitário.*
4. *Se  $A \in \mathbb{R}^{n \times n}$  e  $A^t \neq A$ , então  $A$  possui pelo menos um par de autovalores complexos.*

**Prova.** Vamos provar apenas os itens (1) e (4): (1) Suponhamos que todos os autovalores de  $A$  sejam reais. Então existe uma matriz unitária  $U$  tal que  $U^*AU = D$ , com  $D$  diagonal e real. Assim,

$$U^*AU = D = D^* = (U^*AU)^* = U^*A^*U.$$

Logo,  $A^* = A$ . Portanto,  $A$  é Hermitiana.

(4) Suponhamos, por absurdo, que todos os autovalores de  $A$  sejam reais. Então, pelo item (1),  $A$  seria Hermitiana. Assim,  $A^t = A$ , o que é uma contradição. Portanto,  $A$  possui pelo menos um par de autovalores complexos (conjugado). ■

**Teorema 2.18 (Teorema de Cayley-Hamilton)** *Sejam  $A \in F^{n \times n}$  e  $p(x) \in F[x]$  o polinômio característico de  $A$ , então  $p(A) = 0$ .*

**Prova.** Pelo Teorema de Schur, existe uma matriz unitária  $\mathbf{U}$  tal que  $\mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{T}$ , com  $\mathbf{T}$  uma matriz triangular superior, digamos

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{pmatrix}.$$

Assim,  $p(x) = (x - t_{11})(x - t_{22}) \cdots (x - t_{nn})$ . Logo,

$$p(\mathbf{T}) = (\mathbf{T} - t_{11}\mathbf{I})(\mathbf{T} - t_{22}\mathbf{I}) \cdots (\mathbf{T} - t_{nn}\mathbf{I}).$$

Observe, depois de alguns cálculos, que

$$(\mathbf{T} - t_{11}\mathbf{I})(\mathbf{T} - t_{22}\mathbf{I}) = \begin{pmatrix} 0 & 0 & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

é uma matriz triangular superior com as duas primeiras colunas nulas. Aplicando sucessivamente o produto, obtemos

$$p(\mathbf{T}) = (\mathbf{T} - t_{11}\mathbf{I})(\mathbf{T} - t_{22}\mathbf{I}) \cdots (\mathbf{T} - t_{nn}\mathbf{I}) = \mathbf{0}.$$

Portanto,

$$p(\mathbf{A}) = p(\mathbf{U} \mathbf{T} \mathbf{U}^*) = \mathbf{U} p(\mathbf{T}) \mathbf{U}^* = \mathbf{0},$$

que é o resultado desejado. ■

## Capítulo 3

# Matrizes Circulantes

Matrizes circulantes são predominantes em muitos ramos da matemática, tais como: teoria da codificação, geometria de polígonos circulares, grafos, entre outros. Essas matrizes aparecem naturalmente em áreas da matemática, onde as raízes da unidade desempenham um importante papel, e aqui apresentaremos algumas das razões dessa importância. No entanto, muitos fatos sobre essas matrizes podem ser provadas usando apenas álgebra linear básica. Isso torna a área bastante acessível para alunos de graduação a procura de problemas em pesquisa e/ou professores de matemática em busca de temas com interesse exclusivo para apresentar aos seus alunos. O leitor interessado em mais detalhes, assim como nas provas de certos lemas e proposições, pode consultar [3, 5, 7].

### 3.1 Matrizes de permutações

Nesta seção apresentaremos as principais propriedades de um caso particular de matrizes circulantes, a *matriz de permutação*

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{E}_n & \mathbf{E}_1 & \cdots & \mathbf{E}_{n-1} \end{pmatrix}.$$

Note que  $\mathbf{W}$  é obtida da matriz identidade  $\mathbf{I}$  deslocando uma coluna para a direita ou, equivalentemente, é obtida deslocando ciclicamente os elementos da primeira linha uma unidade para a direita.

Observe que

$$\mathbf{W}\mathbf{E}_j = \begin{pmatrix} \mathbf{E}_2^t \mathbf{E}_j & \mathbf{E}_3^t \mathbf{E}_j & \cdots & \mathbf{E}_1^t \mathbf{E}_j \end{pmatrix}^t = \mathbf{E}_{j-1}, \quad j = 2, \dots, n,$$

e  $\mathbf{W}\mathbf{E}_1 = \mathbf{E}_n$ . Por exemplo,

$$\mathbf{W}\mathbf{E}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \mathbf{E}_1.$$

Assim, a matriz  $\mathbf{W}$  é equivalente a permutação  $\sigma$ , em que  $S = \{1, 2, \dots, n\}$  e  $\sigma : S \rightarrow S$  definida como

$$\sigma(i) = \begin{cases} i-1, & \text{se } i = 2, \dots, n \\ n, & \text{se } i = 1 \end{cases}$$

é claramente bijetora. Neste caso,

$$\mathbf{W}\mathbf{E}_j = \mathbf{E}_{\sigma(j)}.$$

Logo,

$$\begin{aligned} \mathbf{W}^2 &= \begin{pmatrix} \mathbf{W}\mathbf{E}_n & \mathbf{W}\mathbf{E}_1 & \cdots & \mathbf{W}\mathbf{E}_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{E}_{\sigma(n)} & \mathbf{E}_{\sigma(1)} & \cdots & \mathbf{E}_{\sigma(n-1)} \end{pmatrix}. \end{aligned}$$

Em geral,

$$\mathbf{W}^k = \begin{pmatrix} \mathbf{E}_{\sigma^k(n)} & \mathbf{E}_{\sigma^k(1)} & \cdots & \mathbf{E}_{\sigma^k(n-1)} \end{pmatrix}$$

e  $\mathbf{W}^n = \mathbf{I} = \mathbf{W}^0$ . Além disso,  $\mathbf{W}^{-k} = \mathbf{W}^{n-k}$  e  $\mathbf{W}^{n+k} = \mathbf{W}^k$ , para todo  $k \in \mathbb{N}$ , pois  $\mathbf{W}$  é invertível. Um dos principais resultados sobre a matriz  $\mathbf{W}$  é o seguinte:

$$\mathbf{W}\mathbf{W}^t = (\mathbf{E}_i^t \mathbf{E}_j) = (\delta_{ij}) = \mathbf{I} = \mathbf{W}^t \mathbf{W}.$$

Portanto,  $\mathbf{W}$  é uma matriz ortogonal (unitária). Consequentemente, pelo Teorema 2.15,  $\mathbf{W}$  é unitariamente diagonalizável.

Note que se

$$\mathbf{X} = x_1 \mathbf{E}_1 + \cdots + x_n \mathbf{E}_n \neq \mathbf{O}$$

é um vetor em  $F^n$ , então

$$\mathbf{W}\mathbf{X} = x_1 \mathbf{E}_n + \cdots + x_n \mathbf{E}_{n-1}$$

é um vetor obtido de  $\mathbf{X}$  deslocando ciclicamente uma coordenada para a direita. Assim, é fácil verificar que o conjunto

$$\mathcal{B} = \{\mathbf{X}, \mathbf{W}\mathbf{X}, \dots, \mathbf{W}^{n-1}\mathbf{X}\}$$

### 3.1. MATRIZES DE PERMUTAÇÕES

é uma base de  $F^n$ . Além disso, o conjunto

$$G = \{\mathbf{W}^k : k \in \mathbb{Z}\} = \{\mathbf{I}, \mathbf{W}, \dots, \mathbf{W}^{n-1}\}$$

munido com a multiplicação usual de matrizes constitui o que chamamos de *grupo cíclico* gerado por  $\mathbf{W}$  e de ordem  $n$ .

O restante desta seção será dedicado à construção de uma matriz unitária que diagonalize a matriz  $\mathbf{W}$ .

A matriz  $\mathbf{V}_n = (a_{ij}) \in F^{n \times n}$ , com  $a_{ij} = x_i^{j-1}$ :

$$\mathbf{V}_n = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$$

chama-se *matriz de Vandermonde* de ordem  $n$ .

**Lema 3.1 (Determinante de Vandermonde)** *Seja  $\mathbf{V}_n \in F^{n \times n}$  a matriz de Vandermonde. Então*

$$\det(\mathbf{V}_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n (x_j - x_i).$$

**Prova.** Vamos usar indução sobre  $n$ . Se  $n = 2$ , então é claro que

$$\det(\mathbf{V}_2) = x_2 - x_1.$$

Suponhamos que o resultado seja válido para todo  $k$ , com  $1 \leq k \leq n-1$  e  $n > 2$ . Então a sequência de operações elementares de colunas

$$\mathbf{C}_{j+1} \rightarrow \mathbf{C}_{j+1} - x_1 \mathbf{C}_j, \quad j = n-1, \dots, 2, 1,$$

implicam que  $\mathbf{V}_n$  é equivalente por colunas à matriz

$$\mathbf{V}'_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{pmatrix}.$$

Assim, pela Fórmula de Laplace em relação à primeira linha, obtemos

$$\det(\mathbf{V}'_n) = (x_2 - x_1) \cdots (x_n - x_1) \det \begin{pmatrix} 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} \end{pmatrix}.$$

Portanto,

$$\det(\mathbf{V}_n) = \prod_{1 \leq j \leq n} (x_j - x_1) \det(\mathbf{V}_{n-1}) = \prod_{1 \leq i < j \leq n} (x_j - x_i),$$

que é o resultado desejado. ■

Note, para  $x = x_1$ , que  $f(x) = \det(\mathbf{V}_n)$  é um polinômio de grau  $n - 1$ , cujas raízes são:  $x_2, \dots, x_n$ . Consequentemente,

$$f(x) = a(-1)^{n-1} \prod_{j=2}^n (x - x_j),$$

com  $a = \det(\mathbf{V}_{n-1})$  o coeficiente de  $x^{n-1} = x_1^{n-1}$ , obtido pela Fórmula de Laplace em relação à primeira linha.

**Proposição 3.2** *O polinômio característico de  $\mathbf{W}$  é  $p(x) = x^n - 1$ .*

**Prova.** Pela Fórmula de Laplace em relação à última linha, obtemos

$$\begin{aligned} p(x) = \det(x\mathbf{I} - \mathbf{W}) &= \det \begin{pmatrix} x & -1 & 0 & \cdots & 0 \\ 0 & x & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ -1 & 0 & 0 & \cdots & x \end{pmatrix} \\ &= (-1)^{n+1}(-1)(-1)^{n-1} + (-1)^{2n}x \cdot x^{n-1} \\ &= x^n - 1, \end{aligned}$$

que é o resultado desejado. ■

Pela Proposição 2.2, os autovalores de  $\mathbf{W}$  são as raízes  $n$ -ésimas da unidade, a saber,

$$\omega^k = e^{\frac{2k\pi i}{n}}, \quad k = 0, 1, 2, \dots, n-1$$

Sejam  $\rho \in F$ , com  $\rho \neq 0$ , e

$$\mathbf{X} = \mathbf{E}_1 + \rho\mathbf{E}_2 + \cdots + \rho^{n-1}\mathbf{E}_n$$

um vetor qualquer em  $F^n$ , então é fácil verificar que  $\mathbf{W}\mathbf{X} = \rho\mathbf{X}$  se, e somente se,  $\rho^{n-1} = \rho^{-1}$ . Mas  $\rho^{n-1} = \rho^{-1}$  implica que  $\rho$  é uma raiz  $n$ -ésimas da unidade. Como as raízes  $n$ -ésimas da unidade são distintas temos que um conjunto completo de autovetores de  $\mathbf{W}$  é dado, pondo  $\rho = \omega^k$ , por

$$\mathbf{X}_{k+1} = \mathbf{E}_1 + \omega^k\mathbf{E}_2 + \cdots + \omega^{k(n-1)}\mathbf{E}_n \in \mathbb{C}^n, \quad k = 0, 1, \dots, n-1.$$

Neste caso,

$$\mathcal{B} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\} = \{\mathbf{X}_1, \mathbf{W}\mathbf{X}_1, \dots, \mathbf{W}^{n-1}\mathbf{X}_1\}$$

é uma base de autovetores para  $\mathbb{C}^n$ . Sejam

$$\mathbf{P} = [\mathbf{I}]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_n \end{pmatrix}$$

a matriz mudança de base e  $\mathbf{D}$  a matriz diagonal cujos elementos diagonais sejam os autovalores associados. Então

$$\mathbf{P}^{-1}\mathbf{W}\mathbf{P} = \mathbf{D}. \quad (3.1)$$

Depois de alguns cálculos, obtemos

$$\mathbf{P}\mathbf{P}^* = \mathbf{P}^*\mathbf{P} = n\mathbf{I}, \langle \mathbf{X}_i, \mathbf{X}_j \rangle = \sum_{k=0}^{n-1} \bar{\omega}^{kj} \omega^{ki} = \sum_{k=0}^{n-1} \omega^{k(i-j)} = n\delta_{ij}.$$

Assim,

$$\mathbf{F} = \frac{1}{\sqrt{n}}\mathbf{P} = \left( \frac{1}{\sqrt{n}} \omega^{(i-1)(j-1)} \right), \quad i, j = 1, \dots, n,$$

é uma matriz unitária (matriz de Fourier). Portanto, a equação (3.1) pode ser escrita sob a forma

$$\mathbf{F}^*\mathbf{W}\mathbf{F} = \mathbf{D} \quad \text{ou} \quad \mathbf{W} = \mathbf{F}\mathbf{D}\mathbf{F}^*$$

que é uma diagonalização unitária de  $\mathbf{W}$ . Observe que

$$\mathbf{W} = \left( \begin{array}{c|c} \mathbf{O} & \mathbf{I}_{n-1} \\ \hline \mathbf{I}_1 & \mathbf{O} \end{array} \right) \quad \text{e} \quad \mathbf{W}^k = \left( \begin{array}{c|c} \mathbf{O} & \mathbf{I}_{n-k} \\ \hline \mathbf{I}_k & \mathbf{O} \end{array} \right).$$

## 3.2 Circulantes

Veremos nesta seção que as matrizes circulantes desempenham um papel importante em várias aplicações, dentre as quais a que mais nos chama a atenção é a forma simples como se calcula seus autovalores e autovetores utilizando raízes  $n$ -ésimas da unidade. É pertinente ressaltar que matrizes circulantes foram introduzidas pela primeira vez em 1846 por Catalan.

Um *circulante* de ordem  $n$  é qualquer polinômio na matriz de permutação  $\mathbf{W}$ . Por exemplo, se

$$q(x) = 1 + 2x + 0x^2 \in F[x]$$

Então

$$\mathbf{C} = q(\mathbf{W}) = \mathbf{I} + 2\mathbf{W} + 0\mathbf{W}^2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$$

é uma matriz obtida deslocando ciclicamente os elementos da primeira linha uma unidade para a direita. Isto motiva a seguinte definição. A matriz

$$C = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{pmatrix}.$$

é uma *matriz circulante*, isto é,

$$\mathbf{C} = q(\mathbf{W}) = c_0\mathbf{I} + c_1\mathbf{W} + \cdots + c_{n-1}\mathbf{W}^{n-1},$$

com

$$q(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1} \in F[x]$$

o *polinômio representante* ou *polinômio associado* de  $\mathbf{C}$ . Reciprocamente, qualquer vetor

$$\mathbf{X} = c_0\mathbf{E}_1 + c_1\mathbf{E}_2 + \cdots + c_{n-1}\mathbf{E}_n \in F^n,$$

representa uma matriz circulante. Portanto, se  $C_n(F)$  é o conjunto de todas as matrizes circulantes, então  $C_n(F)$  munido com as operações de adição e multiplicação por escalar é um espaço vetorial sobre  $F$ . Como  $f(x) = g(x)(x^n - 1) + r(x)$ , com  $\partial(r(x)) < n$ , temos que

$$C_n(F) = \{c_0\mathbf{I} + c_1\mathbf{W} + \cdots + c_{n-1}\mathbf{W}^{n-1} : c_i \in F\}.$$

É fácil verificar que  $\mathbf{C} \in C_n(F)$  se, e somente se,  $\mathbf{C} = \mathbf{WCW}^t$ . Dados

$$f(x) = \sum_{i=0}^m a_i, g(x) = \sum_{j=0}^n b_j \in F[x].$$

Já vimos que

$$f(x)g(x) = \sum_{l=0}^{m+n} \left( \sum_{k=0}^l a_k b_{l-k} \right) x^l \quad (3.2)$$

Como

$$(ax^m)(bx^n) = a(x^m(bx^n)) = a(b(x^m x^n)) = abx^{m+n}$$

temos, pelas Leis Distributivas em  $F[x]$ , que o produto (3.2) pode também ser escrito sob a forma

$$f(x)g(x) = \sum_{i=0}^m \sum_{j=0}^n a_i b_j x^i x^j.$$

Assim, se  $\mathbf{C}_1, \mathbf{C}_2 \in C_n(F)$ , então

$$\mathbf{C}_1 \mathbf{C}_2 = q_1(\mathbf{W})q_2(\mathbf{W}) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_i b_j \mathbf{W}^{i+j} \in C_n(F).$$

Portanto,  $C_n(F)$  munido com esse produto é uma *álgebra linear* com identidade e comutativa sobre  $F$  de dimensão  $n$ , pois

$$\mathcal{E} = \{\mathbf{I}, \mathbf{W}, \dots, \mathbf{W}^{n-1}\}$$

é a base canônica.

**Proposição 3.3** *Seja  $\mathbf{C} \in C_n(F)$ , com polinômio representante  $q(x) = c_0 + c_1x + \dots + c_{n-1}x^{n-1} \in F[x]$ . Então  $\mathbf{C}$  é uma matriz normal e o polinômio característico de  $\mathbf{C}$  é*

$$p(x) = (x - q(1))(x - q(\omega)) \cdots (x - q(\omega^{n-1})).$$

Além disso,

1.  $\det(\mathbf{C}) = \prod_{k=0}^{n-1} q(\omega^k).$
2.  $\text{tr}(\mathbf{C}) = \sum_{k=0}^{n-1} q(\omega^k) = nc_0.$

**Prova.** Como

$$\mathbf{C} = q(\mathbf{W}) = c_0\mathbf{I} + c_1\mathbf{W} + \dots + c_{n-1}\mathbf{W}^{n-1}$$

temos que o resultado segue do Teorema 2.10. ■

**Exemplo 3.4** *Dado*

$$\mathbf{C} = \begin{pmatrix} -2 & \beta & \bar{\beta} \\ \bar{\beta} & -2 & \beta \\ \beta & \bar{\beta} & -2 \end{pmatrix} \in C_3(\mathbb{C}), \text{ com } \beta = \cos\left(\frac{2\pi}{9}\right) + i\sin\left(\frac{2\pi}{9}\right).$$

*Calcule os polinômios representante e característico de  $\mathbf{C}$ .*

**Solução.** Note que

$$q(x) = -2 + \beta x + \bar{\beta}x^2 \in \mathbb{C}[x]$$

é o polinômio representante de  $\mathbf{C}$ . Como

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

temos que  $x^3 - 1 = 0$  é a equação característica de  $\mathbf{W}$ . Assim,

$$1, \omega \text{ e } \omega^2 = \bar{\omega}, \text{ com } \omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i,$$

são os autovalores de  $\mathbf{W}$ . Logo,

$$\begin{aligned} q(1) &= -2 + 2 \cos\left(\frac{2\pi}{9}\right) \\ q(\omega) &= -2 + 2 \cos\left(\frac{8\pi}{9}\right) \\ q(\omega^2) &= -2 + 2 \cos\left(\frac{14\pi}{9}\right) \end{aligned}$$

são os autovalores reais de  $\mathbf{C}$  e

$$\mathbf{x}_1 = (1, 1, 1), \mathbf{x}_2 = (1, \omega, \omega^2) \text{ e } \mathbf{x}_3 = (1, \omega^2, \omega)$$

são os autovetores de  $\mathbf{C}$  associados a esses autovalores. Pondo

$$\mathbf{F} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix} \text{ e } \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}.$$

Depois de alguns cálculos, obtemos

$$\mathbf{F}^* \mathbf{W} \mathbf{F} = \mathbf{D} \text{ e } \mathbf{W} \mathbf{W}^* = \mathbf{I}.$$

Observe que o polinômio característico de  $\mathbf{C}$  é

$$p(x) = \det(x\mathbf{I} - \mathbf{C}) = x^3 + 6x^2 + 9x + 3.$$

Além disso,  $\mathbf{C} = q(\mathbf{W})$ . ■

**Proposição 3.5** *Sejam  $\mathbf{C} \in C_n(F)$  e  $q(x) \in F[x]$  o polinômio representante de  $\mathbf{C}$ . Então as seguintes condições são equivalentes:*

1.  $\mathbf{C}$  é uma matriz singular;

2.  $q(\omega^k) = 0$ , para algum  $k = 0, \dots, n-1$ ;

3.  $\text{mdc}(q(x), x^n - 1) \neq 1$ .

Finalizaremos esta seção com uma breve conexão entre a álgebra linear  $C_n(F)$  e as raízes de polinômios na álgebra linear  $F[x]$ .

Sejam  $W$  um subespaço de  $F[x]$  e  $f(x), g(x) \in F[x]$ . Diremos que  $f(x)$  é *equivalente* a  $g(x)$  módulo  $W$  se  $g(x) - f(x) \in W$  e denotaremos por  $g(x) \sim f(x)$ . É fácil verificar que “ $\sim$ ” é uma relação de equivalência sobre  $F[x]$ . O conjunto

$$\overline{f(x)} = f(x) + W = \{g(x) \in F[x] : g(x) \sim f(x)\}$$

é a *classe de equivalência* determinada por  $f(x)$ . Logo, o conjunto

$$\frac{F[x]}{W} = \{\overline{f(x)} : f(x) \in F[x]\}$$

com as operações induzidas por  $F[x]$  é uma álgebra linear sobre  $F$ . Neste caso, a função  $\psi : F[x] \rightarrow F^{n \times n}$  definida como

$$\psi(f(x)) = f(\mathbf{W})$$

preserva as operações das álgebras lineares. Então, pelo Teorema de Caylay-Hamilton,

$$\ker \psi = (x^n - 1) = \{(x^n - 1)f(x) : f(x) \in F[x]\}$$

e

$$\frac{F[x]}{(x^n - 1)} \simeq F[\mathbf{W}] = C_n(F).$$

Note que

$$\{\bar{1}, \bar{x}, \dots, \bar{x}^{n-1}\}$$

é a base canônica de

$$\frac{F[x]}{(x^n - 1)}.$$

Seja

$$F_n[x] = \{f(x) \in F[x] : \partial(f(x)) < n\},$$

com  $\partial(0) = -\infty$ . Então é fácil verificar que  $F_n[x]$  é um subespaço de  $F[x]$  com dimensão  $n$ , pois

$$\{1, x, \dots, x^{n-1}\}$$

é a base canônica de  $F_n[x]$ . Estamos prontos para resolver o seguinte problema: dados  $c_0, \dots, c_{n-1} \in F$  distintos, desejamos encontrar polinômios  $p_i(x) \in F_n[x]$ , com  $i = 0, 1, \dots, n-1$ , tais que

$$p_i(c_j) = \delta_{ij}, j = 0, \dots, n-1.$$

Seja

$$p_i(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in F[x]$$

um tal polinômio. Então

$$\begin{cases} a_0 + a_1c_0 + \dots + a_{n-1}c_0^{n-1} &= 0 \\ \vdots &\vdots \\ a_0 + a_1c_i + \dots + a_{n-1}c_i^{n-1} &= 1 \\ \vdots &\vdots \\ a_0 + a_1c_{n-1} + \dots + a_{n-1}c_{n-1}^{n-1} &= 0 \end{cases}$$

é um sistema linear não homogêneo e na forma matricial:

$$\mathbf{V}_n \mathbf{X}_i = \mathbf{E}_i.$$

Como os  $c_0, \dots, c_{n-1}$  são distintos temos que  $\det(\mathbf{V}_n) \neq 0$ . Assim, pela Regra de Cramer e depois de alguns cálculos, obtemos

$$\begin{aligned} p_i(x) &= \frac{(x-c_0)\dots(x-c_{i-1})(x-c_{i+1})\dots(x-c_{n-1})}{(c_i-c_0)\dots(c_i-c_{i-1})(c_i-c_{i+1})\dots(c_i-c_{n-1})} \\ &= \prod_{j \neq i, j=0}^{n-1} \left( \frac{x-c_j}{c_i-c_j} \right). \end{aligned}$$

Observe que o conjunto

$$\{p_0(x), \dots, p_{n-1}(x)\}$$

é linearmente independente, pois se

$$d_0p_0(x) + \dots + d_{n-1}p_{n-1}(x) = 0,$$

então, avaliando em  $x = c_i$ , obtemos

$$d_i p_i(c_i) = 0 \Rightarrow d_i = 0, i = 0, \dots, n-1.$$

Logo,

$$\{p_0(x), \dots, p_{n-1}(x)\}$$

é também uma base de  $F_n[x]$ . Portanto, dados escalares quaisquer  $\lambda_0, \dots, \lambda_{n-1} \in F$ , existe um único polinômio  $q(x) \in F_n[x]$  tal que

$$q(c_i) = \lambda_i, i = 0, \dots, n-1,$$

a saber,

$$q(x) = \sum_{i=0}^{n-1} \lambda_i p_i(x). \quad (3.3)$$

A expressão (3.3) chama-se *Fórmula de Interpolação de Lagrange*.

Já vimos que para qualquer matriz  $\mathbf{C} \in C_n(F)$  existem dois polinômios bem determinados associados a  $\mathbf{C}$ , a saber, o seu polinômio representante  $q(x)$  e o seu polinômio característico  $p(x)$ . Reciprocamente, dado  $f(x) \in F[x]$  de grau  $n$ . Então existem  $\lambda_0, \dots, \lambda_{n-1} \in \mathbb{C}$  tais que

$$f(\lambda_i) = 0, i = 0, \dots, n-1.$$

Por outro lado, pela Fórmula de Interpolação de Lagrange, existe um único polinômio  $q(x) \in F_n[x]$  tal que

$$q(\omega^i) = \lambda_i, i = 0, \dots, n-1.$$

Assim, a matriz  $\mathbf{C} = q(\mathbf{W})$  possui autovalores  $\lambda_i, i = 0, \dots, n-1$ . Portanto,  $f(x) \in F[x]$  é o polinômio característico de  $\mathbf{C}$ . É importante observar que para qualquer reordenação dos escalares

$$\lambda_0, \dots, \lambda_{n-1},$$

obtemos um polinômio  $q(x)$  diferente. Consequentemente, existem no máximo  $n!$  matrizes circulantes com o mesmo polinômio característico  $f(x)$ .

Note que se

$$f(x) = \prod_{i=0}^{n-1} (x - \lambda_i) \text{ e } f_i(x) = \prod_{j \neq i, j=0}^{n-1} (x - \lambda_j).$$

então

$$f(x) = (x - \lambda_i) f_i(x) \text{ e } p_i(x) = \frac{f_i(x)}{f_i(\lambda_i)}.$$

# Capítulo 4

## Equações Polinomiais

Com o objetivo de tornar a leitura deste Capítulo mais didática, apresentaremos nesta seção um pouco da história das equações. O leitor interessado em mais detalhes pode consultar [1, 2, 7, 8].

### 4.1 História

As equações constituem, pelo menos do ponto de vista prático, a parte mais importante da Matemática. “Modelar um problema”, mesmo entre os leigos, é generalizadamente entendido como colocá-lo dentro de um mecanismo do qual ele sairá resolvido.

Equações algébricas são aquelas em que a variável aparece apenas submetida às chamadas operações algébricas. A equação algébrica na forma

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0, \quad \forall n \in \mathbb{N},$$

chama-se *equação polinomial*.

Dentre os antigos documentos matemáticos que conhecemos, os mais famosos são o Papiro de Ahmes (1650 a.C.), também conhecido como de Rhind e o Papiro de Moscou (1850 a.C.). É importante ressaltar que os documentos matemáticos naquela época não empregavam alta dosagem de simbologia à qual estamos atualmente acostumados. Um dos problemas de Ahmes dizia: “*Uma quantidade, somada a seus  $\frac{2}{3}$ , mais sua metade e mais sua sétima parte perfaz 33. Qual é esta quantidade*”? No simbolismo atual escrevemos:

$$x + \frac{2}{3}x + \frac{1}{2}x + \frac{1}{7}x = 33 \Leftrightarrow \frac{97}{42}x = 33,$$

o que é uma equação do 1.º grau.

## 4.1. HISTÓRIA

---

Os Babilônios, na mesma época, já conseguiam trabalhar com equações do 2.º grau e as resolviam por um método baseado no mesmo raciocínio empregado pelos hindus quase três milênios mais tarde, o chamado “completamento de quadrados”. Embora os resultados fossem corretos, os tabletas que contém soluções de equações do 2.º grau apresentam, como todos os demais, apenas sequências do tipo “faça isto”, “faça aquilo”, “este é o resultado”, sem qualquer justificativa lógica sobre o caminho seguido.

A fórmula geral para a solução das equações do 2.º grau é amplamente conhecida mas merece aqui alguns comentários. Em primeiro lugar, seu encontro fundamentou-se na ideia de buscar uma forma de reduzir o grau da equação do 2.º para o 1.º, através da extração de raízes quadradas. Este foi o engenhoso instrumento que os hindus utilizaram com sucesso para chegar à *fórmula de Bhaskara*. Seja a equação do 2.º grau

$$ax^2 + bx + c = 0, \text{ com } a \neq 0.$$

Então

$$x^2 + \frac{b}{a}x = -\frac{c}{a}.$$

Como

$$x^2 + \left(\frac{b}{a}\right)x$$

não é um quadrado perfeito, a ideia foi somar aos dois membros algo que torna o lado esquerdo um quadrado perfeito. Claramente, a quantidade a ser somada era

$$\left(\frac{b}{2a}\right)^2,$$

ou seja,

$$x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = -\frac{c}{a} + \frac{b^2}{4a^2}. \quad (4.1)$$

Assim, como o primeiro membro da igualdade (4.1) é um quadrado perfeito, podemos reescrevê-la sob forma:

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2} \Rightarrow x + \frac{b}{2a} = \pm \sqrt{\frac{b^2 - 4ac}{4a^2}}.$$

Logo, as soluções são dadas pelas fórmulas

$$x = \frac{-b \pm \sqrt{\Delta}}{2a}, \text{ com } \Delta = b^2 - 4ac,$$

a qual é a célebre *fórmula de Bhaskara*, embora não tenha sido deduzida por ele, imortalizou seu nome. As equações do 2.º grau são a chave para a solução de um problema

clássico: encontrar dois números,  $x$  e  $y$ , conhecendo-se sua soma  $S$  e seu produto  $P$ . Este enunciado corresponde ao sistema:

$$\begin{cases} x + y = S \\ xy = P \end{cases}$$

cujas soluções são

$$x = \frac{S + \sqrt{\Delta}}{2} \text{ e } y = \frac{S - \sqrt{\Delta}}{2}, \text{ com } \Delta = S^2 - 4P.$$

Vencidas as equações do 2.º grau, a inesgotável curiosidade dos matemáticos levou-se a conjecturar sobre as formas de resolver as equações do 3.º grau. Os árabes também tiveram papel importante, embora não tenham encontrado a solução.

Consta, por volta de 1510, que Scipione del Ferro, um professor de Matemática da Universidade de Bolonha, encontrou uma forma geral de resolver as equações do tipo

$$x^3 + px + q = 0.$$

Embora tenha morrido sem publicar sua descoberta, ele revelou para seu aluno, Antônio Maria Fior que, mais tarde, tentou adquirir notoriedade valendo-se da descoberta do mestre. Naquela época era bastante frequente o lançamento de desafios entre os sábios e Fior elegeu Tartaglia, já bastante conhecido por seu talento, como alvo. Tartaglia aceitou o desafio, até porque não levava Fior em grande consideração, mas pouco antes da data marcada, veio a saber que seu oponente estava armado de um método descoberto pelo falecido professor Scipione del Ferro. Sentindo-se ameaçado, conforme mais tarde relatou o próprio Tartaglia, “*mobilizei todo o entusiasmo, a aplicação e a arte de que fui capaz, objetivando encontrar uma regra para a solução daquelas equações, o que consegui a 10 de fevereiro de 1535*”. Mas foi mais longe: além de resolver as equações do tipo

$$x^3 + px + q = 0,$$

também, achou a fórmula geral para as equações do tipo

$$x^3 + x^2 + q = 0,$$

que Fior não conhecia. Desse modo Fior saiu humilhado do episódio e hoje só é lembrado como alguém que recebeu o merecido castigo ao pretender fama às custas de outrem.

Cardano ficou sabendo que Tartaglia achara a solução e resolveu pedir-lhe que a revelasse para que fosse publicada em seu livro PRATICA ARITHMETICAE GENERALIS, Tartaglia não aceitou de imediato, porém, cedeu mais tarde após inúmeros juramentos de

segredo do Cardano. Conforme qualquer um poderia prever, Cardano quebrou todas as promessas e juramentos e, em 1545, fez publicar na ARS MAGNA, a fórmula revelada por Tartaglia. Embora tenha feito diversos elogios a ele, acrescentou que, independentemente e trinta anos antes, Scipione del Ferro chegara aos mesmos resultados. A reação de Tartaglia foi pronta e explosiva: publicou sua versão dos fatos e denunciou Cardano por haver traído um sagrado juramento sobre a Bíblia. Em defesa de Cardano veio seu discípulo Ferrari, o descobridor da solução das equações do 4.º grau. Após debates e longas trocas de insultos, a posteridade foi injusta com o sofrido Tartaglia: a fórmula que ele deduzira e que ensinara ao desleal inimigo, ao invés de receber seu nome, é hoje generalizada como fórmula de Cardano. O que ocorreu com a fórmula de Bhaskara repetiu-se nas equações do 3.º grau.

Já sabemos que Tartaglia solucionara os tipos especiais de equações

$$x^3 + px + q = 0 \text{ e } x^3 + px^2 + q = 0$$

e não a equação geral

$$ax^3 + bx^2 + cx + d = 0, \text{ com } a \neq 0. \quad (4.2)$$

Vale salientar que qualquer equação geral pode ser transformada facilmente em um daqueles tipos especiais, digamos

$$x^3 + px + q = 0,$$

fazendo a substituição  $x = y + m$  na equação (4.2) e calculando  $m$  de modo a anular o termo do 2.º grau. Para isto,

$$a(y + m)^3 + b(y + m)^2 + c(y + m) + d = 0$$

e desenvolvendo, obtemos

$$ay^3 + (b + 3am)y^2 + (3am^2 + 2bm + c)y + (m^3a + bm^2 + cm + d) = 0.$$

Assim, impondo à condição  $b + 3am = 0$ , tem-se

$$m = -\frac{b}{3a}$$

e a nova equação do 3.º grau em  $y$  será do tipo

$$y^3 + py + q = 0.$$

Logo, se soubermos resolvê-la, acharemos  $x$  que é  $y + m$ . Portanto, quando encontrou a solução das equações do tipo

$$x^3 + px + q = 0$$

Tartaglia deu uma resposta geral e não apenas particular ao problema, o que aumentou seu mérito. Agora vamos ao “segredo”. Todas as grandes descobertas, invariavelmente, partem de uma ideia fundamental. Neste caso, a ideia de Tartaglia foi supor que a solução procurada era composta de parcelas. Assim, escreveu:

$$x = u + v \Rightarrow x^3 = (u + v)^3 = u^3 + v^3 + 3uvx,$$

ou seja,

$$x^3 - 3uvx - (u^3 + v^3) = 0.$$

Logo,

$$\begin{cases} -3uv &= p \\ -(u^3 + v^3) &= q, \end{cases}$$

ou ainda,

$$u^3 v^3 = -\frac{p^3}{27} \text{ e } u^3 + v^3 = -q.$$

Logo,  $u^3$  e  $v^3$  são dois números dos quais conhecemos a soma e o produto que é um problema clássico que se resolve com equações do segundo grau. Portanto,

$$u^3 = -\frac{q}{2} + \sqrt{\Delta} \text{ e } v^3 = -\frac{q}{2} - \sqrt{\Delta}, \text{ com } \Delta = \frac{27q^2 + 4p^3}{2^2 3^3}.$$

Como  $x = u + v$  temos que

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}. \quad (4.3)$$

Esta é a fórmula de Cardano, que não foi descoberta por ele, mas sim por Tartaglia. Por exemplo, se

$$x^3 - 6x - 9 = 0,$$

então, depois de alguns cálculos, obtemos a solução da equação

$$x = \sqrt[3]{\frac{9}{2} + \sqrt{\frac{49}{4}}} + \sqrt[3]{\frac{9}{2} - \sqrt{\frac{49}{4}}} = 3.$$

Neste ponto ocorre um fato curioso chegando a ter um aspecto paradoxal. Veremos mais tarde que se  $\Delta \leq 0$ , então as três raízes da equação

$$x^3 + px + q = 0$$

#### 4.1. HISTÓRIA

---

são reais. A fórmula exprime  $x = u + v$  como soma de duas raízes cúbicas de números complexos. Este é o “caso irredutível”, pois ao tentar eliminar os radicais, recai-se noutra equação do terceiro grau. Um importante exemplo é dado pela equação

$$x^3 - 3x + 1 = 0, \text{ com } \Delta = -\frac{3}{4},$$

podendo ter três raízes reais e distintas, com uma delas sendo

$$x = \sqrt[3]{-\frac{1}{2} + \frac{\sqrt{3}}{2}i} + \sqrt[3]{-\frac{1}{2} - \frac{\sqrt{3}}{2}i}.$$

Isto parece um número complexo, mas é um número real, como veremos mais adiante. À primeira vista, achou-se que as equações do 3.º grau estavam vencidas pela fórmula de Cardano, analogamente ao que a fórmula de Bhaskara fizera com as equações do 2.º grau. A mais elementar dúvida que surge naturalmente em quem observa a fórmula (4.3) é a seguinte: se a fórmula de Bhaskara exhibe, de maneira tão simples, as duas raízes das equações do 2.º grau, por que a de Cardano só apresenta uma? Por exemplo, a equação

$$x^3 - 15x - 4 = 0$$

possui  $x = 4$  como uma de suas raízes. Não obstante, se tentarmos resolvê-la pela fórmula (4.3), teremos

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}$$

e caímos não apenas na extração de raízes quadradas de números negativos, mas também na extração de raízes cúbicas de números de natureza desconhecida. O homem que conseguiu atravessar a ponte que levava aos novos números foi Rafael Bombelli, nascido em Bolonha, Itália, em 1530 e engenheiro hidráulico por profissão. Ele era aquele tipo de pessoa, nascido para fazer História: corajoso, pertinaz e sempre disposto a pensar em coisas novas. Os estudos de Bombelli começaram com a tentativa de conciliar o resultado fornecido pela fórmula de Cardano para a equação

$$x^3 - 15x - 4 = 0,$$

com a raiz constatada por simples observação. Conforme ele mesmo revelou em 1572 no livro **L'Algebra parte Maggiore dell'Arinewline thmetica**, seu método baseou-se no “pensamento rude” segundo o qual

$$\sqrt[3]{2 + \sqrt{-121}} \text{ e } \sqrt[3]{2 - \sqrt{-121}}$$

deveriam ser números da forma

$$a + \sqrt{-b} \text{ e } a - \sqrt{-b}.$$

Assim, ele escreveu

$$\sqrt[3]{2 + \sqrt{-121}} = a + \sqrt{-b} \text{ e } \sqrt[3]{2 - \sqrt{-121}} = a - \sqrt{-b}.$$

Depois de alguns cálculos, obtemos  $a = 2$  e  $b = 1$ , pois

$$(2 + \sqrt{-1})^3 = 2 + \sqrt{-121} \text{ e } (2 - \sqrt{-1})^3 = 2 - \sqrt{-121}.$$

Logo,

$$x = (2 + \sqrt{-1}) + (2 - \sqrt{-1}) = 4$$

que é o resultado desejado.

L. Ferrari, nascido em Bolonha 1522 e falecido por volta de 1560, foi o mais famoso dos discípulos de Cardano. Oriundo da mais humilde das condições, foi trabalhar como servo na residência de Cardano quando tinha 15 anos, mas sua brilhante inteligência logo foi reconhecida pelo mestre e disto decorreu uma promoção a secretário. A partir dos 18 anos, Ferrari passou a ensinar por conta própria em Milão e, através da proteção do Cardeal de Mantova, alcançou posições que lhe proporcionaram uma boa renda. Logo após tornar-se professor de Matemática na Universidade de Bolonha, veio a falecer aos 38 anos de idade, provavelmente envenenado por sua irmã.

Dentro do costume então vigente entre os matemáticos, de proporem problemas uns aos outros como forma de desafio, Zuanne de Tonini da Coi submeteu a Cardano uma questão que envolvia a equação

$$x^4 + 6x^2 - 60x + 36 = 0.$$

Após inúmeras tentativas sem êxito, Cardano passou a questão ao jovem Ferrari que, num lampejo de gênio, encontrou o método geral para a solução das equações do 4.º grau. Antes de mostrar o raciocínio seguido por Ferrari, vamos relembrar que a equação geral do 4.º grau

$$ax^4 + bx^3 + cx^2 + dx + e = 0, \text{ com } a \neq 0,$$

sempre pode ser transformada em outra do tipo

$$y^4 + py^2 + qy + r = 0 \tag{4.4}$$

fazendo a substituição  $x = y + m$  e calculando  $m$  de modo a anular o termo de 3.º grau. Ferrari olhou a equação (4.4) e procurou reagrupar os termos de modo que, nos dois

#### 4.1. HISTÓRIA

---

membros da igualdade houvesse polinômios quadrados perfeitos. Se tal reagrupamento fosse possível, seriam extraídas as raízes quadradas, cair-se-ia em equações do 2.<sup>o</sup> grau e o problema estaria resolvido. A equação foi, então, escrita assim:

$$x^4 + (p + a)x^2 + (r + b) = ax^2 - qx + b$$

em que  $a$  e  $b$  são números a serem determinados de forma que os dois lados da igualdade sejam quadrados perfeitos. Para que isso ocorra, é necessário e suficiente que os discriminantes daqueles dois trinômios, ao mesmo tempo, sejam iguais a *zero*, ou seja,

$$(p + a)^2 - 4(r + b) = 0 \text{ e } q^2 - 4ab = 0$$

resolvendo o sistema, obtemos

$$a^3 + 2pa^2 + (p^2 - 4r)a - q^2 = 0$$

que é uma equação do 3.<sup>o</sup> grau em  $a$ . Como tais equações podem ser resolvidas, acha-se  $a$ , em seguida  $b$  e extraem-se as raízes quadradas

$$\sqrt{x^4 + (p + a)x^2 + (r + b)} = \pm \sqrt{ax^2 - qx + b}.$$

Para cada alternativa de sinal  $+$  ou  $-$  tem-se uma equação do 2.<sup>o</sup> grau, ambas com duas soluções. Portanto, para a equação do 4.<sup>o</sup> grau, o método fornece quatro raízes, de uma forma semelhante ao que acontece na fórmula de Bhaskara. Os passos para a solução da equação geral do 4.<sup>o</sup> grau são:

- Toma-se a equação geral e faz-se uma transformação do tipo  $x = y + m$  de modo a cair-se em uma equação do 4.<sup>o</sup> grau em  $y$  sem o termo do 3.<sup>o</sup> grau.
- Reagrupam-se seus termos de modo a fazer com que ambos os lados da igualdade sejam quadrados perfeitos. Cai-se em uma equação do 3.<sup>o</sup> grau em  $a$ . Se ela for completa, faz-se a transformação  $a = c + t$  de modo a obter-se uma equação do 3.<sup>o</sup> grau em  $c$ , sem o termo do 2.<sup>o</sup> grau.
- Resolve-se a equação em  $c$  pelo método de Tartaglia;
- Soma-se  $t$  a  $c$  e obtém-se  $a$ . Obtido  $a$  calcula-se  $b$ .
- Com  $a$  e  $b$ , extraem-se as raízes quadradas dos dois lados da igualdade e obtém-se os quatro valores possíveis de  $y$ . Soma-se  $m$  a  $y$  e obtém-se as quatro raízes da equação geral.

Um método perfeito do ponto de vista teórico, mas bastante trabalhoso. O grande mérito de Ferrari foi haver demonstrado que a solução das equações do 4.º grau era possível apenas com operações algébricas.

**Exemplo 4.1** *Vamos aplicar o método de Ferrari na equação*

$$x^4 - 15x^2 - 10x + 24 = 0.$$

**Solução.** Desejamos determinar  $a$  e  $b$  tais que

$$x^4 - (15 - a)x^2 + (24 + b) = ax^2 + 10x + b$$

tenham ambos os lados da igualdade quadrados perfeitos. Para isto

$$(15 - a)^2 - 4(24 + b) = 0 \text{ e } 100 - 4ab = 0.$$

Da segunda equação, obtemos

$$b = \frac{25}{a}, \text{ com } a \neq 0,$$

nos levando a equação

$$a^3 - 30a^2 + 129a - 100 = 0.$$

Esta equação, sendo do 3.º grau, é solúvel algebricamente e suas raízes são

$$a_1 = 1, \quad a_2 = 4 \text{ e } a_3 = 25.$$

Para  $a_1 = 1$  e  $b_1 = 25$ , teremos as raízes  $x_1 = 4, x_2 = -3, x_3 = -2$  e  $x_4 = 1$ . As outras raízes de  $a$  nos levam as mesmas raízes, a menos da ordem. ■

A tentativa de resolver as equações de grau  $n \geq 5$  fracassou. Entre 1824 e 1826, Abel conseguiu mostrar que a equação geral de grau 5 não é solúvel por radicais. Finalmente, Galois 1811-1832, indicou critérios para uma equação qualquer ser solúvel,

## 4.2 Aplicação do Método Circulante

Para ilustrar a ideia principal desta seção, consideramos o problema de encontrar expressões exatas para as raízes do polinômio

$$f(x) = x^3 + 6x^2 + 9x + 3.$$

## 4.2. APLICAÇÃO DO MÉTODO CIRCULANTE

---

É bastante natural tentarmos a fatoração do polinômio por algum caminho adequado, ou olhar para o teste das raízes racionais, este último sem sucesso pois os divisores de 3 são

$$\{1, -1, 3, -3\}$$

que não são raízes de  $f(x)$ . No entanto, pelo Exemplo 3.4,  $f(x)$  é o polinômio característico da matriz circulante

$$\mathbf{C} = \begin{pmatrix} -2 & \beta & \bar{\beta} \\ \bar{\beta} & -2 & \beta \\ \beta & \bar{\beta} & -2 \end{pmatrix}, \text{ com } \beta = \cos\left(\frac{2\pi}{9}\right) + i\sin\left(\frac{2\pi}{9}\right),$$

em que as raízes de  $f(x)$  são os autovalores de  $\mathbf{C}$ , e estas são obtidas pelas avaliações

$$q(1), \quad q(\omega) \text{ e } q(\bar{\omega}),$$

no polinômio representante

$$q(x) = -2 + \beta x + \bar{\beta} x^2.$$

De modo geral, dado um polinômio  $f(x)$ , já sabemos que é sempre possível encontrar uma matriz circulante  $\mathbf{C}$ , tendo  $f(x)$  como seu polinômio característico.

Esta seção será dedicado ao processo unificado de resolução das equações polinomiais de grau menor do que ou igual a 4, via matrizes circulantes, começando pelo caso mais simples, que é o quadrático.

Consideremos o polinômio

$$f(x) = x^2 + ax + b \in F[x]$$

e a matriz circulante

$$\mathbf{C} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} = \alpha \mathbf{I} + \beta \mathbf{W}.$$

Então o polinômio característico de  $\mathbf{C}$  é

$$p(x) = \det(x\mathbf{I} - \mathbf{C}) = (x - \alpha)^2 - \beta^2.$$

Esta expressão sugere a mudança de variáveis lineares

$$y = x - \alpha$$

a qual transforma  $p(x)$  em um polinômio sem o termo de grau 1 em  $y$ . Portanto, o desenvolvimento via matriz circulante inspira, de uma forma muito natural, um passo preliminar na solução tradicional de equações através da *Transformação de Tschirnhaus*, dada em 1683, ou uma *mudança de variável linear* na eliminação do termo de grau  $n - 1$ , com  $n > 1$ . Mais precisamente, seja

$$f(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1} + x^n \in F[x].$$

Então a mudança de variável linear  $y = x + \frac{c_{n-1}}{n}$  elimina o termo de grau  $n - 1$ , apenas usando operações algébricas elementares, ou seja, obtemos o *polinômio reduzido associado*

$$g(x) = f\left(x - \frac{c_{n-1}}{n}\right) \in F[x].$$

Além disso,  $\alpha$  é uma raiz de  $g(x)$  se, e somente se,  $\alpha - \frac{c_{n-1}}{n}$  é uma raiz de  $f(x)$ . No contexto de matrizes circulantes ganhamos uma nova maneira de pensar sobre este resultado. Para ver isto, seja  $p(x)$  o polinômio característico da matriz circulante  $\mathbf{C}$ . Então a soma das raízes é a soma dos autovalores, isto é, o traço de  $\mathbf{C}$ . Assim, eliminar o termo de grau  $n - 1$  é equivalente ao traço de  $\mathbf{C}$  ser igual zero. Observe que como uma matriz circulante  $\mathbf{C}$  possui a diagonal principal constante, digamos  $\alpha$ . Então

$$\text{tr}(\mathbf{C}) = n\alpha \Rightarrow \alpha = -\frac{c_{n-1}}{n}.$$

Isto nos dá as seguintes conclusões:

- $\alpha$  determina um dos parâmetros da matriz circulante (o termo constante do polinômio associado  $q(x)$ ).
- Como no caso da equação do 2.º grau,

$$y = x - \alpha = x - \frac{c_{n-1}}{n}$$

é a mudança de variável linear que elimina o termo de grau  $n - 1$ .

Estas observações indicam que uma mudança de variável linear pode ser sempre realizada para eliminar o termo de grau  $n - 1$  de um polinômio geral de grau  $n$ . Daqui em diante, suponhamos que tal mudança de variáveis linear já tenha sido feita. Retornemos ao caso quadrático. Sejam o polinômio

$$g(x) = x^2 + p \in F[x]$$

e a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \beta \\ \beta & 0 \end{pmatrix} = \beta \mathbf{W}.$$

## 4.2. APLICAÇÃO DO MÉTODO CIRCULANTE

---

Então o polinômio característico de  $\mathbf{C}$  é

$$p(x) = \det(x\mathbf{I} - \mathbf{C}) = x - \beta^2.$$

Desejamos determinar  $\beta$ , para que este polinômio característico seja igual a  $g(x)$  ou, equivalentemente, resolver o sistema não homogêneo

$$-\beta^2 = p,$$

ou seja,  $\beta = -\sqrt{-p}$  ou  $\beta = \sqrt{-p}$ . Como veremos, não há perda de generalidade, em considerarmos  $\beta$  com o sinal positivo. Assim,

$$\mathbf{C} = \begin{pmatrix} 0 & \sqrt{-p} \\ \sqrt{-p} & 0 \end{pmatrix} \text{ e } q(x) = \sqrt{-p}x.$$

Sendo

$$\mathbf{W} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

temos que  $x^2 - 1 = 0$  é a equação característica de  $\mathbf{W}$ . Logo,

$$1 \text{ e } -1$$

são as raízes da unidade (autovalores de  $\mathbf{W}$ ). Portanto,

$$q(1) = \sqrt{-b} \text{ e } q(-1) = -\sqrt{-b}$$

são as raízes de  $g(x)$  (autovalores de  $\mathbf{C}$ ). Observe que se definirmos  $\beta$  com o sinal negativo, então obtemos as mesmas raízes de  $g(x)$ , embora os valores de  $q(1)$  e  $q(-1)$  são permutados. Note, para o polinômio geral, que

$$f(x) = x^2 + ax + b \text{ e } g(x) = f\left(x - \frac{a}{2}\right) = x^2 + p, \quad p = b - \frac{a^2}{4},$$

as soluções são

$$r_1 = -\frac{a}{2} + q(-1) \text{ e } r_2 = -\frac{a}{2} + q(1),$$

ou ainda,

$$r_1 = -\frac{a}{2} - \sqrt{\frac{a^2}{4} - b} \text{ e } r_2 = -\frac{a}{2} + \sqrt{\frac{a^2}{4} - b}.$$

**Teorema 4.2** *Seja  $f(x) = x^2 + ax + b \in F[x]$ .*

1.  $f(x)$  possui raízes reais se, e somente se,  $a^2 - 4b \geq 0$ .

2. Se  $f(x) \in \mathbb{Q}[x]$ , então  $f(x)$  possui raízes irracionais se, e somente se,  $a^2 - 4b > 0$  e  $a^2 - 4b$  é livre de quadrados.

3.  $f(x)$  possui raízes imaginárias se, e somente se,  $a^2 - 4b < 0$ .

**Prova.** Vamos provar apenas os itens (1) e (3): (1) Pelo item (1) do Lema 2.17,  $f(x)$  possui raízes reais se, e somente se, a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \sqrt{-p} \\ \sqrt{-p} & 0 \end{pmatrix}, \quad p = b - \frac{a^2}{4} = \Delta,$$

é Hermitiana. Assim, se, e somente se,  $\sqrt{-p}$  é real. Portanto, se, e somente se,  $-p \geq 0$ , ou seja,  $a^2 - 4b \geq 0$ .

(3) Pelo item (2) do Lema 2.17,  $f(x)$  possui raízes imaginárias se, e somente se, a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \sqrt{-p} \\ \sqrt{-p} & 0 \end{pmatrix}, \quad p = b - \frac{a^2}{4} = \Delta,$$

é anti-Hermitiana. Assim, se, e somente se,  $\sqrt{-p}$  é complexo puro. Portanto, se, e somente se,  $-p < 0$ , ou seja,  $a^2 - 4b < 0$ . ■

**Exemplo 4.3** Seja o polinômio  $f(x) = x^2 - 5x + 6 \in \mathbb{R}[x]$ .

**Solução.** É claro que 2 e 3, são as raízes de  $f(x)$ . Mas, como uma ilustração utilizaremos o método das circulantes para obtê-las. Consideremos a matriz circulante

$$\mathbf{C} = \begin{pmatrix} a & b \\ b & a \end{pmatrix} = a\mathbf{I} + b\mathbf{W}.$$

tal que  $p(x) = f(x)$ . Assim, as raízes de  $f(x)$  serão os autovalores da matriz  $\mathbf{C}$ . Neste caso,

$$p(x) = \det(x\mathbf{I} - \mathbf{C}) = (x - a)^2 - b^2 = x^2 - 2ax + a^2 - b^2.$$

Logo,

$$x^2 - 2ax + a^2 - b^2 = x^2 - 5x + 6 \Leftrightarrow a = \frac{5}{2} \text{ e } b = \frac{1}{2}.$$

Portanto,

$$\mathbf{C} = \begin{pmatrix} \frac{5}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{5}{2} \end{pmatrix} \text{ e } q(x) = \frac{5}{2} + \frac{1}{2}x,$$

com  $\mathbf{C}$  claramente Hermitiana. Consequentemente,

$$q(1) = \frac{5}{2} + \frac{1}{2} = 3 \text{ e } q(-1) = \frac{5}{2} - \frac{1}{2} = 2$$

## 4.2. APLICAÇÃO DO MÉTODO CIRCULANTE

são as raízes racionais de  $f(x)$ , pois  $a^2 - 4b = 1$  não é livre de quadrados. ■

Sejam

$$g(x) = x^3 + px + q \in F[x]$$

e a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \beta & \gamma \\ \gamma & 0 & \beta \\ \beta & \gamma & 0 \end{pmatrix} = \beta \mathbf{W} + \gamma \mathbf{W}^2.$$

Então o polinômio característico de  $\mathbf{C}$  é

$$p(x) = \det(x\mathbf{I} - \mathbf{C}) = x^3 - 3\beta\gamma x - (\beta^3 + \gamma^3).$$

Desejamos determinar  $\beta$  e  $\gamma$ , para que este polinômio característico seja igual a  $g(x)$  ou, equivalentemente, resolver o sistema não homogêneo

$$\begin{cases} 3\beta\gamma &= -p \\ \beta^3 + \gamma^3 &= -q. \end{cases} \quad (4.5)$$

Elevando ao cubo a primeira equação em (4.5) e vendo  $\beta^3$  e  $\gamma^3$  como variáveis no sistema (4.5), obtemos

$$\begin{cases} \beta^3\gamma^3 &= -\frac{p^3}{27} \\ \beta^3 + \gamma^3 &= -q, \end{cases}$$

de modo que  $\beta^3$  e  $\gamma^3$  são as raízes da equação

$$y^2 + qy - \frac{p^3}{27} = 0,$$

a qual chama-se *resolvente quadrática* da cúbica. Neste caso, as raízes são:

$$-\frac{q}{2} \pm \sqrt{\Delta}, \text{ com } \Delta = \frac{27q^2 + 4p^3}{2^2 3^3}. \quad (4.6)$$

Neste ponto somos levados a escrever:

$$\beta = \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} \text{ e } \gamma = \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}. \quad (4.7)$$

Assim,

$$\mathbf{C} = \begin{pmatrix} 0 & \beta & \gamma \\ \gamma & 0 & \beta \\ \beta & \gamma & 0 \end{pmatrix} \text{ e } q(x) = \beta x + \gamma x^2.$$

Como

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

temos que  $x^3 - 1 = 0$  é a equação característica de  $\mathbf{W}$ . Logo,

$$1, \omega \text{ e } \bar{\omega}$$

são as raízes da unidade (autovalores de  $\mathbf{W}$ ). Portanto,

$$q(1) = \beta + \gamma, \quad q(\omega) = \beta\omega + \gamma\bar{\omega} \text{ e } q(\bar{\omega}) = \beta\bar{\omega} + \gamma\omega$$

são raízes de  $g(x)$  (autovalores de  $\mathbf{C}$ ). Na verdade, as igualdades em (4.7), são perfeitamente válidas quando todas as operações envolvem apenas os números reais. Num domínio maior, por exemplo, os números complexos, existe alguma ambiguidade associada com a extração de raízes quadradas e cúbicas. Neste caso, definimos  $\beta$  por (4.7), usando qualquer um dos valores da raiz quadrada, cúbica e escolhemos  $\gamma$ , de modo que a condição

$$\beta\gamma = -\frac{p}{3}$$

seja satisfeita. Isso produz uma solução de (4.7), nos levando para as raízes de  $g(x)$  como já foi explicado. Todas as opções para  $\beta$  resultam nas mesmas raízes. Note, para o polinômio geral, que

$$f(x) = x^3 + ax^2 + bx + c$$

e

$$g(x) = f\left(x - \frac{a}{3}\right) = x^3 + px + q, \quad p = b - \frac{a^2}{3} \text{ e } q = c - \frac{ab}{3} + \frac{2a^3}{27},$$

as raízes são

$$r_1 = -\frac{a}{3} + q(1), \quad r_2 = -\frac{a}{3} + q(\omega) \text{ e } r_3 = -\frac{a}{3} + q(\bar{\omega}).$$

Por exemplo,

$$r_1 = -\frac{a}{3} + \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}.$$

**Teorema 4.4** *Sejam  $g(x) = x^3 + px + q \in F[x]$  e*

$$\Delta = \frac{27q^2 + 4p^3}{2^2 3^3}.$$

1.  $g(x)$  possui raízes reais distintas se, e somente se,  $\Delta < 0$ .

## 4.2. APLICAÇÃO DO MÉTODO CIRCULANTE

---

2.  $g(x)$  possui raízes reais e duas iguais se, e somente se,  $\Delta = 0$ .
3. Se  $F \subseteq \mathbb{R}$ , então  $g(x)$  possui uma raiz real e duas imaginárias se, e somente se,  $\Delta > 0$ .
4. Se  $g(x) \in \mathbb{Q}[x]$  e  $\mathbf{C} \in \mathbb{Q}^{3 \times 3}$ , então  $g(x)$  possui todas as raízes racionais se, e somente se,  $\Delta = 0$ .

**Prova.** Vamos provar apenas os itens (1), (2) e (3): (1) Pelo item (1) do Lema 2.17,  $g(x)$  possui raízes reais se, e somente se, a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \beta & \gamma \\ \gamma & 0 & \beta \\ \beta & \gamma & 0 \end{pmatrix}$$

é Hermitiana. Assim, se, e somente se,  $\gamma = \bar{\beta}$  ou, equivalentemente,  $\gamma^3 = \bar{\beta}^3$ . Note que  $\beta^3$  e  $\gamma^3$  são complexas conjugadas. Logo, pelo item (3) do Teorema 4.2 se, e somente se,  $\Delta < 0$ . Pondo  $\beta = u + vi$  e  $\gamma = u - vi$ , teremos

$$\begin{aligned} q(1) &= \beta + \gamma &= 2u \\ q(\omega) &= \beta\omega + \gamma\bar{\omega} &= -u - v\sqrt{3} \\ q(\bar{\omega}) &= \beta\bar{\omega} + \gamma\omega &= -u + v\sqrt{3} \end{aligned}$$

são as raízes reais e distintas de  $g(x)$ .

(2) Pelo item (1) se, e somente se,  $\Delta = 0$ . Assim,

$$\begin{aligned} q(1) &= \beta + \gamma &= -\sqrt[3]{4q} \\ q(\omega) &= \beta\omega + \gamma\bar{\omega} &= -\frac{1}{2}\sqrt[3]{4q} \\ q(\bar{\omega}) &= \beta\bar{\omega} + \gamma\omega &= -\frac{1}{2}\sqrt[3]{4q} \end{aligned}$$

são as raízes reais de  $g(x)$ .

(3) Pelo item (4) do Lema 2.17,  $g(x)$  possui uma raiz real e duas imaginárias se, e somente se, a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \beta & \gamma \\ \gamma & 0 & \beta \\ \beta & \gamma & 0 \end{pmatrix}$$

é real e não ortogonal. Assim, pelo item (1) do Teorema 4.2 se, e somente se,  $\Delta > 0$ . Portanto,

$$\begin{aligned} q(1) &= \beta + \gamma \\ q(\omega) &= \beta\omega + \gamma\bar{\omega} \\ q(\bar{\omega}) &= \beta\bar{\omega} + \gamma\omega \end{aligned}$$

é uma raiz real e duas imaginárias de  $g(x)$ . ■

É muito importante de um ponto de vista teórico e didático apresentarmos uma interpretação geométrica de localização das raízes de uma cúbica quando todas são reais. Para isto, pondo  $h = 2\beta = 2b$ , as raízes da cúbica são as partes reais de

$$h, \omega h \text{ e } \bar{\omega}h,$$

respectivamente, e estes são igualmente distribuídos em torno do círculo  $|z| = |h|$ . Então, geometricamente, podemos encontrar as raízes da seguinte forma: dobramos  $b$  e o localizamos no plano complexo. A partir daí construímos um triângulo equilátero tendo  $2b$  como um de seus vértices e centro na origem. Em seguida projetamos os vértices no eixo real. Esta construção está ilustrada na Figura 4.1. Embora esta construção é aplicável no caso especial de um cúbico (sem termo quadrático), o caso geral é essencialmente o mesmo, pois os polinômios do tipo

$$f(x) = x^3 + ax^2 + bx + c$$

podem ser transformados em

$$g(x) = x^3 + px + q,$$

em que as raízes ficam transladadas de  $\frac{a}{3}$ . As raízes de  $f(x)$ , são ainda projeções dos vértices do referido triângulo, mas com centro em  $-\frac{a}{3}$ , em vez da origem.

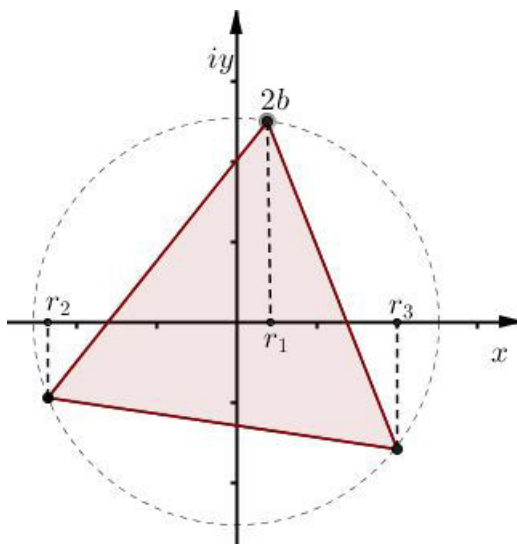


Figura 4.1: Representação gráfica das raízes de  $g(x)$ .

O próximo exemplo ilustra a interpretação geométrica das raízes quando todas são reais.

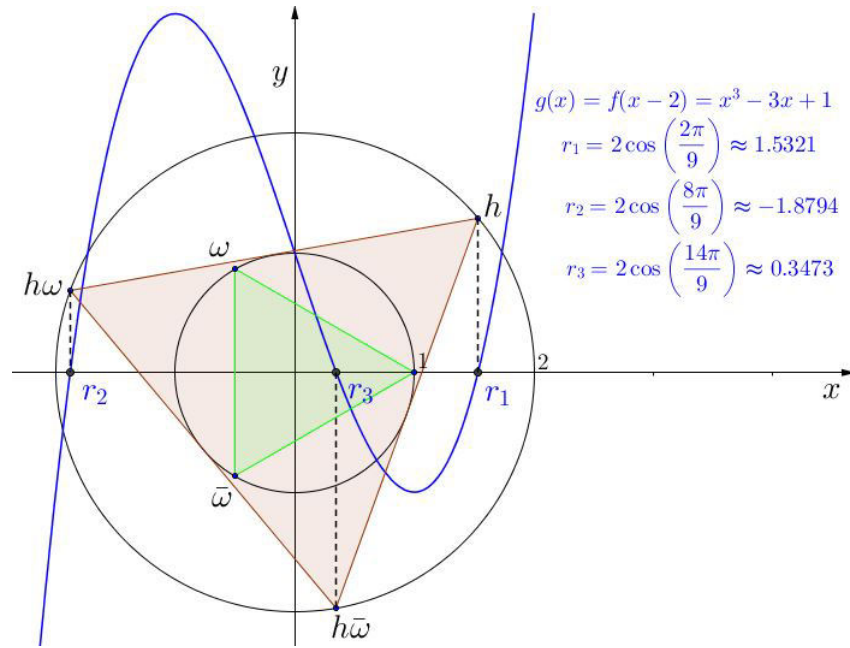


Figura 4.2: Interpretação geométrica das raízes de  $g(x)$ .

**Exemplo 4.5** Seja o polinômio  $f(x) = x^3 + 6x^2 + 9x + 3 \in \mathbb{R}[x]$ .

**Solução.** Já vimos que a mudança de variáveis  $x = y - 2$ , nos levando ao polinômio reduzido

$$g(x) = f(x - 2) = x^3 - 3x + 1 \text{ e } \Delta = -\frac{3}{4} < 0$$

Neste caso, as raízes reais distintas de  $g(x)$  e  $f(x)$  são, respectivamente,

$$\begin{aligned} r_1 &= 2 \cos\left(\frac{2\pi}{9}\right) & r_1 &= -2 + 2 \cos\left(\frac{2\pi}{9}\right) \\ r_2 &= 2 \cos\left(\frac{8\pi}{9}\right) & \text{e } r_2 &= -2 + 2 \cos\left(\frac{8\pi}{9}\right) \\ r_3 &= 2 \cos\left(\frac{14\pi}{9}\right) & r_3 &= -2 + 2 \cos\left(\frac{14\pi}{9}\right), \end{aligned}$$

confira Figura 4.2. ■

**Exemplo 4.6** Seja o polinômio  $f(x) = x^3 + 3x^2 + 9x - 13 \in \mathbb{R}[x]$ .

**Solução.** É fácil verificar que o polinômio reduzido é

$$g(x) = f(x - 1) = x^3 + 6x - 20 \text{ e } \Delta = 2^2 3^3 = 108 > 0.$$

nunca é um quadrado perfeito. Neste caso,

$$C = \begin{pmatrix} 0 & \beta & \gamma \\ \gamma & 0 & \beta \\ \beta & \gamma & 0 \end{pmatrix}, \text{ com } \beta = \sqrt[3]{10 + \sqrt{108}} \text{ e } \gamma = \sqrt[3]{10 - \sqrt{108}}.$$

Note que  $\beta, \gamma \notin \mathbb{Q}$ . Assim,  $\mathbf{C} \notin \mathbb{Q}^{3 \times 3}$ ,  $\mathbf{C} \neq \mathbf{C}^t$  e  $g(x)$  possui um par de raízes complexas. Não obstante,  $q(1) = \beta + \gamma = 2 \in \mathbb{Q}$ . ■

**Exemplo 4.7** Seja o polinômio  $f(x) = x^3 + 6x^2 + 9x + 4 \in \mathbb{R}[x]$ .

**Solução.** É fácil verificar que o polinômio reduzido é

$$g(x) = f(x - 2) = x^3 - 3x + 2 \text{ e } \Delta = 0.$$

Logo,  $\beta = \gamma = -1$  e

$$\mathbf{C} = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$$

é racional e ortogonal. Portanto,  $q(1) = -2$ ,  $q(\omega) = q(\bar{\omega}) = -1$  são as raízes racionais de  $g(x)$ . ■

O restante desta seção será dedicado a resolução da equação quártica. Sejam

$$g(x) = x^4 + px^2 + qx + r \in F[x],$$

e a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \beta & \gamma & \delta \\ \delta & 0 & \beta & \gamma \\ \gamma & \delta & 0 & \beta \\ \beta & \gamma & \delta & 0 \end{pmatrix} = \beta \mathbf{W} + \gamma \mathbf{W}^2 + \delta \mathbf{W}^3.$$

Então o polinômio característico de  $\mathbf{C}$  é

$$\begin{aligned} p(x) &= x^4 - (4\beta\delta + 2\gamma^2)x^2 - 4\gamma(\beta^2 + \delta^2)x \\ &+ \gamma^4 - \beta^4 - \delta^4 - 4\beta\delta\gamma^2 + 2\beta^2\delta^2. \end{aligned}$$

Desejamos determinar  $\beta, \gamma$  e  $\delta$ , para que este polinômio característico seja igual a  $g(x)$  ou, equivalentemente, resolver o sistema não homogêneo

$$\begin{cases} 4\beta\delta + 2\gamma^2 &= -p \\ 4\gamma(\beta^2 + \delta^2) &= -q \\ \gamma^4 - \beta^4 - \delta^4 - 4\beta\delta\gamma^2 + 2\beta^2\delta^2 &= r \end{cases} \quad (4.8)$$

A partir das duas primeiras equações do sistema determinamos  $\beta\delta$  e  $\beta^2 + \delta^2$  em função de  $\gamma$ . Nos inspirando a reescrever a terceira equação sob a forma

$$\gamma^4 - (\beta^2 + \delta^2)^2 + 4(\beta\delta)^2 - 4\beta\delta\gamma^2 = r$$

## 4.2. APLICAÇÃO DO MÉTODO CIRCULANTE

---

e, conseqüentemente, obter uma equação em  $\gamma$ :

$$\gamma^4 - \frac{q^2}{16\gamma^2} + \frac{(p + 2\gamma^2)^2}{4} + (2\gamma^2 + p)\gamma^2 = r$$

e, simplificando, teremos

$$\gamma^6 + \frac{p}{2}\gamma^4 + \left(\frac{p^2}{16} - \frac{r}{4}\right)\gamma^2 - \frac{q^2}{64} = 0, \quad (4.9)$$

qual é uma cúbica em  $\gamma^2$  chamada *resolvente cúbica* e é solúvel pelos métodos já vistos. Desta forma, construímos a matriz circulante

$$\mathbf{C} = \beta\mathbf{W} + \gamma\mathbf{W}^2 + \delta\mathbf{W}^3 = q(\mathbf{W}).$$

Assim,

$$\mathbf{C} = \begin{pmatrix} 0 & \beta & \gamma & \delta \\ \delta & 0 & \beta & \gamma \\ \gamma & \delta & 0 & \beta \\ \beta & \gamma & \delta & 0 \end{pmatrix} \text{ e } q(x) = \beta x + \gamma x^2 + \delta x^3.$$

Como

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

temos que  $x^4 - 1 = 0$  é a equação característica de  $\mathbf{W}$ . Logo,

$$1, -1, i \text{ e } -i$$

são as raízes da unidade (autovalores de  $\mathbf{W}$ ). Portanto,

$$\begin{aligned} q(1) &= \beta + \gamma + \delta, & q(-1) &= -\beta + \gamma - \delta, \\ q(i) &= -\gamma + i(\beta - \delta), & q(-i) &= -\gamma - i(\beta - \delta) \end{aligned}$$

são as raízes de  $g(x)$  (autovalores de  $\mathbf{C}$ ). Note, para o polinômio geral, que

$$f(x) = x^4 + ax^3 + bx^2 + cx + d$$

e

$$g(x) = f\left(x - \frac{a}{4}\right) = x^4 + px^3 + qx + r,$$

com

$$p = b - \frac{3a^2}{8}, q = c - \frac{ac}{2} + \frac{a^3}{8} \text{ e } r = d - \frac{ac}{4} + \frac{2a^2b}{16} + \frac{3a^4}{256},$$

as raízes são

$$r_1 = -\frac{a}{4} + q(1), r_2 = -\frac{a}{4} + q(1), \\ r_3 = -\frac{a}{4} + q(i) \text{ e } r_4 = -\frac{a}{4} + q(-i).$$

**Exemplo 4.8** Seja o polinômio  $f(x) = x^4 - 2x^2 + 8x - 3 \in \mathbb{R}[x]$ .

**Solução.** Como  $p = -2$ ,  $q = 8$  e  $r = -3$  temos que a resolvente cúbica de  $f(x)$  é

$$y^3 - y^2 + y - 1 = 0, \text{ com } y = \gamma^2,$$

a qual possui solução  $\gamma^2 = 1$ , ou ainda,  $\gamma = \pm 1$ . Assim, pela sistema (4.8), obtemos

$$\begin{cases} 4\beta\delta + 2 = 2 \\ -4(\beta^2 + \delta^2) = -8. \end{cases}$$

Logo,  $\beta = 0$ ,  $\gamma = -1$  e  $\delta = \pm\sqrt{2}$ . Neste caso, escolhendo  $\delta = \sqrt{2}$ , obtemos

$$q(x) = -x^2 + \sqrt{2}x^3.$$

Portanto,

$$q(1) = -1 + \sqrt{2}, \quad q(-1) = -1 - \sqrt{2}, \\ q(i) = 1 - \sqrt{2}i, \quad q(-i) = 1 + \sqrt{2}i$$

são as raízes de  $f(x)$  (autovalores de  $\mathbf{C}$ ), em que

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & -1 & \sqrt{2} \\ \sqrt{2} & 0 & 0 & -1 \\ -1 & \sqrt{2} & 0 & 0 \\ 0 & -1 & \sqrt{2} & 0 \end{pmatrix},$$

que é o resultado desejado. ■

Observe que se

$$g(x) = x^4 + px^2 + qx + r$$

e a matriz circulante

$$\mathbf{C} = \begin{pmatrix} 0 & \beta & \gamma & \delta \\ \delta & 0 & \beta & \gamma \\ \gamma & \delta & 0 & \beta \\ \beta & \gamma & \delta & 0 \end{pmatrix},$$

então  $C$  é Hermitiana se, e somente se,  $\delta = \bar{\beta}$  e  $\gamma$  é real. Agora, qualquer solução para (4.8) produz uma circulante com polinômio característico  $g(x)$ . Uma tal solução pode ser construída usando qualquer valor para  $\gamma$  satisfazendo (4.9). Por conseguinte, a fim de que todas as raízes de  $f$  sejam reais e, conseqüentemente, para todos as circulantes correspondentes serem Hermitianas, é necessário que todas as raízes da equação (4.9) sejam reais. A equação (4.9) é uma cúbica em  $\gamma^2$ , à condição necessária para esta redução é que a cúbica.

$$x^3 + \frac{\beta}{2}x^2 + \left(\frac{\beta^2}{16} - \frac{\delta}{4}\right)x - \frac{\gamma^2}{64} = 0, \quad (4.10)$$

tenha todas as raízes reais não negativas. Por outro lado, se (4.10) possui todas as raízes reais não negativas, então  $g(x)$  possui todas as raízes reais. Isso pode ser comprovado utilizando matrizes circulantes através da construção de uma solução do sistema (4.8) para o qual a circulante correspondente é Hermitiana. A conclusão final, em qualquer caso, é a seguinte caracterização: as raízes de  $g(x)$  são todas reais se, e somente se, (4.10) possui todas as raízes reais não negativas.

# Capítulo 5

## Teorema de Fermat

Nosso objetivo neste capítulo é aplicarmos os conhecimentos do Capítulo 4 para provarmos o teorema de Fermat para  $n = 3$ . O leitor interessado em mais detalhes, pode consultar [8, 9].

### 5.1 Teorema de Fermat

Antes de iniciarmos a prova do Teorema de Fermat para  $n = 3$ . Faremos um breve histórico sobre ele.

**Conjectura 5.1 (Fermat-1637)** *Seja  $n \in \mathbb{N}$ . Então a equação*

$$x^n + y^n = z^n, \text{ com } \text{mdc}(x, y) = 1,$$

*não possui soluções não triviais em  $\mathbb{Z}$ , exceto para  $n = 1$  e  $n = 2$ .*

A frase seguinte foi escrita por Fermat sobre a sua conjectura:

*“Eu descobri uma demonstração maravilhosa, mas a margem deste papel é muito pequena para contê-la.”*

- 1660 - Fermat prova para  $n = 4$ .
- 1753 - Euler prova para  $n = 3$ .
- 1825 - Dirichlet e Legendre prova para  $n = 5$ .
- 1839 - Lamé prova para  $n = 7$ .
- 1857 - Kummer prova para  $n \leq 100$ .

- 1983 - Faltings prova que a conjectura pode possuir no máximo uma quantidade finita de soluções.
- 1994 - Wiles prova a conjectura. (358 anos depois - Guinness Book)

Uma *solução trivial* para a conjectura é um termo  $(a, b, c) \in \mathbb{Z}^3$  tal que  $abc = 0$ . Qualquer outra solução é *não trivial*. Observe que se  $(a, b, c) \in \mathbb{Z}^3$  é uma solução não trivial e  $\text{mdc}(a, b) = d$ , então é fácil verificar que  $d$  divide  $c$  e que o terno

$$\left(\frac{a}{d}, \frac{b}{d}, \frac{c}{d}\right)$$

também é uma solução. Assim, podemos supor que

$$\text{mdc}(a, b) = \text{mdc}(a, c) = \text{mdc}(b, c) = 1 \text{ ou } \text{mdc}(a, b, c) = 1.$$

Neste caso, o terno  $(a, b, c)$  chama-se uma *solução primitiva*.

**O caso quando**  $n = 1$ ,  $x + y = z$ . Então é fácil verificar que o terno

$$(k, k + 1, 2k + 1)$$

é uma solução primitiva da equação, para todo  $k \in \mathbb{Z}$ .

**O caso quando**  $n = 2$ ,  $x^2 + y^2 = z^2$ . Então uma solução primitiva  $(a, b, c) \in \mathbb{Z}^3$  também é chamada de *tripla Pitagoriana*, por exemplo,  $(3, 4, 5) \in \mathbb{Z}^3$ , conhecidas desde 582 a.C. Vamos restringir as soluções primitivas a  $\mathbb{N}^3$ . Neste caso,  $x$  e  $y$  não podem ser ambos pares e nem ambos ímpares, pois

$$x^2 \equiv 0 \pmod{4} \text{ e } y^2 \equiv 0 \pmod{4} \Rightarrow z^2 \equiv 0 \pmod{4}$$

e

$$x^2 \equiv 1 \pmod{4} \text{ e } y^2 \equiv 1 \pmod{4} \Rightarrow z^2 \equiv 2 \pmod{4},$$

o que é impossível. Como  $x$  e  $y$  aparecem simetricamente na equação, podemos supor, sem perda de generalidade, que  $y$  é par e  $x$  e  $z$  são ímpares. Note que

$$x^2 + y^2 = z^2 \Leftrightarrow (z + x)(z - x) = y^2 \Leftrightarrow \left(\frac{z + x}{2}\right) \left(\frac{z - x}{2}\right) = \left(\frac{y}{2}\right)^2$$

e a última equação tem sentido, pois  $y$ ,  $z + x$  e  $z - x$  são pares. Assim

$$\text{mdc}\left(\frac{z + x}{2}, \frac{z - x}{2}\right) = 1.$$

De fato, seja

$$d = \text{mdc} \left( \frac{z+x}{2}, \frac{z-x}{2} \right)$$

Então, depois de alguns cálculos,  $d \mid x$  e  $d \mid z$ . Logo,  $d = 1$ , pois  $\text{mdc}(x, z) = 1$ . Pelo Lema 1.15, existem  $a, b \in \mathbb{N}$  tais que

$$\frac{z+x}{2} = b^2, \quad \frac{z-x}{2} = a^2 \quad \text{e} \quad \frac{y}{2} = ab.$$

É fácil verificar que  $a$  e  $b$  possuem paridades distintas, com  $\text{mdc}(a, b) = 1$  e  $b > a > 0$ . Portanto, pelas equações acima, temos que

$$x = b^2 - a^2, \quad y = 2ab \quad \text{e} \quad z = a^2 + b^2, \quad \forall a, b \in \mathbb{N},$$

com  $a$  e  $b$  tendo paridades distintas,  $\text{mdc}(a, b) = 1$  e  $b > a > 0$ , são todas as soluções primitivas da equação. No caso geral,

$$x = \pm(b^2 - a^2), \quad y = \pm 2ab \quad \text{e} \quad z = \pm(a^2 + b^2), \quad \forall a, b \in \mathbb{Z}.$$

É muito importante, de um ponto de vista didático e teórico, uma visão diferente das soluções da conjectura. Como  $y$  é um número par temos que

$$z - x = 2^{2k-1} \quad \text{ou} \quad z - x = 2^{2k-1}d^2, \quad \text{em que } d \mid x \text{ e } \text{mdc}(2, d) = 1.$$

Com efeito, se  $z - x = 2^m$ , para algum  $m \in \mathbb{N}$ , então

$$y^2 = z^2 - x^2 = (x + 2^k) - x^2 = 2^{k+1}(x + 2^{k-1}).$$

Assim,  $x + 2^{k-1}$  é um número ímpar, pois  $x$  é. Logo,  $z - x = 2^{2k-1}$ , para algum  $k$ . Se  $z - x = 2^{2k-1}a$ , para algum  $a \in \mathbb{N}$ , com  $\text{mdc}(2, a) = 1$ , então

$$y^2 = z^2 - x^2 = 2^{2k}a(x + 2^{2k-2}a).$$

Logo, é fácil verificar que  $\text{mdc}(2, a) = 1$ . Pelo Lema 1.15, existe um  $d \in \mathbb{N}$  tal que  $a = d^2$ . Portanto,

$$z - x = 2^{2k-1}d^2, \quad \text{em que } d \mid x \text{ e } \text{mdc}(2, d) = 1.$$

Neste caso,

$$y^2 = z^2 - x^2 = 2 \cdot 2^{2k-1}d^2 \left( \frac{x+z}{2} \right)$$

implica que

$$\frac{x+z}{2} = b^2,$$

para algum  $b \in \mathbb{N}$ . Finalmente, para qualquer  $c \in \mathbb{R}$ , a função  $f : [x, c] \rightarrow \mathbb{R}$  definida como

$$f(t) = t^2$$

é claramente contínua em  $[x, c]$  e derivável em  $(x, c)$ . Assim, pelo Teorema do Valor Médio, existe um  $r \in (x, c)$  tal que

$$c^2 - x^2 = 2r(c - x).$$

Em particular, para  $z = c \in \mathbb{N}$ , obtemos

$$z^2 - x^2 = 2r(z - x) = 2(z - x) \left( x + \frac{z - x}{2} \right).$$

Pondo

$$\theta = \frac{1}{2} \text{ e } h = z - x,$$

teremos a fórmula clássica do Teorema do Valor Médio

$$z^2 - x^2 = 2r(z - x) = 2(z - x)(x + \theta h).$$

Portanto, todas triplas Pitagóricas  $(x, y, z) \in \mathbb{N}^3$ , com  $y$  um número par, satisfazem a equação

$$y^2 = z^2 - x^2 = 2r(z - x) = 2(z - x)(x + \theta h).$$

Por exemplo,

$$z - x = 2^3 \text{ e } \frac{x+z}{2} = 3^2 \Rightarrow 5^2 + 12^2 = 13^2$$

e

$$z - x = 2 \cdot 7^2 \text{ e } \frac{x+z}{2} = 8^2 \Rightarrow 15^2 + 112^2 = 113^2.$$

**O caso quando  $n = 3$ ,  $x^3 + y^3 = z^3$ .** Se  $(a, b, c)$  é uma solução, então  $(-a, -b, -c)$  também o é. Portanto, não há perda de generalidade, em considerarmos as soluções primitivas em  $\mathbb{N}$ , uma vez que o nosso problema é resolver a equação

$$x^3 + y^3 = z^3, \text{ com } \text{mdc}(x, y) = 1, \quad (5.1)$$

**Lema 5.2** *Seja  $(a, b, c) \in \mathbb{N}^3$  uma solução primitiva da equação (5.1).*

1. Se  $\text{mdc}(c, 3) = 1$ , então

$$\text{mdc}(c - a, (c - a)^2 + 3ac) = \text{mdc}(c - b, (c - b)^2 + 3bc) = 1.$$

2. 3 divide exatamente um dos números  $a, b$  e  $c$ .

**Prova.** Vamos provar apenas o item (2). Note que

$$(r + 3q)^3 \equiv r^3 \pmod{9}.$$

Assim, basta considerar  $r = 0, r = 1$  ou  $r = 2$ . Logo,

$$x^3 \equiv 0, 1 \text{ ou } 8 \pmod{9},$$

para todo  $x \in \mathbb{Z}$ . Se  $\text{mdc}(a, 3) = \text{mdc}(b, 3) = 1$ , então

$$a^3 \equiv 1 \text{ ou } 8 \pmod{9} \text{ e } b^3 \equiv 1 \text{ ou } 8 \pmod{9}.$$

Portanto,

$$c^3 = a^3 + b^3 \equiv 0 \pmod{9}.$$

Consequentemente, 3 divide  $c$ . ■

Observe que

$$x^3 + y^3 = z^3 \Leftrightarrow (z - y)[(z - y)^2 + 3yz] = x^3.$$

Assim, se  $\text{mdc}(z, y) = \text{mdc}(z, 3) = 1$ , então, pelo item (1) do Lema 5.2,

$$\text{mdc}(z - y, (z - y)^2 + 3yz) = 1.$$

Logo, pelo o Lema 1.15,  $z - y = u^3$ , para algum  $u \in \mathbb{N}$ , com  $u$  um divisor  $x$ . De modo análogo,

$$\text{mdc}(z - x, (z - x)^2 + 3xz) = 1.$$

Novamente, se  $\text{mdc}(z, x) = \text{mdc}(z, 3) = 1$ , então  $z - x = v^3$ , para algum  $v \in \mathbb{N}$ , com  $v$  um divisor  $y$  e  $\text{mdc}(u, v) = 1$ . Portanto,

$$x = z - v^3 \text{ e } y = z - u^3, \text{ com } \text{mdc}(u, v) = 1.$$

Substituindo  $x$  e  $y$  na equação (5.1) e depois de alguns cálculos, obtemos

$$f(z) = z^3 - 3(u^3 + v^3)z^2 + 3(u^6 + v^6)z - (u^9 + v^9) = 0. \quad (5.2)$$

### 5.1. TEOREMA DE FERMAT

---

Agora, basta provar que o polinômio (equação)  $f(z)$  não possui raízes racionais. Para isto, note que

$$g(z) = f(z + u^3 + v^3) = z^3 - 6u^3v^3z - 3u^3v^3(u^3 + v^3) = 0.$$

Assim,  $p = -6u^3v^3$ ,  $q = -3u^3v^3(u^3 + v^3)$  e, depois de alguns cálculos,

$$27q^2 + 4p^3 = \left(3u^3 - \frac{7}{3}v^3\right)^2 + \frac{32}{9}v^6 > 0,$$

nunca é um quadrado perfeito e  $\Delta > 0$ . Portanto, pelo item (3) do Teorema 4.4,  $g(z)$  não possui raízes racionais, de modo que  $f(z)$  não possui raízes racionais. Donde concluímos que o teorema de Fermat para  $n = 3$ , não possui soluções não triviais em  $\mathbb{N}$ .

## Considerações Finais

Muitas vezes, alunos do ensino médio questionam para que servem as matrizes, assim como, para que servem os números complexos, se estes, inclusive, possuem algo imaginário.

O presente trabalho nos trouxe uma forma de resolver equações polinomiais de graus menores que ou igual a 4, com o auxílio desses conteúdos, algo que não é abordado no ensino básico.

Retomamos conceitos básicos inerentes aos números inteiros, relembramos elementos da álgebra linear e ainda fizemos um breve passeio pela história acerca das técnicas utilizadas pra resolver as equações polinomiais.

Vimos que através de uma equação dada, conseguimos associar uma matriz circulante, que por sua vez estabelece um polinômio característico e um polinômio representante.

Após encontrarmos os coeficientes do polinômio característico e por consequência os elementos da matriz circulante, aplicamos o polinômio representante nas raízes  $n$ -ésimas ( $n \leq 4$ ) da unidade (determinadas através da fórmula de Euler), encontrando assim as raízes da equação polinomial em questão.

Tal procedimento pode ser lecionado ao aluno do ensino médio, de modo a oferecer ao mesmo, novas ferramentas que possibilitam associar diversos conteúdos que por várias vezes aparentam não possuir relação.

A aprendizagem matemática é determinada pela compreensão e apreensão do significado. Com este trabalho, acreditamos que o professor possa ir além de suas aulas normais expositivas, e possibilite o aluno a entender a importância de vários assuntos, abrindo lugar para discussões, o que tornará os conteúdos mais significativos. Espera-se, ainda, que o aluno aprenda a apreciar mais essa ciência, com entusiasmo.

# Referências Bibliográficas

- [1] Boyer, C. B.: *História da Matemática*, Edgard Blücher, Ed. da USP, São Paulo, 1974.
- [2] Garbi, G. G.: *O Romance das Equações Algébricas*, Ed. Livraria da Física, São Paulo, 2007.
- [3] Goldberg, J. L.: *Matrix Theory with Applications*, McGraw-Hill, 1991.
- [4] Hoffman, K. e Kunze, R.: *Álgebra Linear*, Editora Polígono, São Paulo, 1971.
- [5] Kalman, D., White, J. E.: “Polynomial Equation and Circulant Matrices.” *The American Mathematical Monthly*, Vol. 108, N.º9, Nov., 2001, pp. 821-840.
- [6] Niven, I., et al.: *An Introduction to the Theory of Numbers*, John-Wiley, 1991.
- [7] Oliveira Júnior, P. J. S. de: *Equações Polinomiais e Matrizes Circulantes*, TCC, PROFMAT-UFPB, 2015.
- [8] Perera, B. B. U., Piyadasa, R. A. D.: “Proof of Fermat’s Last Theorem for  $n = 3$  using Tschirnhaus Transformation.” *Springer Proceedings in Mathematics and Statistics*, Vol. 124, 2016, pp. 133-136.
- [9] Piyadasa, R. A. D., et al.: “A new interpretation of primitive Pythagorean triples and a conjecture related to Fermat’s last theorem.” *J. Sci. Univ. Kelaniya*, Vol. 3, 2007, pp. 93-58.
- [10] Ribenboim, P.: *Fermat’s Last Theorem for Amateurs*. Springer-Verlag, Inc., 1991.
- [11] Silva, A. de A. e, *Introdução à Álgebra Linear*, Editora Universitária-UFPB, 2007.