Apontamentos de

ÁLGEBRA II

Jorge Picado

Departamento de Matemática

Universidade de Coimbra

2006

Índice

Introdução 1

1. Anéis e corpos 3

Exercícios 17

2. Anéis de polinómios 23

Apêndice 1. Apontamentos para estudo complementar 40

Apêndice 2. Critérios de irredutibilidade 43

Exercícios 47

3. Teoria de Galois 51

Motivação 51

Extensões de corpos 55

Aplicações: construções com régua e compasso 65

Extensões de decomposição 80

Grupo de Galois de um polinómio 95

Exercícios 108

4. Corpos finitos 115

Aplicações: Teoria algébrica de códigos 126

Exercícios 145

Soluções de exercícios 149

Bibliografia 171

Introdução

Estas notas incluem com algum pormenor os principais conceitos e resultados apresentados nas aulas teóricas, completados aqui e acolá com alguns exemplos, observações e exercícios. Espera-se que sejam um auxiliar valioso para o curso, que permita uma maior liberdade nas aulas, na explicação teórica dos assuntos, substituindo uma exposição com grande pormenor formal por uma que realce a motivação e os aspectos intuitivos desses mesmos conceitos e respectivas inter-relações, e que por outro lado sejam um estímulo à atenção e participação activa dos estudantes.

Devem ser encaradas como um mero guião das aulas, e portanto não são um seu substituto. Na sua elaboração baseámo-nos fundamentalmente nos livros [2], [12] (para o Capítulo 3), [8] (para as construções com régua e compasso) e [9] (para o Capítulo 4).

Assumem-se alguns preliminares, nomeadamente:

- matéria dada na disciplina de Álgebra I.
- conhecimentos básicos de Teoria dos Números.
- conhecimentos gerais de Álgebra Linear.
- a "maturidade matemática" que se espera de estudantes do terceiro ano da licenciatura em Matemática.

No desenvolvimento do programa seguir-se-à a recomendação de fundo expressa no programa mínimo da disciplina:

"*... que se faça uma abordagem com um grau de abstracção algo apurado, de acordo com o facto de se tratar de uma disciplina do terceiro ano da licenciatura, mas sem esquecer que a álgebra pode apresentar-se com um olhar nas aplicações, que os seus temas, 'clássicos', ou 'modernos', foram e vão sendo originados por problemas concretos, e que alguns dos seus tópicos mais interessantes têm origem em questões complexas da geometria e da análise. Nesta perspectiva, deverá incluir-se no programa a resolução de problemas clássicos sobre as construções com régua e compasso, a resolução de equações através de radicais e diversas aplicações modernas da teoria dos corpos finitos à teoria dos códigos.*"

1. Anéis e corpos

Uma das características da matemática do último século foi a sua tendência para a abstracção. A teoria moderna dos anéis é um dos frutos dessa abstracção e a forma em que é estudada e ensinada hoje em dia, sendo resultado do trabalho de muitos matemáticos no século XX, tem, no entanto, as suas origens no século XIX, em duas fontes distintas: em Richard Dedekind (1831-1916), que introduziu em 1871 a noção de ideal, no seu trabalho de generalização do Teorema Fundamental da Aritmética (da factorização única em primos) a contextos mais abstractos, e no trabalho de David Hilbert (1862-1945), Edmund Lasker (1868-1941) e F. S. Macaulay (1862-1927) em anéis de polinómios.

O pioneiro no tratamento abstracto da teoria dos anéis foi Adolf Fraenkel (1891-1965) com o artigo "*On the divisors of zero and the decomposition of rings*".[1] Este artigo contém a primeira caracterização axiomática da noção de anel, embora não seja a utilizada hoje em dia. O seu objectivo era sair do estudo particular dos corpos, de modo a obter uma teoria suficientemente geral para poder ser aplicada aos inteiros módulo n, aos números p-ádicos e aos sistemas de "números hipercomplexos". A definição actualmente utilizada de anel (comutativo) parece ter aparecido pela primeira vez em 1917, num artigo do matemático japonês Masazo Sono intitulado "*On congruences*".[2]

O matemático que mais contribuiu para o avanço do ponto de vista abstracto na teoria dos anéis foi uma mulher, Emmy Noether (1882-1935). É costume apontar-se o seu artigo "*Ideal theory in rings*"[3] de 1921 como origem da teoria abstracta dos anéis. O seu tratamento axiomático, muito elegante, constituiu uma novidade ao tempo.[4] Neste artigo, Noether estende o trabalho de Hilbert, Lasker e Macaulay nos anéis de polinómios a anéis mais gerais. Num artigo subsequente, [5] faz num anel abstracto o que Dedekind tinha feito para anéis de números algébricos.

A ideia revolucionária de trabalhar de modo abstracto com anéis e seus ideais — devida a Fraenkel, Sono e Noether — conduziu ao contexto "certo" para o estudo da factorização prima e criou a área que hoje é chamada Álgebra Comutativa. Em 1931 o livro famoso de van der Waerden's[6] colocou todas estas ideias

[1] *Journal für die Reine und Angewandte Mathematik* 145 (1914) 139-176.

[2] *Memoirs of the College of Science of Kyoto* 2 (1917) 203-226.

[3] *Mathematische Annalen* 83 (1921) 24-66.

[4] Nas palavras de Kaplansky, "The importance of this paper is so great that it is surely not much of an exaggeration to call her the mother of modern algebra".

[5] Abstract study of ideal theory in algebraic number- and function-fields, *Mathematische Annalen* 96 (1927) 203-226.

[6] *Modern Algebra*, Springer-Verlag, Berlim, 1931.

à disposição de uma nova geração de algebristas.

Porquê $(-1)(-1) = 1$? Mais geralmente, porquê $(-a)(-b) = ab$? E $a \cdot 0 = 0$? Estas são questões que fazem parte do problema geral de justificação lógica das leis de operação com os números negativos e que nos conduzem aos conceitos de anel (e domínio de integridade).

ANEL

Um *anel* $(A, +, \cdot)$ é um conjunto A com duas operações binárias, que denotaremos por $+$ e $\cdot$, tais que:

(1) $(A, +)$ é um grupo abeliano.

(2) $\cdot$ é associativa; ou seja,

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \text{ para quaisquer } a, b, c \in A.$$

(3) $\cdot$ é distributiva relativamente a $+$; ou seja,

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

e

$$(b + c) \cdot a = b \cdot a + c \cdot a$$

para quaisquer $a, b, c \in A$.

Usaremos simplesmente a letra A para designar um anel arbitrário $(A, +, \cdot)$. Um anel A diz-se *comutativo* se $\cdot$ é comutativa e chama-se *anel com identidade* (ou *anel unitário*) se a operação $\cdot$ possui um elemento neutro (chamado *identidade*) — ou seja, se existe um elemento 1 em A tal que $a \cdot 1 = 1 \cdot a = a$ para qualquer $a \in A$.

Designação	Notação	O que representa
Zero do anel	0	neutro de $+$
Simétrico de $a \in A$	$-a$	inverso de a no grupo $(A, +)$
Múltiplo de $a \in A$	na	$a + a + \cdots + a$ ($n \in \mathbb{Z}$ parcelas)
Identidade do anel	1	neutro de $\cdot$, caso exista
Inverso de $a \in A$	a^{-1}	inverso de a em $(A, \cdot)$, caso exista
Potência de $a \in A$	a^n	$a \cdot a \cdot \cdots \cdot a$ ($n \in \mathbb{Z}^+$ factores)
	a^{-n}	$a^{-1} \cdot a^{-1} \cdot \cdots \cdot a^{-1}$ ($n \in \mathbb{Z}^+$ factores)

Exercício. Verifique, por indução, que, para quaisquer $a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_m$ em A, se tem:

(a) $a(b_1 + b_2 + \cdots + b_m) = ab_1 + ab_2 + \cdots ab_m$.

(b) $(a_1 + a_2 + \cdots + a_n)(b_1 + b_2 + \cdots + b_m) = a_1b_1 + a_1b_2 + \cdots + a_1b_m + a_2b_1 + a_2b_2 + \cdots + a_2b_m + \cdots + a_nb_1 + a_nb_2 + \cdots + a_nb_m$.

Exemplos de anéis:

(1) $(\mathbb{Z}, +, .)$, $(\mathbb{Q}, +, .)$, $(\mathbb{R}, +, .)$ e $(\mathbb{C}, +, .)$.

(2) $(n\mathbb{Z}, +, \cdot)$ $(n = 1, 2, \ldots)$. [para $n \geq 2$ não é unitário]

(3) $(\mathbb{Z}_n, \oplus_n, \otimes_n)$ $(n = 1, 2, \ldots)$. [$\mathbb{Z}_n = \{0\}$ para $n = 1$]

(4) O conjunto $M_n(\mathbb{Z})$ das matrizes quadradas de ordem n ($n \in \mathbb{N}$) com elementos inteiros, munido das operações de adição e multiplicação de matrizes.

[para $n \geq 2$ não é comutativo]

Mais geralmente, $M_n(A)$ para qualquer anel A.

(5) $(\mathcal{P}(X), \Delta, \cap)$ para qualquer conjunto $X \neq \emptyset$.

[recorde: $A\Delta B := (A \cup B) - (A \cap B)$] [$0 = \emptyset$, $1 = X$]

[anel comutativo com identidade]

[observe: $A\Delta A = \emptyset$, $A \cap A = A$]

Proposição 1.1 *Seja A um anel. Para quaisquer $a, b \in A$ tem-se:*

(a) $a \cdot 0 = 0 \cdot a = 0$.

(b) $(-a)b = a(-b) = -(ab)$.

(c) $(-a)(-b) = ab$.

Demonstração. (a) $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, o que implica, pela lei do cancelamento válida em qualquer grupo, $a \cdot 0 = 0$. Analogamente, $0 \cdot a = 0$.
(b) Usando a alínea (a), $ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$, donde $(-a)b = -(ab)$. Analogamente, $a(-b) = -(ab)$.
(c) Pela alínea (b) tem-se $(-a)(-b) = -(a(-b)) = -(-(ab))$. Mas, em qualquer grupo, $-(-(ab)) = ab$. Logo $(-a)(-b) = ab$. ∎

Assumiremos sempre que num anel com identidade $1 \neq 0$. Com efeito, por 1.1(a), se $0 = 1$ então, para qualquer $a \in A$, $a = a \cdot 1 = a \cdot 0 = 0$ e o anel A reduz-se ao caso trivial $A = \{0\}$.
Em

- $\mathbb{Z}$: $ab = 0 \Rightarrow a = 0$ ou $b = 0$
- $\mathbb{Z}_6$: $2 \cdot 3 = 2 \otimes_6 3 = 0$
- $M_2(\mathbb{Z})$: $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Um elemento $a \in A$, diferente de zero, diz-se *divisor de zero* caso exista $b \in A$, diferente de zero, tal que $ab = 0$ ou $ba = 0$. No primeiro caso diremos, mais especificamente, que o divisor de zero é um *divisor de zero à esquerda*, e no segundo caso que é um *divisor de zero à direita*.

[Portanto, $\mathbb{Z}$ não tem divisores de zero, enquanto $\mathbb{Z}_6$ e $M_2(\mathbb{Z})$ têm]

Quando é que a *lei do cancelamento para o produto*

$$\forall a, b, c \in A \; [c \neq 0 \text{ e } (ac = bc \text{ ou } ca = cb) \Rightarrow a = b]$$

é válida num anel? Precisamente quando A não tem divisores de zero.

[Exercício: Verifique]

DOMÍNIO DE INTEGRIDADE

Um *domínio de integridade* é um anel comutativo com identidade $A \neq \{0\}$ sem divisores de zero (ou equivalentemente, onde a lei do cancelamento para o produto é válida).

Em

- $\mathbb{Z}$: só 1 e -1 são invertíveis para a operação $\cdot$
- $\mathbb{Q}$: todos os elementos $\neq 0$ têm inverso.

Chama-se *unidade* do anel a qualquer elemento que tenha inverso. Designando por A^* o conjunto das unidades de A, é evidente que $(A^*, \cdot)$ constitui um grupo.

[Exercício: Verifique]

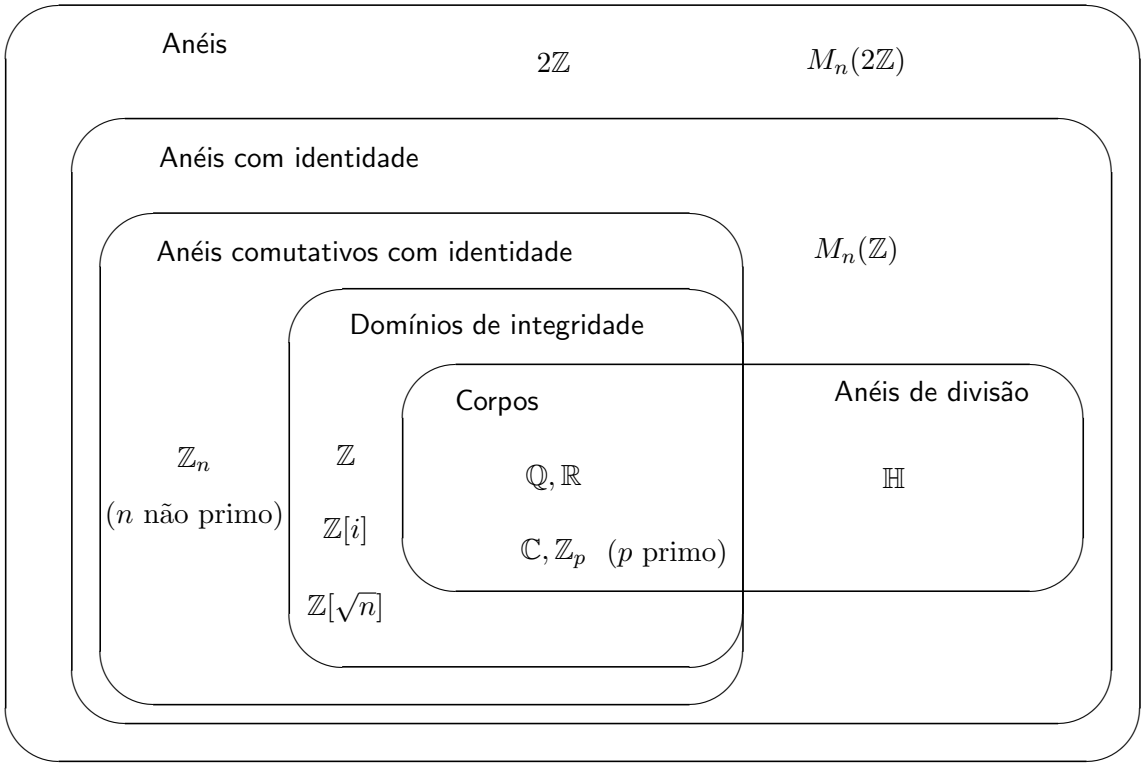
ANEL DE DIVISÃO E CORPO

Um *anel de divisão* é um anel A com identidade tal que $A^* = A - \{0\}$. A um anel de divisão comutativo chama-se *corpo*. Portanto, um corpo é um anel comutativo com identidade onde todo o elemento $\neq 0$ possui inverso.

Todo o corpo é um domínio de integridade. Com efeito, se a tem inverso então não é divisor de zero:

$$ab = 0 \Leftrightarrow a^{-1}(ab) = a^{-1} \cdot 0 \Leftrightarrow b = 0.$$

Em conclusão:



$\mathbb{Z}$ é um exemplo de domínio de integridade que não é corpo. Nenhum exemplo destes pode ser finito:

Teorema 1.2 *Todo o domínio de integridade finito é um corpo.*

Demonstração. Seja $D = \{0, d_1, d_2, \ldots, d_n\}$ um domínio de integridade finito. Para cada $i \in \{1, 2, \ldots, n\}$ consideremos os produtos $d_i d_1, d_2 d_2, \ldots, d_i d_n$. São distintos dois a dois: $d_i d_j = d_i d_k \Leftrightarrow d_i(d_j - d_k) = 0$; como $d_i \neq 0$ e D não tem divisores de zero, necessariamente $d_j - d_k = 0$, isto é, $d_j = d_k$.

Assim, os produtos $d_i d_1, d_2 d_2, \ldots, d_i d_n$ percorrem todos os elementos não nulos de D; em particular, existe j tal que $d_i d_j = 1$, o que significa que d_i é invertível. Portanto, todo o elemento não nulo de D é invertível, logo D é um corpo. ∎

SUBANEL

$S \subseteq A$ é um *subanel* de A se S é fechado para $+$ e $\cdot$ e forma um anel para estas operações.

Exemplos: $2\mathbb{Z}$, $3\mathbb{Z}$, $4\mathbb{Z}$, ... são subanéis de $(\mathbb{Z}, +, \cdot)$.

Qualquer anel A possui sempre os *subanéis triviais* $\{0\}$ e o próprio A. Qualquer outro subanel de A diz-se *subanel próprio*.

Proposição 1.3 *Um subconjunto S de um anel A é um subanel se e só se as seguintes condições se verificam:*

(1) $S \neq \emptyset$.

(2) *Para cada* $x, y \in S$, $x - y \in S$.

(3) *Para cada* $x, y \in S$, $xy \in S$.

Demonstração. Exercício. ∎

Mais exemplos:

- $\mathbb{Z}[\sqrt{-5}] := \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$ é um subanel de $(\mathbb{C}, +, \cdot)$.
- $\left\{ \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} : a \in \mathbb{Z} \right\}$ é um subanel de $M_2(\mathbb{Z})$.

IDEAL

Um subanel I de A diz-se um *ideal* se, para cada $a \in A$ e cada $x \in I$, ax e xa pertencem a I.

Exemplos:

- $\mathbb{Z}$ é um subanel de $\mathbb{Q}$ mas não é um ideal ($1 \cdot \frac{1}{2} = \frac{1}{2} \notin \mathbb{Z}$)
- $n\mathbb{Z}$ é um ideal de $\mathbb{Z}$ ($n \in \mathbb{N}_0$).

`[Observe o paralelismo com a teoria dos grupos: os subanéis correspondem aos subgrupos e os ideais correspondem aos subgrupos normais]`

Da proposição anterior decorre imediatamente que:

Proposição 1.4 *Um subconjunto I de um anel A é um ideal se e só se as seguintes condições se verificam:*

(1) $I \neq \emptyset$.

(2) *Para cada* $x, y \in I$, $x - y \in I$.

(3) *Para cada* $a \in A$ *e* $x \in I$, $ax \in I$ *e* $xa \in I$. ∎

Mais exemplos: Seja A um anel comutativo e $a \in A$.

- $\{xa \mid x \in A\}$ é um ideal de A. `[pode não conter a]`
- O menor ideal de A contendo a é o ideal

$$\langle a \rangle := \{xa + na \mid x \in A, n \in \mathbb{Z}\}.$$

Diz-se o *ideal principal gerado* por a. Se A for também unitário,

$$\langle a \rangle = \{xa \mid x \in A\}.$$

Seja A um anel comutativo. Um ideal I de A diz-se *principal* se existe algum $a \in A$ tal que $I = \langle a \rangle$.

Exemplo: Em Álgebra I observaram que os subconjuntos $n\mathbb{Z}$, $n = 0, 1, 2, \ldots$, são os únicos subgrupos de $(\mathbb{Z}, +)$. Portanto, $n\mathbb{Z}$, $n = 0, 1, 2, \ldots$, são os únicos ideais de $(\mathbb{Z}, +, \cdot)$. Como $n\mathbb{Z} = \langle n \rangle$, são todos principais.

`[ℤ diz-se um domínio de ideais principais]`

Seja I um ideal de um anel $(A, +, \cdot)$. Como $(I, +)$ é um subgrupo normal do grupo abeliano $(A, +)$, sabemos da Álgebra I que o conjunto A/I das classes laterais $a + I := \{a + x \mid x \in A\}$, $a \in A$, forma um grupo abeliano (o chamado *grupo quociente*) para a operação

$$(a + I) + (b + I) := (a + b) + I.$$

Exercício. Dois elementos a e b de A dizem-se *congruentes* módulo I (e escreve-se $a \equiv b \mod I$) se pertencem à mesma classe lateral, ou seja, $a + I = b + I$. Mostre que $a \equiv b \mod I$ implica $a + x \equiv b + x \mod I$, $ax \equiv bx \mod I$, e $xa \equiv xb \mod I$ para qualquer $x \in A$ e $na \equiv nb \mod I$ para qualquer $n \in \mathbb{Z}$.

`[Recorde:` $a + I = b + I$ `sse` $a - b \in I$`]`

Mas agora, no contexto dos anéis, temos mais estrutura em A/I:

$$(a + I)(b + I) := ab + I \tag{1.4.1}$$

define outra operação em A/I. Com efeito, se $a+I=c+I$ e $b+I=d+I$ então

$$\left.\begin{array}{l} a+I=c+I \Leftrightarrow a-c\in I \overset{(*)}{\Rightarrow} (a-c)b\in I \Leftrightarrow ab-cb\in I \\ \\ b+I=d+I \Leftrightarrow b-d\in I \overset{(*)}{\Rightarrow} c(b-d)\in I \Leftrightarrow cb-cd\in I \end{array}\right\} \Rightarrow ab-cd\in I,$$

isto é, $ab+I=cd+I$.

[Observe: a condição 3 na definição de ideal é decisiva no passo (*): se I for somente um subanel, (1.4.1) pode não definir uma operação em A/I]

Proposição 1.5 *A/I forma um anel relativamente às operações*

$$(a+I)+(b+I):=(a+b)+I,$$

$$(a+I)(b+I):=ab+I.$$

Demonstração. $(A/I,+)$ é um grupo abeliano (Álgebra I) e decorre imediatamente da definição do anel A que a operação $\cdot$ de A/I é associativa e é distributiva relativamente à adição. ∎

O anel $(A/I,+,\cdot)$ chama-se *anel quociente* de A por I. É evidente que se A é comutativo então A/I também é comutativo e se A tem identidade 1 então A/I também tem identidade (o elemento $1+I$).

Exemplo: $\mathbb{Z}/\langle 5\rangle$ tem 5 elementos:

$$0+\langle 5\rangle, 1+\langle 5\rangle, 2+\langle 5\rangle, 3+\langle 5\rangle, 4+\langle 5\rangle, 5+\langle 5\rangle = 0+\langle 5\rangle, 6+\langle 5\rangle = 1+\langle 5\rangle, \ldots$$

$$-1+\langle 5\rangle = 4+\langle 5\rangle, -2+\langle 5\rangle = 3+\langle 5\rangle, \ldots$$

Identifiquemo-los simplesmente por $[0],[1],[2],[3]$ e $[4]$, respectivamente.
As tabelas das operações do anel $\mathbb{Z}/\langle 5\rangle$ são então:

$+$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[0]$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[1]$	$[1]$	$[2]$	$[3]$	$[4]$	$[0]$
$[2]$	$[2]$	$[3]$	$[4]$	$[0]$	$[1]$
$[3]$	$[3]$	$[4]$	$[0]$	$[1]$	$[2]$
$[4]$	$[4]$	$[0]$	$[1]$	$[2]$	$[3]$

$\cdot$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$	$[3]$	$[4]$
$[2]$	$[0]$	$[2]$	$[4]$	$[1]$	$[3]$
$[3]$	$[0]$	$[3]$	$[1]$	$[4]$	$[2]$
$[4]$	$[0]$	$[4]$	$[3]$	$[2]$	$[1]$

[É um corpo]

Mais geralmente, para cada $n \in \mathbb{N}$, os elementos de $\mathbb{Z}/\langle n\rangle$ são

$$[0] := 0 + \langle n\rangle\,, [1] := 1 + \langle n\rangle\,, \ldots, [n-1] := n - 1 + \langle n\rangle\,.$$

Em geral, é um anel comutativo com identidade [1]. É um corpo se e só se n é primo.

[Recorde: $(\mathbb{Z}_n - \{0\}, \otimes_n)$ é um grupo sse n é primo]

Por exemplo, para $n = 6$ existem divisores de zero: $[2] \cdot [3] = [0]$. Este exemplo mostra que as propriedades do anel A não são necessariamente herdadas pelo anel quociente: $\mathbb{Z}$ é um domínio de integridade mas $\mathbb{Z}/\langle 6\rangle$ não é.

Seja A um anel comutativo com identidade. Vejamos quais ideais dão origem a anéis quociente que são domínios de integridade ou corpos.

IDEAL PRIMO

Um ideal $P \neq A$ do anel A chama-se *primo* se, para quaisquer $a, b \in A$, $ab \in P$ implica $a \in P$ ou $b \in P$.

Exemplos: Seja $A = \mathbb{Z}$. O ideal $\langle 6\rangle$ não é um ideal primo: $3 \cdot 2 = 6 \in \langle 6\rangle$ mas $3 \notin \langle 6\rangle$ e $2 \notin \langle 6\rangle$. Por outro lado, $\langle 5\rangle$ é um ideal primo:

$$ab \in \langle 5\rangle \Leftrightarrow 5|ab \Rightarrow 5|a \text{ ou } 5|b \Leftrightarrow a \in \langle 5\rangle \text{ ou } b \in \langle 5\rangle\,.$$

[Caso geral: para $n \geq 1$, $\langle n\rangle$ é primo sse n é primo]

$\langle 0\rangle = \{0\}$ é evidentemente um ideal primo de $\mathbb{Z}$. Com efeito, é óbvio que num anel A comutativo com identidade, $\langle 0\rangle$ é primo se e só se A não tem divisores de zero.

IDEAL MAXIMAL

Um ideal $M \neq A$ do anel A chama-se *maximal* se, para qualquer ideal I de A, a propriedade $M \subseteq I$ implica $I = M$ ou $I = A$.

Exemplos: No anel dos inteiros $\mathbb{Z}$, $\langle 0\rangle$ e $\langle 10\rangle$ não são maximais:

$$\langle 0\rangle \subset \langle 10\rangle \subset \langle 5\rangle \subset \mathbb{Z}.$$

[Observe: O exemplo $\langle 0\rangle$ mostra que, em geral, primo $\not\Rightarrow$ maximal]

Por outro lado, $\langle 5\rangle$ é maximal:

$$\langle 5\rangle \subseteq \langle m\rangle \subseteq \mathbb{Z} \Leftrightarrow m|5 \Rightarrow m = 1 \text{ ou } m = 5 \Leftrightarrow \langle m\rangle = \mathbb{Z} \text{ ou } \langle m\rangle = \langle 5\rangle\,.$$

[Caso geral: para $n \geq 1$, $\langle n\rangle$ é maximal sse n é primo]

Finalmente, temos:

Teorema 1.6 *Seja A um anel comutativo com identidade e I um ideal de A. Então:*

(a) *A/I é um domínio de integridade se e só se I é primo.*

(b) *A/I é um corpo se e só se I é maximal.*

(c) *Todo o ideal maximal de A é primo.*

Demonstração. Já sabemos que A/I é um anel comutativo com identidade $1+I$.
(a) Portanto, A/I será um domínio de integridade sse

$$\begin{cases} 1+I \neq 0+I & (*) \\ \\ (a+I)(b+I) = I \text{ implica } a \in I \text{ ou } b \in I. & (**) \end{cases}$$

Mas
$(*) \Leftrightarrow 1 \notin I \Leftrightarrow I \neq A$

`[Verifique: para qualquer ideal` I`,` $1 \in I \Leftrightarrow I = A$`]`

$(**) \Leftrightarrow ab + I = I$ implica $a \in I$ ou $b \in I \Leftrightarrow ab \in I$ implica $a \in I$ ou $b \in I$,

pelo que $(*)$ e $(**)$ significam precisamente que I é primo.
(b) Agora, A/I será um corpo sse

$$\begin{cases} 1+I \neq 0+I & (*) \\ \\ \text{qualquer } a+I \neq I \text{ é invertível.} & (**) \end{cases}$$

Mas
$(**) \Leftrightarrow$ para cada $(a+I) \neq I$ existe $(b+I) \neq I$ tal que $(a+I)(b+I) = 1+I \Leftrightarrow$ para cada $a \in A - I$ existe $b \in A - I$ tal que $ab + I = 1 + I \Leftrightarrow$ para cada $a \in A - I$ existe $b \in A - I$ tal que $ab - 1 \in I$.

Bastará agora observarmos que esta última condição é equivalente a

$$J \text{ ideal de } A,\ I \subset J \subseteq A \Rightarrow J = A,$$

para concluirmos que $(*)$ e $(**)$ significam que I é maximal:

("$\Rightarrow$") Seja então $a \in J - I$. Por hipótese, existe $b \in A - I$ tal que $ab - 1 \in I \subset J$. Como $ab \in J$, então $1 \in J$, logo $J = A$.

("$\Leftarrow$") Reciprocamente, para cada $a \in A - I$ consideremos o menor ideal que contém $I \cup \{a\}$ (o chamado *ideal gerado* por $I \cup \{a\}$), ou seja, o ideal

$$J_a := \{xa + y \mid x \in A, y \in I\}.$$

[Verifique: $\{xa + y \mid x \in A, y \in I\}$ é um ideal de A]

É evidente que $I \subset J_a \subseteq A$ logo, por hipótese, $J_a = A$. Em particular, $1 \in J_a$, ou seja, 1 é um dos elementos $xa + y$ de J_a. Mas $1 = xa + y \Leftrightarrow xa - 1 = -y \in I$. Provámos assim que, para cada $a \in A - I$, existe $b \in A - I$ tal que $ab - 1 \in I$.
(c) É consequência imediata de (b) e (a): Se I é maximal, A/I é um corpo e, em particular, um domínio de integridade, logo I é primo. ∎

Exemplo de aplicação do Teorema: No caso $A = \mathbb{Z}$, $I = \langle 5 \rangle$ é, como vimos, maximal; daí o facto de $\mathbb{Z}/\langle 5 \rangle$ ser um corpo, como tínhamos observado anteriormente.

Outras aplicações: No próximo capítulo, aos anéis de polinómios.

A definição das operações no anel quociente A/I garante que a passagem de A a A/I preserva as operações do anel. Com efeito, a aplicação

$$\begin{array}{rcl} p : A & \to & A/I \\ a & \mapsto & a + I \end{array}$$

satisfaz, pela maneira como definimos as operações em A/I, as propriedades

$$p(a + b) = p(a) + p(b)$$

$$p(ab) = p(a)p(b),$$

para quaisquer $a, b \in A$.

HOMOMORFISMO DE ANÉIS
Sejam A e B dois anéis. Uma aplicação $f : A \to B$ diz-se um homomorfismo de anéis se, para quaisquer $a, b \in A$, $f(a + b) = f(a) + f(b)$ e $f(ab) = f(a)f(b)$.

Portanto, $p : A \to A/I$ é um homomorfismo, claramente sobrejectivo.

APLICAÇÃO 1: Critérios de divisibilidade para os inteiros

Vejamos outro exemplo de homomorfismo. Consideremos a aplicação $f_m : \mathbb{Z} \to \mathbb{Z}_m$ do anel $(\mathbb{Z}, +, \cdot)$ no anel $(\mathbb{Z}_m, \oplus_m, \otimes_m)$ que a cada inteiro a faz corresponder a mod m, isto é, o resto da divisão de a por m.

[Verifique: f_m é um homomorfismo de anéis]

Seja $a = a_n a_{n-1} \cdots a_1 a_0$ um inteiro com $n+1$ algarismos, escrito na base decimal. Como $a = 10^n a_n + 10^{n-1} a_{n-1} + \cdots + 10a_1 + a_0$, então, usando o facto de que f_m é um homomorfismo de anéis, temos

$$f_m(a) = f_m(10^n) \otimes f_m(a_n) \oplus f_m(10^{n-1}) \otimes f_m(a_{n-1}) \oplus \cdots \oplus f_m(10) \otimes f_m(a_1) \oplus f_m(a_0)$$

No caso $m = 9$, como $f_9(10^n) = 1$, para qualquer natural n, obtemos

$$\begin{aligned} f_9(a) &= f_9(a_n) \oplus f_9(a_{n-1}) \oplus \cdots \oplus f_9(a_1) \oplus f_9(a_0) \\ &= f_9(a_n + a_{n-1} + \cdots + a_1 + a_0), \end{aligned}$$

o que mostra que $a \equiv a_n + a_{n-1} + \cdots + a_1 + a_0 \pmod 9$. Portanto,

um inteiro é divisível por 9 sse a soma dos seus algarismos o é.

Como também $f_3(10^n) = 1$, o mesmo critério vale para o 3:

um inteiro é divisível por 3 sse a soma dos seus algarismos o é.

Temos agora uma receita para obter critérios úteis de divisibilidade por m, desde que $f_m(10^n)$ seja dado por uma expressão simples:

m=11:

$$f_{11}(10^n) = \begin{cases} 1 & \text{se } n \text{ é par} \\ -1 & \text{se } n \text{ é ímpar} \end{cases}$$

pelo que

$a_n a_{n-1} \cdots a_1 a_0$ é divisível por 11 sse $(-1)^n a_n + (-1)^{n-1} a_{n-1} + \cdots - a_1 + a_0$ o é.

m=2,5: nestes casos $f_m(10^n) = 0$ logo

$a_n a_{n-1} \cdots a_1 a_0$ é divisível por 2 (resp. 5) sse a_0 o é.

m=4:

$$f_4(10^n) = \begin{cases} 2 & \text{se } n = 1 \\ 0 & \text{se } n \geq 2 \end{cases}$$

logo

$a_n a_{n-1} \cdots a_1 a_0$ é divisível por 4 sse $2a_1 + a_0$ o é.

m=6: $f_6(10^n) = 4$ logo

$a_n a_{n-1} \cdots a_1 a_0$ é divisível por 6 sse $4a_n + 4a_{n-1} + \cdots + 4a_1 + a_0$ o é.

Estes exemplos ilustram bem a ideia de como um homomorfismo de anéis, bem escolhido, permite transferir um problema num determinado anel (no caso presente, saber se um inteiro é divisível por um determinado m) para outro anel, onde se torna mais fácil de resolver.

APLICAÇÃO 2: Prova dos nove

Consideremos novamente o homomorfismo $f_9 : \mathbb{Z} \to \mathbb{Z}_9$. Como se trata de um homomorfismo, então

$$a \cdot b = c \Rightarrow f_9(a) \otimes_9 f_9(b) = f_9(c). \tag{1.6.1}$$

Portanto, se $f_9(a) \otimes_9 f_9(b) \neq f_9(c)$, necessariamente $a \cdot b \neq c$. Por exemplo, 27×12 não é igual a 334 pois $f_9(334) = 1$ (ou seja, "334 noves fora" é igual a 1) enquanto $f_9(27) = 0$ e $f_9(12) = 3$ (ou seja, "27 noves fora" é igual a 0 e "12 noves fora" é igual a 3). De facto, $27 \times 12 = 324$. Esta é a "prova dos nove" ensinada na escola primária.

[Cuidado: O recíproco de (1.6.1) não é válido (por exemplo, $f_9(378) = 0$ mas $27 \times 12 \neq 378$); portanto, se a prova dos nove numa multiplicação der certa não significa que a multiplicação esteja certa.]

As funções também permitem transferir a estrutura de uma álgebra para um conjunto sem estrutura. Por exemplo, seja f a função do anel quociente $\mathbb{Z}/\langle p\rangle$ no conjunto $\mathbb{Z}_p = \{0, 1, 2, \ldots, p-1\}$ que a cada $a + I$ faz corresponder $a \mod p$.

[Verifique: f é uma bijecção]

Então $\mathbb{Z}_p$ herda a estrutura de $\mathbb{Z}/\langle p\rangle$ se definirmos em $\mathbb{Z}_p$ as operações

$$a \oplus b = f(a+I) \oplus f(b+I) := f((a+I)+(b+I)) = f(a+b+I) = (a+b) \mod p$$

(isto é, a adição módulo p) e

$$a \otimes b = f(a+I) \otimes f(b+I) := f((a+I)(b+I)) = f(ab+I) = ab \mod p$$

(a multiplicação módulo p). $\mathbb{Z}_p$ com esta estrutura herdada de $\mathbb{Z}/\langle p\rangle$ é um corpo finito e f é um homomorfismo bijectivo.

[Veremos no último capítulo do curso que todo o corpo finito é necessariamente de ordem p^n para algum primo p e algum natural n e que para cada p^n existe precisamente um corpo (a menos de isomorfismo) de ordem p^n. Este corpo chama-se *corpo de Galois* de ordem p^n e denota-se por $\mathbb{F}_{p^n}$. Assim, $\mathbb{F}_p = \mathbb{Z}_p$.]

ISOMORFISMO DE ANÉIS

A um homomorfismo de anéis bijectivo chama-se *isomorfismo.*

Portanto, f é um isomorfismo de corpos.
Por exemplo, por f, as tabelas das operações em $\mathbb{Z}/\langle 5\rangle$ são transformadas em

$\oplus_5$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\otimes_5$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

e $(\mathbb{Z}/\langle 5\rangle, +, \cdot)$ é um corpo isomorfo a $(\mathbb{Z}_5, \oplus_5, \otimes_5)$.

CARACTERÍSTICA

Seja A um anel com identidade. Se existir algum $n \in \mathbb{N}$ tal que $n1 = 0$, ao menor deles chama-se *característica* de A e diz-se que A tem *característica positiva.* Se tal n não existe, diz-se que A tem característica 0.

(Uma vez que $n1 = 0$ sse $na = 0$ para qualquer $a \in A$, podemos dizer que a característica de A é igual ao menor natural n, caso exista algum, tal que $na = 0$ para todo o $a \in A$, ou, caso contrário, igual a 0; como esta condição alternativa não depende da identidade, toma-se para definição de característica no caso geral de um anel sem necessariamente identidade.)

[Verifique: $n1 = 0$ sse $na = 0$ para qualquer $a \in A$]

Proposição 1.7 *Todo o domínio de integridade com característica positiva tem característica prima.*

Demonstração. Seja D um domínio de integridade com característica positiva $n \geq 1$. Como $1 \neq 0$, $n \geq 2$. Se n não fosse um primo então $n = rs$ para algum par de inteiros satisfazendo $1 < r, s < n$, o que implicaria $0 = n1 = (rs)1 = (r1)(s1)$. Como D não tem divisores de zero, seria $r1 = 0$ ou $s1 = 0$, um absurdo uma vez que n é o menor natural tal que $n1 = 0$. ∎

[Observe: a comutatividade do anel não é relevante para esta prova]

Corolário 1.8 *Todo o domínio de intregridade finito tem característica prima.*

Demonstração. Seja C um domínio de intregridade finito. Pela proposição anterior, bastará provarmos que a característica de C é positiva. Para isso, consideremos os elementos

$$1, 1+1, 1+1+1, \ldots$$

de C. Como C é finito, esta lista é finita, pelo que $r1 = s1$ para alguns naturais r, s tais que $1 \leq r < s$. Consequentemente, $(s-r)1 = 0$, o que mostra que a característica de C não é zero. ∎

Proposição 1.9 *Seja A um anel comutativo de característica prima p. Então, para quaisquer $a, b \in A$ e $n \in \mathbb{N}$:*

(a) $(a+b)^{p^n} = a^{p^n} + b^{p^n}$.

(b) $(a-b)^{p^n} = a^{p^n} - b^{p^n}$.

Demonstração. (a) Provaremos só o caso $n = 1$ (uma simples indução sobre n completa a prova). Pela fórmula do Teorema Binomial, válido em qualquer anel comutativo,

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

Como cada $\binom{p}{i}$, $0 < i < p$, que é um inteiro, é igual a

$$\frac{p(p-1)\cdots(p-i+1)}{1 \cdot 2 \cdot \cdots \cdot i}$$

então $1 \cdot 2 \cdot \cdots \cdot i$ divide $p(p-1)\cdots(p-i+1)$. Mas p é primo e $i < p$ logo $1 \cdot 2 \cdot \cdots \cdot i$ divide $(p-1)\cdots(p-i+1)$. Assim, $\binom{p}{i} \equiv 0 \mod p$. Em conclusão, $(a+b)^p = a^p + b^p$.
(b) Basta observar que, pela alínea (a), $a^{p^n} = ((a-b)+b)^{p^n} = (a-b)^{p^n} + b^{p^n}$. ∎

Exercícios

1.1. Averigúe se os seguintes conjuntos têm estrutura de anel para as operações indicadas. Em caso afirmativo, verifique se têm identidade, divisores de zero e estrutura de corpo.

(a) $(\mathbb{Z}_n, \oplus_n, \otimes_n)$, onde $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$, com n número natural fixo, e $\oplus_n$ e $\otimes_n$ denotam respectivamente a adição e multiplicação módulo n.

(b) $(\mathcal{M}_n(\mathbb{K}), +, \times)$, onde $\mathcal{M}_n(\mathbb{K})$, com n número natural fixo, é o conjunto das matrizes quadradas de ordem n com elementos num corpo $\mathbb{K}$, e $+$ e $\times$ denotam a adição e multiplicação usuais de matrizes, respectivamente.

(c) $(\mathcal{P}(X), \cup, \cap)$.

(d) $(\mathcal{P}(X), \Delta, \cap)$, onde $\mathcal{P}(X)$ é o conjunto das partes de um conjunto não vazio X e

$$A\Delta B = (A \cup B) - (A \cap B),\ \forall A, B \in \mathcal{P}(X).$$

(e) $(\mathbb{Q} - \{0\}, \times, +)$, sendo $\times$ e $+$ a multiplicação e adição usuais de números racionais.

(f) $(A, \oplus, \otimes)$, sendo $(A, +, .)$ um anel com identidade (que denotamos por 1) e

$$a \oplus b = a + b + 1,\ \forall a, b \in A,$$

$$a \otimes b = a + b + a.b,\ \forall a, b \in A.$$

(g) $(\mathbb{Z}[i], +, \times)$, sendo $\mathbb{Z}[i] = \{a + ib \mid a, b \in \mathbb{Z}\}$ o conjunto dos inteiros de Gauss e $+$ e $\times$ a adição e a multiplicação usuais de números complexos.

1.2. Quais das seguintes propriedades são válidas num anel arbitrário A? E num anel comutativo arbitrário?

(a) $a^m a^n = a^{m+n},\ \forall a \in A,\ \forall m, n \in \mathbb{N}$.

(b) $(a^m)^n = a^{mn},\ \forall a \in A,\ \forall m, n \in \mathbb{N}$.

(c) $(ab)^m = a^m b^m,\ \forall a, b \in A,\ \forall m \in \mathbb{N}$.

1.3. Seja A um anel com identidade 1 e não tendo divisores de zero. Para $a, b \in A$ verifique que:

(a) $ab = 1$ se e só se $ba = 1$.

(b) Se $a^2 = 1$ então ou $a = 1$ ou $a = -1$.

1.4. Sejam a e b dois elementos de um anel comutativo A com identidade. Se $n \in \mathbb{Z}^+$, deduza a expressão binomial

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i, \quad \text{onde } \binom{n}{i} = \frac{n!}{i!(n-i)!}\ .$$

1.5. Sendo A um anel e $a \in A - \{0\}$, prove que

$$a \text{ não é um divisor de zero à esquerda} \Leftrightarrow \forall b, c \in A(ab = ac \Rightarrow b = c).$$

1.6. Seja D um domínio de integridade. Para as afirmações seguintes, escreva uma prova se a afirmação é verdadeira, senão apresente um contra-exemplo:

(a) $a^2 = 1 \Rightarrow a = 1$ ou $a = -1$.

(b) $-1 \neq 1$.

(c) $a \neq 0, ab = ac \Rightarrow b = c$.

1.7. Um elemento a de um anel A diz-se *idempotente* se $a^2 = a$ e *nilpotente* se $a^n = 0$ para algum $n \in \mathbb{N}$. Mostre que:

(a) Um elemento idempotente diferente de zero não pode ser nilpotente.

(b) Qualquer elemento nilpotente diferente de zero é um divisor de zero.

1.8. Dados $a, b \in \mathbb{Z}_5$, resolva em $\mathbb{Z}_5$ o sistema

$$\begin{cases} x + 2y = a \\ -3x + 3y = b. \end{cases}$$

1.9. Averigúe quais dos seguintes conjuntos são subanéis ou ideais dos anéis indicados e, sempre que possível, determine o anel quociente.

(a) O conjunto dos inteiros pares em $(\mathbb{Z}, +, \times)$.

(b) O conjunto dos inteiros ímpares em $(\mathbb{Z}, +, \times)$.

(c) O conjunto dos números reais de forma $a + b\sqrt{2}$, com $a, b \in \mathbb{Z}$, em $(\mathbb{R}, +, \times)$.

(d) O conjunto dos números complexos da forma ib, com $b \in \mathbb{R}$, em $(\mathbb{C}, +, \times)$.

(e) O conjunto dos números inteiros em $(\mathbb{Q}, +, \times)$.

1.10. Verifique que $\mathbb{Z} \times \{0\}$ é um subanel de $(\mathbb{Z} \times \mathbb{Z}, +, \times)$ e que $\mathbb{Z} \times \{0\}$ tem identidade diferente da identidade de $(\mathbb{Z} \times \mathbb{Z}, +, \times)$.

1.11. Determine os ideais do anel $\mathbb{Z}_n$ para

(a) $n = 4$; (b) $n = 11$; (c) $n = 12$; (d) $n = 16$.

1.12. Chama-se *centro* de um anel A ao conjunto $\{x \in A \mid xa = ax, \forall a \in A\}$. Mostre que o centro de A é um subanel do anel A. Será um ideal?

1.13.

(a) Qual é o menor subanel de $\mathbb{Z}$ que contém o 3? E o menor ideal?

(b) Qual é o menor subanel de $\mathbb{R}$ que contém o $\frac{1}{2}$? E o menor ideal?

1.14. Considere o anel $\mathbb{Z}$ dos números inteiros.

(a) Prove que o ideal gerado por $p \in \mathbb{N} - \{1\}$ é um ideal primo se e só se p é um número primo.

(b) Determine o ideal gerado por $\{a, b\} \subset \mathbb{N}$, com $m.d.c.(a, b) = 1$.

1.15. Sejam D um domínio de integridade e a e b elementos de D. Mostre que $\langle ab \rangle \subseteq \langle a \rangle$ e indique uma condição necessária e suficiente para que $\langle ab \rangle = \langle a \rangle$.

1.16. Seja A o anel $(\mathbb{R}^{\mathbb{R}}, +, \cdot)$ das funções reais de variável real, onde

$$(f+g)(x) = f(x) + g(x) \text{ e } (f \cdot g)(x) = f(x) \cdot g(x).$$

(a) Determine os divisores de zero de A.

(b) Mostre que $I = \{f \in A \mid f(5) = 0\}$ é um ideal de A. É primo?

1.17. Considere os ideais $\langle 2 \rangle$, $\langle 4 \rangle$ e $\langle 5 \rangle$ do anel $\mathbb{Z}$. Determine o anel quociente respectivo e diga se é um corpo.

1.18. Seja A o anel $(\mathbb{Q}^{\mathbb{Q}}, +, \cdot)$ das funções racionais de variável racional, onde

$$(f+g)(x) = f(x) + g(x) \text{ e } (f \cdot g)(x) = f(x) \cdot g(x).$$

(a) Determine a identidade de A e averigúe se A é um domínio de integridade. Qual é a característica de A?

(b) Considere o ideal $I = \{f \in A \mid f(2) = 0\}$ de A. Determine o anel quociente A/I e diga se I é maximal.

1.19. Prove que se A é um anel, I e J são ideais de A e P é um ideal primo de A, então

$$IJ \subseteq P \Rightarrow I \subseteq P \text{ ou } J \subseteq P.$$

(Observação: IJ denota o conjunto $\{ab \mid a \in I, b \in J\}$.)

1.20.

(a) Mostre que $\mathcal{P}(S)$ é um ideal de $(\mathcal{P}(X), \Delta, \cap)$ (Exercício 1.1(c)) para qualquer subconjunto S de X.

(b) Determine o anel quociente $\mathcal{P}(X)/\mathcal{P}(S)$ e compare-o com o anel $(\mathcal{P}(X - S), \Delta, \cap)$.

1.21. Seja $(A, +, \cdot)$ um anel comutativo com identidade.

(1) Quando é que se diz que um ideal M de A é maximal?

(2) Seja M um ideal próprio de A. Prove que M é maximal se e só se

$$\forall\, a \in A - M \;\; \exists\, x \in A \,:\, 1 - ax \in M.$$

1.22. Seja $(A, +, \cdot)$ um anel. Prove que:

(a) Se P é um ideal primo de A e I e J são ideais de A então $I \subseteq P$ ou $J \subseteq P$ sempre que $IJ \subseteq P$.

(b) Se M é um ideal maximal de A então M é o único ideal de A que é primo e contém M^2.

1.23. Quais das seguintes funções são homomorfismos de anéis?

(a) $\begin{array}{ccc} \mathbb{Z} & \to & \mathbb{Z} \\ a & \mapsto & a^2 \end{array}$ (b) $\begin{array}{ccc} \mathbb{Z}_3 & \to & \mathbb{Z}_3 \\ a & \mapsto & a^3 \end{array}$

(d) $\begin{array}{ccc} \mathbb{Z} & \to & \mathbb{Z} \\ a & \mapsto & 5a \end{array}$ (e) $\begin{array}{ccc} \mathbb{Z} & \to & \mathbb{Z}_n \\ a & \mapsto & \text{resto da divisão de } a \text{ por } n \end{array}$

(f) $\begin{array}{ccc} \mathbb{Z}[i] & \to & \mathbb{Z} \\ a+ib & \mapsto & a^2+b^2, \end{array}$ sendo $\mathbb{Z}[i]$ o anel dos inteiros de Gauss (Exercício 1.1(g)).

1.24. $\theta : \{a+b\sqrt{2} \mid a, b \in \mathbb{Q}\} \longrightarrow \{a+b\sqrt{3} \mid a, b \in \mathbb{Q}\}$, definida por $\theta(a+b\sqrt{2}) = a+b\sqrt{3}$, é um homomorfismo de anéis?

1.25. Seja A um domínio de integridade de característica $n \neq 0$. Prove que a aplicação $\varphi : A \to A$, definida por $\varphi(x) = x^n$ para qualquer $x \in A$, é um homomorfismo.

1.26. Dado um anel $(A, +, \cdot)$, seja $\mathcal{F} = (A^A, +, \cdot)$ o anel das aplicações de A em A com a adição e multiplicação definidas do seguinte modo:

$$\forall\, f, g \in \mathcal{F} \;\; \forall\, x \in A \quad (f+g)(x) = f(x) + g(x), \quad (f \cdot g)(x) = f(x) \cdot g(x).$$

Para cada $(a, b) \in A \times A$ considere o conjunto $\mathcal{F}_{(a,b)} = \{f \in \mathcal{F} \mid f(a) = b\}$.

(a) Prove que $\mathcal{F}_{(a,b)}$ é um subanel de $\mathcal{F}$ se e só se $b = 0$.

(b) Mostre que $\mathcal{F}_{(a,0)}$ é um ideal de $\mathcal{F}$.

(c) Prove que o anel quociente $\mathcal{F}/\mathcal{F}_{(a,0)}$ é isomorfo a A.

1.27. Seja $A = (\mathbb{Q}, +, *)$, onde $+$ denota a adição usual de racionais e $*$ é definida por $a * b = 2ab$.

(a) Mostre que A é um anel comutativo com identidade.

(b) Determine um subanel de A que seja isomorfo ao anel usual $(\mathbb{Z}, +, \times)$ dos inteiros, descrevendo o isomorfismo (e justificando que se trata de facto de um isomorfismo).

1.28. Seja $A = (\mathbb{Q}, +, *)$, onde $+$ denota a adição usual de racionais e $*$ é definida por $a * b = ab/3$.

(a) Mostre que A é um corpo.

(b) Determine um subanel de A que seja isomorfo ao anel usual $(\mathbb{Z}, +, \cdot)$ dos inteiros, descrevendo o isomorfismo.

1.29. Determine a característica dos anéis com identidade do Exercício 1.1.

1.30. Considere no conjunto $C = \{0, 1, \alpha, \beta\}$ as operações $+$ e $\cdot$ definidas pelas tabelas

$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

$\cdot$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

(a) Prove que $(C, +, \cdot)$ é um corpo.

(b) Determine todos os subcorpos de C. Verifique se são ideais.

(c) Indique a característica de C.

2. Anéis polinomiais

A aritmética de polinómios de coeficientes reais é governada por regras familiares. Como generalizá-la a um anel arbitrário?

Na Análise têm trabalhado com polinómios com coeficientes reais, definidos como *funções* $p : \mathbb{R} \to \mathbb{R}$ da forma

$$p(x) = \sum_{i=0}^{n} p_i x^i,$$

onde os números reais p_i são os coeficientes do polinómio. A coeficientes distintos correspondem polinómios (funções polinomiais) distintos. Não podemos definir de modo análogo os polinómios com coeficientes num anel arbitrário A, se desejarmos que polinómios com coeficientes distintos sejam necessariamente polinómios distintos. De facto, desde que A tenha mais de um elemento ($a \neq 0$), existe uma infinidade de possibilidades distintas para os coeficientes de um possível polinómio (por ex., $a, ax, ax^2, ax^3, \ldots$), mas, no caso de A ser finito, existe apenas um número finito de funções $f : A \to A$, pelo que não podem ser usadas para definir todos os polinómios com coeficientes em A.

Por exemplo, se A for o anel $\mathbb{Z}_2$, só existem quatro funções $f : \mathbb{Z}_2 \to \mathbb{Z}_2$

$$\begin{array}{cccc} f_1 & f_2 & f_3 & f_4 \\ 0 \mapsto 0 & 0 \mapsto 0 & 0 \mapsto 1 & 0 \mapsto 1 \\ 1 \mapsto 0 & 1 \mapsto 1 & 1 \mapsto 0 & 1 \mapsto 1 \end{array}$$

mas se quisermos que polinómios com coeficientes distintos sejam de facto polinómios distintos, existe um número infinito de polinómios com coeficientes em $\mathbb{Z}_2$:

$$0, 1, x, 1+x, x^2, 1+x^2, x+x^2, 1+x+x^2, x^3, 1+x^3, x+x^3, x^2+x^3, 1+x+x^3, 1+x^2+x^3,$$

$$x + x^2 + x^3, 1 + x + x^2 + x^3, \ldots$$

[Observe: os polinómios $1 + x$ e $1 + x + x^2 + x^3$ definem ambos f_3]

Resolvemos este problema identificando um polinómio com a sucessão dos seus próprios coeficientes, esquecendo a sua relação com funções de tipo especial.

No que se segue A designa um anel comutativo com identidade.

POLINÓMIO

Uma sucessão

$$\begin{array}{rcl} p : \mathbb{N}_0 & \to & A \\ i & \mapsto & p(i) := p_i \end{array}$$

em A diz-se um *polinómio* se existe $n \in \mathbb{N}_0$ tal que $p(i) = 0$ para todo o $i > n$. O menor número $n \in \mathbb{N}_0$ nessas condições chama-se *grau* do polinómio (no caso em que o polinómio não é o polinómio nulo $(0, 0, 0, \ldots)$; quando se trata do polinómio nulo, convenciona-se que o seu grau é $-\infty$). Os termos $p(i) := p_i$ dizem-se os *coeficientes* do polinómio. Denotaremos por $A[x]$ o conjunto de todos os polinómios com coeficientes no anel A.

Exemplos:

$\mathbf{0} := (0, 0, 0, \ldots)$ é o *polinómio zero* ou *nulo*.

$\mathbf{1} := (1, 0, 0, \ldots)$ é o *polinómio um* ou *identidade*.

$\boldsymbol{a} := (a, 0, 0, \ldots)$ diz-se um *polinómio constante* ($a \in A$).

A soma e produto de polinómios com coeficientes reais (isto é, em $\mathbb{R}[x]$) é-nos seguramente familiar e baseiam-se nas operações de soma e produto dos coeficientes reais. Reconhecendo que essas operações sobre os coeficientes são possíveis em qualquer anel, podemos estender essas operações a qualquer $A[x]$. Note que a soma assim introduzida não passa da soma usual de sucessões, mas o produto já não é o habitual. Quando há risco de ambiguidade, referimo-nos ao produto definido abaixo como o *produto de convolução*, e representamo-lo por $p \star q$ em lugar de pq.

SOMA E PRODUTO (DE CONVOLUÇÃO) DE POLINÓMIOS

Sendo $p, q : \mathbb{N}_0 \to A$ polinómios, a *soma* $\boldsymbol{p} + \boldsymbol{q}$ e o *produto (de convolução)* $\boldsymbol{p} \star \boldsymbol{q}$ são os polinómios dados por

$$(\boldsymbol{p} + \boldsymbol{q})_i = p_i + q_i$$

$$(\boldsymbol{p} \star \boldsymbol{q})_i = \sum_{j=0}^{i} p_j q_{i-j}.$$

Exemplos: (1) Se $\boldsymbol{a} = (a, 0, 0, \ldots)$ é um polinómio constante e

$$\boldsymbol{p} = (p_0, p_1, \ldots, p_n, 0, 0, \ldots)$$

é um polinómio arbitrário, o produto $\boldsymbol{a} \star \boldsymbol{p}$ é o polinómio

$$(ap_0, ap_1, ap_2, \ldots, ap_n, 0, 0, \ldots),$$

porque a soma $\sum_{j=0}^{i} a_j p_{i-j}$ se reduz sempre à parcela com $j = 0$.

(2) Se $\boldsymbol{a} = (a, 0, 0, \ldots)$ e $\boldsymbol{b} = (b, 0, 0, \ldots)$ são polinómios constantes, a sua soma e o seu produto são dados por $\boldsymbol{a} + \boldsymbol{b} = (a + b, 0, 0, \ldots)$ e $\boldsymbol{a} \star \boldsymbol{b} = (ab, 0, 0, \ldots)$. Portanto, o conjunto dos polinómios constantes com as operações acima indicadas é um anel isomorfo a A.

`[Confirme: o isomorfismo é dado pela aplicação` $a \mapsto (a, 0, 0, \ldots)$`]`

(3) Em $\mathbb{Z}_2[x]$, se $\boldsymbol{p} = (1, 1, \ldots, 1, 0, 0, \ldots)$ é de grau $n \geq 0$, então

$$\boldsymbol{pp} = (1, 0, 1, 0, \ldots, 1, 0, 0, \ldots),$$

de grau $2n$, pois

$$(\boldsymbol{pp})_i = \sum_{j=0}^{i} p_j p_{i-j} = \sum_{j=0}^{i} 1 = (i + 1) \bmod 2.$$

O resultado seguinte é evidente, pelo que a sua demonstração fica como exercício.

Proposição 2.1 *Se A é um anel comutativo com identidade, $(A[x], +, \star)$ é também um anel comutativo com identidade. Além disso, $(A[x], +, \star)$ é um domínio de integridade se e só se A é um domínio de integridade.* ∎

O anel $A[x]$ chama-se *anel polinomial* sobre A.

Observámos no exemplo (2) acima que o anel $A[x]$ contém um subanel isomorfo a A (o conjunto dos polinómios constantes), o que justifica que se possa usar o mesmo símbolo a para designar um dado elemento do anel A e o correspondente polinómio constante $(a, 0, 0, \ldots)$. Dizemos então que $A[x]$ é uma *extensão* de A.

Designemos por $\boldsymbol{x}$ (a que chamaremos a *indeterminada* $\boldsymbol{x}$) o polinómio

$$(0, 1, 0, 0, \ldots).$$

É evidente que $\boldsymbol{x}^2 = (0, 0, 1, 0, \ldots)$, $\boldsymbol{x}^3 = (0, 0, 0, 1, 0, \ldots)$, etc. Alargamos esta observação ao caso $n = 0$, convencionando $\boldsymbol{x}^0 = (1, 0, 0, \ldots) = \mathbf{1}$.

Mais geralmente, se $\boldsymbol{p} = (p_0, p_1, \ldots, p_n, 0, 0, \ldots)$ é um polinómio arbitrário de grau n, o produto $\boldsymbol{px}$ é o polinómio de grau $n+1$ que se obtém de $\boldsymbol{p}$ por translação de todos os seus coeficientes para a direita, ou seja

$$\boldsymbol{px} = (0, p_0, p_1, \ldots, p_n, 0, 0, \ldots),$$

porque

$$(\boldsymbol{px})_0 = p_0 x_0 = 0,$$

$$(\boldsymbol{px})_{i+1} = \sum_{j=0}^{i+1} p_j x_{i+1-j} = p_i.$$

Então, identificando, como fizemos anteriormente, cada polinómio constante $\boldsymbol{a}$ pelo correspondente elemento a de A, podemos finalmente obter a forma a que estávamos habituados para representar um polinómio:

$$\begin{aligned}
\boldsymbol{p} &= (p_0, p_1, \ldots, p_n, 0, 0, \ldots) \\
&= (p_0, 0, 0, \ldots) + (0, p_1, 0, 0, \ldots) + (0, 0, p_2, 0, 0, \ldots) + \cdots + (0, \ldots, 0, p_n, 0, 0, \ldots) \\
&= p_0 + p_1\boldsymbol{x} + p_2\boldsymbol{x}^2 + \cdots + p_n\boldsymbol{x}^n \\
&= \sum_{i=0}^{n} p_i \boldsymbol{x}^i.
\end{aligned}$$

A soma à direita é a *forma canónica* do polinómio $\boldsymbol{p}$. Como é habitual, um coeficiente é omitido se for igual a 1.

Temos assim duas formas perfeitamente equivalentes de representar os elementos de $A[x]$: como sucessões

$$\boldsymbol{p} = (p_0, p_1, \ldots, p_n, 0, 0, \ldots)$$

ou como somas formais

$$\boldsymbol{p} = p_0 + p_1\boldsymbol{x} + p_2\boldsymbol{x}^2 + \cdots + p_n\boldsymbol{x}^n = \sum_{i=0}^{n} p_i \boldsymbol{x}^i. \tag{2.1.1}$$

A (2.1.1) chama-se a *forma canónica* do polinómio $\boldsymbol{p}$.

[Confirme: em termos da forma canónica, as operações + e ⋆ do anel $A[x]$ correspondem exactamente às operações de polinómios a que estávamos habituados]

Portanto, para somar e multiplicar estes polinómios, procedemos exactamente como estamos habituados com os polinómios com coeficientes reais.

Exemplo: Em $\mathbb{Z}_4[x]$, para $\boldsymbol{p} = 1 + \boldsymbol{x} + 2\boldsymbol{x}^2$ e $\boldsymbol{q} = 1 + 2\boldsymbol{x}^2$, temos:

$$\begin{aligned} \boldsymbol{p} + \boldsymbol{q} &= (1 + \boldsymbol{x} + 2\boldsymbol{x}^2) + (1 + 2\boldsymbol{x}^2) \\ &= (1 + 1) + (1 + 0)\boldsymbol{x} + (2 + 2)\boldsymbol{x}^2 \\ &= 2 + \boldsymbol{x}, \end{aligned}$$

$$\begin{aligned} \boldsymbol{pq} &= (1 + \boldsymbol{x} + 2\boldsymbol{x}^2)(1 + 2\boldsymbol{x}^2) \\ &= (1 + \boldsymbol{x} + 2\boldsymbol{x}^2)1 + (1 + \boldsymbol{x} + 2\boldsymbol{x}^2)2\boldsymbol{x}^2 \\ &= (1 + \boldsymbol{x} + 2\boldsymbol{x}^2) + (2\boldsymbol{x}^2 + 2\boldsymbol{x}^3 + 0\boldsymbol{x}^4) \\ &= 1 + \boldsymbol{x} + 2\boldsymbol{x}^3. \end{aligned}$$

GRAU

Se $\boldsymbol{p} \neq 0$ é um polinómio, o *grau* de $\boldsymbol{p}$ é o inteiro $gr(\boldsymbol{p})$ definido por

$$gr(\boldsymbol{p}) = \max\{n \in \mathbb{N}_0 \mid p_n \neq 0\}.$$

Se $\boldsymbol{p} = 0$, convencionamos que $gr(\boldsymbol{p}) = -\infty$.

Um polinómio $\boldsymbol{p}$ de grau $n \geq 0$ diz-se *mónico* se o coeficiente p_n do termo de maior grau for igual a 1.

Assim, os polinómios constantes têm grau ≤ 0. O exemplo acima de produto de polinómios em $\mathbb{Z}_4[x]$ mostra que, por causa da possível existência de divisores de zero, nem sempre o grau do produto de dois polinómios é a soma dos graus dos polinómios factores. O próximo resultado esclarece completamente as propriedades do grau relativamente à soma e ao produto de polinómios. Para evitar frequentes excepções envolvendo o polinómio nulo, convencionamos que $gr(\boldsymbol{p}) + gr(\boldsymbol{q}) = -\infty$ sempre que $\boldsymbol{p} = 0$ ou $\boldsymbol{q} = 0$.

Proposição 2.2 *Sejam $\boldsymbol{p}, \boldsymbol{q} \in A[x]$. Então:*

(a) $gr(\boldsymbol{p} + \boldsymbol{q}) \leq \max\{gr(\boldsymbol{p}), gr(\boldsymbol{q})\}$.

(b) $gr(\boldsymbol{pq}) \leq gr(\boldsymbol{p}) + gr(\boldsymbol{q})$.

(c) *Se A é um domínio de integridade, $gr(\boldsymbol{pq}) = gr(\boldsymbol{p}) + gr(\boldsymbol{q})$.*

Demonstração. A prova de (a) é muito simples e deixa-se como exercício. Quanto a (b) e (c) basta observar o seguinte: se $\boldsymbol{p}$ é de grau n e $\boldsymbol{q}$ é de grau m, então $\boldsymbol{pq} = p_0q_0 + (p_0q_1 + p_1q_0)\boldsymbol{x} + \cdots + p_nq_m\boldsymbol{x}^{n+m}$, pelo que $gr(\boldsymbol{pq}) \leq n + m = gr(\boldsymbol{p}) + gr(\boldsymbol{q})$;

não existindo divisores de zero em A, tem-se necessariamente $p_n q_m \neq 0$, donde, neste caso, $gr(\boldsymbol{pq}) = n + m = gr(\boldsymbol{p}) + gr(\boldsymbol{q})$. ∎

Quais são as unidades de $A[x]$? Se A possui divisores de zero, $A[x]$ contém polinómios invertíveis de grau maior que zero — por exemplo, em $\mathbb{Z}_4[x]$,

$$(1 + 2\boldsymbol{x})(1 + 2\boldsymbol{x}) = 1;$$

no entanto, se A é um domínio de integridade, as unidades de $A[x]$ são precisamente os polinómios de grau zero, $\boldsymbol{p} = a$, onde a é uma unidade de A; então, se A é um corpo, as unidades de $A[x]$ são os polinómios de grau zero.

[Verifique: se A é um domínio de integridade, as unidades de $A[x]$ coincidem com as unidades de A]

Vamos agora estudar em pormenor o anel dos polinómios $A[x]$. Na base deste estudo está o algoritmo usual da divisão de polinómios de coeficientes reais. Será que podemos continuar a aplicá-lo num anel A arbitrário? Daqui em diante passamos a adoptar a seguinte convenção: o polinómio $\boldsymbol{p}$ é representado pelo símbolo $p(\boldsymbol{x})$, e o valor do polinómio $\boldsymbol{p}$ no ponto a é representado por $p(a)$. Continuamos a supor que A é um anel comutativo unitário.

Seja $A = \mathbb{Z}_6$. A divisão de $p(\boldsymbol{x}) = \boldsymbol{x}^4 + 2\boldsymbol{x}^3 + 3\boldsymbol{x}^2 + \boldsymbol{x} + 4$ por $d(\boldsymbol{x}) = \boldsymbol{x}^2 + 2\boldsymbol{x} + 2$ é possível, resultando no quociente $q(\boldsymbol{x}) = \boldsymbol{x}^2 + 1$, com resto $r(\boldsymbol{x}) = 5\boldsymbol{x} + 2$:

$$\begin{array}{rl|l} \boldsymbol{x}^4 + 2\boldsymbol{x}^3 + 3\boldsymbol{x}^2 + \boldsymbol{x} + 4 & & \boldsymbol{x}^2 + 2\boldsymbol{x} + 2 \\ \cline{3-3} -\boldsymbol{x}^4 - 2\boldsymbol{x}^3 - 2\boldsymbol{x}^2 & & \boldsymbol{x}^2 + 1 \\ \cline{1-1} \boldsymbol{x}^2 + \boldsymbol{x} + 4 & & \\ -\boldsymbol{x}^2 - 2\boldsymbol{x} - 2 & & \\ \cline{1-1} 5\boldsymbol{x} + 2 & & \end{array}$$

É claro que se o coeficiente d_2 de $d(\boldsymbol{x})$ fosse 2 a divisão já não seria possível: não existe nenhum elemento q_2 em $\mathbb{Z}_6$ tal que $2q_2 = 1$ para podermos prosseguir com o algoritmo! (Tudo porque 2, sendo um divisor de zero, não é invertível.) Quando o polinómio divisor é mónico ou A é um domínio de integridade, a divisão é sempre possível. Mais geralmente:

Teorema 2.3 [Algoritmo de Divisão]
Sejam $p(\boldsymbol{x})$ e $d(\boldsymbol{x}) \neq 0$ elementos de $A[x]$, de graus n e m, respectivamente. Se d_m é uma unidade de A então existem polinómios únicos $q(\boldsymbol{x})$ e $r(\boldsymbol{x})$, com $gr(r(\boldsymbol{x})) < gr(d(\boldsymbol{x}))$, tais que $p(\boldsymbol{x}) = q(\boldsymbol{x})d(\boldsymbol{x}) + r(\boldsymbol{x})$.

Demonstração.
Existência: O caso $n < m$ é evidente: podemos tomar $q(\boldsymbol{x}) = 0$ e $r(\boldsymbol{x}) = p(\boldsymbol{x})$.

Suponhamos então $n \geq m$. Demonstramos a existência de $q(\boldsymbol{x})$ e $r(\boldsymbol{x})$ por indução sobre n:

- Se $n = 0$ então $m = 0$. Portanto $d(\boldsymbol{x}) = d_0$ e d_0 é invertível pelo que bastará tomar $q(\boldsymbol{x}) = d_0^{-1}p(\boldsymbol{x})$ e $r(\boldsymbol{x}) = 0$.
- Vamos agora supor que o resultado é verdadeiro para qualquer polinómio de grau inferior a n. Precisamos de provar que ele também é válido para polinómios de grau n. Seja então $p(\boldsymbol{x}) = p_n\boldsymbol{x}^n + p_{n-1}\boldsymbol{x}^{n-1} + \cdots + p_1\boldsymbol{x} + p_0$, onde $p_n \neq 0$ e comecemos a fazer a divisão de $p(\boldsymbol{x})$ por $d(\boldsymbol{x})$:

$$\begin{array}{l|l} p_n\boldsymbol{x}^n + p_{n-1}\boldsymbol{x}^{n-1} + \cdots + p_1\boldsymbol{x} + p_0 & d_m\boldsymbol{x}^m + d_{m-1}\boldsymbol{x}^{m-1} + \cdots + d_1\boldsymbol{x} + d_0 \\ -p_n\boldsymbol{x}^n - p_n d_m^{-1}d_{m-1}\boldsymbol{x}^{n-1} - \cdots & \\ \hline \underbrace{(p_{n-1} - p_n d_m^{-1}d_{m-1})\boldsymbol{x}^{n-1} + \cdots}_{\tilde{p}(\boldsymbol{x})} & p_n d_m^{-1}\boldsymbol{x}^{n-m} \end{array}$$

Considerando agora o polinómio $\tilde{p}(\boldsymbol{x}) = p(\boldsymbol{x}) - p_n d_m^{-1}\boldsymbol{x}^{n-m}d(\boldsymbol{x})$, é claro que $gr(\tilde{p}(\boldsymbol{x})) < n$, logo, pela hipótese de indução, existem polinómios $\tilde{q}(\boldsymbol{x})$ e $\tilde{r}(\boldsymbol{x})$ satisfazendo $\tilde{p}(\boldsymbol{x}) = \tilde{q}(\boldsymbol{x})d(\boldsymbol{x}) + \tilde{r}(\boldsymbol{x})$, onde $gr(\tilde{r}(\boldsymbol{x})) < gr(d(\boldsymbol{x}))$. Então

$$p(\boldsymbol{x}) = p_n d_m^{-1}x^{n-m}d(\boldsymbol{x}) + \tilde{p}(\boldsymbol{x}) = (\underbrace{p_n d_m^{-1}x^{n-m} + \tilde{q}(\boldsymbol{x})}_{q(\boldsymbol{x})})d(\boldsymbol{x}) + \underbrace{\tilde{r}(\boldsymbol{x})}_{r(\boldsymbol{x})}.$$

Unicidade: Se $p(\boldsymbol{x}) = q_1(\boldsymbol{x})d(\boldsymbol{x}) + r_1(\boldsymbol{x}) = p(\boldsymbol{x}) = q_2(\boldsymbol{x})d(\boldsymbol{x}) + r_2(\boldsymbol{x})$, então $(q_1(\boldsymbol{x}) - q_2(\boldsymbol{x}))d(\boldsymbol{x}) = r_2(\boldsymbol{x}) - r_1(\boldsymbol{x})$. Se $q_2(\boldsymbol{x})$ é diferente de $q_1(\boldsymbol{x})$ obtém-se uma contradição analisando os graus dos polinómios: por um lado,

$$gr(r_2(\boldsymbol{x}) - r_1(\boldsymbol{x})) \leq \max\{gr(r_1(\boldsymbol{x})), gr(r_2(\boldsymbol{x}))\} < gr(d(\boldsymbol{x})),$$

mas, por outro lado,

$$\begin{aligned} gr(r_2(\boldsymbol{x}) - r_1(\boldsymbol{x})) &= gr((q_1(\boldsymbol{x}) - q_2(\boldsymbol{x}))d(\boldsymbol{x})) \\ &= gr(q_1(\boldsymbol{x}) - q_2(\boldsymbol{x})) + gr(d(\boldsymbol{x})) \quad \text{(pois } d_m \text{ não é div. de zero)} \\ &\geq gr(d(\boldsymbol{x})). \end{aligned}$$

Assim $q_1(\boldsymbol{x}) = q_2(\boldsymbol{x})$, o que implica imediatamente $r_1(\boldsymbol{x}) = r_2(\boldsymbol{x})$. ∎

Tal como no caso dos inteiros, os polinómios $q(\boldsymbol{x})$ e $r(\boldsymbol{x})$ dizem-se respectivamente *quociente* e *resto* da divisão de $p(\boldsymbol{x})$ por $d(\boldsymbol{x})$. O caso em que $r(\boldsymbol{x}) = 0$

corresponde, claro está, ao caso em que $d(\boldsymbol{x})$ é *divisor* (ou *factor*) de $p(\boldsymbol{x})$. Neste caso escrevemos $d(\boldsymbol{x})|p(\boldsymbol{x})$.

O argumento de prova da existência, no teorema anterior (Algoritmo de Divisão), pode ser facilmente transformado num algoritmo de cálculo do quociente e do resto (onde, dado um polinómio $p(\boldsymbol{x}) = p_n\boldsymbol{x}^n + p_{n-1}\boldsymbol{x}^{n-1} + \cdots + p_0$, de grau n, designamos por $p^{\text{top}}(\boldsymbol{x}) = p_n\boldsymbol{x}^n$ o termo de grau máximo):

ALGORITMO DA DIVISÃO

Dados: $p(\boldsymbol{x}) = p_n\boldsymbol{x}^n + p_{n-1}\boldsymbol{x}^{n-1} + \cdots + p_0$, $d(\boldsymbol{x}) = d_m\boldsymbol{x}^m + d_{m-1}\boldsymbol{x}^{m-1} + \cdots + d_0$ tal que d_m é invertível.

Para dividir $p(\boldsymbol{x})$ por $d(\boldsymbol{x})$ procede-se por iteração, do seguinte modo:

Começando com $q_0(\boldsymbol{x}) = 0$ e $r_0(\boldsymbol{x}) = p(\boldsymbol{x})$, faz-se em cada passo

$$q_i(\boldsymbol{x}) = q_{i-1}(\boldsymbol{x}) + d_m^{-1}\frac{r_{i-1}^{\text{top}}(\boldsymbol{x})}{\boldsymbol{x}^m}, \quad r_i(\boldsymbol{x}) = r_{i-1}(\boldsymbol{x}) - d_m^{-1}\frac{r_{i-1}^{\text{top}}(\boldsymbol{x})}{\boldsymbol{x}^m}d(\boldsymbol{x}) :$$

$$\begin{array}{rl|l}
 & p_n\boldsymbol{x}^n + p_{n-1}\boldsymbol{x}^{n-1} + \cdots + p_1\boldsymbol{x} + p_0 & d_m\boldsymbol{x}^m + d_{m-1}\boldsymbol{x}^{m-1} + \cdots + d_1\boldsymbol{x} + d_0 \\
 & -p_n\boldsymbol{x}^n - d_m^{-1}p_nd_{m-1}\boldsymbol{x}^{n-1} - \cdots & \underbrace{d_m^{-1}p_n\boldsymbol{x}^{n-m}}_{q_1(\boldsymbol{x})} + d_m^{-1}(p_{n-1} - d_m^{-1}p_nd_{m-1})x^{n-m-1} + \cdots \\
\hline
r_1(\boldsymbol{x}): & (p_{n-1} - d_m^{-1}p_nd_{m-1})\boldsymbol{x}^{n-1} + \cdots & \underbrace{\qquad\qquad\qquad\qquad}_{q_2(\boldsymbol{x})} \\
 & -(p_{n-1} - d_m^{-1}p_nd_{m-1})\boldsymbol{x}^{n-1} + \cdots & \vdots \\
\hline
r_2(\boldsymbol{x}): & \cdots & \underbrace{\qquad\qquad\qquad\qquad}_{q_i(\boldsymbol{x})} \\
\vdots & \vdots & \\
r_i(\boldsymbol{x}): & \cdots &
\end{array}$$

A iteração termina quando $gr(r_i(\boldsymbol{x})) < m$.

Então faz-se $r(\boldsymbol{x}) = r_i(\boldsymbol{x})$ e $q(\boldsymbol{x}) = q_i(\boldsymbol{x})$.

`[Observe: a analogia entre o algoritmo da divisão nos anéis` $A[x]$
`e o algoritmo da divisão em` $\mathbb{Z}$`]`

O resultado seguinte é um corolário imediato do Algoritmo de Divisão:

Corolário 2.4 *Seja C um corpo. Para quaisquer $p(\boldsymbol{x})$ e $d(\boldsymbol{x}) \neq 0$ em $C[x]$, existem polinómios únicos $q(\boldsymbol{x})$ e $r(\boldsymbol{x})$ tais que $p(\boldsymbol{x}) = q(\boldsymbol{x})d(\boldsymbol{x}) + r(\boldsymbol{x})$, com $gr(r(\boldsymbol{x})) < gr(d(\boldsymbol{x}))$.* ∎

Observámos anteriormente que não é de todo conveniente definir os polinómios com coeficientes em A como *funções* de determinado tipo, com domínio e valores em A. No entanto, nada nos impede de definir funções de A em A a partir de polinómios em $A[x]$.

FUNÇÃO POLINOMIAL

Se $p(\boldsymbol{x}) = \sum_{i=0}^{n} p_i \boldsymbol{x}^i$ é um polinómio em $A[x]$, a função $p : A \to A$ definida por $p(a) = \sum_{i=0}^{n} p_i a^i$ diz-se *função polinomial associada a* $p(\boldsymbol{x})$.

Exemplo: Seja $A = \mathbb{Z}_2$ e $p(\boldsymbol{x}) = 1 + \boldsymbol{x} + \boldsymbol{x}^2$. A função polinomial associada ao polinómio $p(\boldsymbol{x})$ é $p : \mathbb{Z}_2 \to \mathbb{Z}_2$ dada por $p(a) = 1 + a + a^2$, para qualquer $a \in \mathbb{Z}_2$. Neste caso, temos $p(0) = p(1) = 1$, e portanto p é uma função constante, apesar de $p(\boldsymbol{x})$ não ser um polinómio constante. Em particular, se $q(\boldsymbol{x}) = 1$, temos $p(\boldsymbol{x}) \neq q(\boldsymbol{x})$ e $p = q$.

O resultado seguinte é outro corolário do Algoritmo de Divisão.

Corolário 2.5 [Teorema do resto]
Se $p(\boldsymbol{x}) \in A[x]$ e $a \in A$, o resto da divisão de $p(\boldsymbol{x})$ por $(\boldsymbol{x} - a)$ é o polinómio constante $r(\boldsymbol{x}) = p(a)$. Portanto, $p(\boldsymbol{x})$ é um múltiplo de $(\boldsymbol{x} - a)$ se e só se $p(a) = 0$.

Demonstração. Como $(\boldsymbol{x} - a)$ é mónico, podemos realizar a divisão de $p(\boldsymbol{x})$ por $(\boldsymbol{x} - a)$, obtendo $p(\boldsymbol{x}) = q(\boldsymbol{x})(\boldsymbol{x} - a) + r(\boldsymbol{x})$ com $gr(r(\boldsymbol{x})) < 1$ (ou seja, $r(\boldsymbol{x})$ é um polinómio constante $r(\boldsymbol{x}) = b$). Então a identidade de polinómios $p(\boldsymbol{x}) = q(\boldsymbol{x})(\boldsymbol{x} - a) + b$ implica $p(a) = b$, donde $r(\boldsymbol{x}) = p(a)$. ∎

RAIZ DE UM POLINÓMIO

Um elemento $a \in A$ diz-se *raiz* de um polinómio $p(\boldsymbol{x}) = \sum_{i=0}^{n} p_i \boldsymbol{x}^i$ de $A[x]$ caso $p(a) = 0$. Portanto, $p(\boldsymbol{x})$ é um múltiplo de $(\boldsymbol{x} - a)$ se e só se a é uma raiz de $p(\boldsymbol{x})$.

Outra das consequências do Algoritmo de Divisão (ou mais directamente do Corolário 2) é o resultado clássico sobre o número máximo de raízes de um polinómio não-nulo, que é válido quando A é um domínio de integridade.

Proposição 2.6 *Seja D um domínio de integridade. Se $p(\boldsymbol{x}) \in D[x]$ e $gr(p(\boldsymbol{x})) = n \geq 0$ então $p(\boldsymbol{x})$ tem no máximo n raízes em D.*

Demonstração. Faremos uma demonstração por indução sobre n. O caso $n = 0$ é óbvio: $p(\boldsymbol{x})$ será um polinómio constante não-nulo pelo que não terá raízes em D.

Suponhamos agora, por hipótese de indução, que o resultado vale para qualquer polinómio de grau n. Nessas condições, seja $p(\boldsymbol{x})$ um polinómio de grau $n+1$. Se $p(\boldsymbol{x})$ não tiver raízes em D, não há nada a provar. Caso contrário, se tem uma raiz $a \in D$ então, pelo Corolário 2, $p(\boldsymbol{x}) = q(\boldsymbol{x})(\boldsymbol{x} - a)$. Como D é um domínio de integridade, $gr(q(\boldsymbol{x})) = n$. Logo, pela hipótese de indução, $q(\boldsymbol{x})$ tem no máximo n raízes. Isto implica que $p(\boldsymbol{x})$ tem no máximo $n+1$ raízes (porque se $b \neq a$ é raiz de $p(\boldsymbol{x})$ então é raiz de $q(\boldsymbol{x})$ pois $0 = p(b) = q(b)(b-a)$ implica $q(b) = 0$). ∎

Mas cuidado: no caso geral em que A não é um domínio de integridade, não há relação nenhuma entre o número de raízes e o grau do polinómio. Por exemplo, em $\mathbb{Z}_4[x]$, o polinómio $2\boldsymbol{x} + 2\boldsymbol{x}^2$ é de grau 2 mas tem 4 raízes: 0, 1, 2 e 3. Por outro lado, $1 + \boldsymbol{x}^2$ é de grau 3 mas só tem uma raiz: 3.

MULTIPLICIDADE DA RAIZ

Seja D um domínio de integridade. Se $a \in D$ é raiz de um polinómio $p(\boldsymbol{x}) \neq 0$ de $D[x]$, o maior natural m tal que $p(\boldsymbol{x})$ é múltiplo de $(x-a)^m$ diz-se a *multiplicidade* da raiz a.

[Exercício: Prove que a soma das multiplicidades das raízes de $p(\boldsymbol{x})$ é $\leq gr(p(\boldsymbol{x}))$]

Exemplos: $1 + \boldsymbol{x}^2$ é de grau 2 e não tem raízes em $\mathbb{R}$ (e, por maioria de razão, em $\mathbb{Q}$ e $\mathbb{Z}$). Em $\mathbb{C}$ tem exactamente 2 raízes, i e $-i$, de multiplicidade 1.
$1 - 2\boldsymbol{x} + 2\boldsymbol{x}^2 - 2\boldsymbol{x}^3 + \boldsymbol{x}^4$ é de grau 4 e tem exactamente uma raiz em $\mathbb{R}$, 1, de multiplicidade 2. Por outro lado, em $\mathbb{C}$ tem exactamente 3 raízes (1, i e $-i$), sendo a primeira de multiplicidade 2 e as outras de multiplicidade 1 (portanto, neste caso a soma das multiplicidades iguala o grau do polinómio).

[No próximo capítulo analisaremos melhor esta diferença entre os corpos $\mathbb{C}$ e $\mathbb{R}$: em $\mathbb{C}[x]$ a soma das multiplicidades das raízes de qualquer polinómio de grau n é exactamente n; em $\mathbb{R}[x]$ a soma das multiplicidades das raízes de qualquer polinómio de grau n não excede n, podendo ser menor que n]

[Dir-se-à que $\mathbb{C}$ é, ao contrário de $\mathbb{R}$, um corpo *algebricamente fechado*]

O facto do algoritmo da divisão em $A[x]$, no caso de A ser um corpo, ser sempre aplicável, tem, como em $\mathbb{Z}$, outra consequência importante:

Teorema 2.7 *Seja C um corpo. Em $C[x]$ todo o ideal é principal.*

Demonstração. Seja I um ideal de $C[x]$. Se $I = \{0\}$, então $I = \langle 0\rangle$ é um ideal principal. Podemos pois admitir que $I \neq \{0\}$. Neste caso, provaremos mais do que é exigido no enunciado do resultado, nomeadamente que existe um polinómio mónico $m(\boldsymbol{x}) \in C[x]$, único, tal que $I = \langle m(\boldsymbol{x})\rangle$.

Consideremos então o conjunto

$$N = \{n \in \mathbb{N}_0 \mid \text{ existe } s(\boldsymbol{x}) \in I, gr(s(\boldsymbol{x})) = n\}.$$

É claro que, como $I \neq \{0\}$, N é não-vazio, pelo que tem um mínimo. Seja $m(\boldsymbol{x})$um polinómio em I de grau igual a esse mínimo (podemos supor que $m(\boldsymbol{x})$ é mónico; com efeito, se não fosse, isto é, se o coeficiente do termo de maior grau fosse igual a $a \neq 1$, poderíamos sempre considerar o polinómio $n(\boldsymbol{x}) = a^{-1}m(\boldsymbol{x}) \in I$).

Provemos que $I = \langle m(\boldsymbol{x})\rangle$. Como $m(\boldsymbol{x}) \in I$, é óbvio que $\langle m(\boldsymbol{x})\rangle \subseteq I$. Por outro lado, se $p(\boldsymbol{x}) \in I$, usando o algoritmo de divisão temos $p(\boldsymbol{x}) = q(\boldsymbol{x})m(\boldsymbol{x}) + r(\boldsymbol{x})$, onde $gr(r(\boldsymbol{x})) < gr(m(\boldsymbol{x}))$. Dado que I é um ideal, podemos concluir que $r(\boldsymbol{x}) = p(\boldsymbol{x}) - q(\boldsymbol{x})m(\boldsymbol{x}) \in I$. Mas então $r(\boldsymbol{x})$ só pode ser igual a 0 pois, com excepção do polinómio nulo, não pode haver nenhum polinómio em I de grau inferior a $gr(m(\boldsymbol{x}))$. Assim, $p(\boldsymbol{x})$ é um múltiplo de $m(\boldsymbol{x})$ pelo que pertence ao ideal $\langle m(\boldsymbol{x})\rangle$.

Para provar a unicidade de $m(\boldsymbol{x})$, suponhamos $I = \langle n(\boldsymbol{x})\rangle$, onde $n(\boldsymbol{x}) \in C[x]$ é mónico. Da igualdade $\langle m(\boldsymbol{x})\rangle = \langle n(\boldsymbol{x})\rangle$ segue

$$\begin{cases} m(\boldsymbol{x}) = p_1(\boldsymbol{x})\, n(\boldsymbol{x}) \\ n(\boldsymbol{x}) = p_2(\boldsymbol{x})\, m(\boldsymbol{x}) \end{cases} \tag{2.7.1}$$

para alguns polinómios $p_1(\boldsymbol{x}), p_2(\boldsymbol{x})$, donde $m(\boldsymbol{x}) = p_1(\boldsymbol{x})p_2(\boldsymbol{x})m(\boldsymbol{x})$. Como $C[x]$ é um domínio de integridade, podemos cancelar $m(\boldsymbol{x}) \neq 0$ à esquerda e concluir que $p_1(\boldsymbol{x})p_2(\boldsymbol{x}) = 1$.

[Num domínio de integridade, a lei do cancelamento para o produto vale para elementos $\neq 0$: se $ba = ca$ ou $ab = ac$, com $a \neq 0$, então $b = c$ (pois $ba = ca \Leftrightarrow (b-c)a = 0 \Rightarrow b - c = 0 \Leftrightarrow b = c$)]

Então $gr(p_1(\boldsymbol{x}))+gr(p_2(\boldsymbol{x})) = 0$ e, consequentemente, $p_1(\boldsymbol{x})$ e $p_2(\boldsymbol{x})$ são polinómios constantes. Como $m(\boldsymbol{x})$ e $n(\boldsymbol{x})$ são mónicos, então de (2.7.1) segue $p_1(\boldsymbol{x}) = p_2(\boldsymbol{x}) = 1$ e $n(\boldsymbol{x}) = m(\boldsymbol{x})$. ∎

[Observe mais esta analogia entre os anéis $C[x]$ e $\mathbb{Z}$: $C[x]$ é, tal como $\mathbb{Z}$, um *domínio de ideais principais*]

Exemplos: $\mathbb{Z}[x]$ não é um domínio de ideais principais; por exemplo, o ideal $\langle 2, \boldsymbol{x}\rangle$ não é principal.

[Verifique]

Mais geralmente, se A é um anel comutativo com identidade, a demonstração acima de que um ideal I de $A[x]$ é principal consegue fazer-se desde que o coeficiente do termo de maior grau do polinómio $m(\boldsymbol{x})$ (que agora não é necessariamente mónico) seja invertível em A. Este não é o caso do ideal $\langle 2, \boldsymbol{x}\rangle$ em $\mathbb{Z}[x]$: qualquer polinómio $m(\boldsymbol{x}) \in \langle 2, x\rangle$ de grau mínimo é uma constante $\neq 1, -1$.

Corolário 2.8 *Sejam $p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})$ polinómios em $C[\boldsymbol{x}]$, onde pelo menos um é não-nulo. Então existe um único polinómio mónico $d(\boldsymbol{x}) \in C[x]$ tal que:*

(1) $d(\boldsymbol{x}) \mid p_i(\boldsymbol{x})$ $(i = 1, 2, \ldots, n)$.

(2) *Se $c(\boldsymbol{x}) \in C[x]$ e $c(\boldsymbol{x}) \mid p_i(\boldsymbol{x})$ $(i = 1, 2, \ldots, n)$ então $c(\boldsymbol{x}) \mid d(\boldsymbol{x})$.*

Além disso, $d(\boldsymbol{x})$ pode ser escrito na forma

$$d(\boldsymbol{x}) = r_1(\boldsymbol{x})p_1(\boldsymbol{x}) + \cdots + r_n(\boldsymbol{x})p_n(\boldsymbol{x}) \tag{2.8.1}$$

com $r_1(\boldsymbol{x}), \ldots, r_n(\boldsymbol{x}) \in C[x]$.

Demonstração. Consideremos o ideal $\langle p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})\rangle$, que é não-nulo. Pela demonstração do Teorema, existe um polinómio mónico $d(\boldsymbol{x})$, único, tal que

$$\langle p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})\rangle = \langle d(\boldsymbol{x})\rangle .$$

Como cada $p_i(\boldsymbol{x}) \in \langle d(\boldsymbol{x})\rangle$, a condição (1) é óbvia, enquanto (2.8.1) é consequência imediata do facto de $d(\boldsymbol{x})$ pertencer a $\langle p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})\rangle$. Quanto a (2), é consequência de (2.8.1). ∎

Por outras palavras, $d(\boldsymbol{x})$ é um *divisor comum* de $p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})$, e é *múltiplo* de qualquer outro divisor comum destes n polinómios.

MÁXIMO DIVISOR COMUM

O polinómio $d(\boldsymbol{x})$ diz-se o *máximo divisor comum* de $p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})$ e escreve-se $d(\boldsymbol{x}) = \text{mdc}\,(p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x}))$.

Analogamente, também existe um único polinómio mónico $m(\boldsymbol{x})$ tal que

$$(p_1(\boldsymbol{x})) \cap \cdots \cap (p_n(\boldsymbol{x})) = m(\boldsymbol{x}).$$

Neste caso:

(1) $p_i(\boldsymbol{x}) \mid m(\boldsymbol{x})$ $(i = 1, 2, \ldots, n)$.

(2) Se $c(\boldsymbol{x}) \in C[x]$ e $p_i(\boldsymbol{x}) \mid c(\boldsymbol{x})$ $(i = 1, 2, \ldots, n)$ então $m(\boldsymbol{x}) \mid c(\boldsymbol{x})$.

Portanto, $m(\boldsymbol{x})$ é *múltiplo comum* de $p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})$, e é *divisor* de qualquer outro polinómio que seja múltiplo comum destes n polinómios.

MÍNIMO MÚLTIPLO COMUM

O polinómio $m(\boldsymbol{x})$ diz-se o *mínimo múltiplo comum* de $p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x})$ e escreve-se $m(\boldsymbol{x}) = \operatorname{mmc}(p_1(\boldsymbol{x}), \ldots, p_n(\boldsymbol{x}))$.

Uma vez que, tal como nos inteiros,

$$p_1(\boldsymbol{x}) = q(\boldsymbol{x})p_2(\boldsymbol{x}) + r(\boldsymbol{x}) \Rightarrow \langle p_1(\boldsymbol{x}), p_2(\boldsymbol{x})\rangle = \langle p_2(\boldsymbol{x}), r(\boldsymbol{x})\rangle\,,$$

o algoritmo de Euclides para o cálculo do máximo divisor comum mantém a sua validade em $C[x]$.

ALGORITMO DE EUCLIDES

Sejam $p_1(\boldsymbol{x}), p_2(\boldsymbol{x}) \in C[x]$, com $p_2(\boldsymbol{x}) \neq 0$.
Se $p_2(\boldsymbol{x}) \mid p_1(\boldsymbol{x})$, então $\operatorname{mdc}(p_1(\boldsymbol{x}), p_2(\boldsymbol{x})) = p_2(\boldsymbol{x})$.
Se $p_2(\boldsymbol{x}) \nmid p_1(\boldsymbol{x})$, usamos o algoritmo da divisão repetidamente do seguinte modo:

$$\begin{array}{ll}
p_1(\boldsymbol{x}) = q_1(\boldsymbol{x})p_2(\boldsymbol{x}) + r_1(\boldsymbol{x}) & 0 \leq gr(r_1(\boldsymbol{x})) < gr(p_2(\boldsymbol{x})) \\
p_2(\boldsymbol{x}) = q_2(\boldsymbol{x})r_1(\boldsymbol{x}) + r_2(\boldsymbol{x}) & 0 \leq gr(r_2(\boldsymbol{x})) < gr(r_1(\boldsymbol{x})) \\
r_1(\boldsymbol{x}) = q_3(\boldsymbol{x})r_2(\boldsymbol{x}) + r_3(\boldsymbol{x}) & 0 \leq gr(r_3(\boldsymbol{x})) < gr(r_2(\boldsymbol{x})) \\
\vdots & \vdots \\
r_{t-2}(\boldsymbol{x}) = q_t(\boldsymbol{x})r_{t-1}(\boldsymbol{x}) + r_t(\boldsymbol{x}) & 0 \leq gr(r_t(\boldsymbol{x})) < gr(r_{t-1}(\boldsymbol{x})) \\
r_{t-1}(\boldsymbol{x}) = q_{t+1}(\boldsymbol{x})r_t(\boldsymbol{x}). &
\end{array}$$

Como $gr(p_2(\boldsymbol{x}))$ é finito, o processo terá que parar ao cabo de um número finito de passos. Seja a o coeficiente de maior grau do último resto não-nulo $r_t(\boldsymbol{x})$. Então $\operatorname{mdc}(p_1(\boldsymbol{x}), p_2(\boldsymbol{x})) = a^{-1}r_t(\boldsymbol{x})$.

Exemplo: O algoritmo de Euclides aplicado aos polinómios

$$p_1(\boldsymbol{x}) = 2\boldsymbol{x}^6 + \boldsymbol{x}^3 + \boldsymbol{x}^2 + 2 \in \mathbb{F}_3[x], \qquad p_2(\boldsymbol{x}) = \boldsymbol{x}^4 + \boldsymbol{x}^2 + 2\boldsymbol{x} \in \mathbb{F}_3[x]$$

dá:

$$\begin{aligned} 2\boldsymbol{x}^6 + \boldsymbol{x}^3 + \boldsymbol{x}^2 + 2 &= (2\boldsymbol{x}^2 + 1)(\boldsymbol{x}^4 + \boldsymbol{x}^2 + 2\boldsymbol{x}) + (\boldsymbol{x} + 2) \\ \boldsymbol{x}^4 + \boldsymbol{x}^2 + 2\boldsymbol{x} &= (\boldsymbol{x}^3 + \boldsymbol{x}^2 + 2\boldsymbol{x} + 1)(\boldsymbol{x} + 2) + 1 \\ \boldsymbol{x} + 2 &= (\boldsymbol{x} + 2)1 + 0. \end{aligned}$$

Portanto $\text{mdc}\,(p_1(\boldsymbol{x}), p_2(\boldsymbol{x})) = 1$ e $p_1(\boldsymbol{x})$ e $p_2(\boldsymbol{x})$ são *primos entre si.*

Além disso, a partir da penúltima divisão, obtemos sucessivamente:

$$\begin{aligned} 1 &= (\boldsymbol{x}^4 + \boldsymbol{x}^2 + 2\boldsymbol{x}) - (\boldsymbol{x}^3 + \boldsymbol{x}^2 + 2\boldsymbol{x} + 1)(\boldsymbol{x} + 2) \\ &= p_2(\boldsymbol{x}) - (\boldsymbol{x}^3 + \boldsymbol{x}^2 + 2\boldsymbol{x} + 1)(p_1(\boldsymbol{x}) - (2\boldsymbol{x}^2 + 1)p_2(\boldsymbol{x})) \\ &= -(\boldsymbol{x}^3 + \boldsymbol{x}^2 + 2\boldsymbol{x} + 1)p_1(\boldsymbol{x}) + (1 + 2\boldsymbol{x}^2 + 1)p_2(\boldsymbol{x}) \\ &= (2\boldsymbol{x}^3 + 2\boldsymbol{x}^2 + \boldsymbol{x} + 2)p_1(\boldsymbol{x}) + (2\boldsymbol{x}^2 + 2)p_2(\boldsymbol{x}). \end{aligned}$$

Seja $q(\boldsymbol{x})$ um factor de $p(\boldsymbol{x})$. Se $p(\boldsymbol{x}) = a(\boldsymbol{x})q(\boldsymbol{x})$ onde nem $a(\boldsymbol{x})$ nem $q(\boldsymbol{x})$ são invertíveis, $q(\boldsymbol{x})$ diz-se um *factor próprio* de $p(\boldsymbol{x})$.

POLINÓMIO IRREDUTÍVEL

Um polinómio $p(\boldsymbol{x})$ de $A[x]$ diz-se *irredutível* em $A[x]$ quando não tem factores próprios (em $A[x]$) e não é invertível (em $A[x]$). Caso contrário, $p(\boldsymbol{x})$ diz-se *redutível.*

Portanto, $p(\boldsymbol{x})$ é irredutível quando não é invertível e $p(\boldsymbol{x}) = q_1(\boldsymbol{x})q_2(\boldsymbol{x})$ implica que um dos polinómios $q_1(\boldsymbol{x})$ ou $q_2(\boldsymbol{x})$ seja invertível. Assim, quando C é um corpo, um polinómio $p(\boldsymbol{x}) \neq 0$ em $C[x]$ é irredutível se e só se $gr(p(\boldsymbol{x}) \geq 1$ e $p(\boldsymbol{x}) = q_1(\boldsymbol{x})q_2(\boldsymbol{x})$ implica $gr(q_1(\boldsymbol{x})) = 0$ ou $gr(q_2(\boldsymbol{x})) = 0$. Em particular, todo o polinómio de grau 1 é irredutível.

Exemplos: (1) Para qualquer anel A, $p(\boldsymbol{x}) = \boldsymbol{x}$ é irredutível.

(2) Se $A = \mathbb{Z}$, $p(\boldsymbol{x}) = 2\boldsymbol{x} - 3$ é irredutível mas $q(\boldsymbol{x}) = 2\boldsymbol{x} + 6$ é redutível (porque $2\boldsymbol{x} + 6 = 2(\boldsymbol{x} + 3)$ e 2 e $\boldsymbol{x} + 3$ não são invertíveis em $\mathbb{Z}[x]$).

(3) A redutibilidade ou irredutibilidade de um dado polinómio depende fortemente do anel em consideração. Por exemplo, o polinómio $\boldsymbol{x}^2 - 2 \in \mathbb{Q}[x]$ é irredutível em

$\mathbb{Q}[x]$, mas $\boldsymbol{x}^2 - 2 = (\boldsymbol{x} + \sqrt{2})(\boldsymbol{x} - \sqrt{2})$ é redutível em $\mathbb{R}[x] \supset \mathbb{Q}[x]$; por outro lado, $\boldsymbol{x}^2 + 1$ é irredutível em $\mathbb{Q}[x]$ ou $\mathbb{R}[x]$ mas é redutível em $\mathbb{C}[x] \supset \mathbb{R}[x] \supset \mathbb{Q}[x]$.

(4) Seja D um domínio de integridade. Um polinómio redutível em $D[x]$ não tem necessariamente raízes. É o caso de $\boldsymbol{x}^4 + 2\boldsymbol{x}^2 + 1$, que é redutível em $\mathbb{Z}[x]$, porque $\boldsymbol{x}^4 + 2\boldsymbol{x}^2 + 1 = (\boldsymbol{x}^2 + 1)^2$, e que não tem raízes em $\mathbb{Z}$.

(5) Se $gr(p(\boldsymbol{x})) \geq 2$ e $p(\boldsymbol{x})$ tem pelo menos uma raiz em D, então, pelo Teorema do Resto, $p(\boldsymbol{x})$ é redutível em $D[x]$.

(6) Se $p(\boldsymbol{x})$ é mónico e tem grau 2 ou 3, então $p(\boldsymbol{x})$ é redutível em $D[x]$ se e só se tem pelo menos uma raiz em D.

[Porquê?]

(7) Em $\mathbb{R}[x]$ os únicos polinómios irredutíveis são os polinómios de grau 1 e os polinómios $p(\boldsymbol{x}) = a\boldsymbol{x}^2 + b\boldsymbol{x} + c$ de grau 2 com *discriminante* $\Delta = b^2 - 4ac$ negativo.

[É consequência do seguinte facto: se $c \in \mathbb{C}$ é raiz de $p(\boldsymbol{x}) \in C[x]$, o complexo conjugado de c é também raiz de $p(\boldsymbol{x})$]

É possível em certos casos descrever todos os polinómios irredutíveis em $D[x]$, como em $\mathbb{R}[x]$. Noutros casos, este problema torna-se muito complexo e é praticamente impossível fazê-lo, conhecendo-se somente resultados parciais (alguns critérios que permitem em alguns casos concluir da redutibilidade ou irredutibilidade de um dado polinómio). É o caso de $\mathbb{Z}[x]$ e $\mathbb{Q}[x]$.

[Alguns desses critérios serão dados na aula prática]

Proposição 2.9 *Sejam $I = \langle p(\boldsymbol{x}) \rangle$ e $J = \langle q(\boldsymbol{x}) \rangle$ ideais de $C[x]$. Então:*

(1) *$I \subseteq J$ se e só se $q(\boldsymbol{x}) \mid p(\boldsymbol{x})$.*

(2) *Se $I = J$ e $p(\boldsymbol{x})$ e $q(\boldsymbol{x})$ são mónicos ou nulos então $p(\boldsymbol{x}) = q(\boldsymbol{x})$.*

(3) *I é maximal se e só se $p(\boldsymbol{x})$ é irredutível.*

Demonstração. (1) $I \subseteq J \Leftrightarrow p(\boldsymbol{x}) \in \langle q(\boldsymbol{x}) \rangle \Leftrightarrow q(\boldsymbol{x}) \mid p(\boldsymbol{x})$.

(2) O caso em que um dos polinómios é nulo é óbvio. Suponhamos então que são ambos mónicos. Por (1), $I = J$ se e só se $p(\boldsymbol{x}) \mid q(\boldsymbol{x})$ e $q(\boldsymbol{x}) \mid p(\boldsymbol{x})$. Então

$$\begin{cases} q(\boldsymbol{x}) = a(\boldsymbol{x})\, p(\boldsymbol{x}) \\ p(\boldsymbol{x}) = b(\boldsymbol{x})\, q(\boldsymbol{x}) \end{cases}$$

para alguns polinómios $a(\boldsymbol{x}), b(\boldsymbol{x}) \in C[x]$. Daqui segue (como já observámos na demonstração da unicidade no Teorema 2.7) que $p(\boldsymbol{x}) = q(\boldsymbol{x})$.

(3) Provaremos que $p(\boldsymbol{x})$ é redutível se e só se I não é maximal. Suponhamos que $p(\boldsymbol{x})$ é redutível. Então ou é invertível ou tem um factor próprio. No primeiro caso tem-se $1 = (p(\boldsymbol{x}))^{-1}p(\boldsymbol{x}) \in I$, donde $I = C[x]$ não é maximal. No segundo caso tem-se $p(\boldsymbol{x}) = q_1(\boldsymbol{x})q_2(\boldsymbol{x})$ com $gr(q_1(\boldsymbol{x})) \geq 1$ e $gr(q_2(\boldsymbol{x})) \geq 1$. Então $1 \leq gr(q_1(\boldsymbol{x})) < gr(p(\boldsymbol{x}))$, pelo que

$$\langle p(\boldsymbol{x})\rangle \subset \langle q_1(\boldsymbol{x})\rangle \subset C[x],$$

o que mostra que, também neste caso, I não é maximal.

Reciprocamente, suponhamos que I não é maximal, ou seja, que existe um ideal $J = \langle q(\boldsymbol{x})\rangle$ (recorde que $C[x]$ é um domínio de ideais principais) tal que $I \subset J \subset C[x]$. Então $p(\boldsymbol{x}) = r(\boldsymbol{x})q(\boldsymbol{x})$ para algum $r(\boldsymbol{x}) \in C[x]$. É claro que $gr(r(\boldsymbol{x})) \geq 1$ (pois se $r(\boldsymbol{x})$ fosse constante, $q(\boldsymbol{x})$ pertenceria a $\langle p(\boldsymbol{x})\rangle$ e teríamos $J = I$). Por outro lado, também $gr(q(\boldsymbol{x})) \geq 1$ (caso contrário, $J = C[x]$). Assim, a factorização $p(\boldsymbol{x}) = r(\boldsymbol{x})q(\boldsymbol{x})$ mostra que $p(\boldsymbol{x})$ é redutível em $C[x]$. ∎

Proposição 2.10 *Se um polinómio irredutível $p(\boldsymbol{x}) \in C[x]$ divide um produto $r_1(\boldsymbol{x})r_2(\boldsymbol{x})\cdots r_m(\boldsymbol{x})$ de polinómios em $C[x]$, então pelo menos um dos factores $r_i(\boldsymbol{x})$ é divisível por $p(\boldsymbol{x})$.*

Demonstração. Consideremos o ideal principal $I = \langle p(\boldsymbol{x})\rangle$. Pelo Teorema 1.6, $C[x]/I$ é um corpo (logo não tem divisores de zero). Mas

$$(r_1(\boldsymbol{x}) + I)\cdot(r_2(\boldsymbol{x}) + I)\cdot\ \cdots\ \cdot(r_m(\boldsymbol{x}) + I) = r_1(\boldsymbol{x})r_2(\boldsymbol{x})\cdots r_m(\boldsymbol{x}) + I = I,$$

uma vez que, por hipótese, $r_1(\boldsymbol{x})r_2(\boldsymbol{x})\cdots r_m(\boldsymbol{x}) \in I$. Então, necessariamente um dos factores é nulo, isto é, $r_i(\boldsymbol{x})+I = I$ para algum $i \in \{1, 2, \ldots, m\}$. Isto significa precisamente que $r_i(\boldsymbol{x}) \in I$, ou seja, $p(\boldsymbol{x}) \mid r_i(\boldsymbol{x})$. ∎

O teorema seguinte mostra a importância dos polinómios irredutíveis no anel $C[x]$.

Teorema 2.11 [Factorização única em $C[x]$] *Todo o polinómio $r(\boldsymbol{x}) \in C[x]$ de grau positivo pode ser escrito na forma*

$$r(\boldsymbol{x}) = cp_1(\boldsymbol{x})^{n_1}p_2(\boldsymbol{x})^{n_2}\cdots p_t(\boldsymbol{x})^{n_t} \tag{2.11.1}$$

onde $c \in C - \{0\}$, $p_1(\boldsymbol{x}), p_2(\boldsymbol{x}), \ldots, p_t(\boldsymbol{x})$ são polimónios mónicos irredutíveis em $C[x]$, todos distintos, e $n_1, n_2, \ldots, n_t \in \mathbb{N}$.

E mais: esta factorização é única a menos da ordem pela qual se escrevem os factores.

[Observe mais uma vez o paralelismo com $\mathbb{Z}$:
os polinómios irredutíveis correspondem aos inteiros primos;
este teorema corresponde ao Teorema Fundamental da Aritmética]

Referir-nos-emos a (2.11.1) como a *factorização canónica* de $r(\boldsymbol{x})$ em $C[x]$.

Demonstração. Comecemos por demonstrar a existência da factorização, por indução sobre $n = gr(r(\boldsymbol{x}))$.

O caso $n = 1$ é evidente: $r(\boldsymbol{x})$ sendo de grau 1 é irredutível. Seja c o coeficiente do termo de grau 1. Então $r(\boldsymbol{x}) = c(c^{-1}r(\boldsymbol{x}))$, onde $c^{-1}r(\boldsymbol{x})$ é um polinómio mónico irredutível.

Suponhamos, por hipótese de indução, que o resultado é válido para todos os polinómios não constantes de grau $< n$. Seja $r(\boldsymbol{x})$ um polinómio de grau n. Se $r(\boldsymbol{x})$ é irredutível nada há a provar (basta considerar a factorização canónica como no caso $n = 1$). Se $r(\boldsymbol{x})$ é redutível então $r(\boldsymbol{x}) = r_1(\boldsymbol{x})r_2(\boldsymbol{x})$, onde $1 \leq gr(r_1(\boldsymbol{x})) < n$ e $1 \leq gr(r_2(\boldsymbol{x})) < n$. Por hipótese de indução, $r_1(\boldsymbol{x})$ e $r_2(\boldsymbol{x})$ podem ser factorizados na forma (2.11.1), logo $r(\boldsymbol{x})$ também.

Quanto à unicidade da factorização, sejam

$$cp_1(\boldsymbol{x})^{n_1}p_2(\boldsymbol{x})^{n_2}\cdots p_t(\boldsymbol{x})^{n_t} = dq_1(\boldsymbol{x})^{m_1}q_2(\boldsymbol{x})^{m_2}\cdots q_k(\boldsymbol{x})^{m_k}$$

duas factorizações canónicas de $r(\boldsymbol{x})$. No polinómio da esquerda, c é o coeficiente do termo de maior grau, enquanto que no da direita esse coeficiente é d. Portanto $c = d$. Daqui segue imediatamente que

$$p_1(\boldsymbol{x})^{n_1}p_2(\boldsymbol{x})^{n_2}\cdots p_t(\boldsymbol{x})^{n_t} = q_1(\boldsymbol{x})^{m_1}q_2(\boldsymbol{x})^{m_2}\cdots q_k(\boldsymbol{x})^{m_k}. \tag{2.11.2}$$

Então $p_1(\boldsymbol{x}) \mid q_1(\boldsymbol{x})^{m_1}q_2(\boldsymbol{x})^{m_2}\cdots q_k(\boldsymbol{x})^{m_k}$ donde, pela Proposição 2.10, $p_1(\boldsymbol{x}) \mid q_i(\boldsymbol{x})$ para algum $i \in \{1, 2, \ldots, k\}$. Como $q_i(\boldsymbol{x})$ é irredutível, então $q_i(\boldsymbol{x}) = ap_1(\boldsymbol{x})$ o que implica $a = 1$ (pois quer $q_i(\boldsymbol{x})$ quer $p_1(\boldsymbol{x})$ são mónicos), ou seja $q_i(\boldsymbol{x}) = p_1(\boldsymbol{x})$. Então (2.11.2) equivale a

$$p_1(\boldsymbol{x})^{n_1-m_i} = p_2(\boldsymbol{x})^{-n_2}\ldots p_t(\boldsymbol{x})^{-n_t}q_1(\boldsymbol{x})^{m_1}\ldots q_{i-1}(\boldsymbol{x})^{m_{i-1}}q_{i+1}(\boldsymbol{x})^{m_{i+1}}\ldots q_k(\boldsymbol{x})^{m_k},$$

o que implica $n_1 = m_i$ (senão, $p_1(\boldsymbol{x}) = q_i(\boldsymbol{x})$ dividiria algum $p_j(\boldsymbol{x})$, $j \neq 1$, ou algum $q_j(\boldsymbol{x})$, $j \neq i$, o que é manifestamente impossível pois $p_1(\boldsymbol{x})$ é diferente de qualquer outro dos polinómios $p_j(\boldsymbol{x})$ e $q_i(\boldsymbol{x})$ é diferente de qualquer outro dos polinómios $q_j(\boldsymbol{x})$).

Cancelando $q_i(\boldsymbol{x})$ e $p_1(\boldsymbol{x})$ em (2.11.2) obtemos

$$p_2(\boldsymbol{x})^{n_2}\cdots p_t(\boldsymbol{x})^{n_t} = q_1(\boldsymbol{x})^{m_1}q_2(\boldsymbol{x})^{m_2}\cdots q_{i-1}(\boldsymbol{x})^{m_{i-1}}q_{i+1}(\boldsymbol{x})^{m_{i+1}}\cdots q_k(\boldsymbol{x})^{m_k}.$$

Repetindo o raciocínio, chegaremos à conclusão que $p_2(\boldsymbol{x}) = q_j(\boldsymbol{x})$ para algum $j \in \{1, 2, \ldots, i-1, i+1, \ldots, n\}$ e $n_2 = m_j$. Continuando assim, após um número finito de passos, temos provada a unicidade da factorização (2.11.1), a menos da ordem pela qual se escrevem os factores. ∎

Apêndice 1: apontamentos para estudo complementar

[O Teorema da Factorização Única é tão importante que é natural averiguar se se pode generalizar a outros anéis. Por outro lado, o estudo que acabámos de fazer dos anéis polinomiais $C[x]$ exibe tantas semelhanças com o anel $\mathbb{Z}$ dos inteiros que é bem possível que não sejam mera coincidência, e sejam sim casos particulares de resultados válidos num contexto muito mais geral.]

Como sabemos, um inteiro $p \neq 0$ não invertível é primo se $p|ab$ implica $p = a$ ou $p = b$. É claro que podemos adaptar esta definição a $C[x]$ e, mais geralmente, a $D[x]$. Do mesmo modo, podemos adaptar a definição de polinómio irredutível ao domínio dos inteiros:

DOMÍNIO	$\mathbb{Z}$	$C[x]$
unidades	$\mathcal{U}_{\mathbb{Z}} = \{-1, 1\}$	$\mathcal{U}_{C[x]} = \{p(\boldsymbol{x}) \in C[x] : gr(p(\boldsymbol{x})) = 0\}$
primo	$p \neq 0,\ p \notin \mathcal{U}_{\mathbb{Z}}$ $p\|ab \Rightarrow p\|a$ ou $p\|b$	$p(\boldsymbol{x}) \neq 0,\ p(\boldsymbol{x}) \notin \mathcal{U}_{C[x]}$ $p(\boldsymbol{x})\|a(\boldsymbol{x})b(\boldsymbol{x}) \Rightarrow p(\boldsymbol{x})\|a(\boldsymbol{x})$ ou $p(\boldsymbol{x})\|b(\boldsymbol{x})$
irredutível	$p \neq 0,\ p \notin \mathcal{U}_{\mathbb{Z}}$ $p = ab \Rightarrow a \in \mathcal{U}_{\mathbb{Z}}$ ou $b \in \mathcal{U}_{\mathbb{Z}}$ isto é $p = ab \Rightarrow a = 1$ ou $a = -1$ ou $b = 1$ ou $b = -1$	$p(\boldsymbol{x}) \neq 0,\ p(\boldsymbol{x}) \notin \mathcal{U}_{C[x]}$ $p(\boldsymbol{x}) = a(\boldsymbol{x})b(\boldsymbol{x}) \Rightarrow a(\boldsymbol{x}) \in \mathcal{U}_{C[x]}$ ou $b(\boldsymbol{x}) \in \mathcal{U}_{C[x]}$ isto é $p(\boldsymbol{x}) = a(\boldsymbol{x})b(\boldsymbol{x}) \Rightarrow gr(a(\boldsymbol{x})) = 0$ ou $gr(b(\boldsymbol{x})) = 0$

DOMÍNIO	$D[x]$
unidades	$\mathcal{U}_{D[x]} = \{p(\boldsymbol{x}) \in D[x] : p(\boldsymbol{x}) = c \in \mathcal{U}_D\}$
primo	$p(\boldsymbol{x}) \neq 0,\ p(\boldsymbol{x}) \notin \mathcal{U}_{D[x]}$ $p(\boldsymbol{x})\|a(\boldsymbol{x})b(\boldsymbol{x}) \Rightarrow p(\boldsymbol{x})\|a(\boldsymbol{x})$ ou $p(\boldsymbol{x})\|b(\boldsymbol{x})$
irredutível	$p(\boldsymbol{x}) \neq 0,\ p(\boldsymbol{x}) \notin \mathcal{U}_{D[x]}$ $p(\boldsymbol{x}) = a(\boldsymbol{x})b(\boldsymbol{x}) \Rightarrow a(\boldsymbol{x}) \in \mathcal{U}_{D[x]}$ ou $b(\boldsymbol{x}) \in \mathcal{U}_{D[x]}$ isto é $p(\boldsymbol{x}) = a(\boldsymbol{x})b(\boldsymbol{x}) \Rightarrow a(\boldsymbol{x}) = c \in \mathcal{U}_D$ ou $b(\boldsymbol{x}) = d \in \mathcal{U}_D$

É claro que podemos estender estas duas noções a um domínio de integridade D qualquer:

- $p \in D$ é *primo* se $p \neq 0$, $p \notin \mathcal{U}_D$ e $p|ab \Rightarrow p|a$ ou $p|b$;
- $p \in D$ é *irredutível* se $p \neq 0$, $p \notin \mathcal{U}_D$ e $p = ab \Rightarrow a \in \mathcal{U}_D$ ou $b \in \mathcal{U}_D$.

Portanto, os elementos irredutíveis são os que apenas admitem factorizações triviais e um elemento $p \neq 0$ é primo se e só se o respectivo ideal principal $\langle p \rangle$ é primo. É fácil verificar que nos anéis $\mathbb{Z}$ e $C[x]$ os elementos primos no sentido da definição acima são exactamente os elementos irredutíveis, e é apenas por razões históricas que usamos o termo "primo" em $\mathbb{Z}$ e o termo "irredutível" em $C[x]$. Não é esse o caso em todos os domínios de integridade, mas é possível identificar extensas classes de domínios onde estas duas noções são equivalentes, e onde é possível estabelecer uma generalização apropriada do Teorema Fundamental da Aritmética e do Teorema da Factorização Única em $C[x]$.

No caso geral, a única implicação que é válida é a seguinte:

$$\text{primo} \ \Rightarrow \ \text{irredutível}.$$

De facto, se $p \in D$ é primo e $p = ab$, então $p|a$ ou $p|b$. Se, por exemplo, $p|a$, então existe $x \in D$ tal que $a = px$. Concluímos então que $p = ab = pxb$, e como $p \neq 0$, $1 = xb$, ou seja, b é invertível. De igual forma, se $p|b$ concluímos que a é invertível.

A implicação recíproca é, em geral, falsa. Por exemplo, no domínio dos inteiros pares, 18 é irredutível mas não é primo, uma vez que $18|(6 \times 6)$ mas $18 \nmid 6$ (note

que neste caso não há factorização única: $36 = 6 \times 6 = 2 \times 18$). Outro exemplo: no domínio

$$\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\},$$

$9 = 3 \cdot 3 = (2 + \sqrt{-5})(2 - \sqrt{-5})$, donde $3|(2 + \sqrt{-5})(2 - \sqrt{-5})$. No entanto, 3, que é irredutível, não divide $2 + \sqrt{-5}$ nem $2 - \sqrt{-5}$, pelo que não é primo (note que também neste exemplo não há factorizações únicas).

No entanto, a demonstração, na Proposição 2.10, de que todo o polinómio irredutível em $C[x]$ é primo pode imediatamente ser adaptada a qualquer domínio de ideais principais D. Portanto:

Proposição 2.12 *Num domínio de ideais principais, um elemento é irredutível se e só se é primo.* ∎

Um elemento a de um domínio de integridade D diz-se *associado* de b (e escreve-se $a \sim b$) se $a|b$ e $b|a$. Um domínio D diz-se um *domínio de factorização única* (abreviadamente, d.f.u.) se as seguintes duas condições são satisfeitas:

- Para cada $d \in D$ ($d \neq 0$, $d \notin \mathcal{U}$), existem elementos irredutíveis $p_1, p_2, \ldots, p_n$ tais que $d = p_1 p_2 \cdots p_n$.
- Se $p_1, p_2, \ldots, p_n$ e $q_1, q_2, \ldots, q_m$ são irredutíveis, e $p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m$, então $n = m$ e existe uma permutação $\pi \in S_n$ tal que $p_i \sim q_{\pi(i)}$.

Por outras palavras, num domínio de factorização única, todo o elemento não-nulo e não invertível possui uma factorização num produto de elementos irredutíveis, e esta factorização é única a menos da ordem dos factores e da multiplicação de cada factor por uma unidade convenientemente escolhida. Por exemplo, em $\mathbb{Z}$, $1 \times 5 = 5 \times 1 = (-1) \times (-5) = (-5) \times (-1)$ são as únicas factorizações do primo 5 e $1 \times (-5) = (-5) \times 1 = (-1) \times 5 = 5 \times (-1)$ são as únicas factorizações do primo -5. Pelo Teorema Fundamental da Aritmética, $\mathbb{Z}$ é um domínio de factorização única. Pelo Teorema da Factorização Única em $C[x]$, $C[x]$ é um domínio de factorização única. Outro exemplo de domínio de factorização única é o anel dos *inteiros de Gauss*,

$$\mathbb{Z}[i] = \{a + ib \mid a, b \in \mathbb{Z}\}.$$

Mais exemplos: $D[x]$ é um d.f.u. sempre que D o é. Em particular, $\mathbb{Z}[x]$ é um d.f.u., assim como $D[x][y]$.

Pode ainda provar-se o seguinte:

Teorema 2.13 *Todo o domínio de ideais principais é um domínio de factorização única.*

O recíproco é falso, como o exemplo $\mathbb{Z}[x]$ mostra.

Observe-se que a factorização indicada na definição de d.f.u. pode equivalentemente ser expressa em potências de elementos irredutíveis, mas neste caso pode ser necessário incluir uma unidade u na factorização, que passa a ser da forma

$$d = up_1^{m_1} \cdots p_n^{m_n},$$

como enunciámos no teorema da factorização única em $C[x]$.

Mais pormenores:
[R. L. Fernandes e M. Ricou, Introdução à Álgebra, IST Press, 2004]
[M. Sobral, Álgebra, Universidade Aberta, 1996]

Apêndice 2: critérios de irredutibilidade (para as aulas práticas)

Como vimos, em $\mathbb{C}[x]$ e $\mathbb{R}[x]$ sabemos quais são os polinómios irredutíveis:

(1) Em $\mathbb{C}[x]$ os polinómios irredutíveis são os polinómios de grau 1.

[Pelo Teorema Fundamental da Álgebra, qualquer polinómio não constante, de coeficientes em $\mathbb{C}$, tem pelo menos uma raiz complexa α. Então, em $\mathbb{C}[x]$, qualquer polinómio de grau ≥ 2 factoriza-se sempre na forma $(x-\alpha)q(x)$, com $gr(q(x)) \geq 1$, pelo que é redutível.]

(2) Em $\mathbb{R}[x]$ os polinómios irredutíveis são os de grau 1 e os de grau 2 com binómio discriminante negativo ($a\boldsymbol{x}^2 + b\boldsymbol{x} + c$ tal que $b^2 - 4ac < 0$).

[Também pelo Teorema Fundamental da Álgebra: em $\mathbb{C}[x]$, qualquer polinómio $p(x)$ de grau ≥ 3 factoriza-se na forma $(x-\alpha)q_1(x)$, onde agora $gr(q_1(x)) \geq 2$; mas se α é raiz de $p(x)$, também o seu conjugado $\overline{\alpha}$ o é e, se $\alpha = a + ib$, $(x-\alpha)(x-\overline{\alpha}) = x^2 - 2ax + a^2 + b^2 \in \mathbb{R}[x]$. Portanto, $p(x) = (x^2 - 2ax + a^2 + b^2)q_2(x)$, onde $gr(q_2(x)) \geq 1$, é uma factorização de $p(x)$ em $\mathbb{R}[x]$, o que mostra que este polinómio é redutível.

No caso em que $p(x)$ tem grau 2 com discriminante não negativo, as suas duas raízes α_1 e α_2 são reais, pelo que se factoriza na forma $(x-\alpha_1)(x-\alpha_2)$ e é redutível.]

A situação é diferente em $\mathbb{Q}[x]$:

(3) Em $\mathbb{Q}[x]$ a identificação dos irredutíveis é mais difícil. Neste caso apenas apenas se conhecem condições suficientes de irredutibilidade e não se consegue indicar explicitamente os polinómios irredutíveis, como fizemos nos dois casos anteriores.

Em primeiro lugar vejamos que todo o polinómio de coeficientes inteiros que seja irredutível em $\mathbb{Z}[x]$ também o é em $\mathbb{Q}[x]$ (contudo, o recíproco é falso: $2x$ é irredutível em $\mathbb{Q}[x]$ mas é redutível em $\mathbb{Z}[x]$ — pois quer 2 quer x não são unidades de $\mathbb{Z}[x]$):

Lema 2.14 [Lema de Gauss] *Se um polinómio $p(\boldsymbol{x}) \in \mathbb{Z}[x]$ se pode escrever como produto de dois polinómios $a(\boldsymbol{x})$ e $b(\boldsymbol{x})$ de $\mathbb{Q}[x]$, com graus inferiores ao de $p(\boldsymbol{x})$, então existem $a_1(\boldsymbol{x})$ e $b_1(\boldsymbol{x})$ em $\mathbb{Z}[x]$ tais que $p(\boldsymbol{x}) = a_1(\boldsymbol{x})b_1(\boldsymbol{x})$, sendo $a_1(\boldsymbol{x})$ associado de $a(\boldsymbol{x})$ e $b_1(\boldsymbol{x})$ associado de $b(\boldsymbol{x})$.*

Deste lema conclui-se que

um polinómio de coeficientes inteiros é irredutível em $\mathbb{Q}[x]$ se e só se não pode decompor-se num produto de polinómios de grau ≥ 1 em $\mathbb{Z}[x]$.

É claro que a todo o polinómio de coeficientes racionais se pode associar um polinómio de coeficientes inteiros: basta multiplicá-lo pelo mínimo múltiplo comum dos denominadores dos coeficientes. Também é simples calcular as raízes racionais (logo os factores lineares) de polinómios de coeficientes inteiros:

Proposição 2.15 *Se o número racional $\frac{c}{d}$ (escrito na forma reduzida, ou seja, tal que $mdc(c,d) = 1$) é raiz do polinómio de coeficientes inteiros*

$$a_0 + a_1\boldsymbol{x} + a_2\boldsymbol{x}^2 + \cdots + a_n\boldsymbol{x}^n, \text{ com } n \geq 1,$$

então c divide a_0 e d divide a_n.

(Este resultado é muito útil. Por exemplo, se quisermos saber *se o polinómio $2\boldsymbol{x}^7 + 1 \in \mathbb{Z}_3[x]$ tem raízes no corpo $\mathbb{Z}_3$*, como $\mathbb{Z}_3$ tem apenas três elementos, é possível calcular o valor da respectiva função polinomial em cada um deles,

concluindo-se que 1 é a única raiz do polinómio. No entanto, se substituirmos $\mathbb{Z}_3$ por $\mathbb{Q}$, já não é possível calcular o valor da função polinomial em todos os elementos de $\mathbb{Q}$. Contudo, a proposição acima reduz o nosso campo de procura a um conjunto finito. Os elementos de $\mathbb{Q}$ que podem ser raízes do polinómio são 1, -1, 1/2 e -1/2. É fácil ver que *estes números não são raízes do polinómio.* Portanto ele não tem raízes racionais.)

Deste modo, determinar os factores lineares, quando existam, de um polinómio de coeficientes inteiros é simples. O problema é mais complicado para factores de ordem superior. O critério seguinte dá-nos uma condição suficiente de irredutibilidade em $\mathbb{Q}[x]$:

Teorema 2.16 [Critério de Eisenstein] *Seja* $a(\boldsymbol{x}) = a_0 + a_1\boldsymbol{x} + \cdots + a_n\boldsymbol{x}^n$ *um polinómio de coeficientes inteiros. Se existe um inteiro primo* p *tal que*

(1) $p|a_i$ *para* $i = 0, 1, \ldots, n-1$,

(2) $p \nmid a_n$,

(3) $p^2 \nmid a_0$,

então $a(\boldsymbol{x})$ *é irredutível em* $\mathbb{Q}[x]$.

Utilizando este critério, podemos concluir que são irredutíveis sobre $\mathbb{Q}$, por exemplo, os polinómios

$$\frac{1}{2}\boldsymbol{x}^4 - 2\boldsymbol{x}^2 + 1 = \frac{1}{2}(\boldsymbol{x}^4 - 4\boldsymbol{x}^2 + 2),$$
$$\boldsymbol{x}^7 + 11\boldsymbol{x}^4 - 22\boldsymbol{x} + 11,$$
$$\boldsymbol{x}^5 + 9\boldsymbol{x}^3 + 27\boldsymbol{x}^2 + 3$$

e muitos outros. Mas nada podemos concluir sobre, por exemplo, $\boldsymbol{x}^5 - 3\boldsymbol{x}^2 + 6\boldsymbol{x} + 5$. Como proceder neste caso?

É fácil concluir que o polinómio não tem factores lineares. Suponhamos então que

$$\boldsymbol{x}^5 - 3\boldsymbol{x}^2 + 6\boldsymbol{x} + 5 = (a_1\boldsymbol{x}^2 + b_1\boldsymbol{x} + c_1)(a_2\boldsymbol{x}^3 + b_2\boldsymbol{x}^2 + c_2\boldsymbol{x} + d_2)$$

é uma factorização desse polinómio em $\mathbb{Z}[x]$. Verifica-se com relativa facilidade que o sistema

$$\begin{cases} a_1a_2 = 1 \\ a_1b_2 + b_1a_2 = 0 \\ a_1c_2 + b_1b_2 + c_1a_2 = 0 \\ a_1d_2 + b_1c_2 + c_1b_2 = -3 \\ b_1d_2 + c_1c_2 = 6 \\ c_1d_2 = 5 \end{cases}$$

não tem soluções inteiras. Logo, o polinómio é irredutível em $\mathbb{Q}[x]$.

Este tipo de problemas pode resolver-se de modo mais rápido com a ajuda de outros critérios.

Dado um homomorfismo de anéis $\phi : A \to B$, é evidente que existe um homomorfismo $\overline{\phi} : A[x] \to B[x]$ tal que $\overline{\phi}_{|A} = \phi$, definido por

$$\overline{\phi}\Big(\sum_{i=0}^{n} a_i \boldsymbol{x}_i\Big) = \sum_{i=0}^{n} \phi(a_i)\boldsymbol{x}^i.$$

Teorema 2.17 *Sejam A um corpo, B um domínio de integridade, $\phi : A \to B$ um homomorfismo e $a(\boldsymbol{x}) \in A[x]$. Se $\overline{\phi}(a(\boldsymbol{x}))$ tem o mesmo grau de $a(\boldsymbol{x})$ e é irredutível em $B[x]$, então $a(\boldsymbol{x})$ é irredutível em $A[x]$.*

No caso mais geral de A ser um domínio de integridade, este resultado ainda é válido para polinómios mónicos:

Teorema 2.18 *Sejam A e B domínios de integridade, $\phi : A \to B$ um homomorfismo e $a(\boldsymbol{x}) \in A[x]$ mónico. Se $\overline{\phi}(a(\boldsymbol{x}))$ tem o mesmo grau de $a(\boldsymbol{x})$ e é irredutível em $B[x]$, então $a(\boldsymbol{x})$ é irredutível em $A[x]$.*

Exemplo: Consideremos o polinómio $a(\boldsymbol{x}) = \boldsymbol{x}^5 - 3\boldsymbol{x}^2 + 6\boldsymbol{x} + 5$ e o homomorfismo $\phi : \mathbb{Z} \to \mathbb{Z}_2$ que a cada inteiro faz corresponder o resto da sua divisão por 2. A imagem de $a(\boldsymbol{x})$ pelo homomorfismo $\overline{\phi} : \mathbb{Z}[x] \to \mathbb{Z}_2[x]$ é $\overline{\phi}(a(\boldsymbol{x})) = \boldsymbol{x}^5 + \boldsymbol{x}^2 + 1$. Como é fácil verificar, este polinómio não tem nenhuma raiz em $\mathbb{Z}_2$, pelo que $\overline{\phi}(a(\boldsymbol{x}))$ não tem factores lineares em $\mathbb{Z}_2[x]$. Suponhamos que

$$\boldsymbol{x}^5 + \boldsymbol{x}^2 + 1 = (a_1\boldsymbol{x}^2 + b_1\boldsymbol{x} + c_1)(a_2\boldsymbol{x}^3 + b_2\boldsymbol{x}^2 + c_2\boldsymbol{x} + d_2)$$

é uma factorização desse polinómio em $\mathbb{Z}_2[x]$. Verifica-se facilmente que o sistema

$$\begin{cases} a_1a_2 = 1 \\ a_1b_2 + b_1a_2 = 0 \\ a_1c_2 + b_1b_2 + c_1a_2 = 0 \\ a_1d_2 + b_1c_2 + c_1b_2 = 1 \\ b_1d_2 + c_1c_2 = 0 \\ c_1d_2 = 1 \end{cases}$$

não tem solução em $\mathbb{Z}_2$. Então $\overline{\phi}(a(\boldsymbol{x}))$ é irredutível em $\mathbb{Z}_2[x]$ e, consequentemente, pelo Teorema e pelo Lema de Gauss, $a(\boldsymbol{x})$ é irredutível em $\mathbb{Q}[x]$.

Se considerarmos o homomorfismo $\phi : \mathbb{Z} \to \mathbb{Z}_5$, que a cada inteiro faz corresponder o seu resto na divisão por 5, vem $\overline{\phi}(a(\boldsymbol{x})) = \boldsymbol{x}^5 + 2\boldsymbol{x}^2 + \boldsymbol{x}$, que não é irredutível em $\mathbb{Z}_5[x]$, pelo que neste caso já não podemos usar o teorema acima. Deste teorema podemos concluir que um polinómio $a(\boldsymbol{x})$ de coeficientes inteiros é irredutível sobre $\mathbb{Q}$ sempre que exista um homomorfismo $\phi : \mathbb{Z} \to B$ nas condições do teorema e $a(\boldsymbol{x})$ seja irredutível em $B[x]$. Em particular, se considerarmos, para algum primo p, o homomorfismo $\phi : \mathbb{Z} \to \mathbb{Z}_p$, que a cada inteiro faz corresponder o seu resto na divisão por p, temos:

Corolário 2.19 *Se $\overline{\phi}(a(\boldsymbol{x}))$ é irredutível em $\mathbb{Z}_p[x]$ e p não divide o coeficiente de maior grau de $a(\boldsymbol{x}) \in \mathbb{Z}[x]$, então $a(\boldsymbol{x})$ é um polinómio irredutível em $\mathbb{Q}[x]$.*

Exercícios

2.1. Determine o produto dos polinómios $f(x)$ e $g(x)$ do anel $A[x]$, sendo:

(a) $f(x) = 2x^5 + 1$, $g(x) = 2x^5 + 1$ e $A = \mathbb{Z}_4$.

(b) $f(x) = 2x^2 + 2x - 2$, $g(x) = 3x - 3$ e $A = \mathbb{Z}_6$.

(c) $f(x) = 2x^2 - 4x + 3$, $g(x) = 4x - 5$ e $A = \mathbb{Z}_8$.

2.2. Mostre que:

(a) Se A é um subanel de um anel B, então $A[x]$ é um subanel de $B[x]$.

(b) O conjunto dos *polinómios homogéneos* sobre um anel A, $\left\{\sum_{i=1}^{n} a_i x^i \mid n \in \mathbb{N},\ a_i \in A\right\}$, é um ideal de $A[x]$.

2.3. Averigúe se os ideais $\langle x \rangle$ e $\langle 2, x \rangle$ do domínio $\mathbb{Z}[x]$ são principais, primos ou maximais.

2.4. Sejam A um anel comutativo e a um elemento fixo de A. Considere a aplicação

$$\begin{array}{rccc} \phi_a : & A[x] & \longrightarrow & A \\ & f & \longmapsto & f(a) \end{array},$$

onde $f(a)$ denota o valor da função polinomial associada a f em a.

(1) Mostre que ϕ_a é um homomorfismo de anéis.

(2) Determine o núcleo de ϕ_a.

2.5. Sejam D um domínio de integridade e $f(x)$ um elemento não nulo de $D[x]$. Prove que $f(x)$ é invertível se e só se $\mathrm{gr}(f(x)) = 0$ e $f(x)$ for invertível considerado como elemento de D. Conclua que se $\mathbb{K}$ for um corpo, então os únicos elementos invertíveis de $\mathbb{K}[x]$ são os polinómios de grau zero. O resultado da alínea anterior é válido se D for um anel comutativo qualquer?

2.6. Sejam D um domínio de integridade e $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in D[x]$. Chama-se *derivada* de $p(x)$ ao polinómio $p(x)' = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1$. Prove que, para quaisquer $p(x), q(x) \in D[x]$ e para qualquer $\alpha \in D$:

(a) $(p(x) + q(x))' = p(x)' + q(x)'$ e $(p(x)q(x))' = p(x)'q(x) + p(x)q(x)'$.

(b) α é raiz de $p(x)$ de multiplicidade > 1 se e só se é simultaneamente raiz de $p(x)$ e $p(x)'$.

2.7. Sendo $f(x)$ e $g(x)$ elementos de $\mathbb{K}[x]$, determine o quociente e o resto da divisão de $f(x)$ por $g(x)$, para:

(a) $f(x) = x^4 + 4x^2 + 4$, $g(x) = x^2$ e $\mathbb{K} = \mathbb{Q}$.

(b) $f(x) = x^3 + 2x^2 - x + 2$, $g(x) = x + 2$ e $\mathbb{K} = \mathbb{Z}_3$.

(c) $f(x) = x^7 - 4x^6 + x^3 - 3x + 5$, $g(x) = 2x^3 - 2$ e $\mathbb{K} = \mathbb{Z}_7$.

2.8. Determine todos os primos ímpares p para os quais $x - 2$ divide $x^4 + x^3 + x^2 + x$ em $\mathbb{Z}_p[x]$.

2.9. Em cada uma das alíneas seguintes determine, em $\mathbb{R}[x]$, $d(x) = \text{mdc}(f(x), g(x))$ e $u(x)$, $v(x) \in \mathbb{R}[x]$ tais que $d(x) = u(x)f(x) + v(x)g(x)$.

(a) $f(x) = x^3 + 1$ e $g(x) = x^4 + x^3 + 2x^2 + x + 1$.

(b) $f(x) = x^3 + 2x^2 + 4x - 5$ e $g(x) = x^2 + x - 2$.

(c) $f(x) = x^3 + 3x^2 + 2x + 8$ e $g(x) = x^4 - 4$.

2.10. O anel quociente $\mathbb{Q}[x]/\left\langle 2x^5 - 6x^3 + 9x^2 - 15 \right\rangle$ é um corpo?

2.11. Sejam p um inteiro positivo primo e $f(x)$ um polinómio irredutível de $\mathbb{Z}_p[x]$ de grau n. Prove que o corpo $\mathbb{Z}_p[x]/\langle f(x) \rangle$ tem exactamente p^n elementos.

2.12. Dê exemplos de polinómios redutíveis sobre um corpo mas que não tenham nenhuma raiz nesse corpo.

2.13. Sendo C um corpo, prove que se $f(x) \in C[x]$ é de grau 2 ou 3 e não tem raízes em C então $f(x)$ é irredutível sobre C. Mostre que a recíproca é válida para polinómios de grau ≥ 2.

2.14. Seja C um corpo finito. Mostre que $C[x]$ contém polinómios irredutíveis de grau tão grande quanto se queira. [Sugestão: Imite a prova de Euclides da existência de um número infinito de primos].

2.15. Demonstre a Proposição 2.15.

2.16. Averigúe quais dos seguintes polinómios de $\mathbb{Z}[x]$ são irredutíveis sobre $\mathbb{Q}$ (em caso negativo, factorize-os como produto de polinómios irredutíveis):

(a) $x^3 - x + 1$.

(b) $x^3 - 2x - 1$.

(c) $x^3 - 2x^2 + x + 15$.

(d) $x^7 + 11x^3 + 33x + 22$.

(e) $x^5 + 2$.

(f) $x^3 + 2x^2 + 10$.

(g) $2x^5 - 6x^3 + 9x^2 - 15$.

2.17. Determine todas as raízes racionais dos seguintes elementos de $\mathbb{Q}[x]$:

(a) $x^{50} - x^{20} + x^{10} - 1$.

(b) $2x^2 - 3x + 4$.

(c) $\frac{1}{2}x^3 - 5x + 2$.

(d) $x^3 - 7x + 3$.

2.18. Mostre que, para quaisquer inteiros a e b, o polinómio $x^3 + (2a+1)x + (2b+1)$ é irredutível sobre $\mathbb{Q}$.

2.19.

(a) Calcule o produto $(2x^2 + x + 1)(2x^2 + 3x + 2)$ em $\mathbb{Z}_m[x]$, para $m = 2, 3, 6$.

(b) $x^4 + 2x^3 + 2x + 2$ é irredutível em $\mathbb{Z}_3[x]$?

2.20. Usando o critério de Eisenstein, prove que, se $n > 1$ e $p_1, p_2, \ldots, p_k$ são números primos distintos dois a dois, então $\sqrt[n]{p_1 p_2 \ldots p_k}$ é um número irracional. Será indispensável exigir que os números $p_1, p_2, \ldots, p_k$ sejam todos distintos?

2.21. Para cada $n \in \mathbb{Z}$, considere o polinómio $p_n(x) = x^2 + 100x + n$.

(a) Indique um conjunto infinito de inteiros n para os quais $p_n(x)$ é redutível sobre $\mathbb{Q}$, e prove esta redutibilidade.

(b) Indique um conjunto infinito de inteiros n para os quais $p_n(x)$ é irredutível sobre $\mathbb{Q}$, e prove esta irredutibilidade.

2.22. Determine $\mathbb{K}[x]/\langle f(x)\rangle$ e escreva as respectivas tabelas de anel para:

(a) $\mathbb{K} = \mathbb{Z}_2$ e $f(x) = x$.

(b) $\mathbb{K} = \mathbb{Z}_2$ e $f(x) = x^2 + x + 1$.

(c) $\mathbb{K} = \mathbb{Z}_3$ e $f(x) = x^2 + 2$.

2.23. Quais dos seguintes subconjuntos de $\mathbb{Q}[x]$ são ideais de $\mathbb{Q}[x]$? (Em caso afirmativo, calcule $p(x)$ mónico tal que $J = \langle p(x) \rangle$.) Quais desses ideais são maximais?

(a) $\{f(x) \in \mathbb{Q}[x] \mid f(1) = f(7) = 0\}$.

(b) $\{f(x) \in \mathbb{Q}[x] \mid f(2) = 0 \text{ e } f(5) \neq 0\}$.

(c) $\{f(x) \in \mathbb{Q}[x] \mid f(\sqrt{3}) = 0\}$.

(d) $\{f(x) \in \mathbb{Q}[x] \mid f(4) = 0 \text{ e } f(0) = f(1)\}$.

2.24. Se $p > 2$ é um número primo, mostre que há exactamente dois elementos $a \in \mathbb{Z}_p$ tais que $a^2 = 1$.

2.25. Seja p um inteiro primo. Prove que o *polinómio ciclotómico*

$$x^{p-1} + x^{p-2} \cdots + x + 1 = \frac{x^p - 1}{x - 1}$$

é irredutível em $\mathbb{Q}[x]$.

3. Teoria de Galois

Motivação

O desenvolvimento da Álgebra está intimamente ligado à resolução de equações polinomiais de coeficientes reais (ou complexos). Uma *equação polinomial* é uma equação do tipo

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0. \tag{3.0.1}$$

Ao primeiro membro chama-se, como vimos no capítulo anterior, um polinómio na indeterminada x.

Resolver a equação (3.0.1) é determinar as suas soluções (ou seja, as raízes do polinómio), isto é, os valores numéricos para x que transformam a equação numa identidade verdadeira.

A equação do primeiro grau, ou linear,

$$ax + b = 0 \quad (a \neq 0)$$

tem uma só solução, óbvia,

$$x = -\frac{b}{a}.$$

A solução de uma equação quadrática era já conhecida pelos matemáticos da Babilónia, que sabiam como "completar o quadrado", e foi popularizada no mundo ocidental durante o Renascimento, por traduções em latim do livro do matemático islâmico Muhammad al-Khowarizmi[7], *Al-jabr wa'l muqābalah*[8], publicado na primeira metade do século IX. Todos sabemos hoje que a equação do segundo grau

$$ax^2 + bx + c = 0 \quad (a \neq 0)$$

tem soluções dadas pela fórmula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Será possível encontrar uma fórmula semelhante para resolver equações do terceiro grau

$$ax^3 + bx^2 + cx + d = 0 \quad (a \neq 0)?$$

E de grau superior?

[7]Nome que deu origem às palavras *algarismo* — para designar cada um dos dígitos de numeração árabe — e *algoritmo* — o termo moderno que designa um procedimento sistemático para resolver problemas matemáticos.

[8]A partir de al-Khowarizmi, o termo *al-jabr* tornou-se sinónimo de resolver equações (*álgebra*).

Vejamos em primeiro lugar o que significa "fórmula semelhante". O que se pretende saber é se existe um processo geral para calcular as raízes de equações de grau superior a dois, a partir dos coeficientes, aplicando as operações racionais (adição, subtracção, multiplicação e divisão) e a extracção de raízes, um número finito de vezes. Soluções obtidas desta forma chamam-se *soluções por radicais*.

Em segundo lugar, observemos que na procura das raízes de um polinómio

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

é evidente que podemos, sem perda de generalidade, supor $a_n = 1$. Além disso, basta considerar o caso $a_{n-1} = 0$. Com efeito, supondo já $a_n = 1$, a mudança de variável

$$x = y - \frac{a_{n-1}}{n} \tag{3.0.2}$$

transforma o polinómio dado num polinómio em y em que o coeficiente de y^{n-1} é zero, sendo as raízes do primeiro polinómio facilmente calculáveis a partir das raízes deste novo polinómio.

`[confirme]`

No século XVI, matemáticos italianos descobriram uma fórmula para resolver as equações do terceiro e quarto graus (vale a pena referir que a descoberta destas fórmulas e a luta pela prioridade da sua descoberta tem uma história bastante curiosa e divertida). Geronimo Cardano (1501-1576), também conhecido por Cardan, inclui no seu livro *Ars Magna*, publicado em 1545, fórmulas para a resolução de equações do terceiro e quarto graus, atribuídas pelo autor, respectivamente, a Nicolo Tartaglia (1500-1565) e Ludovico Ferrari (1522-1565).

A "fórmula de Cardan", como é hoje conhecida, para resolver a equação cúbica da forma

$$y^3 + py = q,$$

escrita em linguagem actual, é a seguinte:

$$y = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

Usando (3.0.2), o caso geral de uma equação do terceiro grau

$$x^3 + bx^2 + cx + d = 0$$

pode ser reduzido a este caso pela translação $x = y - b/3$. A verificação, por substituição, de que a fórmula de Cardan fornece uma solução da equação deverá dar uma ideia do grau de dificuldade envolvido neste tipo de problema.

A equação do quarto grau pode também ser reduzida à solução de uma cúbica. Com efeito, podemos sempre supor, eventualmente após uma translação (3.0.2), que a quártica é da forma

$$x^4 + ax^2 + bx + c = 0.$$

Completando o quadrado, obtemos

$$x^4 + ax^2 + bx + c = 0 \Leftrightarrow (x^2 + a)^2 = ax^2 - bx - c + a^2.$$

O truque consiste em observar que então, para qualquer y, temos

$$\begin{aligned}(x^2 + a + y)^2 &= ax^2 - bx - c + a^2 + 2y(x^2 + a) + y^2 \\ &= (a + 2y)x^2 - bx + (a^2 - c + 2ay + y^2). \qquad (3.0.3)\end{aligned}$$

Como esta última equação é quadrática em x, podemos escolher y de forma a que seja um quadrado perfeito. Isto consegue-se precisamente, impondo que o discriminante $b^2 - 4(a + 2y)(a^2 - c + 2ay + y^2)$ seja zero, o que dá uma equação cúbica em y,

$$-8y^3 - 20ay^2 + (-16a^2 + 8c)y + (b^2 - 4a^3 + 4ac) = 0,$$

que pode ser resolvida com recurso à fórmula de Cardan. Para este valor de y, o membro direito de (3.0.3) fica igual ao quadrado perfeito

$$\Big(x - \frac{b}{2(a + 2y)}\Big)^2,$$

de forma que, extraindo as raízes em ambos os membros de (3.0.3), obtemos uma equação quadrática que pode ser resolvida.

Nos três séculos que se seguiram, muitos esforços foram feitos para obter uma fórmula resolvente para a equação quíntica. No princípio do século XIX, Niels Henrik Abel (1802-1829), na sequência de trabalhos de matemáticos eminentes como Joseph Lagrange (1736-1813) e Paolo Ruffini (1765-1833), provou que existem equações do quinto grau cujas soluções não podem ser obtidas por radicais. Este facto levantou de imediato um novo problema: dada uma equação desse grau como reconhecer se ela é ou não resolúvel por radicais?

Foi Évariste Galois (1811-1832) quem obteve uma condição necessária e suficiente para a resolubilidade por radicais de uma equação polinomial de qualquer grau e mostrou a impossibilidade de resolução da equação algébrica geral de grau maior ou igual a cinco. Este matemático, com uma vida breve e aventurosa, é considerado o criador da Álgebra tal como ela é entendida nos nossos dias e o seu trabalho teve consequências muito para além do problema original da resolução de

equações algébricas por radicais. Galois associou a cada equação um grupo, hoje chamado *grupo de Galois*; as propriedades desse grupo revelam a resolubilidade por radicais da equação. O feito de Galois é tanto mais notável quanto a noção de grupo era ainda incipiente nessa altura.

Para ilustrarmos as ideias de Galois, consideremos a equação quártica com coeficientes racionais

$$x^4 + x^3 + x^2 + x + 1 = 0.$$

Esta equação tem as raízes $r_k = e^{i\frac{2\pi k}{5}}$ $(k = 1, 2, 3, 4)$. Pensemos agora em todas as possíveis equações polinomiais, com coeficientes racionais, que são satisfeitas por estas raízes. Estas incluem, entre outras, as equações

$$\begin{aligned}
&r_1 + r_2 + r_3 + r_4 - 1 = 0,\\
&(r_1 + r_4)^2 + r_1 + r_4 - 1 = 0,\\
&r_1 r_4 = 1,\\
&(r_1)^5 - 1 = 0,\\
&(r_4)^5 - 1 = 0,\\
&\ldots
\end{aligned}$$

A observação chave é a seguinte: se considerarmos todas as permutações de $\{r_1, r_2, r_3, r_4\}$ que transformam equações deste tipo ainda em equações deste tipo, obtemos o chamado *grupo de Galois* G da equação. Por exemplo, a permutação $(14)(23)$ transforma todas as equações listadas em cima em equações dessa lista. Pode provar-se que, neste exemplo, $G = \{id, (1243), (14)(23), (1342)\}$. Galois descobriu que a estrutura deste grupo é a chave para a resolução desta equação (mas antes Galois teve de inventar o próprio conceito de grupo, inexistente até à data!).

Consideremos por exemplo o subgrupo $H = \{id, (14)(23)\}$. É simples verificar que as expressões polinomiais nas raízes, com coeficientes racionais, que são fixas pelos elementos de H são precisamente os polinómios em $y_1 = r_1 + r_4$ e $y_2 = r_2 + r_3$. Mas y_1 e y_2 são as soluções da equação quadrática

$$x^2 + x - 1 = 0.$$

Assim, e supondo que não conhecíamos as expressões das soluções da equação original, poderíamos descobri-las resolvendo primeiro esta equação quadrática, obtendo

$$r_1 + r_4 = \frac{-1 + \sqrt{5}}{2}, \qquad r_2 + r_3 = \frac{-1 - \sqrt{5}}{2},$$

e de seguida a equação quadrática

$$(x - r_1)(x - r_4) = x^2 - (r_1 + r_4)x + r_1 r_4 = 0,$$

já que de facto esta equação tem como coeficientes expressões polinomiais em y_1 e y_2 (pois $r_1 r_4 = 1$).

Note-se que o grupo de Galois pode ser caracterizado como o *grupo de simetrias* da equação original: são as transformações que levam soluções (raízes) em soluções preservando a estrutura algébrica das soluções. Este é precisamente o ponto de partida na exposição moderna da Teoria de Galois: constrói-se o corpo[9] $\mathbb{Q}(r_1, \ldots, r_n)$ gerado pelas raízes da equação, e os elementos do grupo de Galois aparecem como automorfismos destes corpos. Nesta linguagem, a Teoria de Galois consiste em transformar questões sobre a estrutura destes corpos em questões sobre a estrutura do grupo de automorfismos associado.

Extensões de corpos

As sucessivas extensões do conceito de número, dos naturais para os inteiros, racionais, reais e, finalmente, complexos foram impostas pela necessidade de *resolver equações polinomiais* ou, o que é equivalente, de *determinar raízes de polinómios*.

Os números irracionais surgiram com a necessidade de resolver a equação polinomial $x^2 - 2 = 0$, imposta pelo Teorema de Pitágoras. É bem conhecido que $x^2 + 1 = 0$ não tem solução no corpo dos reais. Para resolver uma tal equação foi necessária a introdução do número "imaginário" $i = \sqrt{-1}$. Portanto, estes problemas foram resolvidos com a construção de sucessivas extensões do conceito de número.

Nos nossos dias todos estes números nos são familiares mas é claro que não foi sempre assim. Atribui-se ao matemático do século XIX Leopold Kronecker (1823-1891) a seguinte frase:

Deus criou os números inteiros e tudo o resto é obra do homem.

Na resolução da equação do segundo grau, é com a maior tranquilidade que trabalhamos com o caso em que o binómio discriminante $b^2 - 4ac$ é negativo. Os números complexos são-nos perfeitamente familiares o que não sucedia no século XVI. De facto foi Cardan quem primeiro introduziu números da forma $a + \sqrt{-b}$, com a e b inteiros positivos. No entanto, fê-lo com sérias reservas e um forte sentimento de culpa.

[9] A noção de corpo só foi formalizada por Dedekind em 1879, mais de 50 anos depois da morte trágica de Galois.

É curioso notar que foi a determinação das soluções da equações de terceiro grau que levou à construção dos números complexos. As equações de grau dois e binómio discriminante negativo eram simplesmente classificadas como insolúveis mas, para a equação de terceiro grau, o caso muda de figura pois soluções reais são obtidas passando por números complexos. Por exemplo, a equação $x^3-15x-4=0$, pela regra de Cardan dá

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

e, consequentemente, é considerada sem solução. No entanto, ela tem três raízes reais: $4, -2+\sqrt{3}$ e $-2-\sqrt{3}$. Isto foi constatado por Bombelli, também matemático italiano do século XVI. Ele foi o primeiro a introduzir uma notação para o que hoje denotamos por i e $-i$ (a que ele chamou "più di meno" e "meno di meno") e a trabalhar com esses símbolos utilizando as regras bem conhecidas $i \times i = -1$, $-i \times i = 1$, etc. Às sucessivas extensões do conceito de número, dos naturais para os inteiros, racionais e reais, algumas bem conturbadas, tornava-se inevitável juntar mais uma: os números complexos.

A invenção de novos números se, por um lado, foi inevitável - por exemplo para resolver equações de terceiro grau, como já foi referido - não foi um processo pacífico nem facilmente aceite pela comunidade matemática como o revelam nomes tais como "irracionais" ou "imaginários".

O estudo que fizemos sobre anéis e corpos dá-nos, como veremos, um processo sistemático de "inventar raízes de polinómios". Neste processo os polinómios irredutíveis desempenham um papel determinante.

Sendo L um corpo, $K \subseteq L$ é um *subcorpo* de L quando K é um subconjunto não-vazio de L tal que $(K,+)$ é um subgrupo de $(L,+)$ e $(K\setminus\{0\},\cdot)$ é um subgrupo de $(L\setminus\{0\},\cdot)$.

[Observe: $K \subseteq L$ é um subcorpo de L sse
(1) $0, 1 \in K$
(2) $a - b \in K$ para quaisquer $a, b \in K$
(3) $ab^{-1} \in K$ para quaisquer $a \in K$, $b \in K \setminus \{0\}$]

EXTENSÃO DE UM CORPO

Diz-se que um corpo L é uma *extensão* de um corpo K, se K é um subcorpo de L. A extensão é *própria* quando $L \neq K$.

Consideremos o corpo de Galois de ordem p (prima), $\mathbb{F}_p = (\mathbb{Z}_p, \oplus_p, \otimes_p)$. Qualquer subcorpo K de $\mathbb{F}_p$ contém a identidade 1 logo contém os elementos

$$1+1, 1+1+1, \ldots, -1, -1-1, \ldots$$

Portanto $\mathbb{F}_p \subseteq K$, pelo que $K = \mathbb{F}_p$. Isto mostra que $\mathbb{F}_p$ não contém subcorpos próprios (isto é, $\neq \mathbb{F}_p$). Diz-se que $\mathbb{F}_p$ é um *corpo primo*. Portanto, os corpos primos são, em certo sentido, os *menores* corpos que existem. Outro exemplo de corpo primo é o corpo dos racionais: sendo K um subcorpo de $\mathbb{Q}$, se $1 \in K$ então imediatamente $\mathbb{Z} \subseteq K$, donde qualquer $\frac{n}{m} = nm^{-1}$ $(n \in \mathbb{Z}, m \in \mathbb{Z} \setminus \{0\})$ também pertence a K, isto é, $K = \mathbb{Q}$. Por outro lado, $\mathbb{R}$ e $\mathbb{C}$ não são primos.

Aproveitaremos agora para mostrar que os corpos $\mathbb{F}_p$ e $\mathbb{Q}$ são, a menos de isomorfismo, os *únicos* corpos primos que existem.

É fácil verificar que a intersecção de qualquer família de subcorpos de um corpo L é ainda um subcorpo de L.

`[Este facto decorre imediatamente do correspondente facto para grupos, provado em Álgebra I]`

Em particular, a intersecção de todos os subcorpos de L é um subcorpo P de L.

SUBCORPO PRIMO

A este subcorpo P chama-se *subcorpo primo* de L. Evidentemente, trata-se de um corpo primo.

Teorema 3.1 *O subcorpo primo de um corpo L é isomorfo a $\mathbb{F}_p$ ou a $\mathbb{Q}$, consoante a característica de L seja p ou 0.*

Demonstração. Consideremos a aplicação $\phi : \mathbb{Z} \to L$ definida por $\phi(n) = n1_L$, onde 1_L designa a identidade do corpo L. É evidente que ϕ é um homomorfismo de anéis:

- $\phi(n+m) = (n+m)1_L = n1_L + m1_L = \phi(n) + \phi(m)$.
- $\phi(nm) = (nm)1_L = (n1_L)(m1_L) = \phi(n)\phi(m)$.

Consideremos o *núcleo* de ϕ:

$$Nuc\,\phi = \{n \in \mathbb{Z} \mid \phi(n) = 0\}.$$

[Em Álgebra I foi observado que $Nuc\,\phi$ é um subgrupo de $\mathbb{Z}$.
Observe agora que $Nuc\,\phi$ é um ideal de $\mathbb{Z}$]

Pelo Teorema do Isomorfismo para anéis, $\phi(\mathbb{Z}) \cong \mathbb{Z}/Nuc\,\phi$.

[Este teorema é uma generalização imediata para anéis
do Teorema do Isomorfismo para grupos, estudado em Álgebra I:
*Se $\phi : A \to B$ é um homomorfismo de grupos (anéis), e N é o
núcleo de ϕ, então os grupos (anéis) $\phi(A)$ e A/N são
isomorfos.*]

Como qualquer subcorpo de L contém 1_L, também contém $\phi(\mathbb{Z})$. Logo $\phi(\mathbb{Z})$ está contido no subcorpo primo P de L. Por outro lado,

$$Nuc\,\phi = \{n \in \mathbb{Z} \mid n1_L = 0\} = \begin{cases} p\mathbb{Z} & \text{se } car(L) = p \\ \\ \{0\} & \text{se } car(L) = 0 \end{cases}$$

No primeiro caso, tem-se $\phi(\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z} \cong \mathbb{Z}_p$. Como $\mathbb{Z}_p$ é um corpo, $\phi(\mathbb{Z})$ é um corpo, donde necessariamente coincide com P.

No segundo caso, tem-se $\phi(\mathbb{Z}) \cong \mathbb{Z}/\{0\} \cong \mathbb{Z}$, donde $\mathbb{Z} \cong \phi(\mathbb{Z}) \subset P$. Portanto P contém uma cópia isomorfa de $\mathbb{Z}$. Estendendo o homomorfismo $\phi : \mathbb{Z} \to \phi(\mathbb{Z})$ a $\mathbb{Q}$, definindo $\overline{\phi} : \mathbb{Q} \to P$ por $\overline{\phi}(\frac{n}{m}) = \phi(n)\phi(m)^{-1}$, obtemos um isomorfismo de anéis, o que mostra que, neste caso, $P \cong \mathbb{Q}$.

[Alternativamente, podia observar-se, como fizemos para
$\mathbb{Q}$, que um corpo P que contenha (uma cópia de) $\mathbb{Z}$,
terá que conter necessariamente (uma cópia de) $\mathbb{Q}$, pois
$n, m \in P \Rightarrow \frac{n}{m} = nm^{-1} \in P$]

■

Exemplos: $\mathbb{Q}$ é o subcorpo primo de $\mathbb{R}$ e $\mathbb{C}$. Da mesma forma, $\mathbb{Q}$ é também o subcorpo primo de $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$.

Qualquer extensão L de um corpo K pode ser vista como um espaço vectorial sobre K, tomando para adição vectorial $\oplus$ a própria adição no corpo e para multiplicação escalar $*$ a multiplicação em L:

$$\text{Adição vectorial:} \quad a \oplus b := a + b, \ \forall a, b \in L$$

$$\text{Multiplicação escalar:} \quad \kappa * a := \kappa a, \ \forall \kappa \in K, \forall a \in L$$

[Exercício: Verifique]

Este resultado é fundamental para o desenvolvimento da teoria dos corpos, porque nos permite aplicar as ferramentas da álgebra linear.

GRAU DE UMA EXTENSÃO

Seja L uma extensão de K. O *grau* da extensão L sobre K, que denotaremos por $[L : K]$, é a dimensão do espaço vectorial L sobre K. A extensão L diz-se *finita* se $[L : K]$ for finita, e diz-se uma *extensão infinita*, caso contrário.

Vamos ver mais adiante técnicas para calcular o grau $[L : K]$ em certos casos importantes. Para já começamos com um resultado geral, que tem um papel nesta teoria análogo ao do Teorema de Lagrange na teoria dos grupos (finitos).

Teorema 3.2 [Teorema da Torre]
Sejam $M \supseteq L \supseteq K$ extensões sucessivas de um corpo K. Então

$$[M : K] = [M : L][L : K].$$

[Note que o produto à direita é simplesmente uma multiplicação de cardinais; no caso de algum dos graus ser infinito, a fórmula significa que $[M : K] = \infty$ se e só se $[M : L] = \infty$ ou $[L : K] = \infty$]

Demonstração. Seja $\{a_i\}_{i\in I}$ uma base do espaço vectorial L sobre K e seja $\{b_j\}_{j\in J}$ uma base do espaço vectorial M sobre L. Bastará provar que $\{a_i b_j\}_{i\in I, j\in J}$ é uma base do espaço vectorial M sobre K.

É claro que cada elemento $a_i b_j$ pertence a M, pois $a_i \in L \subseteq M$ e $b_j \in M$. Provemos que se trata de um conjunto de vectores linearmente independente sobre K:

Se

$$\sum_{i\in I, j\in J} \kappa_{ij} a_i b_j = 0,$$

com $\kappa_{ij} \in K$, isto significa que $\sum_{j\in J} \Big(\sum_{i\in I} \kappa_{ij} a_i \Big) b_j = 0$. Como cada $\sum_{i\in I} \kappa_{ij} a_i$ pertence a L e os b_j são linearmente independentes sobre L, então $\sum_{i\in I} \kappa_{ij} a_i = 0$ para qualquer $j \in J$. Mas os a_i são linearmente independentes sobre K e, portanto, $\kappa_{i,j} = 0$ para qualquer $i \in I$ e $j \in J$.

Finalmente, vejamos que se trata de um conjunto de geradores de M sobre K:

Seja $c \in M$. Então podemos escrever $c = \sum_{j\in J} l_j b_j$, onde $l_j \in L$, porque $\{b_j\}_{j\in J}$ é uma base de M sobre L. Mas, por sua vez, cada l_j é uma combinação

linear $l_j = \sum_{i\in I} \kappa_{ij} a_i$, porque $\{a_i\}_{i\in I}$ é uma base de L sobre K. Consequentemente, $c = \sum_{i,j} \kappa_{ij} a_i b_j$. ∎

Note que $[L : K] = 1$ se e só se $L = K$. De facto, se $[L : K] = 1$, seja $\{a\}$ uma base do espaço L sobre K; como $1 \in L$, podemos escrever $1 = \kappa a$ para algum $\kappa \in K$, o que mostra que $a = \kappa^{-1} \in K$ e, consequentemente, que $L \subseteq K$. O recíproco é óbvio.

EXTENSÃO GERADA E EXTENSÃO SIMPLES

Seja L uma extensão de K. Se $S \subseteq L$ é um subconjunto, designamos por $K(S)$ a extensão de K *gerada* por S, ou seja, o *menor* subcorpo de L que contém $K \cup S$. É claro que $K(S)$ é uma extensão de K contida em L. Se $S = \{\theta_1, \ldots, \theta_n\}$ ou $S = \{\theta\}$, escrevemos simplesmente $K(\theta_1, \ldots, \theta_n)$ ou $K(\theta)$ em vez de $K(S)$. Neste último caso, $K(\theta)$ diz-se uma *extensão simples* de K.

Exemplos: (1) $\mathbb{R}(i) = \mathbb{C}$: Por definição, $\mathbb{R}(i)$ é o menor subcorpo de $\mathbb{C}$ que contém $\mathbb{R} \cup \{i\}$, em particular, $\mathbb{R}(i) \subseteq \mathbb{C}$. Como $\mathbb{R}(i)$ é um corpo terá que conter necessariamente todos os elementos da forma $a + ib$, com $a, b \in \mathbb{R}$. Portanto $\mathbb{C} \subseteq \mathbb{R}(i)$.

Se $z \in \mathbb{C}$ então z escreve-se na forma $a + ib$ com a e b únicos, o que implica que $\{1, i\}$ é uma base de $\mathbb{C}$ sobre $\mathbb{R}$. Logo $[\mathbb{C} : \mathbb{R}] = 2$. Como 2 é primo, segue do Teorema da Torre que se K é tal que $\mathbb{R} \subseteq K \subseteq \mathbb{C}$ então ou $[K : \mathbb{R}] = 1$ ou $[\mathbb{C} : K] = 1$, ou seja, $K = \mathbb{R}$ ou $K = \mathbb{C}$.

(2) $\mathbb{Q}(i) = \{a + ib : a, b \in \mathbb{Q}\} \subset \mathbb{C}$: Como $\mathbb{Q}(i)$ é um corpo, por definição, terá que conter necessariamente todos os elementos da forma $a + ib$, com $a, b \in \mathbb{Q}$. Quanto à inclusão recíproca, bastará assegurarmos que $\{a + ib : a, b \in \mathbb{Q}\}$ é um subcorpo de $\mathbb{C}$. Sejam $a + ib, c + id$ com $a, b, c, d \in \mathbb{Q}$. Não é difícil mostrar que $(a+ib) - (c+id)$ ainda pertence a $\{a+ib : a, b \in \mathbb{Q}\}$. Suponhamos que $c+id \neq 0$ (isto é, $c \neq 0$ ou $d \neq 0$). Então $c - id \neq 0$, pelo que

$$(a + ib)(c + id)^{-1} = \frac{a + ib}{c + id} = \frac{a + ib}{c + id}\,\frac{c - id}{c - id} = \frac{ac - bd}{c^2 + d^2} + i\frac{bc - ad}{c^2 + d^2}$$

ainda pertence a $\{a + ib : a, b \in \mathbb{Q}\}$.

É claro que, tal como no exemplo anterior, $[\mathbb{Q}(i) : \mathbb{Q}] = 2$, sendo $\{1, i\}$ a base de $\mathbb{Q}(i)$ sobre $\mathbb{Q}$.

(3) Do mesmo modo que no exemplo anterior, pode provar-se que

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

e $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$. Neste caso a base é $\{1, \sqrt{2}\}$.

(4) Note que para o elemento $\sqrt[3]{2}$ ainda se tem $\{a + b\sqrt[3]{2} : a, b \in \mathbb{Q}\} \subseteq \mathbb{Q}(\sqrt[3]{2})$, mas desta vez não temos igualdade (o elemento $\sqrt[3]{4} = (\sqrt[3]{2})^2$ pertence a $\mathbb{Q}(\sqrt[3]{2})$ mas não pertence a $\{a + b\sqrt[3]{2} : a, b \in \mathbb{Q}\}$). Neste caso,

$$\mathbb{Q}(\sqrt[3]{2}) = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} : a, b, c \in \mathbb{Q}\}$$

e $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$.

ELEMENTOS ALGÉBRICOS E TRANSCENDENTES

Seja L uma extensão de K e seja $\theta \in L$. Dizemos que θ é *algébrico sobre* K se existe um polinómio não-nulo $p(x) \in K[x]$ tal que $p(\theta) = 0$. Caso contrário, dizemos que θ é *transcendente sobre* K.

Exemplos: (1) Se $\theta \in K$ então θ é raiz de $x - \theta \in K[x]$ e portanto θ é algébrico sobre K.

(2) $\sqrt{2}$ e i são algébricos sobre $\mathbb{Q}$: $\sqrt{2}$ é raiz de $x^2 - 2 \in \mathbb{Q}[x]$ e i é raiz de $x^2 + 1 \in \mathbb{Q}[x]$.

(3) É um facto bem conhecido que os números reais π e e são ambos transcendentes sobre $\mathbb{Q}$, isto é, não existe nenhum polinómio $p(x) \in \mathbb{Q}[x]$ que tenha π ou e por raiz. As demonstrações destes factos envolvem análise infinitesimal e devem-se originalmente a Lindemann (1882) e a Hermite (1873), respectivamente.

Mas é claro que π e e já são algébricos sobre $\mathbb{R}$.

EXTENSÕES ALGÉBRICAS E TRANSCENDENTES

Uma extensão L de K diz-se uma *extensão algébrica de* K se <u>todos</u> os elementos de L são algébricos sobre K. Caso contrário, dizemos que L é uma *extensão transcendente de* K.

Proposição 3.3 *Seja L uma extensão finita de K. Então L é algébrica sobre K.*

Demonstração. Suponhamos que $[L : K] = n \in \mathbb{N}$. Para cada $\theta \in L$, $\{1, \theta, \theta^2, \ldots, \theta^n\}$ é um conjunto linearmente dependente de L sobre K (pois tem $n + 1$ vectores). Isso significa que existem $a_0, a_1, a_2, \ldots, a_n \in K$, não todos nulos, tais que

$$a_0 + a_1\theta + a_2\theta^2 + \cdots + a_n\theta^n = 0.$$

Então o polinómio

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots a_nx^n \in K[x]$$

tem a raiz θ, o que mostra que θ é algébrico sobre K. ∎

Portanto, uma extensão transcendente é necessariamente de dimensão infinita.

Seja L uma extensão de K e seja $\theta \in L$ um elemento algébrico sobre K. Consideremos o conjunto

$$I = \{p(x) \in K[x] : p(\theta) = 0\}.$$

[Exercício: I é um ideal de $K[x]$]

Como I é um ideal de $K[x]$, pela demonstração do Teorema 2.7, podemos concluir que existe um polinómio mónico $m_\theta(x) \in K[x]$, único, tal que $I = \langle m_\theta(x)\rangle$.

Este polinómio satisfaz as seguintes propriedades:

Proposição 3.4 *Seja $\theta \in L$ um elemento algébrico sobre K. Então:*

(1) *$m_\theta(x)$ é irredutível sobre K.*

(2) *Para cada $p(x) \in K[x]$, $p(\theta) = 0$ se e só se $m_\theta(x) \mid p(x)$.*

(3) *$m_\theta(x)$ é o polinómio mónico não-nulo em $K[x]$ de menor grau que tem θ por raiz.*

Demonstração. (1) Como $m_\theta(x)$ tem uma raiz, tem de ser de grau ≥ 1 necessariamente. Suponhamos que $m_\theta(x)$ era redutível, isto é, que $m_\theta(x) = p_1(x)p_2(x)$, com

$$1 \leq gr(p_1(x)), gr(p_2(x)) < gr(m_\theta(x)). \tag{3.4.1}$$

Então $0 = m_\theta(\theta) = p_1(\theta)p_2(\theta)$, donde $p_1(\theta) = 0$ ou $p_2(\theta) = 0$. Qualquer uma destas possibilidades contradiz (3.4.1): se $p_i(\theta) = 0$ ($i = 1$ ou $i = 2$), então $p_i(x) \in I$, ou seja, $m_\theta(x) \mid p_i(x)$, donde $gr(p_i(x)) \geq gr(m_\theta(x)$.

(2) É evidente: $m_\theta(x) \mid p(x) \Leftrightarrow p(x) \in \langle m_\theta(x)\rangle = I \Leftrightarrow p(\theta) = 0$.

(3) É consequência imediata de (2): seja $p(x)$ mónico; se $p(\theta) = 0$ então $m_\theta(x) \mid p(x)$, logo $p(x) = m_\theta(x)$ ou $gr(p(x)) > gr(m_\theta(x))$. ∎

POLINÓMIO MÍNIMO

O polinómio $m_\theta(x)$ chama-se o *polinómio mínimo* de θ sobre K.

Exemplos: x^2+1 é o polinómio mínimo de i sobre $\mathbb{R}$, x^2-2 é o polinómio mínimo de $\sqrt{2}$ sobre $\mathbb{Q}$ e $x-\sqrt{2}$ é o polinómio mínimo de $\sqrt{2}$ sobre $\mathbb{R}$.

Teorema 3.5 *Seja θ algébrico sobre K, com polinómio mínimo $m_\theta(x)$ sobre K. Então cada elemento $\lambda \in K(\theta)$ tem uma expressão única na forma $\lambda = p(\theta)$ onde $p(x) \in K[x]$ é tal que $gr(p(x)) < gr(m_\theta(x))$.*

[Por outras palavras: se $gr(m_\theta(x)) = n$ então existem $a_0, a_1, \ldots, a_{n-1} \in K$, únicos, tais que $\lambda = a_0 + a_1\theta + \cdots + a_{n-1}\theta^{n-1}$]

Demonstração. Comecemos por provar que todo o elemento λ de $K(\theta)$ se pode escrever na forma $p(\theta)$ para algum $p(x) \in K[x]$ tal que $gr(p(x)) < n$. É evidente que

$$K \cup \{\theta\} \subseteq \{p(\theta) : p(x) \in K[x]\} \subseteq K(\theta).$$

Mas $\mathcal{S} := \{p(\theta) : p(x) \in K[x]\}$ é um subcorpo de $K(\theta)$:

- Se $p(\theta), q(\theta) \in \mathcal{S}$, é evidente que $p(\theta) - q(\theta) \in \mathcal{S}$, pois $p(x) - q(x) \in K[x]$.
- Se $p(\theta), q(\theta) \in \mathcal{S}$, com $q(\theta \neq 0$ então, como θ não é raiz de $q(x)$, pela propriedade (2) na Proposição, $m_\theta(x) \nmid q(x)$, donde $\text{mdc}(m_\theta(x), q(x)) = 1$, uma vez que $m_\theta(x)$ é irredutível sobre K. Isto significa que existem polinómios $a(x), b(x) \in K[x]$ tais que $1 = a(x)m_\theta(x) + b(x)q(x)$. Mas então $1 = a(\theta)m_\theta(\theta) + b(\theta)q(\theta) = b(\theta)q(\theta)$, o que mostra que $b(\theta)$ é o inverso de $q(\theta)$ em $K(\theta)$. Portanto, $p(\theta)q(\theta)^{-1} = p(\theta)b(\theta)$, que ainda pertence a $\mathcal{S}$, porque $p(x)q(x) \in K[x]$.

Logo, $\{p(\theta) : p(x) \in K[x]\} = K(\theta)$.

Observemos agora que

$$\{p(\theta) : p(x) \in K[x]\} = \{p(\theta) : p(x) \in K[x],\ gr(p(x)) < n\},$$

uma vez que, para cada $p(x) \in K[x]$, $p(x) = q(x)m_\theta(x) + r(x)$, com $gr(r(x)) < gr(m_\theta(x))$, donde $p(\theta) = q(\theta)m_\theta(\theta) + r(\theta) = r(\theta)$.

Em conclusão, $K(\theta) = \{p(\theta) : p(x) \in K[x],\ gr(p(x)) < n\}$, o que mostra que todo o elemento se pode escrever na forma desejada. Finalmente, provemos a unicidade: se $\lambda = p(\theta) = q(\theta)$, com $p(x), q(x) \in K[x]$ ambos de grau $< n$,

então $gr(p(x) - q(x)) < n$. Mas $p(\theta) - q(\theta) = 0$. Se $p(x) \neq q(x)$, o polinómio $p(x) - q(x)$ seria um polinómio não-nulo de grau $< n$ com a raiz θ, o que contradiz a propriedade (3) da Proposição 3.4. ∎

Daqui decorre imediatamente que toda a extensão algébrica simples é finita:

Corolário 3.6 *Se θ é algébrico sobre K e $gr(m_\theta(x)) = n$, então $[K(\theta) : K] = n$ e $\{1, \theta, \theta^2, \ldots, \theta^{n-1}\}$ é uma base do espaço vectorial $K(\theta)$ sobre K.* ∎

[Agora entende-se porque se chama *grau* da extensão à dimensão $[K(\theta) : K]$: este número coincide com o grau do polinómio mínimo $m_\theta(x)$]

Exemplos: (1) O que fizemos nos exemplos da página 60 pode agora ser feito de modo muito mais rápido: por este corolário, segue imediatamente que, para qualquer inteiro primo p, $[\mathbb{Q}(\sqrt{p}) : \mathbb{Q}] = 2$ e $\{1, \sqrt{p}\}$ é uma base de $\mathbb{Q}(\sqrt{p})$ sobre $\mathbb{Q}$; basta para isso observar que $x^2 - p$ é o polinómio mínimo de $\sqrt{p}$ sobre $\mathbb{Q}$.

(2) Consideremos a extensão $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ de $\mathbb{Q}$. Podemos olhar para $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ como a extensão simples $\mathbb{Q}(\sqrt{2})(\sqrt{3})$ de $\mathbb{Q}(\sqrt{2})$. Pelo Teorema da Torre,

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})]\ [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})].$$

Qual é o polinómio mínimo de $\sqrt{3}$ sobre $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$? $\sqrt{3}$ é raiz de $x^2 - 3 \in \mathbb{Q}[x] \subset \mathbb{Q}(\sqrt{2})[x]$. Será que este polinómio é irredutível sobre $\mathbb{Q}(\sqrt{2})$? Sim, pois as suas duas raízes $\pm\sqrt{3}$ não pertencem a $\mathbb{Q}(\sqrt{2})$:

Com efeito, $\pm\sqrt{3} = a + b\sqrt{2}$ para algum par a, b de racionais implicaria $3 = a^2 + 2b^2 + 2ab\sqrt{2}$, ou seja,

$$\sqrt{2} = \frac{3 - a^2 - 2b^2}{2ab} \in \mathbb{Q} \quad \text{(no caso } a, b \neq 0)$$

ou $3 = 2b^2$ (no caso $a = 0$) ou $3 = a^2$ (no caso $b = 0$), uma contradição, em qualquer um dos três casos.

Portanto, $x^2 - 3$ é o polinómio mínimo de $\sqrt{3}$ sobre $\mathbb{Q}(\sqrt{2})$, pelo que

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] = 2,$$

sendo $\{1, \sqrt{3}\}$ uma base de $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ sobre $\mathbb{Q}(\sqrt{2})$.

Em conclusão, $[\mathbb{Q}(\sqrt{2},\sqrt{3}):\mathbb{Q}]=4$ e, pela demonstração do Teorema da Torre, $\{1,\sqrt{2},\sqrt{3},\sqrt{6}\}$ constitui uma base de $\mathbb{Q}(\sqrt{2},\sqrt{3})$ sobre $\mathbb{Q}$. Assim,

$$\mathbb{Q}(\sqrt{2},\sqrt{3})=\{a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6}\mid a,b,c,d\in\mathbb{Q}\}.$$

[Por vezes, uma extensão está escrita de tal maneira que ''esconde'' a sua simplicidade. Por exemplo, a extensão $\mathbb{Q}(\sqrt{2},\sqrt{3})$ é simples porque coincide com $\mathbb{Q}(\sqrt{2}+\sqrt{3})$, como facilmente se pode verificar]

Aplicações

Construções com régua e compasso

Nesta altura do curso já podemos tirar dividendos dos nossos esforços: o grau de uma extensão algébrica é uma ferramenta muito poderosa. Antes mesmo de entrarmos a sério na Teoria de Galois, podemos aplicar o grau à resolução de vários problemas geométricos famosos, inventados pelos Gregos.

Os matemáticos da Grécia Antiga exprimiam de forma geométrica muitos dos seus conceitos e ideias. Mas, segundo Platão, as únicas figuras geométricas perfeitas eram a recta e a circunferência. Isto tinha o efeito de restringir os instrumentos disponíveis para efectuar construções geométricas a dois: em geral, só admitiam como válidas construções geométricas que pudessem ser obtidas pelo uso exclusivo do compasso e da régua não graduada (isto é, sem escala).

Apesar da sua grande habilidade, há algumas construções aparentemente simples para as quais não conseguiram descobrir um método de construção. Não é surpreendente que os Gregos tenham achado essas construções tão difíceis; são impossíveis de realizar! Mas os Gregos não tinham nem os métodos para provar essa impossibilidade nem, ao que parece, nenhuma suspeita de que as construções eram de facto impossíveis[10]

Esses problemas ficaram pois em aberto e só viriam a ser resolvidos nos finais do século XIX, com a ajuda da Álgebra, depois de convenientemente reformulados em questões da Teoria dos Corpos (mais concretamente, extensões de corpos).

[10]Sabiam, no entanto, que, sem essas imposições "platónicas", os problemas podiam ser resolvidos.

Entre os mais famosos desses problemas contam-se quatro que ficaram conhecidos por:

(I) **Problema da duplicação de um cubo**;

(II) **Problema da trissecção de um ângulo arbitrário**;

(III) **Problema da quadratura do círculo**;

(IV) **Problema da inscrição de um heptágono regular numa circunferência**.

Descrição dos problemas

O Problema I consiste em construir um cubo com o dobro do volume de um cubo dado. Se tomarmos um cubo de aresta 1, o problema consiste em construir um segmento de comprimento $\sqrt[3]{2}$.

O Problema II questiona a existência de um método geral de divisão de qualquer ângulo em três partes iguais (há vários ângulos que podem ser trissecados com régua e compasso; a questão está em saber se todos o são).

O Problema III está ligado ao cálculo da área do círculo. Consiste em saber se é possível construir um quadrado cuja área é igual à de um círculo dado. Partindo de um círculo de raio unitário a questão resume-se a construir um segmento de comprimento $\sqrt{\pi}$.

Quanto ao Problema IV, consiste em inscrever um heptágono regular numa circunferência dada.

História dos problemas

Uma referência ao Problema I aparece num documento antigo, supostamente escrito por Eratóstenes ao Rei Ptolomeu III cerca do ano 240 a.C.:

> *Diz-se que um dos antigos poetas trágicos descreveu Minos preparando um túmulo cúbico para Glaucus e declarando, quando observou que cada lado media 100 pés: "O túmulo que escolhestes é pequeno demais para túmulo real. Duplica-o [em volume] sem lhe modificar a forma. Conseguirás isso se duplicares cada lado do túmulo." Mas estava errado. Quando se duplicam os lados, a área aumenta quatro vezes e o volume oito vezes. Tornou-se um assunto de investigação entre os geómetras o modo como se poderá duplicar o volume dado sem modificar a forma. E este problema foi chamado de duplicação do cubo, pois dado um cubo pretendia-se duplicá-lo ...*

As origens do Problema II são obscuras. Os Gregos preocupavam-se com a construção de polígonos regulares, e é bem provável que o problema da trissecção tenha surgido neste contexto, pois a construção de um polígono regular com nove lados necessita da trissecção de um ângulo.

A história do Problema III está ligada ao cálculo da área de um círculo. O *Papiro de Rhind*[11] contém informação acerca disto. O manuscrito foi copiado pelo escriba Ahmes, por volta de 1650 a.C., a partir de um trabalho mais antigo.

Ao longo dos anos estes problemas foram abordados por muitos matemáticos. Curiosamente têm também fascinado muitos matemáticos amadores. No tempo dos gregos usava-se a palavra especial $\tau\varepsilon\tau\rho\alpha\gamma\omega\nu\iota\zeta\varepsilon\iota\nu$[12] para denominar estes curiosos. Em 1775, a Academia de Paris achou por bem proteger os seus funcionários da perda de tempo e energia com a examinação das "soluções" destes problemas apresentadas por matemáticos amadores; decretou que mais nenhuma solução destes problemas seria analisada.

Estes problemas foram finalmente resolvidos no século XIX. Em 1837, Wantzel resolveu os Problemas I, II e IV. Em 1882, Lindemann solucionou o terceiro, ao provar a transcendência de π sobre o corpo dos racionais.

Porque é que decorreram tantos séculos até estes problemas serem resolvidos? Por dois tipos de razões:

- as construções requeridas são impossíveis;
- Embora os problemas sejam geométricos, foi recorrendo a técnicas algébricas que essa impossibilidade foi demonstrada. Essas técnicas, nomeadamente a construção de extensões do corpo dos números racionais, só começaram a desenvolver-se no século XIX.

Descrição das regras impostas pelos Gregos

Todos aprendemos a efectuar construções geométricas com compasso e régua não graduada, isto é utilizando a régua apenas para traçar o segmento que une (ou a recta que passa por) dois pontos. Por exemplo, sabemos bissecar um ângulo, construir a mediatriz dum segmento, traçar por um ponto uma recta paralela a uma recta dada, etc. No entanto, com os mesmos instrumentos e regras, há várias construções que são impossíveis de realizar, tais como as dos problemas famosos acima referidos, como veremos.

[11] O manuscrito matemático mais antigo que se conhece.
[12] Significa *preencher o tempo com a quadratura.*

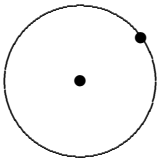
As regras para estas construções foram impostas pelos geómetras gregos e são muito estritas. Usando somente uma régua e um compasso, podemos realizar uma grande variedade de construções[13]. Em todos estes problemas são-nos dados alguns pontos, alguns segmentos de recta passando por esses pontos e, eventualmente, algumas circunferências. A partir deles podemos construir, usando a régua e o compasso como adiante se descreve, novos segmentos e circunferências. Note que *a régua é usada como mero instrumento auxiliar para traçar linhas direitas mas não para medir ou marcar distâncias.* Obtemos novos pontos onde o novo segmento de recta ou a nova circunferência intersecta outro segmento ou circunferência já existentes.

As regras de utilização da régua e do compasso são então as seguintes:

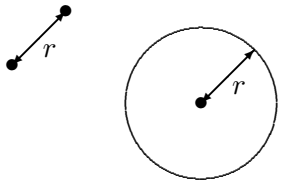
(1) A régua pode ser usada para traçar uma nova linha, com a extensão que quisermos, através de quaisquer dois pontos previamente na figura;

(2) O compasso pode ser usado para traçar novas circunferências, de dois modos:

   (a) Coloque uma das extremidades do compasso num dos pontos dados e a outra extremidade noutro dos pontos dados e trace a circunferência (ou um arco de circunferência):



   (b) Coloque o compasso como em (a), mas de seguida mova (sem alterar a abertura do compasso) uma das extremidades para um terceiro ponto na figura dada. Trace aí a circunferência (ou arco de circunferência), com este terceiro ponto como centro:



Observação. Em rigor, o nosso uso do compasso é mais versátil que o permitido pelos Gregos. De facto, o compasso imaginado pelos Gregos só podia ser utilizado segundo a regra 2(a) (não admitiam a regra 2(b)). Presumivelmente, os Gregos

[13] Algumas destas construções estão descritas com pormenor em muitos livros de Geometria Plana.

olhavam o seu compasso como não tendo existência logo que fosse levantado da folha de papel e portanto não podia ser utilizado directamente para transferir comprimentos, como em 2(b). Contudo, ao admitirmos a regra 2(b) não estamos a alterar o jogo em nada, pois pode-se provar que qualquer construção que se possa fazer seguindo as regras 1, 2(a) e 2(b) pode também ser realizada somente com as regras 1 e 2(a). A única diferença é que esta última construção poderá eventualmente envolver mais passos do que a primeira.

Não é difícil descrever construções, nas condições referidas, que levem, por exemplo, à divisão de um segmento de recta num número qualquer de partes iguais, ao traçado de uma paralela ou de uma perpendicular a uma recta dada, passando por um ponto dado, à bissecção de um ângulo dado, etc. Por exemplo:

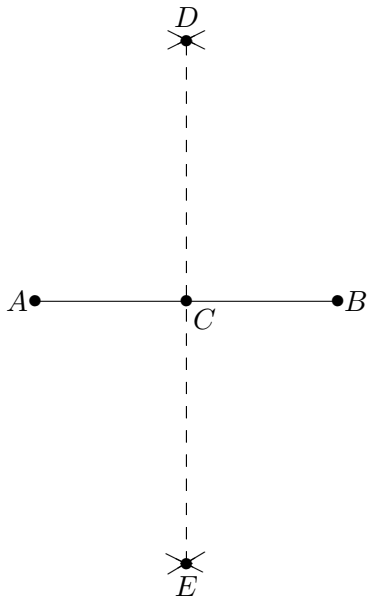
Problema [Bissecção de um segmento de recta]: Dados dois pontos A e B, construa o ponto médio C do segmento de recta $[AB]$.

Método de construção:

(1) Ponha o compasso em A e estenda a outra extremidade do compasso até que esteja exactamente em B. Desenhe então um arco na região acima de $[AB]$ e um outro na região abaixo de $[AB]$.

(2) Ponha o compasso em B e estenda a outra extremidade até que esteja exactamente em A. Desenhe arcos que intersectem os arcos de (1). Designe os pontos de intersecção por D e E, respectivamente.

(3) Com o auxílio da régua trace o segmento $[DE]$. O ponto C requerido é o ponto de intersecção de $[DE]$ com $[AB]$:



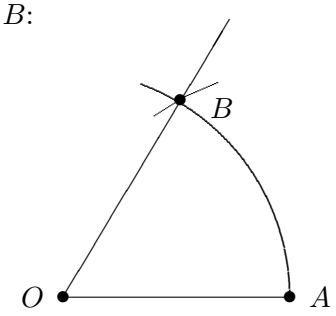
[É claro que é preciso provar que C é de facto o ponto médio de $[AB]$, o que pode ser feito sem grande dificuldade]

■

Outros exemplos:

Problema [Construção de um ângulo de $60°$]: Dados dois pontos O e A, construa o ponto B tal que $\widehat{AOB} = 60°$.

Método de construção: Trace arcos de raio $[OA]$ e centros em O e A. Designe o seu ponto de intersecção por B:

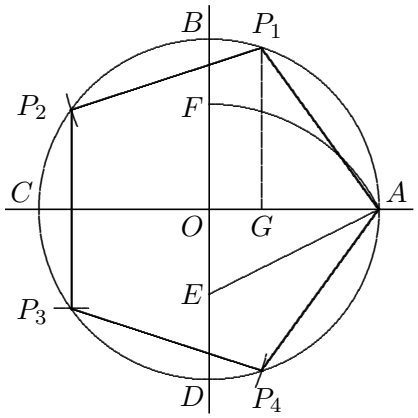


[O ângulo $\widehat{AOB}$ mede $60°$, uma vez que o triângulo $[AOB]$ é equilátero]

■

Problema [Inscrição de um pentágono regular numa circunferência (unitária)]: Dados os pontos $A = (1, 0)$, $B = (0, 1)$, $C = (-1, 0)$ e $D = (0, -1)$ numa circunferência unitária, construa um pentágono regular inscrito nessa circunferência.

Método de construção: Dividindo o segmento $[OD]$ em duas partes iguais, marque o ponto E. Com o compasso centrado em E obtenha o arco $[AF]$. Obtenha o ponto G no eixo horizontal, de forma a que $\overline{OG} = \overline{OF}/2$. Finalmente obtenha o vértice P_1 do pentágono por intersecção da circunferência com a recta vertical que passa por G. Os restantes vértices P_2, P_3 e P_4 podem construir-se sequencialmente, a partir de P_1, com o compasso com uma abertura igual a $\overline{AP_1}$:



[$[AP_1]$ é, de facto, lado de um pentágono regular inscrito na circunferência: basta observar que $P_1 = (\cos\frac{2\pi}{5}, \sin\frac{2\pi}{5})$, pois, como $\overline{EA} = \sqrt{5}/2$, então $\overline{OF} = \frac{-1+\sqrt{5}}{2}$ e $\overline{OG} = \frac{\overline{OF}}{2} = \frac{-1+\sqrt{5}}{4} = \cos\frac{2\pi}{5}$]

■

Por volta de 300 a.C., nos diversos volumes dos "Elementos", Euclides sistematizou uma grande variedade de construções possíveis de realizar com régua e compasso:[14]

- *Livro 1, Proposição 1.* Dado um segmento de recta, construir um triângulo equilátero em que um dos lados seja esse segmento.
- *Livro 1, Proposição 2.* Com extremo num ponto dado, traçar um segmento de recta igual a um segmento de recta dado.
- *Livro 1, Proposição 9.* Bissecar um ângulo dado.

[14] Veja, por exemplo, [T. Heath, *The Thirteen Books of Euclid's Elements*, Dover, 1956].

- *Livro 1, Proposição 42.* Construir um paralelogramo com área igual à de um ângulo dado e que tenha um ângulo igual a um ângulo dado.

- *Livro 1, Proposição 44.* Construir um paralelogramo com área igual à de um triângulo dado, que tenha um ângulo igual a um ângulo dado e um lado igual a um segmento de recta dado.

- *Livro 1, Proposição 45.* Construir um paralelogramo com área igual à de um polígono dado e que tenha um ângulo igual a um ângulo dado.

- *Livro 2, Proposição 14.* Construir um quadrado com área igual à de um polígono dado.

- *Livro 4, Proposição 2.* Inscrever, numa circunferência dada, um triângulo equiangular a um triângulo dado.

- *Livro 4, Proposição 6.* Inscrever um quadrado numa circunferência dada.

- *Livro 4, Proposição 11.* Inscrever um pentágono regular numa circunferência dada.

- *Livro 4, Proposição 15.* Inscrever um hexágono regular numa circunferência dada.

- *Livro 4, Proposição 16.* Inscrever um polígono regular com 15 lados numa circunferência dada.

Para mais exemplos de construções, consulte o livro [A. Jones, S. A. Morris e K. R. Pearson, *Abstract Algebra and Famous Impossibilities*, Springer, 1994]. Aí pode ver, entre muitas outras coisas, que se podem construir, sem grande dificuldade, somas, produtos, quocientes e raízes quadradas:

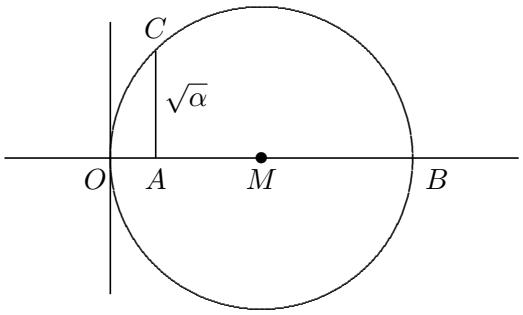
- (*Soma*) Dados dois segmentos de recta de comprimentos α e β, é possível construir segmentos de recta de comprimentos $\alpha \pm \beta$.

- (*Produto*) Dados dois segmentos de recta de comprimentos α e β, é possível construir um segmento de recta de comprimento $\alpha\beta$.

- (*Quociente*) Dados dois segmentos de recta de comprimentos α e $\beta \neq 0$, é possível construir um segmento de recta de comprimento α/β.

- (*Raiz quadrada*) Dado um segmento de recta de comprimento $\alpha > 0$, é possível construir um segmento de recta de comprimento $\sqrt{\alpha}$.

A construção neste caso pode ser realizada do seguinte modo: partindo dos extremos $A = (1, 0)$ e $B = (1 + \alpha, 0)$ do segmento, e da origem $O = (0, 0)$, construímos o ponto $(1, 1)$ e o ponto médio M do segmento $[OB]$. A intersecção da circunferência de centro em M e raio $\overline{MB}$ com a recta vertical definida pelos pontos A e $(1, 1)$ dá-nos um ponto C que está à distância $\sqrt{\alpha}$ de A, uma vez que $\overline{AM} = \frac{\alpha+1}{2} + 1$ e $\overline{MC} = \frac{\alpha+1}{2}$:



Portanto, começando com um segmento de comprimento 1, conseguimos construir todos os comprimentos racionais e alguns irracionais.

Todas estas construções devem seguir rigorosamente as regras do jogo. São portanto consideradas "ilegais" as construções que usem régua graduada ou curvas auxiliares, as construções aproximadas ou as construções com régua e compasso num número infinito de passos.

Retornemos aos quatro problemas famosos. O Problema I consiste em construir, com régua e compasso, um cubo com volume duplo de um dado cubo. Se o lado deste cubo medir 1 unidade de comprimento, o seu volume mede $1^3 = 1$, pelo que o volume do cubo a construir deverá medir 2 e, portanto, o seu lado deverá medir $\sqrt[3]{2}$. O problema resume-se pois a construir, a partir de um segmento de comprimento 1, um segmento de comprimento $\sqrt[3]{2}$. Como veremos, se tal fosse possível, então um determinado espaço vectorial teria a dimensão errada! Isto resolverá o Problema I.

Quanto ao Problema II, será suficiente apresentar um exemplo de um ângulo que não possa ser trissecado. Um tal exemplo é o ângulo de 60°. Mostraremos que este ângulo só poderá ser trissecado caso o ponto $(\cos 20°, 0)$ seja construtível, o que não é o caso uma vez que $\cos 20°$ é raiz do polinómio $8x^3 - 6x - 1 = 0$ que é irredutível sobre $\mathbb{Q}$. Mais uma vez veremos que isto pode ser justificado de modo rigoroso considerando as dimensões possíveis para um determinado espaço vectorial.

Como também veremos, as soluções de III e IV também se baseiam na discussão da dimensão de um espaço vectorial. Por exemplo, a impossibilidade de quadrar o

círculo é consequência do facto do espaço vectorial $\mathbb{Q}(\pi)$ sobre o corpo dos racionais ter dimensão infinita o que, por sua vez, é consequência de, como Lindemann provou, π ser transcendente sobre $\mathbb{Q}$.

A solução algébrica

Comecemos por formular a geometria das construções com régua e compasso em termos algébricos. A fim de enquadrarmos convenientemente o problema, consideremos o corpo $\mathbb{R}$ dos números reais e seja $\mathcal{P}$ uma parte qualquer de $\mathbb{R}^2$ de cardinal maior que 1.

PONTOS DO PLANO CONSTRUTÍVEIS

Um ponto P do plano diz-se *construtível num passo a partir de* $\mathcal{P}$ se P for a intersecção de duas rectas, uma recta e uma circunferência ou duas circunferências construídas a partir de pontos de $\mathcal{P}$, usando régua e compasso, de acordo com as regras (1) e (2).

Mais geralmente, um ponto P do plano diz-se *construtível a partir de* $\mathcal{P}$ se existirem pontos $P_1, P_2, \ldots, P_n = P$ tais que P_1 é construtível num passo a partir de $\mathcal{P}$ e, para cada $i = 2, 3, \ldots, n$, P_i é construtível num passo a partir de $\mathcal{P}_{i-1} := \mathcal{P} \cup \{P_1, P_2, \ldots, P_{i-1}\}$.

Por exemplo, no problema da bissecção de um segmento de recta, D e E são construtíveis num passo a partir de $\mathcal{P} = \{A, B\}$, e C é construtível a partir de $\mathcal{P}$ (em dois passos).

Seja K_0 o subcorpo de $\mathbb{R}$ gerado pelo conjunto

$$\{x, y \in \mathbb{R} \mid (x, y) \in \mathcal{P}\},$$

e seja $K_i = K_{i-1}(x_i, y_i)$, onde $P_i = (x_i, y_i)$. Desta construção resulta obviamente que

$$K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_n \subseteq \mathbb{R}.$$

[Observe: quando $\mathcal{P} = \{(0,0), (1,0)\}$, $K_0 = \mathbb{Q}$]

Por exemplo, no problema da bissecção de um segmento de recta, supondo $A = (0,0)$ e $B = (1,0)$, temos $K_0 = \mathbb{Q}$ e $K_1 = \mathbb{Q}(\sqrt{3}) = K_2$, pois $D = (1/2, \sqrt{3}/2)$, $E = (1/2, -\sqrt{3}/2)$ e $C = (1/2, 0)$.

O lema seguinte resulta do facto de as rectas e as circunferências utilizadas para a construção dos pontos $P_1, P_2, \ldots, P_n$ serem definidas por equações de graus

1 e 2 pois, como é bem sabido, uma recta de $\mathbb{R}^2$ pode ser definida, relativamente a um referencial ortonormado, por uma equação do tipo

$$ax + by + c = 0 \qquad (a, b, c \in \mathbb{R}),$$

e uma circunferência pode ser definida por uma equação do tipo

$$x^2 + y^2 + ax + by + c = 0 \qquad (a, b, c \in \mathbb{R}).$$

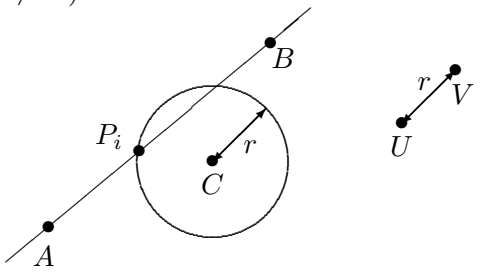
Lema 3.7 *Os números reais x_i e y_i são raízes em K_i de polinómios de coeficientes em K_{i-1} de grau 1 ou 2; em particular $[K_i : K_{i-1}] \in \{1, 2, 4\}$.*

Demonstração. Como $P_i = (x_i, y_i)$ é construtível a partir de $\mathcal{P}_{i-1}$, então ou é

- a intersecção de duas rectas definidas por pontos de $\mathcal{P}_{i-1}$, ou
- a intersecção de uma recta e uma circunferência definidas por pontos de $\mathcal{P}_{i-1}$, ou
- a intersecção de duas circunferências definidas por pontos de $\mathcal{P}_{i-1}$.

O primeiro caso é óbvio pelo que o deixamos como exercício: neste caso x_i e y_i pertencem mesmo a K_{i-1}, e $[K_i : K_{i-1}] = 1$. Quanto ao terceiro, pode ser deduzido imediatamente a partir do segundo caso, pelo que só provaremos este.

Suponhamos então que P_i é um ponto de intersecção de uma recta l, definida pelos pontos $A = (a_1, a_2)$ e $B = (b_1, b_2)$ de $\mathcal{P}_{i-1}$, e uma circunferência c de centro $C = (c_1, c_2) \in \mathcal{P}_{i-1}$ e raio r dado pela distância entre os pontos $U = (u_1, u_2)$ e $V = (v_1, v_2)$ de $\mathcal{P}_{i-1}$ ($U \neq V$).



A equação de l é

$$\frac{x - a_1}{b_1 - a_1} = \frac{y - a_2}{b_2 - a_2}$$

(onde deixamos os casos $a_1 = b_1$ ou $a_2 = b_2$ como exercício). A equação de c é

$$(x - c_1)^2 + (y - c_2)^2 = r^2.$$

Portanto, (x_i, y_i) é solução do sistema

$$\begin{cases} \dfrac{x-a_1}{b_1-a_1} = \dfrac{y-a_2}{b_2-a_2} \\ \\ (x-c_1)^2 + (y-c_2)^2 = r^2 \end{cases}$$

onde $a_1, a_2, b_1, b_2, c_1, c_2, u_1, u_2, v_1, v_2 \in K_{i-1}$ e, pelo Teorema de Pitágoras,

$$r^2 = (v_1 - u_1)^2 + (v_2 - u_2)^2 \in K_{i-1}.$$

Resolvendo em ordem a x concluímos que x_i é raiz do polinómio quadrático

$$(x-c_1)^2 + \Big(\frac{b_2-a_2}{b_1-a_1}(x-a_1) + a_2 - c_2\Big)^2 - r^2 \in K_{i-1}[x].$$

Se este polinómio for irredutível sobre K_{i-1} então $[K_{i-1}(x_i) : K_{i-1}] = 2$. Senão $[K_{i-1}(x_i) : K_{i-1}] = 1$.

Analogamente, resolvendo em ordem a y, concluímos que y_i é raiz de um polinómio quadrático em $K_{i-1}[y]$, pelo que também $[K_{i-1}(y_i) : K_{i-1}] \in \{1, 2\}$.

Em conclusão, em qualquer um dos três casos, $[K_{i-1}(x_i) : K_{i-1}]$ e $[K_{i-1}(y_i) : K_{i-1}]$, para $i = 1, 2, \ldots, n$, só podem tomar os valores 1 ou 2 e então, como

$$[K_{i-1}(x_i, y_i) : K_{i-1}(x_i)] \leq [K_{i-1}(y_i) : K_{i-1}],$$

também $[K_{i-1}(x_i, y_i) : K_{i-1}(x_i)] \in \{1, 2\}$. Consequentemente,

$$[K_i : K_{i-1}] = [K_{i-1}(x_i, y_i) : K_{i-1}(x_i)][K_{i-1}(x_i) : K_{i-1}]$$

só pode ser 1, 2 ou 4. ∎

Do Lema 3.7 segue o teorema fundamental desta secção:

Teorema 3.8 *Se o ponto $P = (x, y) \in \mathbb{R}^2$ é construtível a partir de $\mathcal{P}$ então $[K_0(x) : K_0]$ e $[K_0(y) : K_0]$ são potências de 2.*

Demonstração. Por definição, existe uma sequência finita de pontos de $\mathbb{R}^2$,

$$P_1, \ldots, P_n = P,$$

tais que, para cada $i = 1, \ldots, n$, o ponto $P_i = (x_i, y_i)$ é construtível num passo a partir de $\mathcal{P}_{i-1}$. Pelo lema anterior, $[K_i : K_{i-1}] \in \{1, 2, 4\}$. Ora

$$[K_n : K_0] = [K_n : K_{n-1}][K_{n-1} : K_{n-2}] \ldots [K_1 : K_0]$$

pelo que $[K_n : K_0]$ é uma potência de 2. Finalmente, as igualdades

$$[K_n : K_0] = [K_n : K_0(x)][K_0(x) : K_0]$$

$$[K_n : K_0] = [K_n : K_0(y)][K_0(y) : K_0]$$

provam a tese. ∎

Observação. Este resultado, que é a chave para a prova da impossibilidade dos problemas clássicos de construções com régua e compasso, como veremos adiante, permite-nos ter a certeza da não construtibilidade de muitos números a partir dos racionais.

Note-se que o recíproco deste teorema é falso: para um contra-exemplo consulte o Exemplo 13-18 em [3], que especifica um número θ, algébrico sobre $\mathbb{Q}$, com $[\mathbb{Q}(\theta) : \mathbb{Q}] = 4$ mas que, contudo, não é construtível a partir de $\mathbb{Q}$. Portanto, não podemos usar o Teorema 3.8 para concluir da construtibilidade de números θ tais que $[\mathbb{Q}(\theta) : \mathbb{Q}]$ é da forma 2^n.

No entanto, com a ajuda dos resultados enunciados na página 72, podemos fazer isso para muitos números θ. Por exemplo, para qualquer número α construtível a partir dos racionais, $\sqrt{\alpha}$ é também construtível. Aplicando, repetidamente este resultado, conjuntamente com o facto de que aplicações sucessivas das operações de corpo mantêm a construtibilidade, podemos então concluir que números do tipo

$$\sqrt{p+\sqrt{p}}, \quad \sqrt{5\sqrt{2}-3} \quad \text{ou} \quad \frac{\sqrt{5\sqrt{2}-3}+\sqrt[4]{2}}{5-\sqrt{2\sqrt{3}-4}}$$

são construtíveis a partir de $\mathbb{Q}$.

Com estes resultados, podemos finalmente resolver os quatro problemas geométricos clássicos.

Corolário 3.9 *Não é possível duplicar o cubo.*

Demonstração. Podemos partir de um cubo de lado unitário e, portanto, de volume 1, que tem como uma das arestas o segmento entre $(0,0)$ e $(1,0)$ no eixo OX. Um cubo de volume 2 teria um lado de comprimento α tal que $\alpha^3 = 2$.

A duplicação do cubo é equivalente à construção, a partir de $\mathcal{P} = \{(0,0),(1,0)\}$, de uma aresta de comprimento $\sqrt[3]{2}$, ou, o que é equivalente, à construção do ponto $(\sqrt[3]{2}, 0)$ a partir de $\mathcal{P}$. Como $K_0 = \mathbb{Q}$, se tal fosse possível, então $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}]$ seria uma potência de 2, pelo Teorema. Ora isto é impossível, visto que $\sqrt[3]{2}$ é raiz de

$x^3 - 2$, que é irredutível sobre $\mathbb{Q}$ pelo critério de Eisenstein. Portanto o polinómio mínimo de $\sqrt[3]{2}$ sobre $\mathbb{Q}$ é $x^3 - 2$ pelo que $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$. Logo o cubo não pode ser duplicado. ∎

Corolário 3.10 *Não é possível trissecar um ângulo de amplitude* $60°$.

Demonstração. Comecemos com $\mathcal{P} = \{(0,0),(1,0)\}$. Na nossa notação, $K_0 = \mathbb{Q}$. Construamos a circunferência c de centro $O = (0,0)$ que passa por $A = (1,0)$. Como vimos, é fácil construir o ponto $B \in c$ tal que $\widehat{AOB} = \frac{\pi}{3}$.

Se fosse possível trissecar o ângulo $\widehat{AOB}$, seria possível construir, a partir de $\mathcal{P}$, o ponto $C \in c$ tal que $\widehat{AOC} = \frac{\pi}{9}$ e, portanto, o ponto $(\cos\frac{\pi}{9}, 0) \in [OA]$. Mas então também o ponto $(2\cos\frac{\pi}{9}, 0)$ seria construtível, pelo que $[\mathbb{Q}(2\cos\frac{\pi}{9}) : \mathbb{Q}]$ seria uma potência de 2 o que é falso:

De facto, como para qualquer θ, $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$, temos

$$4\cos^3\frac{\pi}{9} - 3\cos\frac{\pi}{9} = \cos\frac{\pi}{3} = 1/2.$$

Então $\cos\frac{\pi}{9}$ é raiz do polinómio $8x^3 - 6x - 1 = 0$, ou seja, $2\cos\frac{\pi}{9}$ é raiz do polinómio $x^3 - 3x - 1$. Mas $x^3 - 3x - 1 \in \mathbb{Q}[x]$ é irredutível sobre $\mathbb{Q}$, porque não tem raízes racionais. Em conclusão $[\mathbb{Q}(2\cos\frac{\pi}{9}) : \mathbb{Q}] = 3$. ∎

Corolário 3.11 *Não é possível quadrar o círculo.*

Demonstração. Podemos supor que a unidade de medida é tal que o raio do círculo é 1, e então temos de construir um quadrado que tenha lado de medida $\sqrt{\pi}$. Portanto a quadratura do círculo equivale à construção do número $(\sqrt{\pi}, 0)$. Mas se $(\sqrt{\pi}, 0)$ fosse construtível então $[\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}] = 2^n$ para algum $n \in \mathbb{N}_0$, e então $[\mathbb{Q}(\pi) : \mathbb{Q}]$ dividiria 2^n e, em particular, π seria algébrico sobre $\mathbb{Q}$. Isto é absurdo visto que, como Lindemann mostrou em 1882, π é transcendente sobre $\mathbb{Q}$. ∎

Corolário 3.12 *Não é possível inscrever um heptágono regular numa circunferência.*

Demonstração. Se essa construção fosse possível, o ponto $(\cos\frac{2\pi}{7}, \sin\frac{2\pi}{7})$ seria construtível a partir de $\mathcal{P} = \{(0,0),(1,0)\}$. Mas tal não é verdade, pois o polinómio mínimo de $\cos\frac{2\pi}{7}$ sobre $\mathbb{Q}$ é $x^3 + \frac{1}{2}x^2 - \frac{1}{2}x - \frac{1}{8}$, pelo que $[\mathbb{Q}(\cos\frac{2\pi}{7}) : \mathbb{Q}] = 3$. ∎

```
[O Teorema não é verdadeiro na direcção inversa, como se tornará
    claro durante o estudo da Teoria de Galois: existem números
   algébricos cujo grau é uma potência de 2 que não dão origem a
   pontos do plano construtíveis. A Teoria de Galois fornece um
       critério mais eficiente para determinar se um dado par de
                números algébricos define um ponto construtível]
```

Construção de polígonos regulares

Acabámos de observar que, contrariamente ao caso do pentágono, é impossível construir um heptágono regular. E quanto ao caso geral de um polígono com n lados?

POLÍGONOS CONSTRUTÍVEIS

Um polígono diz-se *construtível* se todos os seus vértices são pontos construtíveis de $\mathbb{R}^2$.

Tal como vimos no caso $n = 7$, a construção de um polígono regular com n lados resume-se à construção do ponto $(\cos(2\pi/n), \sin(2\pi/n))$:

Se inscrevermos um polígono regular com n lados no círculo unitário em torno da origem de $\mathbb{R}^2$, com um vértice no ponto $(1, 0)$, então os outros vértices estão nos pontos

$$\Big\{(\cos\Big(\frac{2\pi k}{n}\Big), \sin\Big(\frac{2\pi k}{n}\Big)) \mid 0 < k < n\Big\}.$$

Se conseguirmos construir o ponto $(\cos(2\pi/n), \sin(2\pi/n))$, então conseguimos construir os outros vértices a partir deste. Assim, o polígono é construtível se e só se este ponto é construtível.

Os Gregos foram capazes de construir, com régua e compasso, polígonos regulares com 3 e 5 lados, mas não foram capazes de construir um com 7 lados (que, como acabámos de ver, é uma tarefa impossível).

Nenhum progresso foi feito neste problema durante mais de 2000 anos até que, em 1796, Gauss[15] surpreendeu o mundo matemático com a construção de um polígono regular com 17 lados.

Gauss descobriu mesmo um critério suficiente para que um polígono regular de n lados (um *n-gono*) seja construtível com régua e compasso:

O n-gono regular é construtível com régua e compasso se

$$n = 2^\alpha \quad ou \quad n = 2^\alpha p_1 \dots p_t,$$

[15] Na altura, com 19 anos!

onde $\alpha \in \mathbb{N}_0$, $t \in \mathbb{N}$ e os p_i são primos ímpares distintos da forma $p_i = 2^{2^{r_i}} + 1$ ($r_i \in \mathbb{N}_0$).

E se n não tiver tal forma? A resposta foi dada em 1837 por Pierre Wantzel, que provou o recíproco do Teorema de Gauss: se n não for desta forma, a construção é impossível.[16]

O número $F_r = 2^{2^r} + 1$, $r \in \mathbb{N}_0$, chama-se o *r-ésimo número de Fermat*, enquanto um *primo de Fermat* é um número F_r que seja primo. Aqui está uma tabela dos primeiros cinco números F_r que são primos de Fermat, descobertos pelo próprio Fermat:

r	$2^{2^r}+1$
0	3
1	5
2	17
3	257
4	65537

Fermat conjecturou que qualquer F_r é primo, mas Euler mostrou em 1732 que

$$F_5 = 2^{2^5} + 1 = 4294967297 = 641 \times 6700417.$$

Hoje ainda não se conhece mais nenhum primo de Fermat além dos encontrados por Fermat. Portanto, só se sabe que um polígono regular com p-lados (p primo) é construtível para $p = 2, 3, 5, 17, 257, 65537$. [17]

Extensões de decomposição

Depois do passeio por algumas aplicações do conceito de grau de uma extensão e do Teorema da Torre, voltemos ao estudo das extensões de corpos, começando por observar mais uma consequência do Teorema 3.5.

Sejam K um corpo, L uma extensão de K e $\theta \in L$. Consideremos o homomorfismo de anéis

$$\begin{array}{rccc} \phi: & K[x] & \to & L \\ & \displaystyle\sum_{i=0}^{n} a_i x^i & \mapsto & \displaystyle\sum_{i=0}^{n} a_i \theta^i \end{array}$$

[16] A prova do Teorema de Gauss e desta impossibilidade requere pouco mais do que as ideias que vimos até agora sobre extensões de corpos, e pode ser consultada em, por exemplo, [I. Stewart, *Galois Theory*, 3ª ed., Chapman & Hall, 2004].

[17] Para o polígono com 17 lados é apresentada uma construção em [H.S.M. Coxeter, *Introduction to Geometry*, 2ª ed., Wiley, 1989] e [I. Stewart, *Galois Theory*, 3ª ed., Chapman & Hall, 2004]. No primeiro destes livros podemos encontrar ainda uma demonstração muito elegante e curiosa de que 641 divide $2^{2^5} + 1$.

que a cada polinómio $p(x) = \sum_{i=0}^{n} a_i x^i$ faz corresponder o seu valor em θ. O núcleo $Nuc(\phi)$ deste homomorfismo é um ideal de $K[x]$, logo necessariamente principal. Por outro lado, o contradomínio de ϕ é claramente o subanel

$$K[\theta] := \{a_0 + a_1\theta + \cdots + a_n\theta^n \mid n \in \mathbb{N}, a_i \in K\}$$

de L.

[$K[\theta]$ é um subdomínio de integridade de $K(\theta)$]

Portanto $\phi : K[x] \to K[\theta]$ é um homomorfismo sobrejectivo de anéis, donde, pelo Teorema do Homomorfismo,

$$\frac{K[x]}{Nuc(\phi)} \cong K[\theta]. \tag{3.12.1}$$

Temos então dois casos:

(1) θ é algébrico sobre K: Então $Nuc(\phi) \neq \{0\}$, donde $Nuc(\phi) = \langle m(x)\rangle$, onde $m(x)$ é um polinómio irredutível que tem θ por raiz, e é o de menor grau nessas condições, ou seja, $m(x)$ é o polinómio mínimo de θ sobre K. Pelo Teorema 3.5 sabemos que, neste caso, $K(\theta) = K[\theta]$. Logo, por (3.12.1), temos

$$K(\theta) = K[\theta] \cong \frac{K[x]}{\langle m(x)\rangle}.$$

Poe exemplo, no caso $K = \mathbb{R}$ e $\theta = i$, obtemos $\mathbb{R}(i) \cong \mathbb{R}[x]/\langle x^2 + 1\rangle$. Já vimos que $\mathbb{R}(i) = \mathbb{C}$, logo

$$\mathbb{C} \cong \frac{\mathbb{R}[x]}{\langle x^2 + 1\rangle}.$$

(2) θ é transcendente sobre K: Neste caso, $Nuc(\phi) = \{0\}$, logo

$$K[\theta] \cong \frac{K[x]}{\{0\}} \cong K[x].$$

[Dado um domínio $K[x]$, seja $L := \left\{\frac{p(x)}{q(x)} \mid p(x), q(x) \in K[x], q(x) \neq 0\right\}$, com as operações óbvias de adição e multiplicação de ''fracções''.

Verifique que L é um corpo, o chamado *corpo das fracções* do domínio $K[x]$. Se identificarmos $a \in K$ com o elemento $\frac{a}{1}$ de L e $p(x) \in K[x]$ com o elemento $\frac{p(x)}{1}$ de L, não é difícil mostrar que L coincide com a extensão simples $K(x)$ de K. Uma vez que o polinómio $p(x) = a_n x^n + \cdots + a_0$ satisfaz $p(x) = 0 \in K[x]$ se e só se

$a_n = \cdots = a_0 = 0$, então x não é raiz de nenhum polinómio $p(x) \neq 0$ em $K[x]$, ou seja, x é transcendente sobre K. Portanto, o corpo das fracções $K(x)$ é uma extensão simples e transcendente de K. Assim, no caso (2), quando θ é transcendente sobre K, como $K[\theta] \cong K[x]$, os respectivos corpos de fracções $K(\theta)$ e $K(x)$ são isomorfos]

Em conclusão:

EXTENSÕES SIMPLES DE K
(1) θ é algébrico sobre K: $K(\theta) \cong \frac{K[x]}{\langle m(x) \rangle}$.
(2) θ é transcendente sobre K: $K(\theta) \cong K(x)$.

Quando estamos em subcorpos do corpo $\mathbb{C}$ dos números complexos, não há qualquer problema em determinar extensões em que um dado polinómio possua raízes, devido a uma propriedade fundamental de $\mathbb{C}$:

> **Teorema Fundamental da Álgebra**: qualquer polinómio (de grau ≥ 1) com coeficientes em $\mathbb{C}$ tem pelo menos uma raiz.

Isto implica que, em $\mathbb{C}$, todo o polinómio se decomponha em factores lineares do tipo $x - \theta$.

Mas existem muitos exemplos interessantes de corpos que não são subcorpos de $\mathbb{C}$ (por exemplo, os corpos $\mathbb{Z}_p$, importantes na Teoria dos Números). Para estes corpos não é claro que dado um polinómio com coeficientes nesse corpo, exista uma extensão onde o polinómio possua raízes (e, consequentemente, se possa decompor em factores lineares). Iremos agora abordar esta questão.

CORPO ALGEBRICAMENTE FECHADO

Um corpo K diz-se *algebricamente fechado* se qualquer polinómio $p(x) \in K[x]$, de grau ≥ 1, possui uma raiz em K.

Portanto, $\mathbb{C}$ é um corpo algebricamente fechado.

Proposição 3.13 *As seguintes afirmações são equivalentes:*

(i) *K é um corpo algebricamente fechado.*

(ii) *Todo o polinómio $p(x) = a_x x^n + \cdots + a_1 x + a_0 \in K[x]$ se decompõe num produto de factores lineares $a_n \prod_{i=1}^{n} (x - \theta_i)$.*

(iii) *Todo o polinómio irredutível de $K[x]$ tem grau 1.*

(iv) *Não existem extensões algébricas próprias de K.*

Demonstração.
"(i)⇔(ii)" Por hipótese, $p(x)$ tem uma raiz θ_1 em K, pelo que $p(x) = a_n(x-\theta_1)q_1(x)$. Por sua vez, $q_1(x)$ também tem uma raiz θ_2 em K, donde $p(x) = a_n(x-\theta_1)(x-\theta_2)q_2(x)$. Repetindo este raciocínio indutivamente chegaremos à conclusão que

$$p(x) = a_n \prod_{i=1}^{n}(x-\theta_i).$$

A implicação recíproca é trivial.

"(ii)⇔(iii)" Óbvio.

"(iii)⇒(iv)" Seja L uma extensão algébrica de K e seja $\theta \in L$. Como $[K(\theta) : K]$ é dada pelo grau de um polinómio irredutível, então $[K(\theta) : K] = 1$. Logo $K(\theta) = K$, ou seja, $\theta \in K$, o que mostra que $L = K$.

"(iv)⇒(i)" É óbvio. ∎

[O Teorema Fundamental da Álgebra assegura que $\mathbb{C}$ é algebricamente fechado. Outro facto importante é que qualquer corpo K pode ser imerso num corpo algebricamente fechado. Mais do que isso, existe uma extensão algebricamente fechada L de K, que é menor que todas as outras, no sentido de que, se L' é uma extensão algebricamente fechada de K, L' contém uma cópia isomorfa de L. Uma tal extensão chama-se o *fecho algébrico* de K. Portanto, *todo o corpo tem um fecho algébrico, que é único,* a menos de isomorfismo. As demonstrações deste facto e do Teorema Fundamental da Álgebra podem encontrar-se na bibliografia]

Recordemos a questão que começámos a estudar na aula anterior:

Seja K um corpo e $p(x) \in K[x]$ um polinómio de grau ≥ 1. Existirá uma extensão L de K onde $p(x)$ se decomponha em factores lineares?

É claro que se K for o corpo $\mathbb{Q}$ ou o corpo $\mathbb{R}$ há uma resposta óbvia: o corpo $\mathbb{C}$. E se K for o corpo $\mathbb{Z}_2$? Por exemplo, o polinómio $p(x) = x^2 + x + 1 \in \mathbb{Z}_2[x]$ é irredutível sobre $\mathbb{Z}_2$, uma vez que não tem raízes em $\mathbb{Z}_2$: $p(0) = 1$ e $p(1) = 1$. Existirá uma extensão de $\mathbb{Z}_2$ onde $p(x)$ já tenha raízes e possa ser então decomposto num produto de termos lineares?

A resposta a todas estas questões é afirmativa. Como K não é *a priori* um subcorpo de um corpo algebricamente fechado, tal extensão é, necessariamente, "abstracta". A construção desta extensão é dada no seguinte teorema, e é inspirada na construção de $\mathbb{C}$ a partir de $\mathbb{R}$, como o quociente $\mathbb{R}[x]/\langle x^2+1\rangle$.

Teorema 3.14 [Teorema de Kronecker]
Seja K um corpo e $p(x) \in K[x]$ um polinómio de grau $n \geq 1$. Existe uma extensão L de K onde $p(x)$ se decompõe num produto de termos lineares, da forma $L = K(\theta_1, \ldots, \theta_n)$, onde $\theta_1, \ldots, \theta_n$ são as raízes de $p(x)$ em L.

Demonstração. Como $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = a_n q(x)$, sendo $q(x) = x^n + \frac{a_{n-1}}{a_n}x^{n-1} + \cdots + \frac{a_1}{a_n}x + \frac{a_0}{a_n}$ mónico, é evidente que $p(x)$ se decompõe num produto de termos lineares se e só se $q(x)$ se decompõe num produto de termos lineares. Assim, sem perda de generalidade, podemos assumir que $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ é mónico. Podemos ainda supor que $p(x)$ é irredutível. Com efeito, se $p(x)$ for redutível, sendo $p(x) = p_1(x)p_2(x)\ldots p_t(x)$ a factorização (única) de $p(x)$ em polinómios mónicos irredutíveis, se o resultado for válido para polinómios irredutíveis, provamos imediatamente o caso geral:

$$\begin{array}{cc} p_1(x) = (x-\theta_1^1)\ldots(x-\theta_{m_1}^1) & \text{em } K(\theta_1^1,\ldots,\theta_{m_1}^1), \\ p_2(x) = (x-\theta_1^2)\ldots(x-\theta_{m_2}^2) & \text{em } K(\theta_1^2,\ldots,\theta_{m_2}^2), \\ \vdots & \vdots \\ p_t(x) = (x-\theta_1^t)\ldots(x-\theta_{m_t}^t) & \text{em } K(\theta_1^t,\ldots,\theta_{m_t}^t), \end{array}$$

pelo que

$$p(x) = (x-\theta_1^1)\ldots(x-\theta_{m_1}^1)\ldots(x-\theta_1^t)\ldots(x-\theta_{m_t}^t)$$

em $K(\theta_1^1,\ldots,\theta_{m_1}^1)\ldots(\theta_1^t,\ldots,\theta_{m_t}^t) = K(\theta_1^1,\ldots,\theta_{m_1}^1\ldots\theta_1^t,\ldots,\theta_{m_t}^t)$.

Suponhamos então que $p(x)$ é um polinómio mónico irredutível. Então $I := \langle p(x)\rangle$ é maximal e, como vimos anteriormente, $\psi : K \to K[x]/I$, definida por $\psi(a) = a + I$, é um homomorfismo injectivo,

$$[\psi(a) = \psi(b) \Leftrightarrow a + I = b + I \Leftrightarrow a - b \in I \Rightarrow a = b,$$
$$\texttt{pois } gr(a-b) = 0 \texttt{ e } gr(p(x) \geq 1]$$

donde $K \cong \psi(K) \subseteq K[x]/I$. Portanto, $L := K[x]/I$ é uma extensão de K.

[Cometemos aqui um abuso de linguagem;
em rigor, L é uma extensão de uma cópia isomorfa de K:
$\psi(K) = \{a + I : a \in K\}$ é um subcorpo de L isomorfo a K]

Pelo isomorfismo $K \cong \psi(K)$, podemos identificar dentro do novo corpo L os elementos do corpo inicial K, como os elementos $a+I$ ($a \in K$). Por essa identificação, o polinómio $p(x) \in K[x]$ pode ser visto como um polinómio em $L[x]$:

$$p(x) = x^n + (a_{n-1} + I)x^{n-1} + \cdots + (a_1 + I)x + (a_0 + I).$$

Seja $\theta := x + I \in K[x]/I$. Trata-se de uma raiz de $p(x)$ em L:

$$\begin{aligned} p(\theta) &= \theta^n + (a_{n-1} + I)\theta^{n-1} + \cdots + (a_1 + I)\theta + (a_0 + I) \\ &= (x + I)^n + (a_{n-1} + I)(x + I)^{n-1} + \cdots + (a_1 + I)(x + I) + (a_0 + I) \\ &= (x^n + I) + (a_{n-1} + I)(x^{n-1} + I) + \cdots + (a_1 + I)(x + I) + (a_0 + I) \\ &= x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 + I \\ &= p(x) + I = 0. \end{aligned}$$

Portanto, em L já $p(x)$ se factoriza na forma $(x - \theta)p_1(x)$. Além disso, $p(x)$ é o polinómio mínimo de θ sobre K. Consequentemente, pelo que vimos na página 81,

$$L = \frac{K[x]}{\langle p(x) \rangle} \cong K(\theta).$$

Repetindo o raciocínio para $p_1(x)$, que podemos, sem perda de generalidade (como no início da demonstração), supôr que é irredutível sobre $L \cong K(\theta)$, chegaremos por indução (sobre o grau do polinómio) à solução que procuramos. ∎

Tal extensão chama-se *extensão* (ou *corpo*) *de decomposição* de $p(x)$.

Exemplo: Apliquemos a construção geral dada pelo Teorema ao polinómio $p(x) = x^2 + x + 1$ de $\mathbb{Z}_2[x]$, que é irredutível sobre $\mathbb{Z}_2$, como observámos no início.

Seja L a extensão

$$\begin{aligned} \frac{\mathbb{Z}_2[x]}{\langle p(x) \rangle} &= \{a_0 + a_1x + \langle p(x) \rangle \mid a_0, a_1 \in \mathbb{Z}_2\} \\ &= \{0 + \langle p(x) \rangle, 1 + \langle p(x) \rangle, x + \langle p(x) \rangle, 1 + x + \langle p(x) \rangle\} \end{aligned}$$

constituída pelas classes definidas pelos restos da divisão dos polinómios de coeficientes em $\mathbb{Z}_2[x]$ por $p(x)$. Denotando $0 + \langle p(x) \rangle$ por 0, $1 + \langle p(x) \rangle$ por 1, $x + \langle p(x) \rangle$ por α e $1 + x + \langle p(x) \rangle$ por β, as tabelas das operações de L são as seguintes:[18]

[18] Note que α é um *elemento primitivo* de L, isto é, é um gerador do grupo multiplicativo $(L - \{0\}, \cdot)$.

$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

$\cdot$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

[Por exemplo, $\alpha+\beta=(x+\langle p(x)\rangle)+(1+x+\langle p(x)\rangle)=1+\langle p(x)\rangle=1$ e $\alpha\beta=x(1+x)+\langle p(x)\rangle=x+x^2+\langle p(x)\rangle=1+\langle p(x)\rangle=1$. Observe que $L=\mathbb{Z}_2(\alpha)=\mathbb{Z}_2(\beta)$.]

O Teorema garante-nos que α é uma raiz de $p(x)$. Portanto, em L já o polinómio $p(x)$ é redutível. De facto,

$$x^2+x+1=(x-\alpha)(x-\beta).$$

Consideremos agora o polinómio $q(x)=x^2+\beta x+\beta\in L[x]$. Como $q(0)=\beta$, $q(1)=1$, $q(\alpha)=1$ e $q(\beta)=\beta$, $q(x)$ é irredutível sobre L. O Teorema diz-nos agora que a extensão de decomposição de $q(x)$ é dada pelo corpo

$$M:=\frac{L[x]}{\langle q(x)\rangle}=\{a_0+a_1x+\langle q(x)\rangle \mid a_0,a_1\in L\},$$

que tem 16 elementos:

$$[0],\ [1],\ [\alpha],\ [\beta],\ [x],\ [1+x],\ [\alpha+x],\ [\beta+x],\ [\alpha x],\ [1+\alpha x],$$

$$[\alpha+\alpha x],\ [\beta+\alpha x],\ [\beta x],\ [1+\beta x],\ [\alpha+\beta x],\ [\beta+\beta x]$$

(denotando cada elemento $a_0+a_1x+\langle q(x)\rangle$ por $[a_0+a_1x]$). Simplifiquemos a escrita um pouco mais, denotando os 16 elementos de M por, respectivamente,

$$0,1,\alpha,\beta,c,d,e,f,g,h,i,j,k,l,m,n.$$

As tabelas das operações de M são:

$+$	0	1	α	β	c	d	e	f	g	h	i	j	k	l	m	n
0	0	1	α	β	c	d	e	f	g	h	i	j	k	l	m	n
1	1	0	β	α	d	c	f	e	h	g	j	i	l	k	n	m
α	α	β	0	1	e	f	c	d	i	j	g	h	m	n	k	l
β	β	α	1	0	f	e	d	c	j	i	h	g	n	m	l	k
c	c	d	e	f	0	1	α	β	k	l	m	n	g	h	i	j
d	d	c	f	e	1	0	β	α	l	k	n	m	h	g	j	i
e	e	f	c	d	α	β	0	1	m	n	k	l	i	j	g	h
f	f	e	d	c	β	α	1	0	n	m	l	k	j	i	h	g
g	g	h	i	j	k	l	m	n	0	1	α	β	c	d	e	f
h	h	g	j	i	l	k	n	m	1	0	β	α	d	c	f	e
i	i	j	g	h	m	n	k	l	α	β	0	1	e	f	c	d
j	j	i	h	c	n	m	l	k	β	α	1	0	f	e	d	c
k	k	l	m	n	g	h	i	j	c	d	e	f	0	1	α	β
l	l	k	n	m	h	g	j	i	d	c	f	e	1	0	β	α
m	m	n	k	l	i	j	g	h	e	f	c	d	α	β	0	1
n	n	m	l	k	j	i	h	g	f	e	d	c	β	α	1	0

$\cdot$	0	1	α	β	c	d	e	f	g	h	i	j	k	l	m	n
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	α	β	c	d	e	f	g	h	i	j	k	l	m	n
α	0	α	β	1	g	i	j	h	k	m	n	l	c	e	f	d
β	0	β	1	α	k	n	l	m	c	f	d	e	g	j	h	i
c	0	c	g	k	n	j	f	β	d	1	l	h	i	m	α	e
d	0	d	i	n	j	m	1	c	l	g	f	α	e	β	k	h
e	0	e	j	l	f	1	k	i	h	n	α	c	m	g	1	β
f	0	f	h	m	β	c	i	l	1	e	g	n	α	d	j	k
g	0	g	k	c	d	l	h	1	i	α	e	m	n	f	β	j
h	0	h	m	f	1	g	n	e	α	j	k	d	β	i	l	c
i	0	i	n	d	l	f	α	g	e	k	h	β	j	1	c	m
j	0	j	l	e	h	α	c	n	m	d	β	g	f	k	i	1
k	0	k	c	g	i	e	m	α	n	β	j	f	d	h	1	l
l	0	l	e	j	m	β	g	d	f	i	1	k	h	c	n	α
m	0	m	f	h	α	k	d	j	β	l	c	i	1	n	e	g
n	0	n	d	i	e	h	β	k	j	c	m	1	l	α	g	f

`[Verifique]`

Note que c é um elemento primitivo de M: $c^0 = 1$, $c^1 = c$, $c^2 = n$, $c^3 = e$,

$c^4 = f$, $c^5 = \beta$, $c^6 = k$, $c^7 = i$, $c^8 = l$, $c^9 = m$, $c^{10} = \alpha$, $c^{11} = g$, $c^{12} = d$, $c^{13} = j$, $c^{14} = h$. [19] Podemos então escrever as tabelas de M na forma:

$+$	0	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}
0	0	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}
1	1	0	c^{12}	c^9	c^4	c^3	c^{10}	c^8	c^{13}	c^6	c^2	c^5	c^{14}	c	c^7	c^{11}
c	c	c^{12}	0	c^{13}	c^{10}	c^5	c^4	c^{11}	c^9	c^{14}	c^7	c^3	c^6	1	c^2	c^8
c^2	c^2	c^9	c^{13}	0	c^{14}	c^{11}	c^6	c^5	c^{12}	c^{10}	1	c^8	c^4	c^7	c	c^3
c^3	c^3	c^4	c^{10}	c^{14}	0	1	c^{12}	c^7	c^6	c^{13}	c^{11}	c	c^9	c^5	c^8	c^2
c^4	c^4	c^3	c^5	c^{11}	1	0	c	c^{13}	c^8	c^7	c^{14}	c^{12}	c^2	c^{10}	c^6	c^9
c^5	c^5	c^{10}	c^4	c^6	c^{12}	c	0	c^2	c^{14}	c^9	c^8	1	c^{13}	c^3	c^{11}	c^7
c^6	c^6	c^8	c^4	c^5	c^7	c^{13}	c^{12}	0	c^3	1	c^{10}	c^9	c	c^{14}	c^4	c^{12}
c^7	c^7	c^{13}	c^9	c^{12}	c^6	c^8	c^{14}	c^3	0	c^4	c	c^{11}	c^{10}	c^2	1	c^5
c^8	c^8	c^6	c^{14}	c^{10}	c^{13}	c^7	c^9	1	c^4	0	c^5	c^2	c^{12}	c^{11}	c^3	c
c^9	c^9	c^2	c^7	1	c^{11}	c^{14}	c^8	c^{10}	c	c^5	0	c^6	c^3	c^{13}	c^{12}	c^4
c^{10}	c^{10}	c^5	c^3	c^8	c	c^{12}	1	c^9	c^{11}	c^2	c^6	0	c^7	c^4	c^{14}	c^{13}
c^{11}	c^{11}	c^{14}	c^6	c^4	c^9	c^2	c^{13}	c	c^{10}	c^{12}	c^3	c^7	0	c^8	c^5	1
c^{12}	c^{12}	c	1	c^7	c^5	c^{10}	c^3	c^{14}	c^2	c^{11}	c^{13}	c^4	c^8	0	c^9	c^6
c^{13}	c^{13}	c^7	c^2	c	c^8	c^6	c^{11}	c^4	1	c^3	c^{12}	c^{14}	c^5	c^9	0	c^{10}
c^{14}	c^{14}	c^{11}	c^8	c^3	c^2	c^9	c^7	c^{12}	c^5	c	c^4	c^{13}	1	c^6	c^{10}	0

[19] Há outros elementos primitivos de M, nomeadamente f, g, h, i, j, l, n.

$\cdot$	0	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}
c	0	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1
c^2	0	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c
c^3	0	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2
c^4	0	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3
c^5	0	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4
c^6	0	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5
c^7	0	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6
c^8	0	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6	c^7
c^9	0	c^9	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8
c^{10}	0	c^{10}	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9
c^{11}	0	c^{11}	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}
c^{12}	0	c^{12}	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}
c^{13}	0	c^{13}	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}
c^{14}	0	c^{14}	1	c	c^2	c^3	c^4	c^5	c^6	c^7	c^8	c^9	c^{10}	c^{11}	c^{12}	c^{13}

O Teorema garante-nos que c é uma raiz de $q(x)$ em M. Assim, o corpo M (que coincide com a extensão simples $L(c)$ de L) é, de facto, a extensão de decomposição de $q(x)$:

$$q(x) = x^2 + \beta x + \beta = (x - c)(x - f).$$

`[Verifique]`

O Teorema motiva ainda a seguinte definição:

EXTENSÃO DE DECOMPOSIÇÃO

Seja $p(x)$ um polinómio com coeficientes num corpo K. Uma *extensão de decomposição* de $p(x)$ é uma extensão L de K em que:

(1) $p(x)$ decompõe-se em L num produto de termos de grau 1.

(2) $L = K(\theta_1, \ldots, \theta_n)$, onde $\theta_1, \ldots, \theta_n$ são as raízes de $p(x)$ em L.

Analogamente, dizemos que uma extensão L de K é uma *extensão de decomposição de uma família de polinómios* $\{p_i(x)\}_{i \in I} \subseteq K[x]$ se

(1) cada $p_i(x)$ decompõe-se em L num produto de termos de grau 1.

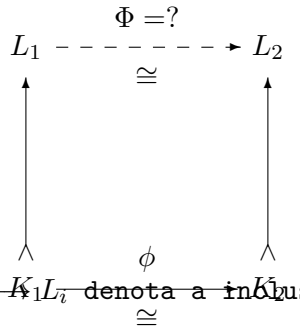
(2) L é gerada pelas raízes destes polinómios.

HOMOMORFISMO DE EXTENSÕES

Seja L_1 uma extensão de K_1 e L_2 uma extensão de K_2. Um homomorfismo de corpos $\Phi : L_1 \to L_2$ diz-se um *homomorfismo de extensões* se $\Phi(K_1) \subseteq K_2$.

Vamos analisar a seguinte questão:

> *Dado um isomorfismo de corpos $\phi : K_1 \to K_2$, é possível prolongar ϕ a um isomorfismo de extensões $\Phi : L_1 \to L_2$?*



[$K_i \to K_1$ L_i denota a inclusão de K_i em L_i $(i = 1, 2)$]

O estudo desta questão servirá, em particular, para provarmos a unicidade (a menos de um isomorfismo) das extensões de decomposição.

Dados um homomorfismo de corpos $\phi : K_1 \to K_2$ e um polinómio

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 \in K_1[x],$$

designaremos por $p^\phi(x)$ o polinómio

$$\phi(a_n)x^n + \cdots + \phi(a_1)x + \phi(a_0)$$

de $K_2[x]$.

Proposição 3.15 *Sejam $\phi : K_1 \to K_2$ um isomorfismo de corpos, L_1 e L_2 extensões de K_1 e K_2, e $\theta \in L_1$ um elemento algébrico sobre K_1 com polinómio mínimo $m(x)$. O isomorfismo ϕ pode ser prolongado a um homomorfismo injectivo de extensões $\Phi : K_1(\theta) \to L_2$ se e só se o polinómio $m^\phi(x)$ tem uma raiz em L_2. O número de prolongamentos é igual ao número de raízes distintas de $m^\phi(x)$ em L_2.*

Demonstração. Suponhamos que $m(x) = a_n x^n + \cdots + a_1 x + a_0$. Se $\Phi : K_1(\theta) \to L_2$ é um prolongamento de ϕ, então $\Phi(\theta) \in L_2$ é uma raiz de $m^\phi(x)$:

$$\begin{aligned} m^\phi(\Phi(\theta)) &= \phi(a_n)\Phi(\theta)^n + \cdots + \phi(a_1)\Phi(\theta) + \phi(a_0) \\ &= \Phi(a_n)\Phi(\theta)^n + \cdots + \Phi(a_1)\Phi(\theta) + \Phi(a_0) \\ &= \Phi(a_n\theta^n + \cdots + a_1\theta + a_0) \\ &= \Phi(m(\theta)) = \Phi(0) = 0. \end{aligned}$$

Reciprocamente, seja λ uma raiz de $m^\phi(x)$ em L_2. É fácil verificar que

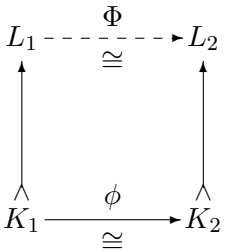
$$\begin{array}{rccc} \Phi_\lambda : & K_1(\theta) & \to & L_2 \\ & a \in K_1 & \mapsto & \phi(a) \\ & \theta & \mapsto & \lambda \end{array}$$

define um homomorfismo injectivo de corpos que prolonga ϕ. Trata-se do único homomorfismo de corpos tal que $\Phi|_{K_1} = \phi$ e $\Phi(\theta) = \lambda$.

É evidente que o número destes prolongamentos é assim igual ao número de raízes distintas de $m^\phi(x)$ em L_2. ∎

A partir da Proposição 3.15 é possível provar, por indução sobre o grau $[L_1 : K_1]$, o seguinte resultado (não o faremos na aula):

Teorema 3.16 *Seja $\phi : K_1 \to K_2$ um isomorfismo de corpos, $p(x) \in K_1[x]$ e $p^\phi(x) \in K_2[x]$. Se L_1 é uma extensão de decomposição de $p(x)$ e L_2 é uma extensão de decomposição de $p^\phi(x)$, existe um isomorfismo $\Phi : L_1 \to L_2$ tal que $\Phi|_{K_1} = \phi$.*



O número de tais prolongamentos é $\leq [L_1 : K_1]$, e é precisamente $[L_1 : K_1]$ quando $p^\phi(x)$ tem raízes distintas em L_2.

[A demonstração é por indução sobre $[L_1 : K_1]$. Se $[L_1 : K_1] = 1$, então $p(x) = a_n \prod_{i=1}^{n}(x - \theta_i)$, onde $\theta_i \in L_1 = K_1$. Como as raízes de um polinómio geram o seu corpo de decomposição, concluímos que $L_2 = K_2$, logo existe apenas 1 $(= [L_1 : K_1])$ prolongamento. Suponhamos que $[L_1 : K_1] > 1$. Então

$p(x)$ possui um factor irredutível $q(x)$ de grau ≥ 1. Seja θ uma raiz de $q(x)$ em L_1. Pela Proposição, o isomorfismo $\phi : K_1 \to K_2$ pode ser prolongado num homomorfismo injectivo $\overline{\phi} : K_1(\theta) \to L_2$ e existem tantos prolongamentos quantas as raízes distintas de $q^{\phi}(x)$ em L_2. Podemos considerar L_1 e L_2 como corpos de decomposição de $p(x)$ e $p^{\phi}(x)$ sobre $K_1(\theta)$ e $\overline{\phi}(K_1(\theta))$, respectivamente. Como $[L_1 : K_1(\theta)] = [L_1 : K_1]/[K_1(\theta) : K_1] = [L_1 : K_1]/gr(q(x)) < [L_1 : K_1]$, podemos utilizar a hipótese de indução para prolongar $\overline{\phi}$ num isomorfismo $\Phi : L_1 \to L_2$, e o número de prolongamentos é $\leq [L_1 : K_1(\theta)]$, sendo precisamente igual a $[L_1 : K_1(\theta)]$ se $p^{\phi}(x)$ tem raízes distintas em L_2. Combinando estes resultados, é fácil de ver que Φ é um prolongamento de ϕ, e o número de prolongamentos de ϕ deste tipo é precisamente $[L_1 : K_1(\theta)] \cdot gr(q(x)) = [L_1 : K_1(\theta)] \cdot [K_1(\theta) : K_1] = [L_1 : K_1]$ se $p^{\phi}(x)$ tem raízes distintas em L_2. Finalmente, observe-se que obtemos todos os prolongamentos de ϕ se prolongarmos primeiro a $K_1(\theta)$ e depois a L_1. Com efeito, se Φ é um prolongamento de ϕ a L_1, então a sua restrição a $K_1(\theta)$ fornece um homomorfismo injectivo $K_1(\theta) \to L_2$, que é necessariamente um dos prolongamentos de ϕ fornecidos pela Proposição ■]

Se neste teorema tomarmos $K_1 = K_2 = K$ e $\phi = id$, obtemos imediatamente:

Corolário 3.17 *Dois quaisquer corpos de decomposição de $p(x) \in K[x]$ são isomorfos (por um isomorfismo que deixa fixos os elementos de K).* ■

Exemplo: O polinómio $x^3 - 2$ é irredutível sobre $\mathbb{Q}$. Formemos a extensão $L = \mathbb{Q}[x]/\langle x^3 - 2\rangle$, e seja $\theta_1 = x + \langle x^3 - 2\rangle$. Já sabemos que L é uma extensão de $\mathbb{Q}$ da forma $\mathbb{Q}(r_1)$, e em L o polinómio $x^3 - 2$ admite uma factorização através do monómio $(x - \theta_1)$, nomeadamente $(x - \theta_1)(x^2 + \theta_1 x + \theta_1^2)$. O polinómio $x^2 + \theta_1 x + \theta_1^2$ é irredutível sobre $\mathbb{Q}(\theta_1)$.

[Verifique]

Podemos então formar uma nova extensão $M = \mathbb{Q}(\theta_1)[x]/\langle x^2 + \theta_1 x + \theta_1^2\rangle$. Designando por θ_2 o elemento $x + \langle x^2 + \theta_1 x + \theta_1^2\rangle$ desta extensão, vemos que $M = \mathbb{Q}(\theta_1, \theta_2)$. Em $\mathbb{Q}(\theta_1, \theta_2)[x]$ temos finalmente a factorização $x^3 - 2 = (x - \theta_1)(x -$

$\theta_2)(x-\theta_3)$ de x^3-2 em factores lineares. Portanto, $M=\mathbb{Q}(\theta_1,\theta_2)=\mathbb{Q}(\theta_1,\theta_2,\theta_3)$ é uma extensão de decomposição (abstracta) de x^3-2, que tem grau $[\mathbb{Q}(\theta_1,\theta_2,\theta_3):\mathbb{Q}]=3\cdot 2=6$.

Podemos construir uma outra extensão de decomposição M_2 considerando o subcorpo de $\mathbb{C}$ gerado por $\mathbb{Q}$ e as três raízes complexas de x^3-2 (que são $\sqrt[3]{2}$, $\sqrt[3]{2}(-1+i\sqrt{3})/2$ e $\sqrt[3]{2}(-1-i\sqrt{3})/2$). Pelos resultados que acabámos de ver, existem isomorfismos $M\to M_2$ que deixam fixos os números racionais e transformam $\theta_1,\theta_2,\theta_3$ em qualquer uma das raízes $\sqrt[3]{2}$, $\sqrt[3]{2}(-1+i\sqrt{3})/2$, $\sqrt[3]{2}(-1-i\sqrt{3})/2$.

A ideia fulcral da Teoria de Galois consiste em substituir um problema de extensões de corpos por um problema de teoria dos grupos. Os grupos em questão são os que agora introduzimos.

AUTOMORFISMOS DE GALOIS

Seja L uma extensão de K. Um automorfismo Φ de L diz-se um *K-automorfismo* (ou *automorfismo de Galois*) se deixa fixos os elementos de K, isto é, $\Phi|_K=id$.

Se Φ_1 e Φ_2 são K-automorfismos de L, então $\Phi_1\circ\Phi_2$ ainda é um K-automorfismo. É evidente então que o conjunto dos K-automorfismos de L, munido da operação usual de composição de funções, forma um grupo.

GRUPO DE GALOIS de uma extensão

Chama-se *grupo de Galois* de uma extensão L de K, que se denota por $Gal(L,K)$, ao grupo dos K-automorfismos de L.

Como observámos anteriormente, os automorfismos de Galois $\Phi:L\to L$ de uma extensão L de K permutam as raízes em L dos polinómios com coeficientes no corpo de base K. De facto, sendo $p(x)\in K[x]$ e θ uma raiz de $p(x)$ em L, então $\Phi(\theta)$ é também uma raiz de $p(x)$:

$$p(\Phi(\theta))=\Phi(p(\theta))=\Phi(0)=0.$$

Exemplos 3.18 (1) Seja $L=\mathbb{Q}(\sqrt{2})$. O elemento $\sqrt{2}$ tem polinómio mínimo x^2-2. Como vimos anteriormente, qualquer $\mathbb{Q}$-automorfismo $\Phi:L\to L$ trans-

forma raízes deste polinómio em raízes. Existem, pois, precisamente dois $\mathbb{Q}$-automorfismos:

$$\begin{array}{rccc} \Phi_{\sqrt{2}}: & \mathbb{Q}(\sqrt{2}) & \to & \mathbb{Q}(\sqrt{2}) \\ & a\in\mathbb{Q} & \mapsto & a \\ & \sqrt{2} & \mapsto & \sqrt{2} \end{array} \quad \text{e} \quad \begin{array}{rccc} \Phi_{-\sqrt{2}}: & \mathbb{Q}(\sqrt{2}) & \to & \mathbb{Q}(\sqrt{2}) \\ & a\in\mathbb{Q} & \mapsto & a \\ & \sqrt{2} & \mapsto & -\sqrt{2}. \end{array}$$

O primeiro é a identidade e o segundo aplica cada elemento $a+b\sqrt{2}$ de $\mathbb{Q}(\sqrt{2})$ em $a-b\sqrt{2}$. Portanto, $Gal(L,\mathbb{Q})=\{id,\Phi_{-\sqrt{2}}\}$, que é um grupo isomorfo a $\mathbb{Z}_2$.

(2) Quanto ao grupo de Galois da extensão $\mathbb{C}$ sobre $\mathbb{R}$, como $\mathbb{C}=\mathbb{R}(i)$, cada $\Phi\in Gal(\mathbb{C},\mathbb{R})$ é completamente determinado por $\Phi(i)$. Mas, como x^2+1 é o polinómio mínimo de i sobre $\mathbb{R}$, tem-se, pela Proposição 3.15, que $\Phi(i)=\pm i$. Assim, $Gal(\mathbb{C},\mathbb{R})=\{id, z\mapsto\overline{z}\}$ é também isomorfo a $\mathbb{Z}_2$.

(3) Seja $L=\mathbb{Q}(\sqrt{2},\sqrt{3})$. Cada $\Phi\in Gal(L,\mathbb{Q})$ é completamente determinado pela sua acção no conjunto $\{\sqrt{2},\sqrt{3}\}$. A restrição $\Phi|_{\mathbb{Q}(\sqrt{2})}:\mathbb{Q}(\sqrt{2})\to L$ é um homomorfismo injectivo que mantém fixos os elementos de $\mathbb{Q}$. Então, pela Proposição 3.15, só há duas possibilidades para esta restrição, como vimos no Exemplo (1): é a identidade ou aplica cada elemento $a+b\sqrt{2}$ de $\mathbb{Q}(\sqrt{2})$ em $a-b\sqrt{2}$. Portanto, Φ prolonga o isomorfismo identidade de $\mathbb{Q}(\sqrt{2})$ ou prolonga o isomorfismo $\Phi_{-\sqrt{2}}$ de $\mathbb{Q}(\sqrt{2})$. Usando novamente a Proposição 3.15, como x^2-3 é o polinómio mínimo de $\sqrt{3}$ sobre $\mathbb{Q}(\sqrt{2})$, estes dois isomorfismos de $\mathbb{Q}(\sqrt{2})$ só podem ser prolongados a $\mathbb{Q}(\sqrt{2},\sqrt{3})$ aplicando $\sqrt{3}$ em $\sqrt{3}$ ou $-\sqrt{3}$. Portanto, só existem 4 possibilidades para Φ: a identidade e

$$\Phi(\sqrt{2})=-\sqrt{2},\ \Phi(\sqrt{3})=\sqrt{3};$$

$$\Phi(\sqrt{2})=\sqrt{2},\ \Phi(\sqrt{3})=-\sqrt{3};$$

$$\Phi(\sqrt{2})=-\sqrt{2},\ \Phi(\sqrt{3})=-\sqrt{3}.$$

O grupo de Galois tem, pois, neste caso, 4 elementos, que designamos respectivamente por $\Phi_0,\Phi_1,\Phi_2,\Phi_3$:

$$\Phi_0(a+b\sqrt{2}+c\sqrt{3})=a+b\sqrt{2}+c\sqrt{3},$$

$$\Phi_1(a+b\sqrt{2}+c\sqrt{3})=a-b\sqrt{2}+c\sqrt{3},$$

$$\Phi_2(a+b\sqrt{2}+c\sqrt{3})=a+b\sqrt{2}-c\sqrt{3},$$

$$\Phi_3(a+b\sqrt{2}+c\sqrt{3})=a-b\sqrt{2}-c\sqrt{3}.$$

A tabela deste grupo é a seguinte:

$\circ$	Φ_0	Φ_1	Φ_2	Φ_3
Φ_0	Φ_0	Φ_1	Φ_2	Φ_3
Φ_1	Φ_1	Φ_0	Φ_3	Φ_2
Φ_2	Φ_2	Φ_3	Φ_0	Φ_1
Φ_3	Φ_3	Φ_2	Φ_1	Φ_0

Em conclusão, $Gal(L, \mathbb{Q})$ é isomorfo a $\mathbb{Z}_2 \oplus \mathbb{Z}_2$.

(4) Seja $L = \mathbb{Q}(\sqrt{3}, \sqrt[3]{2})$. Cada $\Phi \in Gal(L, \mathbb{Q})$ é completamente determinado pela sua acção no conjunto $\{\sqrt{3}, \sqrt[3]{2}\}$. A restrição $\Phi|_{\mathbb{Q}(\sqrt{3})} : \mathbb{Q}(\sqrt{3}) \to L$ é um homomorfismo injectivo que mantém fixos os elementos de $\mathbb{Q}$. Então, pela Proposição 3.15, só há duas possibilidades para esta restrição: é a identidade ou aplica cada elemento $a + b\sqrt{3}$ de $\mathbb{Q}(\sqrt{3})$ em $a - b\sqrt{3}$. Portanto, Φ prolonga o isomorfismo identidade de $\mathbb{Q}(\sqrt{3})$ ou prolonga o isomorfismo $\Phi_{-\sqrt{3}}$ de $\mathbb{Q}(\sqrt{3})$. Pela Proposição 3.15, como $x^3 - 2$ é o polinómio mínimo de $\sqrt[3]{2}$ sobre $\mathbb{Q}(\sqrt{3})$, o número de prolongamentos de Φ a L é igual ao número de raízes distintas de $x^3 - 2$ em L, ou seja, um (que corresponde à unica raiz $\sqrt[3]{2}$). Assim, os dois isomorfismos de $\mathbb{Q}(\sqrt{3})$ só podem ser prolongados a $\mathbb{Q}(\sqrt{3}, \sqrt[3]{2})$ aplicando $\sqrt[3]{2}$ em $\sqrt[3]{2}$, pelo que existem exactamente duas possibilidades para Φ: a identidade ou

$$\Phi(\sqrt{3}) = -\sqrt{3},\ \Phi(\sqrt[3]{2}) = \sqrt[3]{2}.$$

O grupo de Galois tem pois dois elementos:

$$\Phi_0(a + b\sqrt{3} + c\sqrt[3]{2}) = a + b\sqrt{3} + c\sqrt[3]{2},$$

$$\Phi_1(a + b\sqrt{3} + c\sqrt[3]{2}) = a - b\sqrt{3} + c\sqrt[3]{2}.$$

Neste caso, $Gal(L, \mathbb{Q})$ é isomorfo a $\mathbb{Z}_2$.

(5) Seja K um corpo de característica p tal que $K \neq K^p$. Se $a \notin K^p$, o polinómio $q(x) = x^p - a$ é irredutível sobre K. Seja L uma extensão de decomposição de $q(x)$. Em L temos $q(x) = (x - \theta)^p$, logo $L = K(\theta)$. Se $\Phi : L \to L$ é um K-automorfismo, então $\Phi(\theta) = \theta$ e concluímos que $\Phi = id$. Isto mostra que, neste exemplo, o grupo de Galois, $Gal(L, K)$, é trivial.

Do trabalho de Vandermonde (1735-96), Lagrange (1736-1813), Gauss (1777-1855), Ruffini (1765-1822), Abel (1802-29) e, principalmente, de Galois (1811-32), sobre a existência de "fórmulas resolventes" de grau ≤ 5, resultaram muitas das noções que temos vindo a estudar. Vamos agora fazer uma descrição muito concisa (por manifesta falta de tempo) do principal resultado de Galois, numa

reformulação feita por Artin nos anos 30 do século passado, que resolve completamente o problema de saber quando um determinado polinómio é *resolúvel por radicais*, ou seja, quando as suas raízes são números que são combinações finitas de elementos do corpo dos seus coeficientes, usando as operações do corpo e raízes de índice arbitrário.

Como os corpos de decomposição de um polinómio, como vimos, são isomorfos, é natural a seguinte definição:

GRUPO DE GALOIS de um polinómio
Seja $p(x) \in K[x]$. Chama-se *grupo de Galois de* $p(x)$ *sobre* K (ou *grupo de Galois da equação* $p(x) = 0$), que denotaremos por $Gal(p(x), K)$, ao grupo $Gal(L, K)$, onde L é uma qualquer extensão de decomposição de $p(x)$ sobre K.

Os automorfismos de Galois de uma extensão L de K permutam as raízes, nessa extensão, dos polinómios com coeficientes no corpo de base K. De facto, se

$$p(x) = \sum_{i=0}^{n} a_i x^i \in K[x],$$

$\theta \in L$ é uma raiz de $p(x)$ e $\Phi \in Gal(L, K)$, então $\Phi(\theta)$ é também uma raiz de $p(x)$:

$$p(\Phi(\theta)) = \sum_{i=0}^{n} a_i \Phi(\theta)^i = \sum_{i=0}^{n} \Phi(a_i)\Phi(\theta^i) = \sum_{i=0}^{n} \Phi(a_i\theta^i) = \Phi\Big(\sum_{i=0}^{n} a_i\theta^i\Big) = \Phi(0) = 0.$$

Portanto, é natural identificar o grupo de Galois de um polinómio $p(x)$ com um subgrupo de permutações[20] das raízes de $p(x)$:

Se L é uma extensão de decomposição de $p(x)$, e $R = \{\theta_1, \ldots, \theta_n\}$ são as raízes distintas de $p(x)$, então $L = K(\theta_1, \ldots, \theta_n)$. Se soubermos como Φ transforma as raízes de $p(x)$, então sabemos como Φ transforma todo o elemento de $L = K(\theta_1, \ldots, \theta_n)$. Portanto, o automorfismo Φ é completamente descrito pelas imagens das raízes θ_i $(i = 1, 2, \ldots, n)$. Por outro lado, como acabámos de ver, se $\Phi \in Gal(p(x), K)$, então Φ transforma raízes de $p(x)$ em raízes de $p(x)$. Portanto

$$\Phi(\theta_i) = \theta_{\tilde{\Phi}(i)} \text{ para algum } \tilde{\Phi}(i) \in \{1, 2, \ldots, n\}.$$

É evidente que, como Φ é injectiva, $\tilde{\Phi} : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ é uma permutação.

[20] Era assim que Galois concebia o grupo que hoje tem o seu nome, ainda antes de se ter formalizado sequer o conceito de grupo!

Em conclusão, todo o $\Phi \in Gal(L,K)$ fica completamente descrito pela respectiva permutação $\tilde{\Phi} \in \mathcal{S}_n$ e a aplicação $\Phi \mapsto \tilde{\Phi}$ é claramente um homomorfismo injectivo $Gal(p(x),K) \to \mathcal{S}_n$:

$$\theta_{\widetilde{\Phi_1 \circ \Phi_2}(i)} = (\Phi_1 \circ \Phi_2)(\theta_i) = \Phi_1(\theta_{\widetilde{\Phi_2}(i)}) = \theta_{\widetilde{\Phi_1}\widetilde{\Phi_2}(i)} \Rightarrow \widetilde{\Phi_1 \circ \Phi_2} = \widetilde{\Phi_1} \circ \widetilde{\Phi_2}.$$

Podemos assim identificar $Gal(p(x),K)$ com um subgrupo do grupo das permutações de R, e concluir o seguinte:

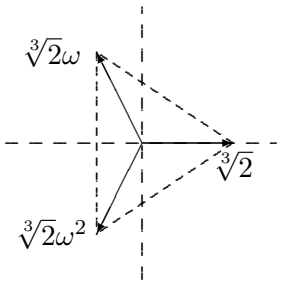
Proposição 3.19 *Se $p(x) \in K[x]$ tem n raízes distintas no seu corpo de decomposição então $Gal(p(x),K)$ é isomorfo a um subgrupo do grupo simétrico $\mathcal{S}_n$.* ∎

Note que, mesmo quando $p(x)$ é irredutível, $Gal(p(x),K)$ pode ser isomorfo a um subgrupo próprio de $\mathcal{S}_n$, como os exemplos (2) e (3) abaixo mostram.

Exemplos 3.20 (1) Vejamos que $Gal(x^3-2,\mathbb{Q}) \cong \mathcal{S}_3$. Da Proposição 3.19 sabemos que o grupo de Galois $Gal(x^3-2,\mathbb{Q})$ é isomorfo a um subgrupo de $\mathcal{S}_3$, pelo que bastará assegurar que $|Gal(x^3-2,\mathbb{Q})| = 6$. Em primeiro lugar, como em $\mathbb{C}$ temos

$$x^3 - 2 = (x - \sqrt[3]{2})(x - \sqrt[3]{2}\omega)(x - \sqrt[3]{2}\omega^2),$$

onde ω é uma raiz cúbica *primitiva* da unidade (isto é, $\omega^3 = 1$ e $\omega^t \neq 1 \; \forall \, 0 < t < 3$),



então $\mathbb{Q}(\omega, \sqrt[3]{2})$ é o corpo de decomposição de x^3-2 em $\mathbb{C}$. Como x^3-2 é o polinómio mínimo de $\sqrt[3]{2}\omega^t$ $(t = 0,1,2)$ sobre $\mathbb{Q}$ e x^2+x+1 é o polinómio mínimo de ω sobre $\mathbb{Q}$, então

$$[\mathbb{Q}(\omega, \sqrt[3]{2}) : \mathbb{Q}] = \underbrace{[\mathbb{Q}(\omega, \sqrt[3]{2}) : \mathbb{Q}(\omega)]}_{\leq 3} \underbrace{[\mathbb{Q}(\omega) : \mathbb{Q}]}_{=2} \leq 6.$$

Por outro lado, $\mathbb{Q}(\omega, \sqrt[3]{2}) = \mathbb{Q}(\sqrt[3]{2}\omega, \omega)$ e

$$[\mathbb{Q}(\sqrt[3]{2}\omega, \omega) : \mathbb{Q}] = \underbrace{[\mathbb{Q}(\sqrt[3]{2}\omega, \omega) : \mathbb{Q}(\sqrt[3]{2}\omega)]}_{\leq 2} \underbrace{[\mathbb{Q}(\sqrt[3]{2}\omega) : \mathbb{Q}]}_{=3} \leq 6.$$

Portanto, $[\mathbb{Q}(\omega, \sqrt[3]{2}) : \mathbb{Q}] \leq 6$ e é divisível por 2 e 3, logo $[\mathbb{Q}(\omega, \sqrt[3]{2}) : \mathbb{Q}] = 6$. Isto significa que $\{1, \sqrt[3]{2}, \sqrt[3]{4}, \omega, \sqrt[3]{2}\omega, \sqrt[3]{4}\omega\}$ constitui uma base da extensão $\mathbb{Q}(\omega, \sqrt[3]{2})$. É facil de ver (de modo análogo aos Exemplos 3.18) que existem precisamente seis $\mathbb{Q}$-automorfismos de $\mathbb{Q}(\omega, \sqrt[3]{2})$.

`[Descreva esses seis automorfismos explicitamente]`

Como $|\mathcal{S}_3| = 3! = 6$, teremos necessariamente $Gal(x^3 - 2, \mathbb{Q}) \cong \mathcal{S}_3$.

(2) Consideremos o polinómio $p(x) = x^4 - 2$, que é irredutível sobre $\mathbb{Q}$. As suas quatro raízes em $\mathbb{C}$ são

$$\theta_1 = \sqrt[4]{2},\ \theta_2 = \sqrt[4]{2}i,\ \theta_3 = -\sqrt[4]{2},\ \theta_4 = -\sqrt[4]{2}i,$$

e $\mathbb{Q}(i, \sqrt[4]{2})$ é o seu corpo de decomposição. Para definir um $\mathbb{Q}$-automorfismo de $\mathbb{Q}(i, \sqrt[4]{2})$, basta fixarmos as imagens das raízes θ_1 e θ_2 (pois as imagens de θ_3 e θ_4 ficam automaticamente definidas). Por exemplo,

$$\begin{array}{ccc} \theta_1 & \mapsto & \theta_2 \\ \theta_2 & \mapsto & \theta_3 \end{array}$$

define um $\mathbb{Q}$-automorfismo $\alpha : \mathbb{Q}(i, \sqrt[4]{2}) \to \mathbb{Q}(i, \sqrt[4]{2})$. É óbvio que $\alpha(\theta_3) = \theta_4$ e $\alpha(\theta_4) = \theta_1$ (e $\alpha(i) = i$). Pelo isomorfismo da Proposição 3.19, a este automorfismo corresponde a permutação $(1\,2\,3\,4)$ de $\mathcal{S}_4$.

Outro exemplo: a $\beta : \mathbb{Q}(i, \sqrt[4]{2}) \to \mathbb{Q}(i, \sqrt[4]{2})$, definido por $\beta(\theta_1) = \theta_1$ e $\beta(\theta_2) = \theta_4$, corresponde a permutação $(2\,4)$.

No entanto, nem todas as 24 permutações de $\mathcal{S}_4$ correspondem a elementos de $Gal(p(x), \mathbb{Q})$, uma vez que este grupo tem, no máximo, 8 elementos:

É evidente que $[\mathbb{Q}(i, \sqrt[4]{2}) : \mathbb{Q}(i)] = 4$ e $[\mathbb{Q}(i) : \mathbb{Q}] = 2$, logo $[\mathbb{Q}(i, \sqrt[4]{2}) : \mathbb{Q}] = 8$. Então, pelo Teorema 3.16, existem, no máximo, oito $\mathbb{Q}$-automorfismos de $\mathbb{Q}(i, \sqrt[4]{2})$, isto é, $|Gal(p(x), \mathbb{Q})| \leq 8$. Portanto, neste caso, $Gal(p(x), \mathbb{Q})$ é isomorfo a um subgrupo próprio de $\mathcal{S}_4$.

Por exemplo, o ciclo $(1\,2)$ não corresponde a nenhum $\mathbb{Q}$-automorfismo

$$\Phi : \mathbb{Q}(i, \sqrt[4]{2}) \to \mathbb{Q}(i, \sqrt[4]{2}),$$

uma vez que Φ, para originar tal ciclo, teria que satisfazer $\Phi(\theta_1) = \theta_2$, $\Phi(\theta_2) = \theta_1$, $\Phi(\theta_3) = \theta_3$ e $\Phi(\theta_4) = \theta_4$, mas tal Φ não é, claramente, um homomorfismo de corpos (com efeito, $\theta_1 + \theta_3 = 0$ mas $\Phi(\theta_1) + \Phi(\theta_3) = \theta_2 + \theta_3 \neq 0$).

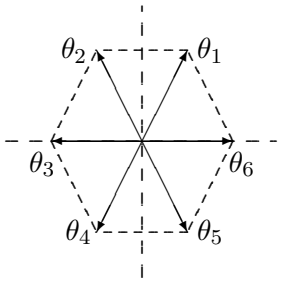
`[Conclua que` $|Gal(p(x)), \mathbb{Q})| = 8$`, observando que, respectivamente,` $\theta_1 \mapsto \theta_1$ `e` $\theta_2 \mapsto \theta_2$`,` $\theta_1 \mapsto \theta_1$ `e` $\theta_2 \mapsto \theta_4$`,` $\theta_1 \mapsto \theta_2$ `e` $\theta_2 \mapsto \theta_1$`,`

$\theta_1 \mapsto \theta_2$ e $\theta_2 \mapsto \theta_3$, $\theta_1 \mapsto \theta_3$ e $\theta_2 \mapsto \theta_2$, $\theta_1 \mapsto \theta_3$ e $\theta_2 \mapsto \theta_4$,
$\theta_1 \mapsto \theta_4$ e $\theta_2 \mapsto \theta_1$, $\theta_1 \mapsto \theta_4$ e $\theta_2 \mapsto \theta_3$,
definem oito $\mathbb{Q}$-automorfismos de $\mathbb{Q}(i, \sqrt[4]{2})$.
Portanto, $Gal(p(x)), \mathbb{Q})$é isomorfo a
$\{id, (2\,4), (1\,2)(3\,4), (1\,2\,3\,4), (1\,3), (1\,3)(2\,4), (1\,4\,3\,2), (1\,4)(2\,3)\}$.
Este grupo G é isomorfo ao grupo diedral D_4 das simetrias de um quadrado, pois é gerado pelos elementos $\sigma = (2\,4)$ e $\tau = (1\,2\,3\,4)$, de ordens 2 e 4, que satisfazem a relação $(\tau\sigma)^2 = id$:
$G = \{id, \sigma, \sigma\tau, \tau, \sigma\tau^2, \tau^2, \tau^3, \sigma\tau^3\}$]

(3) Seja $L \subseteq \mathbb{C}$ a extensão de decomposição sobre $\mathbb{Q}$ do polinómio irredutível $p(x) = x^6 - 2$. As raízes de $p(x)$ são

$$\theta_k = \sqrt[6]{2} e^{\frac{2k\pi i}{6}}, \quad k = 1, \ldots, 6.$$

Neste caso, $|\mathcal{S}_6| = 6! = 720$ mas $|Gal(p(x), \mathbb{Q})| < 720$; por exemplo, não existe um automorfismo do grupo de Galois que corresponda à transposição $(1\,6)$, pois $\theta_3 + \theta_6 = 0$ mas $\theta_3 + \theta_1 \neq 0$, como se observa imediatamente na representação, no plano complexo, das raízes de $p(x)$:



Outro exemplo: como $(\theta_1 + \theta_5)^6 = \theta_6^6 = 2$, não existem automorfismos do grupo de Galois que correspondam às permutações $(1\,3)(5\,6)$ e $(1\,6)(3\,5)$. Muitos outros elementos de $\mathcal{S}_6$ podem ser excluídos; de facto, como veremos mais adiante,

$$|Gal(x^6 - 2, \mathbb{Q})| = 12.$$

EXTENSÃO DE GALOIS

Diz-se que uma extensão finita L de K é uma *extensão de Galois* se L for um corpo de decomposição de algum polinómio de $K[x]$.

Trabalhando a demonstração (que não estudámos) do Teorema 3.16 sobre extensões de isomorfismos a corpos de decomposição, não é difícil provar o seguinte resultado:

Teorema 3.21 *Seja L uma extensão finita de K. Então:*

(1) $|Gal(L,K)| \leq [L:K]$.

(2) *Se L é uma extensão de Galois de K, então $|Gal(L,K)| = [L:K]$.*

[A demonstração pode ser consultada em *Introdução à Álgebra*, R. Loja Fernandes e M. Ricou, IST Press, 2004]

Exemplos 3.22 (1) A observação, no Exemplo 3.20(2), de que $|Gal(p(x),\mathbb{Q})| = 8$, é uma consequência imediata deste teorema, uma vez que $\mathbb{Q}(i,\sqrt[4]{2})$ é uma extensão de Galois de $\mathbb{Q}$ e $[\mathbb{Q}(i,\sqrt[4]{2}):\mathbb{Q}] = 8$.

(2) No Exemplo 3.20(3) de há pouco, $\mathbb{Q}(\sqrt[6]{2}, e^{\frac{2\pi}{3}i})$ é uma extensão de decomposição de $p(x) = x^6 - 2$. Como

$$[\mathbb{Q}(\sqrt[6]{2}, e^{\frac{2\pi}{3}i}):\mathbb{Q}] = [\mathbb{Q}(\sqrt[6]{2}, e^{\frac{2\pi}{3}i}):\mathbb{Q}(\sqrt[6]{2})] \cdot [\mathbb{Q}(\sqrt[6]{2}):\mathbb{Q}] = 2 \cdot 6 = 12,$$

então $|Gal(x^6-2,\mathbb{Q})| = 12$, como tínhamos anunciado.

(3) A extensão $L = \mathbb{Q}(\sqrt{3}, \sqrt[3]{2})$ do Exemplo 3.18(4) não é uma extensão de Galois de $\mathbb{Q}$ uma vez que, como vimos, $|Gal(L,\mathbb{Q})| = 2$ enquanto $[L:\mathbb{Q}] = 6$.

Estamos finalmente em condições de explicar como é que a Teoria de Galois permite substituir problemas sobre polinómios por um problema em princípio mais simples de teoria dos grupos. Galois descobriu que existe uma correspondência entre extensões intermédias e subgrupos do grupo de Galois, que passamos a descrever.

CORRESPONDÊNCIA DE GALOIS

Seja M uma extensão de K. Se L é uma extensão intermédia (isto é, $K \subseteq L \subseteq M$), todo o L-automorfismo de M é obviamente um K-automorfismo de M e, portanto, $Gal(M,L)$ é um subgrupo do grupo $Gal(M,K)$. Por outro lado, se H é um subgrupo de $Gal(M,K)$, o conjunto $Fix(H) := \{a \in M \mid \Phi(a) = a \ \forall \Phi \in H\}$ dos pontos fixos por H é uma extensão intermédia $K \subseteq Fix(H) \subseteq M$. A esta correspondência entre extensões intermédias de $K \subseteq M$ e subgrupos de $Gal(M,K)$ chama-se *correspondência de Galois*.

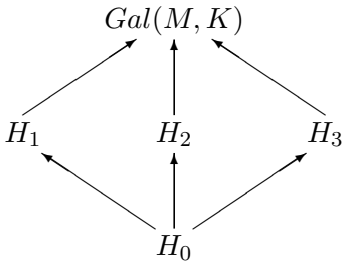
[Esta correspondência não é, em geral, uma bijecção, mas tem boas propriedades:

(1) Se $L_1 \subseteq L_2$ então $Gal(M, L_1) \supseteq Gal(M, L_2)$.
(2) Se $H_1 \subseteq H_2$ então $Fix(H_1) \supseteq Fix(H_2)$.
(3) $Fix(Gal(M, L)) \supseteq L$.
(4) $Gal(M, Fix(H)) \supseteq H$]

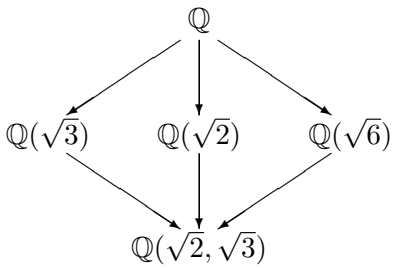
Exemplo: Consideremos a extensão $M = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ de $K = \mathbb{Q}$. Vimos anteriormente que o grupo de Galois desta extensão contém 4 elementos e é isomorfo a $\mathbb{Z}_2 \oplus \mathbb{Z}_2$:

$$Gal(M, K) := \{\Phi_0, \Phi_1, \Phi_2, \Phi_3\}.$$

Este grupo possui, para além do subgrupo trivial $H_0 = \{\Phi_0\}$, os subgrupos $H_1 = \{\Phi_0, \Phi_1\}$, $H_2 = \{\Phi_0, \Phi_2\}$ e $H_3 = \{\Phi_0, \Phi_3\}$. Assim, o conjunto parcialmente ordenado dos subgrupos de $Gal(M, K)$ pode ser representado pelo diagrama



O corpo fixo pelo grupo de Galois $Gal(M, K)$ é o corpo de base $\mathbb{Q}$, enquanto que $Fix(H_0) = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. Por outro lado, é facil de ver que $Fix(H_1) = \mathbb{Q}(\sqrt{3})$, $Fix(H_2) = \mathbb{Q}(\sqrt{2})$, $Fix(H_3) = \mathbb{Q}(\sqrt{6})$. Assim, o conjunto parcialmente ordenado das extensões intermédias de $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$ é dado pelo diagrama



Teorema 3.23 [Teorema Fundamental de Galois]
Seja $K \subseteq L \subseteq M$ uma torre de corpos, onde M é uma extensão de Galois de K. Então $Gal(M, L)$ é um subgrupo normal de $Gal(M, K)$ se e só se L é também uma extensão de Galois de K. Neste caso, $Gal(L, K) \cong Gal(M, K)/Gal(M, L)$:

$$G\left\{\begin{array}{l} H\left\{\begin{array}{l} M \\ \uparrow \\ L \end{array}\right. \\ \qquad \uparrow \qquad \left.\begin{array}{l} \\ \\ \end{array}\right\} G/H \\ \qquad K \qquad \end{array}\right.$$

Demonstração. Faremos somente a prova da implicação "⇐".

Suponhamos então que L é uma extensão de Galois de K, ou seja, $L = K(\theta_1, \ldots, \theta_n) \subseteq M$, onde $\theta_1, \ldots, \theta_n$ são as raízes de algum polinómio $p(x) \in K[x]$. Como cada $\Phi \in Gal(M, K)$ permuta as raízes de p(x) e mantém fixos os elementos de K, então $\Phi(L) \subseteq L$. Podemos assim considerar a aplicação

$$\begin{array}{rccc} h: & Gal(M,K) & \to & Gal(L,K) \\ & \Phi & \mapsto & \Phi|_L \end{array}$$

É evidente que se trata de um homomorfismo de grupos, sendo o seu núcleo precisamente o subgrupo $Gal(M, L)$. Assim, $Gal(M, L)$ é um subgrupo normal de $Gal(M, K)$. O Teorema 3.16 garante que, dado $\Psi \in Gal(L, K)$, existe $\Phi \in Gal(M, K)$ que prolonga Ψ. Portanto, h é sobrejectivo e, pelo Teorema do Homomorfismo estudado em Álgebra I, tem-se $Gal(L, K) \cong Gal(M, K)/Gal(M, L)$.

■

Vamos agora discutir o critério descoberto por Galois que permite decidir se uma equação algébrica é ou não resolúvel por radicais. Até ao final, para simplificar, assumiremos que todos os corpos têm característica 0.

É preciso algum cuidado na formalização da ideia de resolubilidade por radicais. Informalmente, uma extensão por radicais obtem-se por uma sequência de adjunções de raízes (radicais) índice n, para vários n. Por exemplo, a seguinte expressão é radical:

$$\frac{\sqrt[5]{2-\sqrt[3]{2}}+\sqrt{3}}{\sqrt[7]{1-\sqrt[4]{5}}}.$$

Para encontrar uma extensão de $\mathbb{Q}$ que contenha este elemento, podemos juntar, consecutivamente, elementos

$$a_1 = \sqrt[4]{5} \qquad a_2 = \sqrt[7]{1-a_1} \qquad a_3 = \sqrt[3]{2} \qquad a_4 = \sqrt[5]{2-a_3} \qquad a_5 = \sqrt{3}.$$

Isto sugere as seguintes definições:

EXTENSÃO PURA

Uma extensão L de K diz-se *pura* se $L = K(\theta)$, onde $\theta \in L$ é tal que $\theta^m \in K$ para algum $m \in \mathbb{N}$ (isto é, θ é um *radical* de K).

POLINÓMIO RESOLÚVEL POR RADICAIS

Uma extensão L de K diz-se uma *extensão por radicais* se existir uma torre de corpos

$$K = L_0 \subseteq L_1 \subseteq L_2 \subseteq \cdots \subseteq L_t = L$$

tal que cada L_{i+1} é uma extensão pura de L_i, para $i = 0, 1, \ldots, t-1$.

Um polinómio $p(x) \in K[x]$ diz-se *resolúvel por radicais* sobre K se existir uma extensão por radicais L de K onde $p(x)$ se decompõe em factores lineares (isto é, que contém um corpo de decomposição de $p(x)$).

Exemplos 3.24 (1) Suponhamos que uma raiz θ de um polinómio $p(x) \in \mathbb{Q}[x]$ se exprime por meio do radical de há pouco:

$$\theta = \frac{\sqrt[5]{2 - \sqrt[3]{2}} + \sqrt{3}}{\sqrt[7]{1 - \sqrt[4]{5}}}.$$

Considerando $a_1 = \sqrt[4]{5}$, $a_2 = \sqrt[7]{1 - a_1}$, $a_3 = \sqrt[3]{2}$, $a_4 = \sqrt[5]{2 - a_3}$, $a_5 = \sqrt{3}$, temos

$$\mathbb{Q} \subseteq \underbrace{\mathbb{Q}(a_1)}_{L_1} \subseteq \underbrace{\mathbb{Q}(a_1, a_2)}_{L_2 = L_1(a_2)} \subseteq \underbrace{\mathbb{Q}(a_1, a_2, a_3)}_{L_3 = L_2(a_3)} \subseteq \underbrace{\mathbb{Q}(a_1, a_2, a_3, a_4)}_{L_4 = L_3(a_4)} \subseteq \underbrace{\mathbb{Q}(a_1, a_2, a_3, a_4, a_5)}_{L_5 = L_4(a_5)}.$$

Como

$$a_1^4 \in \mathbb{Q}, a_2^7 \in L_1, a_3^3 \in L_2, a_4^5 \in L_3, a_5^2 \in L_4,$$

então L_5 é uma extensão por radicais de $\mathbb{Q}$ que contém $\frac{a_4 + a_5}{a_2} = \theta$.

Este exemplo ilustra como, a partir de um dado elemento θ, expresso por radicais em termos dos elementos de um determinado corpo de base, se pode construir uma extensão por radicais desse corpo contendo o elemento θ.

(2) Consideremos uma equação quadrática $ax^2 + bx + c = 0$ ($a \neq 0$) em $\mathbb{Q}$, arbitrária. A fórmula resolvente dá-nos as suas duas raízes expressas por radicais, em termos dos seus coeficientes a, b, c:

$$r_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \underbrace{-\frac{b}{2a}}_{\in \mathbb{Q}} \underbrace{+\sqrt{\frac{b^2 - 4ac}{4a^2}}}_{\theta},$$

$$r_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} = \underbrace{-\frac{b}{2a}}_{\in \mathbb{Q}} \underbrace{-\sqrt{\frac{b^2 - 4ac}{4a^2}}}_{-\theta}.$$

É evidente que $\mathbb{Q}(\theta)$ é o corpo de decomposição do polinómio $ax^2 + bx + c$, e é uma extensão pura de $\mathbb{Q}$ (pois $\theta^2 \in \mathbb{Q}$), pelo que se trata de uma extensão por radicais de $\mathbb{Q}$. Isto mostra que qualquer polinómio de grau 2 é resolúvel por radicais.

[Do mesmo modo, não é difícil, usando as ''fórmulas resolventes'', provar que todos os polinómios de grau 3 e 4, com coeficientes em corpos de característica 0, também são resolúveis por radicais]

Observe-se bem o significado desta definição: qualquer raiz de $p(x)$ pertence a L e pode ser expressa a partir de elementos de K por uma sequência de operações em K e de extracção de raízes. De facto: numa extensão por radicais L de K, os elementos de L são "combinações polinomiais" de radicais de radicais de ... etc. (em número finito) ... de elementos de K, com coeficientes em K. Por outras palavras, todos os elementos de L são construídos a partir de um número finito de elementos do corpo de base K, e usando as operações $+$, $\cdot$ e $\sqrt[*]{}$. A definição de polinómio resolúvel por radicais é pois equivalente a dizer que as suas raízes, num corpo de decomposição, são "combinações" de radicais de radicais de ... etc. (em número finito) ... de elementos do seu corpo dos coeficientes.

GRUPO RESOLÚVEL

Um grupo G diz-se *resolúvel* se existir uma torre de subgrupos

$$\{1\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq \cdots \subseteq G_{n-1} \subseteq G_n = G$$

tal que, para cada $i \in \{1, 2, \ldots, n\}$, G_{i-1} é um subgrupo normal de G_i e G_i/G_{i-1} é abeliano.

[Tem-se que: (1) Subgrupos de grupos resolúveis são resolúveis.
(2) Quocientes de grupos resolúveis são resolúveis.
(3) Dado um subgrupo normal de um grupo G, G é resolúvel se e só se H e G/H são resolúveis]

Exemplos 3.25 (1) Todo o grupo abeliano G é resolúvel pois $\{e\} \subseteq G$ satisfaz a definição. Em particular, $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{A}_1$, $\mathcal{A}_2$ e $\mathcal{A}_3$ são resolúveis.

(2) $\mathcal{S}_3$ é resolúvel pois $\{id\} \subseteq \{id, (1\,2\,3), (1\,3\,2)\} \subseteq \mathcal{S}_3$ satisfaz a definição.

(3) $\mathcal{S}_4$ e $\mathcal{A}_4$ são resolúveis pois

$$\{id\} \subseteq \{id, (1\,2)(3\,4)\} \subseteq \{id, (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3)\} \subseteq \mathcal{A}_4 \subseteq \mathcal{S}_4$$

satisfaz a definição.

(4) $\mathcal{S}_n$ $(n \geq 5)$ não é resolúvel.

[Demonstração na bibliografia]

(5) Seja $\mathbb{Z}_m^*$ o grupo das unidades do anel $(\mathbb{Z}_m, +, \cdot) = (\mathbb{Z}_m, \oplus_m, \otimes_m)$. O conjunto $\mathbb{Z}_m \times \mathbb{Z}_m^*$, munido da operação

$$(a, b) \cdot (c, d) = (a + bc, bd),$$

é um grupo

[Verifique]

a que se chama *produto semi-directo* de $\mathbb{Z}_m$ e $\mathbb{Z}_m^*$, e que se denota por $\mathbb{Z}_m \rtimes \mathbb{Z}_m^*$.

[Observe: (1) $\mathbb{Z}_3 \rtimes \mathbb{Z}_3^* \cong \mathcal{S}_3$.
(2) $\mathbb{Z}_m$ pode ser visto como um subgrupo normal de $\mathbb{Z}_m \rtimes \mathbb{Z}_m^*$ através da imersão natural $i : x \mapsto (x, 1)$ $(x \in \mathbb{Z}_m)$]

$\mathbb{Z}_m \rtimes \mathbb{Z}_m^*$ é resolúvel pois $\{(0, 1)\} \subseteq i(\mathbb{Z}_m) \subseteq \mathbb{Z}_m \rtimes \mathbb{Z}_m^*$ satisfaz a definição de grupo resolúvel.

O grupo $\mathbb{Z}_m \rtimes \mathbb{Z}_m^*$ é importante neste contexto por causa da proposição seguinte:

Proposição 3.26 *Seja $K \subseteq \mathbb{C}$ e $x^m - a \in K[x]$ $(m \in \mathbb{N})$. O grupo de Galois deste polinómio é isomorfo a um subgrupo de $\mathbb{Z}_m \rtimes \mathbb{Z}_m^*$.*

Demonstração. Se $\theta \in \mathbb{C}$ é uma raiz de índice m de a e ω é uma raiz *primitiva* de índice m da unidade (isto é, $\omega^m = 1$ e $\omega^t \neq 1$, $\forall\, 0 < t < m$; por exemplo, $\omega = \cos\frac{2\pi}{m} + i \sin\frac{2\pi}{m}$), então

$$x^m - a = \prod_{i=0}^{m-1} (x - \theta\omega^i).$$

Resulta daqui que o corpo de decomposição, em $\mathbb{C}$, de $x^m - a$ é $K(\theta, \omega)$. Assim, um elemento Φ de $Gal(x^m - a, K)$ é completamente determinado por $\Phi(\theta)$ e $\Phi(\omega)$. Como os K-automorfismos permutam as raízes de polinómios com coeficientes em K, tem-se $\Phi(\theta) = \theta\omega^{i_\Phi}$ e $\Phi(\omega) = \omega^{j_\Phi}$ para alguns $i_\Phi, j_\Phi \in \{0, 1, \ldots, m-1\}$.

Vejamos que $mdc(j_\Phi, m) = 1$ para qualquer $\Phi \in Gal(x^m - a, K)$. Denotando $mdc(j_\Phi, m)$ por d temos

$$\Phi(\omega^{\frac{m}{d}} = \Phi(\omega)^{\frac{m}{d}} = \omega^{j_\Phi \cdot \frac{m}{d}} = \omega^{m \cdot \frac{j_\Phi}{d}} = 1.$$

Como Φ é injectiva, resulta que $\omega^{\frac{m}{d}} = 1$ e, consequentemente, como ω é uma raiz primitiva índice m da unidade, só pode ser $d = 1$. Assim, a correspondência

$$\begin{array}{ccc} Gal(x^m - a, K) & \to & \mathbb{Z}_m \rtimes \mathbb{Z}_m^* \\ \Phi & \mapsto & (i_\Phi \, mod \, m, j_\Phi \, mod \, m) \end{array}$$

define uma aplicação, que é um homomorfismo injectivo de grupos, como se pode verificar facilmente. ∎

[Este resultado ainda é válido para qualquer subcorpo de um corpo de característica 0]

Corolário 3.27 *$Gal(x^m - a, K)$ é um grupo resolúvel para todo o subcorpo K de um corpo de característica zero, $a \in K$ e $m \in \mathbb{N}$.*

Demonstração. Resulta imediatamente da proposição anterior e do facto de subgrupos de grupos resolúveis serem ainda resolúveis. ∎

Teorema 3.28 [Critério de Galois]
Seja K um subcorpo de um corpo de característica zero e $p(x) \in K[x]$. Então $p(x)$ é resolúvel por radicais se e só se $Gal(p(x), K)$ for um grupo resolúvel.

[*Demonstração.* Esboçaremos somente a prova da implicação ''$\Rightarrow$''.

Seja então $p(x) \in K[x]$ um polinómio resolúvel por radicais, sendo

$$K = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_t = L$$

a correspondente torre de extensões puras tal que $L = L_t$ contém um corpo de decomposição de $p(x)$. Então, para cada $i \in \{1, \ldots, t\}$, $L_i = L_{i-1}(\theta_i)$, onde cada θ_i é um radical de L_{i-1}, ou seja, $\theta_i^{m_i} \in L_{i-1}$ para algum $m_i \in \mathbb{N}$ (portanto, θ_i é raiz de $x^{m_i} - \theta_i^{m_i} \in L_{i-1}[x]$). Seja ω_i uma raiz primitiva de índice m_i da unidade.

Na torre de extensões

$$\underbrace{K = L_0}_{\tilde{L}_0} \subseteq \underbrace{L_0(\theta_1, \omega_1)}_{\tilde{L}_1} \subseteq \underbrace{L_1(\theta_2, \omega_2)}_{\tilde{L}_2} \subseteq \cdots \subseteq \underbrace{L_{t-1}(\theta_t, \omega_t)}_{\tilde{L}_t}$$

cada $\tilde{L}_i$ é uma extensão de Galois de $\tilde{L}_{i-1}$ (porque é o corpo de decomposição do polinómio $x^{m_i} - \theta_i^{m_i} \in L_{i-1}[x]$) e $\tilde{L}_t$ contém um corpo de decomposição de $p(x)$. Com um pouco mais de trabalho pode construir-se uma torre de extensões

$$K = \hat{L}_0 \subseteq \hat{L}_1 \subseteq \cdots \subseteq \hat{L}_s$$

tal que cada $\hat{L}_i$ é uma extensão de Galois de K e $\hat{L}_s$ contém um corpo de decomposição de $p(x)$, que designaremos por L. Seja $G_i := Gal(\hat{L}_s, \hat{L}_{s-i})$. Pelo Teorema Fundamental de Galois podemos concluir que na torre de subgrupos

$$\{1\} = G_0 \subseteq G_1 \subseteq G_2 \subseteq \cdots \subseteq G_{s-1} \subseteq G_s = Gal(\hat{L}_s, K)$$

cada subgrupo é normal e, para cada $i \in \{1, \ldots, s\}$, G_i/G_{i-1} é isomorfo a $Gal(\hat{L}_{s-i+1}, \hat{L}_{s-i}) = Gal(x^{m_{s-i+1}} - \theta_{s-i+1}^{m_{s-i+1}}, \hat{L}_{s-i})$, que é, pelo Corolário, um grupo resolúvel. Como G_0 e G_1/G_0 são resolúveis, G_1 também é; então, como G_2/G_1 é resolúvel, G_2 também é; indutivamente, podemos concluir que G_s é resolúvel. Mas $Gal(p(x), K) = Gal(L, K)$ é isomorfo a $G_s/Gal(\hat{L}_s, L)$, pelo Teorema Fundamental. Uma vez que quocientes de grupos resolúveis são resolúveis, podemos finalmente concluir que $Gal(p(x), K)$ é resolúvel]

Exemplos: Do Corolário anterior podemos concluir imediatamente que, para qualquer $m \in \mathbb{N}$, os polinómios $x^m - a \in \mathbb{Q}[x]$ são resolúveis por radicais.

No entanto, para cada $m > 4$ existem também polinómios de grau m que não são resolúveis por radicais. Por exemplo, no caso $m = 5$:

Corolário 3.29 [Teorema de Abel-Ruffini]
Existem polinómios de grau 5 que não são resolúveis por radicais.

Demonstração. Seja $p(x) = x^5 - 4x + 2$. Não é difícil ver que $p(x)$ tem precisamente 3 raízes reais $\theta_1, \theta_2, \theta_3$ e 2 raízes complexas conjugadas θ_4, θ_5, todas distintas. Então $L = \mathbb{Q}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ é a extensão de decomposição de $p(x)$.

Seja G o grupo de Galois de $p(x)$. Pela Proposição 3.19, pode ser considerado como sendo um subgrupo de $\mathcal{S}_5$. Pelo critério de Eisenstein, $p(x)$ é irredutível sobre $\mathbb{Q}$, logo, para qualquer raiz θ de $p(x)$, $[\mathbb{Q}(\theta) : \mathbb{Q}] = 5$. Consequentemente, $[L : \mathbb{Q}]$ é um múltiplo de 5. Isto significa, pelo Teorema 3.21, que $|G|$ é um múltiplo de 5. Portanto, pelos Teoremas de Sylow estudados em Álgebra I, $G \subseteq \mathcal{S}_5$ contém um elemento de ordem 5, ou seja, um ciclo de comprimento 5.

Por outro lado, a aplicação $z \mapsto \bar{z}$ de $\mathbb{C}$ induz um $\mathbb{Q}$-automorfismo de L que mantém fixas as três raízes reais e permuta as duas raízes complexas, a que corresponde a transposição $(4\,5)$.

Em conclusão, G contém um ciclo de ordem 5 e uma transposição. Mas pode provar-se que um qualquer ciclo de ordem 5 e uma transposição geram $\mathcal{S}_5$, pelo que $G = \mathcal{S}_5$. Como $\mathcal{S}_5$ não é resolúvel, o critério de Galois assegura que $p(x)$ não é resolúvel por radicais. ∎

[Observe que a mesma argumentação vale para qualquer outro polinómio de grau 5 com coeficientes em $\mathbb{Q}$ que seja irredutível e que em $\mathbb{C}$ tenha exactamente 3 raízes reais]

[Pode ver a teoria de Galois na sua forma original em H.M. Edwards, *Galois Theory*, Springer, 1984, e no apêndice 4 de J. Rotman,*Galois Theory*, Springer, 1990. A prova de Abel da inexistência de uma ''fórmula resolvente'' do quinto grau encontra-se no seu artigo *Démonstration de l'impossibilité de la résolution algébrique des équations générales qui passent le quatrième degré*, J. reine angew. Math. 1 (1826) 65-84]

Exercícios

3.1. Sejam K um subcorpo de um corpo L e α, β elementos de L. Prove que $K(\alpha, \beta) = K(\alpha)(\beta)$. Generalize para o caso de n elementos $\alpha_1, \ldots, \alpha_n \in L$.

3.2. Sejam K um subcorpo de um corpo L e θ um elemento de L. Prove que:
(a) se θ é algébrico sobre K, o mesmo sucede a $\theta + c$ e a $c\theta$, qualquer que seja $c \in K$;
(b) se θ é algébrico sobre K, o mesmo sucede a θ^2 e reciprocamente.

3.3. Mostre que $\mathbb{C}$ é uma extensão algébrica de $\mathbb{R}$.

3.4. Averigúe quais dos seguintes elementos são algébricos ou transcendentes sobre o corpo $\mathbb{Q}$:
(a) $\sqrt{7}$ (b) $\sqrt[3]{2}$ (c) π^2 (d) $e + 3$ (e) $1 + i$.

3.5. Determine o inverso de cada um dos seguintes elementos nas extensões simples $\mathbb{Q}(\theta)$ indicadas:

(a) $2+\sqrt[3]{4}$ em $\mathbb{Q}(\sqrt[3]{2})$.

(b) $1-2\theta+3\theta^2$, onde θ é raiz do polinómio x^3-x+1.

(c) $-\theta^2+2\theta-3$, para $\theta=\sqrt[3]{2}$.

(d) $\theta+1$ e $\theta^2-6\theta+8$, onde $\theta\neq 0$ é tal que $\theta^4-6\theta^3+9\theta^2+3\theta=0$.

3.6. Sejam K um subcorpo de um corpo L e θ um elemento de L. Prove que se θ é algébrico sobre K então $K(\theta)=K[\theta]$, justificando pormenorizadamente os seguintes passos:

(a) $K[\theta]$ é um domínio de integridade.

(b) Sendo $f(\theta)$ um elemento não nulo de $K[\theta]$ e $m(x)$ o polinómio mínimo de θ sobre K, então:

- $f(x)$ não é múltiplo de $m(x)$;
- existem $t(x), s(x)\in K[x]$ tais que $t(x)f(x)+s(x)m(x)=1$, donde $t(\theta)f(\theta)=1$.

(c) $K[\theta]$ é um corpo.

3.7. Sejam K e L dois corpos tais que $K\subseteq L$. Sabendo que, se $\alpha,\ \beta\in L$ são elementos algébricos sobre K e $[K(\alpha,\beta):K]$ é finita, prove que os elementos de L que são algébricos sobre K formam um subcorpo de L.

3.8. Seja L uma extensão dum corpo K e $\theta\in L$ um elemento algébrico de grau n sobre K. Prove que todo o elemento de $K(\theta)$ se pode exprimir de modo único na forma $a_0+a_1\theta+\cdots+a_{n-1}\theta^{n-1}$ com $a_i\in K$ $(i=0,...,n-1)$.

3.9. Exprima na forma referida no exercício anterior os seguintes elementos das extensões algébricas $\mathbb{Q}(\theta)$ indicadas:

(a) θ^4, θ^2, θ^5 e $\theta^5-\theta^4+2$, onde θ é raiz do polinómio x^3-6x^2+9x+3.

(b) $(\theta^3+2)(\theta^3+3\theta)$, $\theta^4(\theta^4+3\theta^2+7\theta+5)$ e $\frac{\theta+2}{\theta^2+3}$, sendo θ uma solução da equação $x^5+2x+2=0$.

(c) $\frac{\theta^2}{\theta^2+1}$, onde θ é uma raiz não nula do polinómio $x^4-x^3+x^2-2x$.

3.10. Determine o polinómio mínimo sobre $\mathbb{Q}$ dos seguintes elementos:
(a) $2+\sqrt{3}$. (b) θ^2-1, com $\theta^3=2\theta+2$. (c) $\theta^2+\theta$, com $\theta^3=-3\theta^2+3$.

3.11. Prove que $\sqrt{7}\notin\mathbb{Q}(\sqrt{3})$, $i\notin\mathbb{Q}(\sqrt{5})$ e $\sqrt{5}\notin\mathbb{Q}(i)$.

3.12. Seja L uma extensão finita de K. Prove que:

(a) Se $[L:K]$ é um número primo, então L é uma extensão simples de K.

(b) Se $\theta \in L$, então o grau de θ é um divisor de $[L : K]$. Conclua que se tem $L = K(\theta)$ se e só se o grau de θ coincidir com $[L : K]$.

(c) Se $f(x) \in K[x]$ é irredutível sobre K e o grau de $f(x)$ é um número primo com $[L : K]$ e maior do que 1, então $f(x)$ não tem raízes em L.

3.13. Seja p um número primo e c um elemento do corpo C. Prove que $x^p - c$ é irredutível sobre C se e só se $x^p - c$ não tem raízes em C.

3.14. Sejam C,C_1 e C_2 corpos com $C \subseteq C_i$ $(i = 1, 2)$. Se C_1 e C_2 são extensões finitas de C tais que $[C_1 : C]$ e $[C_2 : C]$ são primos entre si, então $C_1 \cap C_2 = C$.

3.15. Averigúe se os seguintes polinómios são irredutíveis sobre o corpo indicado:
(a) $x^2 + 2$ sobre $\mathbb{Q}(\sqrt{5})$. (b) $x^2 - 2x + 2$ sobre $\mathbb{Q}(\sqrt{-3})$.
(c) $x^3 - 3x + 3$ sobre $\mathbb{Q}(\sqrt[4]{2})$.

3.16. Determine o grau sobre $\mathbb{Q}$ e uma base de cada uma das seguintes extensões de $\mathbb{Q}$:

(a) $\mathbb{Q}(\sqrt{3}, i)$.

(b) $\mathbb{Q}(\sqrt{18}, \sqrt[4]{2})$.

(c) $\mathbb{Q}(\sqrt[3]{2}, \theta)$, onde $\theta^4 + 6\theta + 2 = 0$.

(d) $\mathbb{Q}(\sqrt{7}, \theta)$, onde $\theta^3 + 3 = 0$.

(e) $\mathbb{Q}(\alpha, \beta)$, onde $\alpha^3 - \alpha + 1 = 0$ e $\beta^2 - \beta = 1$.

(f) $\mathbb{Q}(\sqrt{2}, \alpha)$, onde $3\alpha^3 + 7\alpha^2 = 14\alpha - 56$.

(g) $\mathbb{Q}(\sqrt{7}, \theta)$ sendo θ uma raiz do polinómio $x^3 + 2x^2 + 2x - 4$ tal que $[\mathbb{Q}(\theta) : \mathbb{Q}] > 1$.

3.17. Determine o grau e uma base da extensão $\mathbb{Q}(\sqrt{\pi})$ de $\mathbb{Q}(\pi)$.

3.18. Sejam $\alpha^3 = 2$, w uma raiz cúbica da unidade e $\beta = w\alpha$. Determine a dimensão e uma base de $\mathbb{Q}(\alpha, \beta)$ sobre $\mathbb{Q}$.

3.19. Determine para quais dos seguintes polinómios $f(x) \in K[x]$ existem extensões $K(\alpha)$ tais que $f(x)$ é o polinómio mínimo de α:
(a) $x^2 - 4$, $K = \mathbb{Q}$. (b) $x^3 + x + 2$, $K = \mathbb{Z}_3$. (c) $x^2 + 1$, $K = \mathbb{Z}_5$.

3.20. Para cada uma das extensões de $\mathbb{Q}$ indicadas averigúe se θ gera a mesma extensão:

(a) $\theta = 2 + \sqrt[3]{4}$, $\mathbb{Q}(\sqrt[3]{2})$.

(b) $\theta = \sqrt{2} + \sqrt{3}$, $\mathbb{Q}(\sqrt{2})$.

(c) $\theta = u^2 + u + 1$, $\mathbb{Q}(u)$, com $u^2 + 5u - 5 = 0$.

3.21. Considere o polinómio $f(x) = x^3 - x + 1 \in \mathbb{Q}[x]$. Seja θ uma raiz de $f(x)$.

(a) Determine o inverso de $\theta + 1$ em $\mathbb{Q}(\theta)$, escrevendo-o como polinómio em θ de coeficientes racionais.

(b) Considere $u = \theta^2 + 1$. As extensões $\mathbb{Q}(u)$ e $\mathbb{Q}(\theta)$ coincidem?

3.22. Mostre que $x^2 + 1$ é irredutível sobre $\mathbb{Z}_3$. Sendo u uma raíz deste polinómio determine o número de elementos de $\mathbb{Z}_3(u)$ e as tabelas de adição e multiplicação.

3.23. Considere $\mathbb{Z}_5(\alpha)$, sendo $\alpha^2 + 3 = 0$, e determine:

(a) a expressão geral dos elementos desse corpo e o seu cardinal.

(b) o polinómio mínimo de $\beta = \alpha + 1$.

(c) o inverso de β.

3.24. É possível, usando régua (não graduada) e compasso, construir o ponto

$$\left(\sqrt{5\sqrt{2} - 3} + \sqrt{2 - \sqrt[3]{2}}, 0\right)$$

a partir dos pontos $(0, 0)$ e $(1, 0)$?

3.25. Seja p um inteiro primo positivo.

(a) Determine a dimensão e uma base da extensão $\mathbb{Q}(\sqrt{p + \sqrt{p}})$ de $\mathbb{Q}$.

(b) Será possível construir o ponto $(\sqrt{p + \sqrt{p}}, \sqrt{p + \sqrt{p}})$ a partir dos pontos $(0, 0)$ e $(1, 0)$?

3.26. Mostre que é impossível construir com régua e compasso:

(a) um cubo com volume igual ao de uma esfera dada.

(b) o ponto $(\sqrt{5\sqrt{5} - 3} + \sqrt{2 - \sqrt[3]{2}}, 0)$ a partir dos pontos $(0, 0)$ e $(1, 0)$.

3.27. Considere o polinómio $p(x) = 8x^3 - 6x - 1$ sobre $\mathbb{Q}$.

(a) Mostre que $p(x)$ é irredutível sobre $\mathbb{Q}$.

(b) Construa uma extensão de decomposição de $p(x)$ e determine a sua dimensão.

3.28. Determine a extensão de decomposição de:

(a) $x^2 - 5$ sobre $\mathbb{Q}$.

(b) $x^2 + 1$ sobre $\mathbb{R}$.

(c) $x^5 - 2x^4 - 10x^3 + 20x^2 + 25x - 50$ sobre $\mathbb{Q}$.

3.29. Seja L uma extensão de $\mathbb{Q}$. Determine os $\mathbb{Q}$-automorfismos de L para:

(a) $L = \mathbb{Q}(\sqrt{2})$.

(b) $L = \mathbb{Q}(\alpha) \subseteq \mathbb{R}$, com $\alpha^5 = 7$.

(c) $L = \mathbb{Q}(\sqrt{2}, \sqrt{3})$.

(d) L a extensão de decomposição de $x^4 - 4x^2 - 5$.

3.30.

(a) Para as extensões L de $\mathbb{Q}$ do exercício anterior, calcule os respectivos grupos de Galois, $Gal(L, \mathbb{Q})$.

(b) Verifique em quais desses casos a correspondência de Galois entre os subgrupos do grupo de Galois e as extensões intermédias (entre $\mathbb{Q}$ e L) é uma bijecção.

3.31.

(a) Determine os corpos intermédios entre $\mathbb{Q}$ e $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$.

(b) Calcule o respectivo grupo de Galois e compare os resultados.

3.32. Considere a extensão $L = \mathbb{Q}(\sqrt{3}, \sqrt[3]{2}) \subseteq \mathbb{R}$ de $\mathbb{Q}$.

(a) Como se define o grupo de Galois de L (sobre $\mathbb{Q}$)? Determine-o.

(b) Indique todas as extensões intermédias de $\mathbb{Q}$ em L.

(c) L é uma extensão de Galois de $\mathbb{Q}$? Justifique.

3.33. Determine o grupo de Galois associado a cada uma das extensões dos Exercícios 3.18 e 3.16.

3.34. Seja θ uma raiz de $x^2 + x + 1 \in \mathbb{Z}_2[x]$. Mostre que $\Phi : \mathbb{Z}_2(\theta) \to \mathbb{Z}_2(\theta)$ definido por $\Phi(a + b\theta) = a + b + b\theta$, para quaisquer $a, b \in \mathbb{Z}_2$, é um $\mathbb{Z}_2$-automorfismo de $\mathbb{Z}_2(\theta)$.

3.35. Mostre que $Gal(L, K) = 1$ não implica $L = K$.
(Sugestão: Considere $K = \mathbb{Q}$ e L uma extensão de K gerada pela única raiz real de um polinómio irredutível sobre $\mathbb{Q}$, como no Exercício 3.29(b)).

3.36. Seja L uma extensão algébrica simples de K, $\alpha \in L - K$ e $\Phi \in Gal(L, K)$. Mostre que α e $\Phi(\alpha)$ têm o mesmo polinómio mínimo sobre K.

3.37. Calcule o grupo de Galois do polinómio $f(x)$ sobre o corpo K nos seguintes casos:

(a) $f(x) = x^2 + 1$, $K = \mathbb{R}$.

(b) $f(x) = x^4 - 2$, $K = \mathbb{Q}$.

(c) $f(x) = x^3 - x + 1$, $K = \mathbb{Q}$ (veja Exercício 3.5(a)).

(d) $f(x) = x^4 - 4x^2 - 5$, $K = \mathbb{Q}(i)$.

3.38. Sejam K um corpo e L uma extensão de K. Prove que se $\theta \in L$ é algébrico sobre K, de grau n, então $|Gal(K(\theta), K)| \leq n$.

3.39. Sejam n um número natural e K um corpo que contém as raízes de índice n da unidade.

(a) Se n for primo e α raiz do polinómio $x^n - a$, $a \in K$, então $Gal(K(\alpha), K)$ é um grupo cíclico de ordem 1 ou de ordem n.

(b) Se β é raiz de $x^n - a$, $a \in K$, então $Gal(K(\beta), K)$ é cíclico.

3.40.

(a) Sejam p um número primo e K um corpo que contém as raízes de índice p da unidade. Mostre que $x^p - a$, $a \in K$, é irredutível sobre K se e só se não tem raízes sobre K.

(b) Prove que a hipótese de K conter as raízes de índice p da unidade não é necessária.

3.41. Sejam K um corpo de característica diferente de 2, e L uma extensão de K tal que $[L : K] = 2$. Mostre que $L = K(\sqrt{a})$ para algum $a \in K$ e que L é de Galois sobre K.

3.42. Considere um polinómio $f(x)$ irredutível, de grau 3, escrito na sua forma reduzida $x^3 + px + q$, e as suas três raízes complexas distintas a, b e c.

(a) Verifique que $\begin{cases} a+b+c=0 \\ ab+ac+bc=p \\ abc=-q \end{cases}$.

(b) A partir da alínea anterior, mostre que $((a-b)(a-c)(b-c))^2 = -4p^3 - 27q^2$.

(c) Seja D o número $-4p^3 - 27q^2$ da alínea anterior. Prove que se $\sqrt{D} \in \mathbb{Q}$ e $\Phi \in Gal(f(x), \mathbb{Q})$, então $\Phi(\sqrt{D}) = \sqrt{D}$ e portanto $Gal(f(x), \mathbb{Q}) \cong \mathcal{A}_3$.

(d) Prove que se $\sqrt{D} \notin \mathbb{Q}$, então $\mathbb{Q}(\sqrt{D})$ está na extensão de decomposição de $f(x)$ e, portanto, $Gal(f(x), \mathbb{Q}) \cong \mathcal{S}_3$.

3.43. Mostre que se os grupos A e B são resolúveis, então $A \times B$ também é resolúvel. Conclua que se os factores irredutíveis de um polinómio são resolúveis por radicais, então ele também é resolúvel por radicais.

3.44. Para cada um dos seguintes grupos, mostre que são resolúveis e indique um polinómio de coeficientes racionais cuja resolubilidade por radicais resulte desse facto.

(a) $\mathbb{Z}_m \rtimes \mathbb{Z}_m^*$. (b) $\mathbb{Z}_2^n$. (c) $\mathcal{S}_3$. (d) $\mathcal{S}_4$. (e) $\mathbb{Z}_2 \oplus \mathcal{S}_3$.

3.45.

(a) Mostre que, se um grupo resolúvel não tem subgrupos normais próprios, então é um grupo cíclico de ordem prima.

(b) Sabendo que o grupo $\mathcal{A}_5$ não tem subgrupos normais próprios, conclua que ele é resolúvel.

(c) A partir da alínea anterior, mostre que $\mathcal{S}_n$ não é resolúvel para $n \geq 5$.

3.46. Sejam p um número primo, e $f(x) \in \mathbb{Q}[x]$ um polinómio irredutível de grau p. Mostre que:

(a) se $f(x)$ tem exactamente duas raízes complexas não reais, então $Gal(f(x), \mathbb{Q})$ é o grupo simétrico $\mathcal{S}_p$ e portanto $f(x)$ não é resolúvel por radicais.

(b) se $f(x)$ tem exactamente quatro raízes complexas não reais, então não é resolúvel por radicais.

3.47. Mostre que os seguintes polinómios $f(x) \in \mathbb{Q}[x]$ não são resolúveis por radicais:

(a) $f = 2x^5 - 10x + 5$.

(b) $f = 2x^5 - 5x^4 + 20$.

(c) $f = x^5 - 6x^2 + 5$.

(d) $f = x^7 - 10x^5 + 15x + 5$.

3.48. Resolva as seguintes equações por meio de radicais.

(a) $x^5 - 5x^4 + 10x^3 - 10x^2 + 5x - 8 = 0$. (Sugestão: $y = x - 1$.)

(b) $x^3 + 2x^2 - 5x + 9 - \frac{5}{x} + \frac{2}{x^2} + \frac{1}{x^3} = 0$. (Sugestão: $y = x + \frac{1}{x}$.)

3.49. Determine a extensão radical sobre $\mathbb{Q}$ que contém o seguinte número complexo:

(a) $\sqrt[3]{8} + \sqrt{2}$.

(b) $\dfrac{\sqrt[7]{13 + \sqrt{2}}}{\sqrt[3]{5}}$.

3.50. Verifique que, apesar de $x^3 - 3x + 1$ ser resolúvel por radicais, a sua extensão de decomposição não é uma extensão radical. (Veja Exercício 3.42)

4. Corpos finitos

Neste capítulo final vamos estudar as propriedades fundamentais dos corpos finitos e descrever algumas das suas muitas aplicações (à teoria dos códigos, teoria dos números e teoria matemática dos jogos).

O corpo $\mathbb{F}_p = (\mathbb{Z}_p, \oplus_p, \otimes_p)$ dos inteiros módulo p (p primo) é, evidentemente, o exemplo mais familiar de corpo finito. Muitas das suas propriedades generalizam-se aos corpos finitos arbitrários. Os corpos $\mathbb{F}_p$ representam um papel muito importante na teoria dos corpos pois, como vimos, todo o corpo de característica p contém uma cópia isomorfa de $\mathbb{F}_p$ (como seu subcorpo primo) e pode então ser visto como uma extensão de $\mathbb{F}_p$. Esta observação, conjuntamente com o facto óbvio de que todo o corpo finito tem característica finita (prima), é fundamental para a classificação dos corpos finitos.

Além dos corpos $\mathbb{F}_p$, de ordem prima p, já encontrámos outros exemplos de corpos finitos: um corpo de ordem $4 = 2^2$, definido pelas tabelas

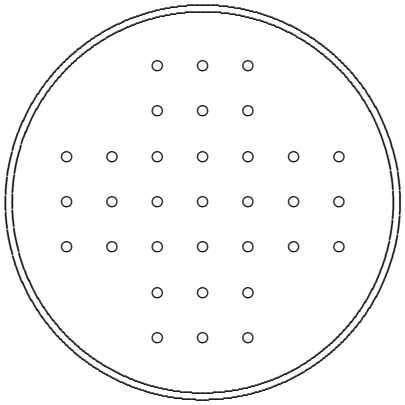
$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

$\cdot$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

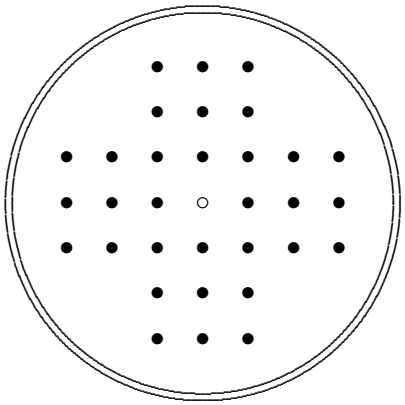
e um corpo de ordem $16 = 2^4$. Haverá algum corpo de ordem 6? Veremos em seguida que não, ao provarmos que a ordem de qualquer corpo finito é necessariamente da forma p^n para algum primo p e algum natural n, e que, para cada número dessa forma existe, a menos de isomorfismo, exactamente um corpo com esse número de elementos.

Antes de avançarmos para a prova desses resultados que permitem classificar os corpos finitos, vejamos uma aplicação do corpo com 4 elementos acima referido, que se pode encontrar em [N. de Bruijn, *A solitaire game and its relation to a finite field*, J. Recreational Math. 5 (1972) 133-137].

O jogo do solitário é jogado num tabuleiro representado na figura seguinte:



Inicialmente, em cada buraco, com excepção do central, coloca-se uma bola (32 bolas no total):



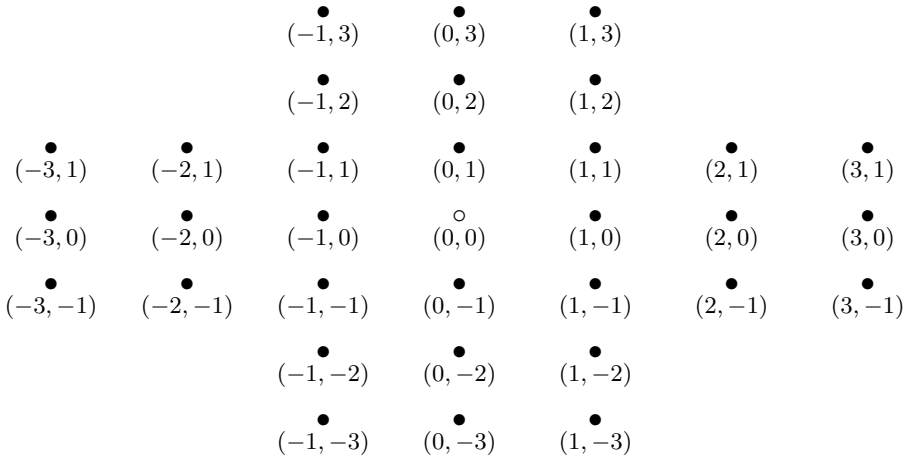
O jogo desenrola-se movimentando uma bola por cima de outra adjacente (na vertical ou na horizontal) para um buraco vazio; a bola sobre a qual se saltou é então removida do jogo. O objectivo do jogador é chegar a uma situação em que só reste uma bola no tabuleiro, idealmente na posição central.[21]

> `[É claro que se conseguirmos terminar com uma única bola na posição central, também conseguiremos terminar com essa bola noutras posições; experimente!]`

Em quais posições é possível terminar o jogo, ganhando?

Depois de jogarmos algumas vezes não será difícil convencermo-nos que talvez não possa ocupar qualquer posição. A ideia de de Bruijn é usar o corpo acima referido para determinar tais posições. Para isso, consideremos os buracos do tabuleiro referenciados por pares de inteiros (i, j), com o buraco central em $(0, 0)$:

[21] Para mais informação sobre este jogo e suas variantes e generalizações veja [E.R. Berlekamp, J.H. Conway e R.K. Guy, *Winning Ways for your Mathematical Plays*, Vol. 4, A K Peters, 2004] e as referências aí incluídas, e [George I. Bell, *A fresh look at peg solitaire*, Mathematics Magazine 80 (2007) 16-28].

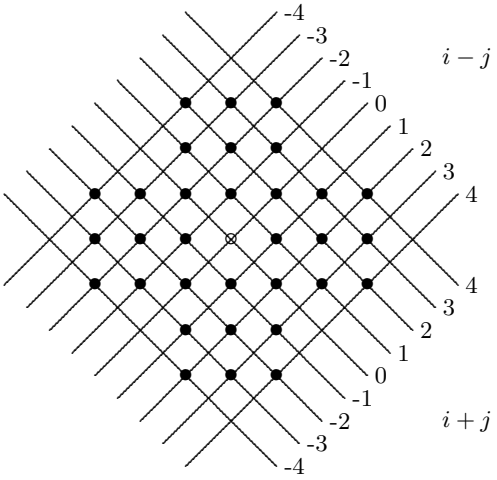


Definamos, para cada conjunto X de bolas colocadas no tabuleiro, os números

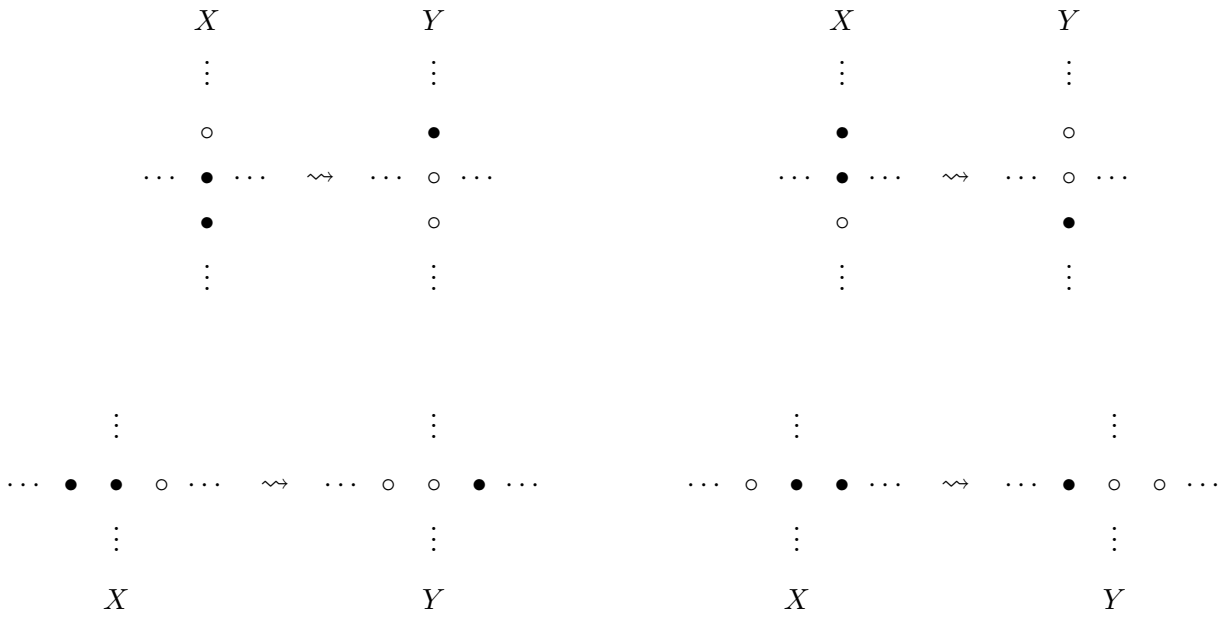
$$A(X) = \sum_{(i,j)\in X} \alpha^{i+j}, \qquad B(X) = \sum_{(i,j)\in X} \alpha^{i-j}.$$

Por exemplo, para a posição inicial X_1 do jogo, é fácil de ver (observe a figura abaixo) que

$$\begin{aligned} A(X_1) = B(X_1) &= 2\alpha^4 + 4\alpha^3 + 5\alpha^2 + 4\alpha^1 + 2\alpha^0 + 4\alpha^{-1} + 5\alpha^{-2} + 4\alpha^{-3} + 2\alpha^{-4} \\ &= 0 + 0 + 5\beta + 0 + 0 + 0 + 5\alpha + 0 + 0 \\ &= \alpha + \beta = 1. \end{aligned}$$



Cada jogada, que transforma um conjunto X de bolas no tabuleiro num conjunto Y, é necessariamente de um dos quatro tipos seguintes:



É fácil de ver que, em qualquer um desses tipos de jogada, se tem $A(Y) = A(X)$ e $B(Y) = B(X)$. Por exemplo, no primeiro tipo, se supusermos que a bola a movimentar está inicialmente na posição (i, j) (e portanto, após a jogada, vai ficar na posição $(i, j + 2)$), então

$$A(X) - A(Y) = \alpha^{i+j} + \alpha^{i+j+1} - \alpha^{i+j+2} = \alpha^{i+j}(1 + \alpha + \alpha^2) = 0,$$

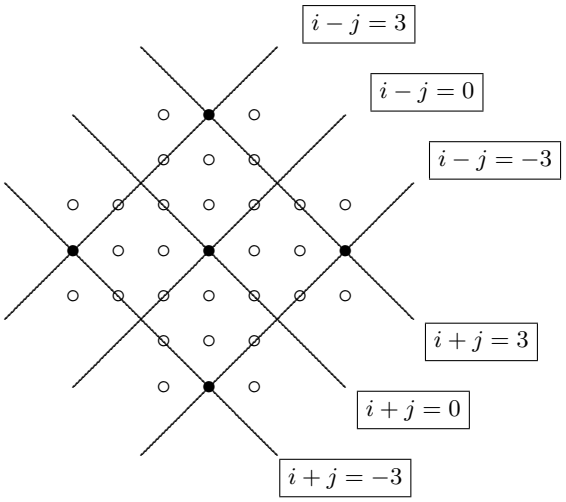
$$B(X) - B(Y) = \alpha^{i-j} + \alpha^{i-j-1} - \alpha^{i-j-2} = \alpha^{i-j}(1 + \beta + \beta^2) = 0.$$

Portanto, o par $(A(X), B(X))$ é invariante ao longo do jogo.

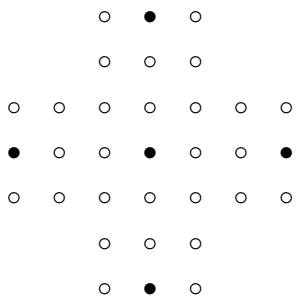
Assim, se o jogo terminar com uma só bola no tabuleiro, na posição (i, j), teremos necessariamente $A(\{(i, j)\}) = 1$ e $B(\{(i, j)\}) = 1$, isto é, $\alpha^{i+j} = 1$ e $\alpha^{i-j} = 1$. Como as sucessivas potências de α são

$$\alpha^{-4} = \beta, \boxed{\alpha^{-3} = 1}, \alpha^{-2} = \alpha, \alpha^{-1} = \beta, \boxed{\alpha^0 = 1}, \alpha^1 = \alpha, \alpha^2 = \beta, \boxed{\alpha^3 = 1}, \alpha^4 = \alpha,$$

a posição (i, j) da bola final terá que satisfazer $i+j \in \{-3, 0, 3\}$ e $i-j \in \{-3, 0, 3\}$:



Em conclusão, as únicas posições finais possíveis são $(-3,0)$, $(0,-3)$, $(0,0)$, $(0,3)$ e $(3,0)$:



Por experimentação, é possível concluir que todas elas podem ser, de facto, obtidas.

[Basta para isso mostrar que se consegue atingir a posição $(0,0)$. A maneira de atingir as outras é depois óbvia.]

Voltemos agora à classificação dos corpos finitos.

Teorema 4.1 *Seja F um corpo finito. Então F tem p^n elementos, onde $p = car(F)$ e n é a dimensão $[F:P]$ de F como extensão do seu subcorpo primo P.*

Demonstração. Como F é finito, F é uma extensão finita do seu subcorpo primo P e a sua característica é um primo p. Já sabemos que $P \cong \mathbb{F}_p$. Suponhamos que $[F:P] = n$ e seja $\{\theta_1, \theta_2, \ldots, \theta_n\}$ uma base do espaço vectorial F sobre o corpo P. Cada elemento de F escreve-se de forma única como combinação linear dos vectores $\theta_1, \theta_2, \ldots, \theta_n$, pelo que

$$F = \Big\{ a_1\theta_1 + a_2\theta_2 + \cdots + a_n\theta_n \mid a_1, a_2, \ldots, a_n \in P \Big\}.$$

É claro que, como P tem p elementos, o número destas combinações lineares é igual a p^n (número de arranjos com repetição de p elementos n a n). Portanto, $|F| = p^n$. ∎

A partir dos corpos primos $\mathbb{F}_p$, podemos construir outros corpos finitos pelo processo de adjunção de raízes descrito no capítulo anterior. Se $p(x) \in \mathbb{F}_p[x]$ é um polinómio de grau n, irredutível sobre $\mathbb{F}_p$, então juntando uma raiz de $p(x)$ a $\mathbb{F}_p$ obtemos um corpo finito com p^n elementos. Contudo, não é claro, nesta altura, que exista, para qualquer natural n, um tal polinómio irredutível de grau n. Assim, de modo a provarmos que para cada primo p e para cada natural n existe um corpo com p^n elementos, seguiremos uma abordagem sugerida pelo seguinte resultado.

Proposição 4.2 *Seja F um corpo com p^n elementos. Então F é isomorfo à extensão de decomposição do polinómio $x^{p^n} - x$ sobre $\mathbb{F}_p$.*

Demonstração. O grupo multiplicativo $(F \setminus \{0\}, \cdot)$ tem ordem $p^n - 1$, pelo que, para qualquer $a \in F$ diferente de 0, $a^{p^n-1} = 1$. Isto significa que $a^{p^n} \cdot a^{-1} = 1$, isto é, $a^{p^n} = a$.

[Este facto será decisivo: em qualquer corpo F com q elementos, cada $a \in F$ satisfaz $a^q = a$]

Portanto, todos os elementos de F são raízes do polinómio $p(x) = x^{p^n} - x \in \mathbb{F}_p[x]$. Como este polinómio tem grau p^n e $|F| = p^n$, isto mostra que F contém todas as suas raízes e

$$x^{p^n} - x = \prod_{a \in F} (x - a).$$

Portanto F contém uma extensão de decomposição de $p(x)$. Mas F é exactamente o conjunto das raízes de $p(x)$, pelo que, necessariamente, F é a extensão de decomposição de $p(x)$. ∎

Corolário 4.3 [E. H. Moore, 1893]
Dois corpos finitos com o mesmo número de elementos são isomorfos.

Demonstração. É consequência imediata da proposição anterior e da unicidade, a menos de isomorfismo, das extensões de decomposição, provada no capítulo anterior. ∎

Estamos agora em condições de provar o recíproco do Teorema 4.1.

Teorema 4.4 [Teorema de Galois]
Para cada primo p e cada $n \in \mathbb{N}$, existe um corpo com p^n elementos, único a menos de isomorfismo.

Demonstração. Provemos somente a existência de tal corpo, estando a unicidade assegurada pelo corolário anterior.

Para $q = p^n$, consideremos o polinómio $p(x) = x^q - x$ de $\mathbb{F}_p[x]$. Seja ainda F a extensão de decomposição de $p(x)$.

[Observe que um elemento a de um corpo K é uma *raiz múltipla* de $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in K[x]$ se e só se é uma raiz de $p(x)$ e da sua derivada $D(p(x)) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$]

Como, neste caso $D(p(x)) = qx^{q-1} - 1 = -1 \neq 0$, todas as raízes de $p(x)$ são simples. Portanto, o conjunto $R = \{a \in F \mid a^q - a = 0\}$ das raízes de $p(x)$ em F tem cardinal q. Mas R é um subcorpo de F.

[Verifique]

Está assim encontrado um corpo com p^n elementos: o corpo R das raízes de $p(x)$ em F, que coincide forçosamente com F, uma vez que $p(x)$ se decompõe em factores lineares em R. ∎

Em conclusão:

CLASSIFICAÇÃO DOS CORPOS FINITOS
• Todo o corpo finito tem p^n elementos, para algum primo p e algum natural n. • Para cada primo p e cada natural n, existe um corpo com p^n elementos. • Qualquer corpo com p^n elementos é isomorfo à extensão de decomposição de $x^{p^n} - x$ sobre $\mathbb{F}_p$.

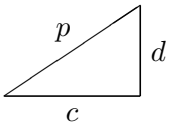
A unicidade no Teorema de Galois justifica que se fale *no* corpo finito (ou *no* corpo de Galois) com q elementos:

CORPO DE GALOIS de ordem q
A este corpo (único, a menos de isomorfismo) chama-se o *corpo de Galois com q elementos*, que se denota por $\mathbb{F}_q$ (ou por $\mathbf{GF}(q)$).

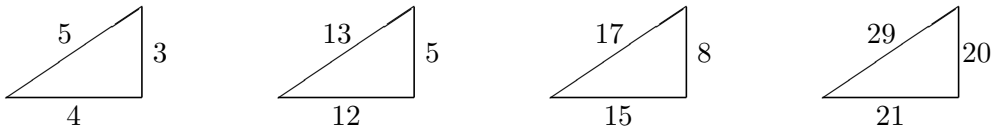
Uma aplicação dos corpos finitos à Teoria dos Números

A seguinte questão constitui um problema clássico da Teoria dos Números:

> Problema: *Seja $p \in \mathbb{N}$, primo. Quando é que p pode ser a hipotenusa de um triângulo rectângulo de catetos c e d inteiros?*



É claro que tal é possível exactamente quando $p^2 = c^2 + d^2$, para algum par c, d de inteiros positivos. Por exemplo, para $p = 5, 13, 17, 29$:



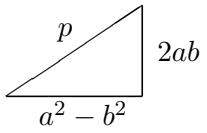
Como ilustração do que se pode fazer com os resultados que vimos até ao momento, vamos agora apresentar uma prova extremamente elegante, retirada dos apontamentos de Álgebra II de A. Machiavelo [DMUP, 1997-99], de um resultado de Fermat que ajuda a resolver este problema.

> Proposição [Fermat]: *Se $p \in \mathbb{N}$ é primo e $p \equiv 1 \pmod 4$ então p é soma de dois quadrados.*

De facto, se p é uma soma $a^2 + b^2$ de dois quadrados então

$$p^2 = (a^2 + b^2)^2 = a^4 + 2a^2b^2 + b^4 = (a^2 - b^2)^2 + (2ab)^2,$$

pelo que, tomando $c = a^2 - b^2$ e $d = 2ab$, obtemos um triângulo nas condições do problema, com hipotenusa p:



Assim, a Proposição de Fermat dá-nos uma condição suficiente para que um primo p seja hipotenusa de um tal triângulo:

$$p \equiv 1 \pmod 4.$$

É o caso de todos os exemplos que apresentámos acima:

$$\begin{aligned}
p = 5: \quad & 5 = 2^2 + 1^2 \Rightarrow 5^2 = (2^2 - 1^2)^2 + (2 \times 2 \times 1)^2 = 3^2 + 4^2; \\
p = 13: \quad & 13 = 3^2 + 2^2 \Rightarrow 13^2 = (3^2 - 2^2)^2 + (2 \times 3 \times 2)^2 = 5^2 + 12^2; \\
p = 17: \quad & 17 = 4^2 + 1^2 \Rightarrow 17^2 = (4^2 - 1^2)^2 + (2 \times 4 \times 1)^2 = 15^2 + 8^2; \\
p = 29: \quad & 29 = 5^2 + 2^2 \Rightarrow 29^2 = (5^2 - 2^2)^2 + (2 \times 5 \times 2)^2 = 21^2 + 20^2.
\end{aligned}$$

Demonstremos então a Proposição de Fermat, usando alguns factos sobre corpos finitos provados anteriormente.

Para isso começamos por determinar todos os primos p para os quais -1 é um quadrado módulo p, ou seja, para os quais $\mathbb{F}_p$ tem uma raiz quadrada de -1.

Quando $p = 2$ a resposta é óbvia: $-1 = 1 = 1^2$. Suponhamos pois $p \neq 2$. Seja F uma extensão de decomposição sobre $\mathbb{F}_p$ do polinómio $x^2 + 1 \in \mathbb{F}_p[x]$, e denotemos por i uma das duas raízes deste polinómio em F. Como vimos na Proposição 4.2, para cada $a \in F$ tem-se que $a \in \mathbb{F}_p$ se e só se $a^p = a$. Assim, em particular, $i \in \mathbb{F}_p$ se e só se $i^p = i$. Mas

$$i^p = (i^2)^{\frac{p-1}{2}} i = (-1)^{\frac{p-1}{2}} i,$$

que é igual a i quando e só quando $(-1)^{\frac{p-1}{2}} = 1$, ou seja, quando e só quando $p - 1$ é um múltiplo de 4. Portanto, a equação $x^2 \equiv -1 \pmod p$ (p primo) tem solução se e só se $p = 2$ ou $p \equiv 1 \pmod 4$.

Seja agora p um primo tal que $p \equiv 1 \pmod 4$. Então, pelo que acabámos de ver, $m^2 \equiv -1 \pmod p$, ou seja, $p|(m^2 + 1)$, para algum inteiro m. Isto implica que, no domínio $\mathbb{Z}[i]$ dos inteiros de Gauss, $p|(m + i)(m - i)$. Mas $p \nmid (m + i)$, pois $m + i = (a + bi)p$ implicaria $pa = m$, ou seja, $m \equiv 0 \pmod p$; analogamente, $p \nmid (m - i)$. Daqui resulta que p não é primo em $\mathbb{Z}[i]$. Mas $\mathbb{Z}[i]$ é um domínio de ideais principais, donde p, não sendo primo, é necessariamente redutível, ou seja, existem inteiros a, b, c, d tais que $p = (a + bi)(c + di)$, onde $a + bi$ e $c + di$ não são unidades de $\mathbb{Z}[i]$ (ou seja, $a + bi, c + di \neq \pm 1, \pm i$). Consequentemente,

$|p| = |a+bi|\,|c+di|$ e, elevando ao quadrado, $p^2 = (a^2+b^2)(c^2+d^2)$. Como p é um inteiro primo, é fácil de ver que isto implica $a^2+b^2 = c^2+d^2 = p$. Em conclusão, $p = a^2+b^2$ como Fermat afirmou.

Exercício:

(1) Seja p um primo ímpar e F uma extensão de decomposição sobre $\mathbb{F}_p$ do polinómio x^2+1. Designando por i uma das raízes em F de x^2+1, use a relação $(1+i)^2 = 2i$ para determinar quais os primos p tais que 2 é um quadrado módulo p.

(2) Use (1) para provar o seguinte resultado de Euler:

> *Se p é um primo tal que $p \equiv 3$ (mod 4) e $2p+1$ é primo, então $(2p+1)|(2^p-1)$.*

[Este resultado de Euler mostra, em particular, que o *número de Mersenne* 2^p-1 não é primo para $p>3$ nas condições enunciadas; por exemplo: $23|2^{11}-1$, $47|2^{23}-1$]

[Mais uma vez, note a utilidade da introdução do conceito de polinómio como função definida em $\mathbb{N}_0$ com suporte finito, distinguindo-os assim das respectivas funções polinomiais. De facto, pelo Teorema pequeno de Fermat (''para cada a não divisível pelo primo p, $a^{p-1} \equiv 1 \pmod p$''), existe apenas um número finito de funções polinomiais $\mathbb{F}_p \to \mathbb{F}_p$ (por exemplo, a função $x \mapsto x^p$ é igual a $x \mapsto x$), enquanto que os polinómios permitem construir uma infinidade de extensões de $\mathbb{F}_p$, para cada primo p, e tais extensões permitem-nos obter resultados não triviais sobre, por exemplo, os números inteiros, como acabámos de ilustrar]

Teorema 4.5 [Critério dos subcorpos]
Seja $\mathbb{F}_q$ o corpo de Galois com $q = p^n$ elementos. Então:

(a) *Todo o subcorpo de $\mathbb{F}_q$ tem ordem p^d, para algum divisor positivo d de n.*

(b) *Reciprocamente, para cada divisor positivo d de n, existe exactamente um subcorpo de $\mathbb{F}_q$ com p^d elementos.*

Demonstração. (a) Seja K um subcorpo de $\mathbb{F}_q$. É evidente que K e $\mathbb{F}_q$ têm o mesmo subcorpo primo P, que é isomorfo a $\mathbb{F}_p$:

$$\mathbb{F}_p \cong P \subseteq K \subseteq \mathbb{F}_q.$$

Então, pelo Teorema 4.1, $|K| = p^d$, onde $d = [K : P]$. Mas

$$n = [\mathbb{F}_q : P] = [\mathbb{F}_q : K]\,[K : P] = [\mathbb{F}_q : K]\,d,$$

logo $d|n$.

(b) Se $d|n$ (isto é, $n = md$ para algum $m \in \mathbb{N}$) então $x^d - 1|x^n - 1$:

$$x^n - 1 = x^{dm} - 1 = (x^d - 1)(x^{d(m-1)} + x^{d(m-2)} + \cdots + x^d + 1). \tag{4.5.2}$$

Em particular, para $x = p$ segue $p^d - 1|p^n - 1$ donde, aplicando (4.5.2) a esta relação, $x^{p^d-1} - 1|x^{p^n-1} - 1$. Multiplicando por x obtemos, ainda,

$$x^{p^d} - x|x^{p^n} - x = x^q - x.$$

Portanto, qualquer raiz de $x^{p^d} - x$ é raiz de $x^q - x \in \mathbb{F}_q[x]$. Por outro lado,

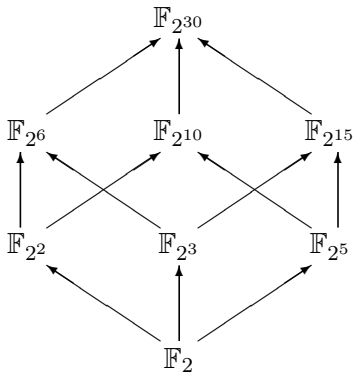
`[Recorde da aula anterior: a extensão de decomposição do polinómio` $x^{p^n} - x$ `sobre` $\mathbb{F}_p$ `tem exactamente` p^n `elementos, e é isomorfa a` $\mathbb{F}_{p^n}$`]`

$\mathbb{F}_q$ é a extensão de decomposição de $x^q - x$ sobre $\mathbb{F}_p$. Então $\mathbb{F}_q$ contém todas as raízes de $x^{p^d} - x$, pelo que contém como subcorpo a extensão de decomposição de $x^{p^d} - x$ sobre $\mathbb{F}_p$. Isto mostra que esta extensão, que tem precisamente p^d elementos, é um subcorpo de $\mathbb{F}_q$, e é precisamente o subcorpo que procurávamos.

A unicidade decorre imediatamente do seguinte facto: se houvesse dois subcorpos distintos de ordem p^d em $\mathbb{F}_q$, juntos teriam mais do que p^d elementos (que são raízes em $\mathbb{F}_q$ de $x^{p^d} - x$), uma contradição, pois $x^{p^d} - x$ só pode ter no máximo p^d raízes. Portanto, o único subcorpo de $\mathbb{F}_{p^n}$ de ordem p^d é o corpo das raízes de $x^{p^d} - x \in \mathbb{F}_p[x]$ em $\mathbb{F}_{p^n}$. ∎

Isto significa que a lista de subcorpos de $\mathbb{F}_{p^n}$, a menos de isomorfismo, coincide precisamente com $\{\mathbb{F}_{p^d} : d|n\}$.

Por exemplo, os subcorpos de $\mathbb{F}_{2^{30}}$ podem ser determinados listando todos os divisores positivos de 30: como $30 = 2 \times 3 \times 5$, os únicos divisores positivos de 30 são 1,2,3,5,6,10,15,30, pelo que existem precisamente 8 subcorpos de $\mathbb{F}_{2^{30}}$:



Neste diagrama indicam-se ainda ainda as relações de inclusão entre os vários subcorpos. Pelo Critério dos Subcorpos, estas relações são equivalentes às relações de divisibilidade entre os divisores positivos de 30. O corpo $\mathbb{F}_2$ é o subcorpo primo de $\mathbb{F}_{2^{30}}$.

Aplicações: Teoria Algébrica dos Códigos

Consideremos o seguinte código binário, a que chamaremos $\mathcal{C}_1$, que permite dar as instruções de comando a um leitor de DVD, através de um comando à distância:

PLAY	REW	FORWARD	STOP
00	01	10	11

Suponhamos que carregamos na tecla PLAY do comando, a que corresponde a palavra 00 do código; o comando transmite esta palavra ao leitor de DVD mas se, porventura, nessa comunicação ocorrer o erro

$$\mathbf{0}0 \overset{erro}{\dashrightarrow} \mathbf{1}0$$

o leitor receberá a palavra 10, e como esta faz parte de $\mathcal{C}_1$ (corresponde à instrução FORWARD), aquele não terá nenhuma maneira de detectar o erro e executará a instrução FORWARD!

O código $\mathcal{C}_1$ é um exemplo de *código binário*, ou seja, um código definido sobre o alfabeto (corpo) $\mathbb{F}_2$, constituído por todas as palavras de comprimento 2 nesse alfabeto. Trata-se de um código muito pobre, pois nem sequer detecta erros *simples* (*singulares*) como o do exemplo acima.

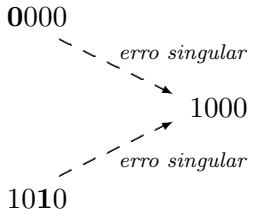
O que fazemos habitualmente quando não entendemos o que outra pessoa nos quer dizer? Pedimos que repita a mensagem. Façamos isso no código $\mathcal{C}_1$, isto é, pensemos no código $\mathcal{C}_2$ que se obtem de $\mathcal{C}_1$ repetindo a informação em cada palavra uma vez:

PLAY	REW	FORWARD	STOP
0000	0101	1010	1111

Agora, ao ser transmitida a instrução PLAY (ou seja, a palavra 0000), se ocorrer o mesmo erro singular de há pouco,

$$\mathbf{0}000 \overset{erro}{\dashrightarrow} \mathbf{1}000$$

como a palavra recebida não faz parte de $\mathcal{C}_2$, o leitor de DVD pode concluir imediatamente que ocorreu algum erro na transmissão. Neste caso, o código $\mathcal{C}_2$ já detecta este erro singular (e é fácil de ver que detecta qualquer outro erro singular). Terá maneira de corrigir esse erro, isto é, de identificar a palavra original (*assumindo que na transmissão só poderão ocorrer, quando muito, erros singulares*)? Não; de facto, há duas palavras em $\mathcal{C}_2$ que poderiam ser as originais:



Consideremos, finalmente, o código $\mathcal{C}_3$, definido pela tabela

PLAY	REW	FORWARD	STOP
000000	010101	101010	111111

Agora, além de qualquer erro singular ser detectável, também pode ser corrigido automaticamente (assumindo novamente que *na transmissão só poderão ocorrer, quando muito, erros singulares*). Por exemplo, o erro singular

$$\mathbf{0}00000 \overset{erro}{\dashrightarrow} \mathbf{1}00000$$

é evidentemente detectado e corrigido; a única palavra de $\mathcal{C}_3$ que poderia ter dado origem à palavra 100000, *na assumpção que só ocorreram erros singulares*, é a palavra 000000:

Palavra de $\mathcal{C}_3$	000000	010101	101010	111111
Palavra recebida	**1**00000	**10000**	10**000**	1**00000**
Número de erros	1	4	2	5

É claro que se puderem ocorrer erros duplos no canal de comunicação, $\mathcal{C}_3$ já não corrige o erro singular acima: a palavra original poderia muito bem ser a palavra 101010.

Assim, esta ideia de construir códigos correctores de erros só funciona se conhecermos *a priori* um limite para o número de erros que pode ocorrer no respectivo canal de comunicação. Ou, então, se adoptarmos o seguinte princípio de bom senso (o chamado *princípio do vizinho mais próximo*):

> *A palavra original correspondente a uma palavra recebida com erros deve ser a palavra do código "mais próxima" da palavra recebida*
>
> (isto é, assumimos que é mais provável que o menor número de erros possível tenha ocorrido na transmissão).

Daqui em diante, assumimos sempre este princípio. (Mais adiante, tornaremos precisa a noção de proximidade implícita no termo "mais próxima".)

Os códigos $\mathcal{C}_1$, $\mathcal{C}_2$ e $\mathcal{C}_3$ são exemplos do tipo de códigos que vamos estudar, e que podem ser formalizados do seguinte modo:

CÓDIGOS SOBRE UM CORPO FINITO $\mathbb{F}_q$. CÓDIGOS LINEARES

Um *código de comprimento n sobre* o corpo $\mathbb{F}_q$ é um subconjunto $\mathcal{C}$ de $(\mathbb{F}_q)^n$. Portanto, $\mathcal{C}$ é formado por palavras de comprimento n, $a_1a_2\ldots a_n$, formadas com o alfabeto $\mathbb{F}_q$ (isto é, cada $a_i \in \mathbb{F}_q$).

Note que $\mathbb{F}_q^n$ é um espaço vectorial sobre $\mathbb{F}_q$, de dimensão n. Assim, as palavras de $\mathcal{C}$ são simplesmente vectores deste espaço. Quando $\mathcal{C}$ é um subespaço linear de $\mathbb{F}_q^n$, de dimensão k, diz-se que $\mathcal{C}$ é um *código (n,k)-linear* ou *(n,k)-código* sobre $\mathbb{F}_q$.

Exemplos: $\mathcal{C}_1 = \mathbb{F}_2^2$, pelo que $\mathcal{C}_1$ é um $(2,2)$-código sobre $\mathbb{F}_2$. Os códigos $\mathcal{C}_2$ e $\mathcal{C}_3$ também são códigos lineares sobre $\mathbb{F}_2$ (binários), como é fácil de ver: $\mathcal{C}_2$ é um $(4,2)$-código enquanto $\mathcal{C}_3$ é um $(6,2)$-código.

Os (n,k)-códigos sobre o corpo $\mathbb{F}_2$ foram o tipo de códigos utilizados pelas sondas que viajaram até Marte, na transmissão das fotografias para a Terra. No caso dos CDs de música, utiliza-se o corpo $\mathbb{F}_{256} = \mathbb{F}_{2^8}$.

Precisemos agora a noção de distância entre duas palavras de $\mathbb{F}_q^n$.

DISTÂNCIA DE HAMMING

A *distância de Hamming* entre duas palavras $\vec{a} = a_1a_2\ldots a_n$ e $\vec{b} = b_1b_2\ldots b_n$ é o número de índices $i \in \{1,2,\ldots,n\}$ tais que $a_i \neq b_i$.

Note que $d(\vec{a},\vec{b})$ indica o número de erros ocorridos se $\vec{a}$ é a palavra transmitida e $\vec{b}$ é a palavra recebida.

Por exemplo, $d(1101, 0111) = 2$.
É muito fácil de ver que a distância de Hamming é uma métrica em $\mathbb{F}_q^n$, isto é, para quaisquer $\vec{a}, \vec{b}, \vec{c} \in \mathbb{F}_q^n$, tem-se:

(1) $d(\vec{a}, \vec{b}) \geq 0$; $d(\vec{a}, \vec{b}) = 0$ se e só se $\vec{a} = \vec{b}$.

(2) $d(\vec{a}, \vec{b}) = d(\vec{b}, \vec{a})$.

(3) $d(\vec{a}, \vec{b}) \leq d(\vec{a}, \vec{c}) + d(\vec{c}, \vec{b})$.

DISTÂNCIA MÍNIMA

Chama-se *distância mínima* de um código $\mathcal{C}$, que se denota por $\delta(\mathcal{C})$, ao número

$$\min_{\vec{a},\vec{b}\in\mathcal{C},\vec{a}\neq\vec{b}} d(\vec{a}, \vec{b}).$$

Este número mede o grau de vizinhança das palavras em $\mathcal{C}$. Por exemplo, $\delta(\mathcal{C}_1) = 1$, $\delta(\mathcal{C}_2) = 2$ e $\delta(\mathcal{C}_3) = 3$.

Quanto maior é o valor de $\delta(\mathcal{C})$, mais eficiente é o código. Portanto, um dos objectivos na construção de um código é que tenha as palavras o mais afastadas entre si. Por outro lado, isto limita o número de palavras do código, logo limita a sua capacidade de armazenar e transmitir informação. Reconciliar estes dois objectivos (isto é, procurar o ponto de equilíbrio entre eles) é um dos problemas da teoria dos códigos.

CÓDIGOS t-DETECTORES E t-CORRECTORES DE ERROS

Seja $t \in \mathbb{N}$. Diz-se que um código $\mathcal{C}$ é *t-detector de erros* se detecta qualquer combinação de t erros em qualquer palavra.
Diz-se que $\mathcal{C}$ é *t-corrector de erros* se corrige qualquer combinação de t erros em qualquer palavra.

Teorema 4.6 *Seja $\mathcal{C}$ um código com distância mínima $\delta(\mathcal{C})$. Então:*

(a) *$\mathcal{C}$ é t-detector de erros se e só se $t \leq \delta(\mathcal{C}) - 1$.*

(b) *$\mathcal{C}$ é t-corrector de erros se e só se $t \leq \frac{\delta(\mathcal{C})-1}{2}$.*

Demonstração. (a) É evidente que em qualquer código $\mathcal{C}$, existindo duas palavras $\vec{a}$ e $\vec{b}$ tais que $d(\vec{a}, \vec{b}) = \delta(\mathcal{C})$, se a palavra transmitida for $\vec{a}$ e acontecerem $\delta(\mathcal{C})$ erros que a transformem em $\vec{b}$, esses erros nunca serão detectados. Portanto, se $\mathcal{C}$ é t--detector de erros então $t < \delta(\mathcal{C})$. Reciprocamente, suponhamos que na transmissão de uma palavra $\vec{a} \in \mathcal{C}$ ocorreram t erros, resultando na palavra $\vec{b}$:

$$\vec{a} \quad \overset{\substack{t \\ \text{erros}}}{\dashrightarrow} \quad \vec{b}$$

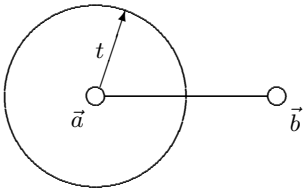
(portanto, $d(\vec{a},\vec{b}) = t$). Para provarmos que o código terá a capacidade de detectar o erro, teremos que garantir que $\vec{b} \notin \mathcal{C}$, o que é fácil: como $d(\vec{a},\vec{b}) = t < \delta(\mathcal{C})$ e $\vec{a} \in \mathcal{C}$ então $\vec{b} \notin \mathcal{C}$.

(b) Se $\mathcal{C}$ é t-corrector de erros, então $2t \leq \delta(\mathcal{C}) - 1$. De facto, $\delta(\mathcal{C}) = 2t$ implicaria a existência de duas palavras $\vec{a}$ e $\vec{b}$ diferindo exactamente em $2t$ posições; acontecendo t erros em metade dessas $2t$ posições na transmissão de $\vec{a}$, nunca seria possível corrigir esses erros pois poderia ter sido a palavra $\vec{b}$ a palavra emitida (tendo os t erros ocorrido na outra metade dessas $2t$ posições).

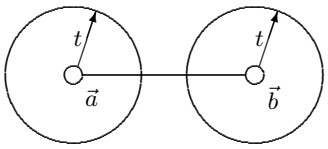
Reciprocamente, suponhamos que na transmissão de uma palavra $\vec{a} \in \mathcal{C}$ ocorreram t erros, resultando na palavra recebida $\vec{b}$ (portanto, $d(\vec{a},\vec{b}) = t$). Agora, para provarmos que o código terá a capacidade de corrigir o erro, bastará garantir que mais nenhuma palavra em $\mathcal{C}$ além de $\vec{a}$ pode ter dado origem à palavra errada $\vec{b}$, ou seja, que qualquer outra palavra $\vec{c} \in \mathcal{C}$ está a uma distância de $\vec{b}$ maior do que t, o que também é fácil: pela desigualdade triangular da distância,

$$d(\vec{b},\vec{c}) \geq d(\vec{a},\vec{c}) - d(\vec{a},\vec{b}) \geq \delta(\mathcal{C}) - t \geq 2t + 1 - t = t + 1. \qquad \blacksquare$$

Portanto, um código consegue detectar t erros se quaisquer duas palavras do código estiverem a uma distância de Hamming pelo menos $t + 1$:



Por sua vez, um código consegue corrigir t erros se quaisquer duas palavras do código estiverem a uma distância de Hamming pelo menos $2t + 1$:



Nos exemplos que vimos anteriormente, tem-se:

Código	$\delta(\mathcal{C})$	No. erros que detecta	No. erros que corrige
$\mathcal{C}_1$	1	0	0
$\mathcal{C}_2$	2	1	0
$\mathcal{C}_3$	3	2	1

Portanto $\mathcal{C}_2$ é 1-detector de erros e $\mathcal{C}_3$ é 1-corrector de erros e 2-detector de erros.

A definição de código t-corrector implica que quaisquer bolas de raio t, centradas em palavras distintas, sejam disjuntas. Se, além disso, estas bolas cobrirem a totalidade do espaço (uma propriedade rara mas interessante), o código diz-se *perfeito*. Assim, um código t-corrector $\mathcal{C}$ sobre $\mathbb{F}_q$ diz-se *perfeito* se

$$\bigcup_{\vec{a}\in\mathcal{C}} B(\vec{a},t) = \mathbb{F}_q^n.$$

Suponhamos que, num determinado sistema de comunicação, necessitamos de um código com, no máximo, q^k palavras. Poderemos então usar todas as palavras $a_1a_2\cdots a_k \in \mathbb{F}_q^k$ de comprimento k. Este código será muito pouco eficiente, uma vez que a distância mínima entre palavras é igual a 1.

O Teorema 4.6 diz-nos que, se quisermos aumentar a eficiência deste código, teremos de aumentar a distância mínima entre as suas palavras. Como poderemos fazer isso? Muito simplesmente, acrescentando a cada palavra $a_1a_2\cdots a_k$ um bloco $c_{k+1}\cdots c_n \in \mathbb{F}_q^{n-k}$ tal que, sempre que

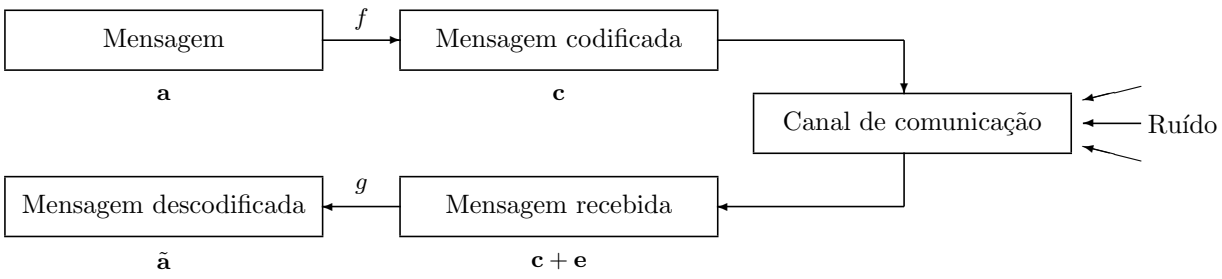
$$d(a_1a_2\cdots a_k, a'_1a'_2\cdots a'_k) = 1$$

então $d(c_{k+1}\cdots c_n, c'_{k+1}\cdots c'_n)$ é máxima, ou seja, igual a $n-k$. Se, além disso, tivermos o cuidado de garantir que $d(c_{k+1}\cdots c_n, c'_{k+1}\cdots c'_n) = n-k+1-i$ sempre que $d(a_1a_2\cdots a_k, a'_1a'_2\cdots a'_k) = i$, teremos um código $\mathcal{C}$ com distância mínima $\delta(\mathcal{C}) = n-k+1$.

Os primeiros k símbolos de cada palavra

$$\mathbf{c} = a_1a_2\cdots a_kc_{k+1}\cdots c_n$$

são a *mensagem original* e os $n-k$ símbolos adicionais são os *símbolos de controle*. A função $\mathbb{F}_q^k \to \mathbb{F}_q^n$ que aplica a palavra $a_1a_2\cdots a_k$ na palavra $a_1a_2\cdots a_kc_{k+1}\cdots c_n$ chama-se um *esquema de codificação*. Estes esquemas de codificação fazem parte de qualquer sistema de comunicação actual, que pode ser descrito do seguinte modo:



A função f é um esquema de codificação. À função $g : \mathbb{F}_q^n \to \mathbb{F}_q^k$ chama-se *esquema de descodificação*. Os esquemas de codificação podem ser apresentados do seguinte modo. Seja H uma matriz $(n-k) \times n$, com entradas em $\mathbb{F}_q$, do tipo $H = [A, I_{n-k}]$, onde A é uma matriz $(n-k) \times k$ e I_{n-k} é a matriz identidade de ordem $n-k$. Os símbolos de controle $c_{k+1}, \ldots, c_n$ podem então ser determinados a partir do sistema de equações $H\mathbf{c}^T = \mathbf{0}$, onde $\mathbf{0}$ denota o vector nulo de $\mathbb{F}_q^{n-k}$.

Exemplo: Seja H a seguinte matriz 3×7 sobre $\mathbb{F}_2$:

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

O código definido por H será constituído pelas palavras $\mathbf{c} = a_1a_2a_3a_4c_5c_6c_7$, onde os símbolos de controle c_5, c_6, c_7 podem ser calculados resolvendo o sistema $H\mathbf{c}^T = \mathbf{0}$, dados a_1, a_2, a_3, a_4:

$$\begin{cases} a_1 + a_3 + a_4 + c_5 = 0 \\ a_1 + a_2 + a_4 + c_6 = 0 \\ a_1 + a_2 + a_3 + c_7 = 0 \end{cases}$$

Portanto,

$$\begin{cases} c_5 = a_1 + a_3 + a_4 \\ c_6 = a_1 + a_2 + a_4 \\ c_7 = a_1 + a_2 + a_3 \end{cases}$$

pelo que $\mathbf{c} = (a_1, a_2, a_3, a_4, a_1 + a_3 + a_4, a_1 + a_2 + a_4, a_1 + a_2 + a_3)$. Assim, neste exemplo o esquema de codificação é a função linear de $\mathbb{F}_2^4$ em $\mathbb{F}_2^7$, definida por

$$(a_1, a_2, a_3, a_4) \mapsto (a_1, a_2, a_3, a_4, a_1 + a_3 + a_4, a_1 + a_2 + a_4, a_1 + a_2 + a_3),$$

e $\mathcal{C}$ é formado pelas 16 palavras

$$(a_1, a_2, a_3, a_4, a_1 + a_3 + a_4, a_1 + a_2 + a_4, a_1 + a_2 + a_3) \qquad a_1, a_2, a_3, a_4 \in \mathbb{F}_2.$$

Em geral, quando os esquemas de codificação são dados por aplicações lineares, usa-se a seguinte terminologia:

CÓDIGOS (n,k)-LINEARES

Seja $H = [A, I_{n-k}]$ uma matriz $(n-k) \times n$ com entradas em $\mathbb{F}_q$. O conjunto $\mathcal{C}$ dos vectores n-dimensionais $\mathbf{c} \in \mathbb{F}_q^n$ tais que $H\mathbf{c}^T = \mathbf{0}$ diz-se um *código (n,k)-linear* sobre $\mathbb{F}_q$. A matriz H diz-se a *matriz de controle* de $\mathcal{C}$. No caso $q = 2$, $\mathcal{C}$ diz-se um *código binário*.

[Note que o conjunto $\mathcal{C}$ das soluções do sistema $H\mathbf{c}^T = \mathbf{0}$ de equações lineares é um subespaço de dimensão k do espaço vectorial $\mathbb{F}_q^n$]

Exemplos: Os códigos $\mathcal{C}_2$ e $\mathcal{C}_3$ são exemplos de códigos lineares. O código $\mathcal{C}_2$ é um código $(4,2)$-linear sobre $\mathbb{F}_2$, com matriz de controle

$$H_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},$$

e $\mathcal{C}_3$ é um código $(6,2)$-linear sobre $\mathbb{F}_2$, com matriz de controle

$$H_3 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Teorema 4.7 *Um código (n,k)-linear com matriz de controle H tem distância mínima $\delta(\mathcal{C}) = s$ se e só se quaisquer $s-1$ colunas de H são linearmente independentes e existem s colunas de H que são linearmente dependentes.*

Demonstração. Por definição,

$$\delta(\mathcal{C}) < s \Leftrightarrow \exists\, \mathbf{c}, \mathbf{d} \in \mathcal{C}, \mathbf{c} \neq \mathbf{d}, d(\mathbf{c}, \mathbf{d}) = t < s.$$

Como $\mathcal{C}$ é linear, $\mathbf{e} = \mathbf{c} - \mathbf{d} \in \mathcal{C}$ e, obviamente, $d(\mathbf{c}, \mathbf{d}) = d(\mathbf{e}, \mathbf{0})$. Portanto,

$$\delta(\mathcal{C}) < s \Leftrightarrow \exists\, \mathbf{e} \in \mathcal{C}, \mathbf{e} \neq \mathbf{0}, d(\mathbf{e}, \mathbf{0}) = t < s. \qquad (*)$$

Sejam $e_{i_1}, e_{i_2}, \ldots, e_{i_t}$ as t $(t < s)$ coordenadas (letras) da palavra $\mathbf{e}$ que não são nulas, isto é, $\mathbf{e} = (0, \ldots, 0, e_{i_1}, 0, \ldots, 0, e_{i_2}, 0, \ldots, 0, e_{i_t}, 0, \ldots, 0)$. Denotando por H_i a i-ésima coluna de H, a condição $\mathbf{e} \in \mathcal{C}$ significa que $H\mathbf{e}^T = \mathbf{0}$, ou seja,

$$H_{i_1} e_{i_1} + H_{i_2} e_{i_2} + \cdots + H_{i_t} e_{i_t} = \mathbf{0},$$

o que mostra que as t colunas $H_{i_1}, H_{i_2}, \ldots, H_{i_t}$ de H são linearmente dependentes.

Portanto, a condição $(*)$ significa que existem $t \leq s-1$ colunas em H que são linearmente dependentes e, por maioria de razão, existem $s-1$ colunas de H linearmente dependentes. Provámos assim que

- $\delta(\mathcal{C}) < s$ se e só se existem $s-1$ colunas de H linearmente dependentes,

o que é evidentemente equivalente a dizer que

- $\delta(\mathcal{C}) \geq s$ se e só se quaisquer $s-1$ colunas de H são linearmente independentes.

Concluindo: $\delta(\mathcal{C}) = s$, isto é, $\delta(\mathcal{C}) \geq s$ e $\delta(\mathcal{C}) < s+1$, se e só se quaisquer $s-1$ colunas de H são linearmente independentes e existem s colunas de H que são linearmente dependentes. ∎

Exemplos: Na matriz H_2 acima, $s = 1$, uma vez que há duas colunas linearmente dependentes (a primeira e a terceira, por exemplo). Na matriz H_3, quaisquer duas colunas são linearmente independentes mas as colunas 1, 3 e 5 são linearmente dependentes, pelo que $s = 2$.

Vimos já que, depois de recebida uma palavra $\mathbf{y}$ pelo receptor, a sua *descodificação*, isto é, a determinação da palavra exacta $\mathbf{c}$ que lhe deu origem (isto é, a palavra enviada pelo emissor), pode ser feita determinando a palavra de $\mathcal{C}$ que está mais próxima de $\mathbf{y}$ (princípio do vizinho mais próximo). Claro que isto pode ser feito por "força bruta", determinando a distância de Hamming entre $\mathbf{y}$ e todas as palavras de $\mathcal{C}$. Mas isto é impraticável quando $|\mathcal{C}|$ é muito grande!

Em vez da força bruta, pode usar-se uma abordagem através da matriz H. Para isso, consideremos o espaço vectorial $\mathbb{F}_q^n/\mathcal{C}$ formado por todas as classes

$$\mathbf{a} + \mathcal{C} := \{\mathbf{a} + \mathbf{c} : \mathbf{c} \in \mathcal{C}\}$$

com $\mathbf{a} \in \mathbb{F}_q^n$. Cada classe contém q^k palavras e $\mathbb{F}_q^n$ pode particionar-se em $l+1 = q^{n-k}$ classes de $\mathcal{C}$:

$$\mathbb{F}_q^n = (\mathbf{0} + \mathcal{C}) \cup (\mathbf{a}^{(1)} + \mathcal{C}) \cup \cdots \cup (\mathbf{a}^{(l)} + \mathcal{C}).$$

A palavra recebida $\mathbf{y}$ tem que estar nalguma das classes, digamos $\mathbf{a}^{(i)} + \mathcal{C}$, pelo que $\mathbf{y} = \mathbf{a}^{(i)} + \mathbf{d}$ para algum $\mathbf{d} \in \mathcal{C}$. Se $\mathbf{c}$ foi a palavra transmitida, então o erro é dado por $\mathbf{e} = \mathbf{y} - \mathbf{c} = \mathbf{a}^{(i)} + \mathbf{d} - \mathbf{c} \in \mathbf{a}^{(i)} + \mathcal{C}$. Portanto, o erro $\mathbf{e}$ pertence à mesma classe da palavra $\mathbf{y}$ recebida. Assim, pelo princípio do vizinho mais próximo, para determinar o erro $\mathbf{e}$, e consequentemente a palavra original $\mathbf{y} - \mathbf{e}$, bastará determinar o *líder* da classe de $\mathbf{y}$:

PESO DE UMA PALAVRA; LÍDER DE UMA CLASSE

O *peso* (de Hamming) de $\mathbf{c} \in \mathbb{F}_q^n$ é o número de coordenadas não-nulas de $\mathbf{c}$. Por outras palavras, o peso de $\mathbf{c} \in \mathbb{F}_q^n$ é a distância $d(\mathbf{c}, \mathbf{0})$.

Um elemento de peso mínimo numa classe $\mathbf{a} + \mathcal{C}$ chama-se *líder* de $\mathbf{a} + \mathcal{C}$.

É claro que se houver mais do que um líder na classe de $\mathbf{y}$ o erro não poderá ser corrigido, uma vez que o receptor não conseguirá decidir qual dos líderes será o vector erro $\mathbf{e}$. Por exemplo, no código $(4, 2)$-linear binário $\mathcal{C}$ com matriz de controle

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

a lista das 4 classes deste código é a seguinte:

$$\begin{array}{rl} \text{classe } \mathbf{0}+\mathcal{C}=\mathcal{C}: & 0000 \quad 1010 \quad 0111 \quad 1101 \\ \text{outras classes:} & \left\{ \begin{array}{llll} 1000 & 0010 & 1111 & 0101 \\ 0100 & 1110 & 0011 & 1001 \\ 0001 & 1011 & 0110 & 1100 \end{array} \right. \end{array}$$

A classe na segunda linha tem dois líderes: 1000 e 0010. Por exemplo, se a palavra recebida for a palavra $\mathbf{y} = 1111$ que está na segunda classe, o vector erro tanto pode ser 1000 como 0010, ou seja, a palavra original pode bem ter sido a palavra 0111 ou 1101. Isto acontece porque $\delta(\mathcal{C}) = 2$ e, portanto, o código não corrige todos os erros singulares. Se a palavra $\mathbf{y}$ recebida for a palavra 1110 na terceira classe, o erro só poderá ser igual a 0100 e, portanto, o receptor descobre imediatamente o erro: a palavra original só pode ter sido a palavra 1010.

[Se no canal de comunicação só ocorrerem no máximo t erros e $\delta(\mathcal{C}) \geq 2t+1$ (portanto $\mathcal{C}$ corrige sempre os t eventuais erros), não poderão existir dois líderes $\mathbf{e}_1$ e $\mathbf{e}_2$ na mesma classe; de facto, se tal fosse possível, $\mathbf{c}_1 := \mathbf{y} - \mathbf{e}_1$ e $\mathbf{c}_2 := \mathbf{y} - \mathbf{e}_2$ seriam palavras de $\mathcal{C}$ tais que $d(\mathbf{c}_1, \mathbf{c}_2) \leq d(\mathbf{c}_1, \mathbf{y}) + d(\mathbf{y}, \mathbf{c}_2) = d(\mathbf{e}_1, \mathbf{0}) + d(\mathbf{e}_2, \mathbf{0}) \leq t + t$, uma contradição com o facto $\delta(\mathcal{C}) \geq 2t+1$]

A classe de cada $\mathbf{y}$ pode ser determinada calculando a sua síndrome:

SÍNDROME DE UMA PALAVRA

O vector $S(\mathbf{c}) = H\mathbf{c}^T$ de comprimento $n-k$ chama-se a *síndrome* de $\mathbf{c} \in \mathbb{F}_q^n$.

Proposição 4.8 (1) $S(\mathbf{c}) = \mathbf{0}$ *se e só se* $\mathbf{c} \in \mathcal{C}$.

(2) $S(\mathbf{c}) = S(\mathbf{d})$ *se e só se* $\mathbf{c}+\mathcal{C} = \mathbf{d}+\mathcal{C}$.

Demonstração. (1) É imediato da definição de $\mathcal{C}$ em termos de H.

(2) $S(\mathbf{c}) = S(\mathbf{d}) \Leftrightarrow H\mathbf{c}^T = H\mathbf{d}^T \Leftrightarrow H(\mathbf{c}-\mathbf{d})^T = \mathbf{0} \Leftrightarrow \mathbf{c}-\mathbf{d} \in \mathcal{C} \Leftrightarrow \mathbf{c}+\mathcal{C} = \mathbf{d}+\mathcal{C}$.

■

No exemplo anterior,

$$\begin{array}{lll} \text{palavras de } \mathcal{C}: & 0000 \quad 1010 \quad 0111 \quad 1101 & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \text{outras classes:} & \left\{ \begin{array}{llll} 1000 & 0010 & 1111 & 0101 \\ 0100 & 1110 & 0011 & 1001 \\ 0001 & 1011 & 0110 & 1100 \end{array} \right. & \begin{array}{l} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\text{Síndromes}} \end{array} \end{array}$$

ALGORITMO DE DESCODIFICAÇÃO
Dados: palavra $\mathbf{y}$ recebida.

(1) Calcular $S(\mathbf{y})$.

(2) Determinar o líder $\mathbf{e}$ tal que $S(\mathbf{e}) = S(\mathbf{y})$.

(3) A palavra original é a palavra $\mathbf{c} = \mathbf{y} - \mathbf{e}$.

Exemplo: Consideremos o código do exemplo anterior. Se $\mathbf{y} = 1110$ é recebida, começamos por determinar $S(\mathbf{y}) = H\mathbf{y}^T = \binom{1}{1}$. O erro $\mathbf{e}$ será então igual ao líder da respectiva classe, ou seja, a 0100. A palavra original era então igual a $\mathbf{y} - \mathbf{e} = 1010$.

Em códigos lineares muito grandes é praticamente impossível listar todas as classes e determinar os respectivos líderes; por exemplo, um código $(50, 20)$-linear binário tem aproximadamente 10^9 classes. Nesse caso, determina-se directamente o líder da classe da palavra $\mathbf{y}$, determinando a palavra $\mathbf{e}$ de menor peso tal que $H\mathbf{e}^T = S(\mathbf{y})$. No exemplo acima,

$$H\mathbf{e}^T = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Leftrightarrow \begin{cases} e_1 + e_3 = e_4 \\ e_2 = 1 + e_4 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \mathbf{e} = (0100) \vee \mathbf{e} = (1110) \vee \mathbf{e} = (0011) \vee \mathbf{e} = (1001).$$

O vector (0100) é o que tem menor peso, pelo que $\mathbf{e} = (0100)$.

Já vimos maneiras de codificar mensagens de modo a que, no caso de ocorrerem alguns erros na sua transmissão, o receptor possa ser capaz de corrigir esses erros. Esses códigos, chamados códigos lineares (ou códigos de Hamming), baseavam-se em definir as palavras codificadas como vectores de soluções em $\mathbb{F}_q$ de sistemas de equações lineares.

Terminamos com exemplos de outra classe de códigos, os chamados códigos BCH, descobertos em 1960 por Bose, Chaudhuri e Hocquenghem. As palavras destes códigos serão vectores definidos pelos coeficientes de polinómios em $\mathbb{F}_q[x]$. Estes polinómios terão como raízes certas potências de um *elemento primitivo* de alguma extensão apropriada do corpo $\mathbb{F}_q$.

Comecemos com um exemplo que usa o corpo $\mathbb{F}_8$ com 8 elementos. Este corpo pode obter-se como extensão de $\mathbb{F}_2[x]$, de modo análogo aos Exemplos da página 85. Com efeito, seja

$$m(x) = x^3 + x + 1 \in \mathbb{F}_2[x].$$

É fácil ver que se trata de um polinómio irredutível sobre $\mathbb{F}_2$, pelo que o quociente $\mathbb{F}_2[x]/(m(x))$ é uma extensão de $\mathbb{F}_2$ com 8 elementos:

$$\begin{aligned}\frac{\mathbb{Z}_2[x]}{(m(x))} &= \{a_0 + a_1x + a_2x + (p(x)) \mid a_0, a_1, a_2 \in \mathbb{Z}_2\} \\ &= \Big\{0 + (m(x)), 1 + (m(x)), x + (m(x)), x + 1 + (m(x)), x^2 + (m(x)), \\ &\qquad x^2 + 1 + (m(x)), x^2 + x + (m(x)), x^2 + x + 1 + (m(x))\Big\}\end{aligned}$$

Denotando estes elementos por, respectivamente, $0, 1, \alpha, \beta, \gamma, \delta, \varepsilon, \varphi$, as tabelas das operações deste corpo são as seguintes:

$+$	0	1	α	β	γ	δ	ε	φ
0	0	1	α	β	γ	δ	ε	φ
1	1	0	β	α	δ	γ	φ	ε
α	α	β	0	1	ε	φ	γ	δ
β	β	α	1	0	φ	ε	δ	γ
γ	γ	δ	ε	φ	0	1	α	β
δ	δ	γ	φ	ε	1	0	β	α
ε	ε	φ	γ	δ	α	β	0	1
φ	φ	ε	δ	γ	β	α	1	0

$\cdot$	0	1	α	β	γ	δ	ε	φ
0	0	0	0	0	0	0	0	0
1	0	1	α	β	γ	δ	ε	φ
α	0	α	γ	ε	β	1	φ	δ
β	0	β	ε	δ	φ	γ	1	α
γ	0	γ	β	φ	ε	α	δ	1
δ	0	δ	1	γ	α	φ	β	ε
ε	0	ε	φ	1	δ	β	α	γ
φ	0	φ	δ	α	1	ε	γ	β

Neste corpo já o polinómio $m(x)$ tem uma raiz (que é o elemento α). Observe que todos os seus elementos podem ser vistos como polinómios em α, onde $\alpha^3+\alpha+1 =$

0, e que α é um *elemento primitivo* de $\mathbb{F}_8$, isto é, α é um gerador do grupo multiplicativo $(\mathbb{F}_8 \setminus \{0\}, \cdot)$:

0	0	0
1	1	1
α	α	α
β	$\alpha + 1$	α^3
γ	α^2	α^2
δ	$\alpha^2 + 1$	α^6
ε	$\alpha^2 + \alpha$	α^4
φ	$\alpha^2 + \alpha + 1$	α^5

[Pode provar-se que, em qualquer corpo finito $\mathbb{F}_q$, o grupo multiplicativo $(\mathbb{F}_q \setminus \{0\}, \cdot)$ é cíclico. Consulte a bibliografia]

As duas colunas mais à direita desta tabela retêm toda a informação sobre as operações do corpo. Esta é a maneira mais eficiente de trabalhar neste corpo: os seus elementos são potências de α, donde a multiplicação passa a ser imediata (basta reter que $\alpha^7 = 1$)

$\cdot$	0	1	α	α^2	α^3	α^4	α^5	α^6
0	0	0	0	0	0	0	0	0
1	0	1	α	α^2	α^3	α^4	α^5	α^6
α	0	α	α^2	α^3	α^4	α^5	α^6	1
α^2	0	α^2	α^3	α^4	α^5	α^6	1	α
α^3	0	α^3	α^4	α^5	α^6	1	α	α^2
α^4	0	α^4	α^5	α^6	1	α	α^2	α^3
α^5	0	α^5	α^6	1	α	α^2	α^3	α^4
α^6	0	α^6	1	α	α^2	α^3	α^4	α^5

enquanto a adição é simplesmente igual a

$+$	0	1	α	α^2	α^3	α^4	α^5	α^6
0	0	1	α	α^2	α^3	α^4	α^5	α^6
1	1	0	α^3	α^6	α	α^5	α^4	α^2
α	α	α^3	0	α^4	1	α^2	α^6	α^5
α^2	α^2	α^6	α^4	0	α^5	α	α^3	1
α^3	α^3	α	1	α^5	0	α^6	α^2	α^4
α^4	α^4	α^5	α^2	α	α^6	0	1	α^3
α^5	α^5	α^4	α^6	α^3	α^2	1	0	α
α^6	α^6	α^2	α^5	1	α^4	α^3	α	0

Vamos agora construir um código usando este corpo, do seguinte modo:

Seja $(a, b, c, d) \in \mathbb{F}_2^4$ uma palavra que queremos transmitir. Formemos o respectivo polinómio

$$p_{\mathcal{C}}(x) = ax^6 + bx^5 + cx^4 + dx^3.$$

Dividindo $p_{\mathcal{C}}(x)$ por $m(x)$ (em $\mathbb{F}_2[x]$) obtemos $p_{\mathcal{C}}(x) = q(x)m(x) + r_{\mathcal{C}}(x)$, onde o resto $r_{\mathcal{C}}(x)$ tem grau inferior a 3, isto é, $r_{\mathcal{C}}(x) = rx^2+sx+t$ para alguns $r, s, t \in \mathbb{F}_2$. Então

$$\begin{aligned} q(x)m(x) &= p_{\mathcal{C}}(x) - r_{\mathcal{C}}(x) \\ &= p_{\mathcal{C}}(x) + r_{\mathcal{C}}(x) \\ &= ax^6 + bx^5 + cx^4 + dx^3 + rx^2 + sx + t. \end{aligned}$$

Este polinómio, que denotaremos por $p(x)$, quando calculado em α, uma raiz de $m(x)$, dá $p(\alpha) = m(\alpha)q(\alpha) = 0$. Codificaremos a palavra inicial (a, b, c, d) pelo vector $(a, b, c, d, r, s, t) \in \mathbb{F}_2^7$ definido pelos coeficientes de $p(x)$. Este vector tem 4 dígitos de informação e 3 dígitos de controle e é caracterizado pela seguinte propriedade:

> *Corresponde ao único polinómio de grau inferior a 7 com coeficientes de maior grau a, b, c, d e tendo α por raiz.*

Na descodificação, quando o receptor recebe a palavra (A, B, C, D, R, S, T), forma o polinómio

$$r(x) = Ax^6 + Bx^5 + Cx^4 + Dx^3 + Rx^2 + Sx + T.$$

Suponhamos que aconteceu no máximo um erro singular. Então o erro

$$e(x) = p(x) - r(x)$$

é o polinómio nulo ou consiste num único termo x^e (onde $e \in \{6,5,4,3,2,1,0\}$ corresponde ao coeficiente onde aconteceu o erro):

$$e(x) = \begin{cases} 0 & \text{se não ocorreram erros} \\ \\ x^e & \text{se ocorreu um erro na posição } e. \end{cases}$$

Por exemplo, se o erro aconteceu no coeficiente c, ou seja, $C \neq c$, então $e(x) = (c - C)x^4 = x^4$. Para detectar e corrigir o erro basta ao receptor calcular $r(\alpha)$:

- *Caso 1*: Se $r(\alpha) = 0$, então, como $p(\alpha) = 0$, $e(\alpha) = 0$. Como $\mathcal{O}(\alpha) = 7$, $e(x)$ só pode ser o polinómio nulo e não ocorreram erros.
- *Caso 2*: Se $r(\alpha) \neq 0$, então, como $p(\alpha) = 0$, $e(\alpha) \neq 0$. Portanto, $e(x) = Ex^e$, donde $E\alpha^e = e(\alpha) = r(\alpha)$. O receptor pode assim descobrir o valor de e onde aconteceu o erro e corrigir automaticamente o erro.

Portanto, calculando $r(x)$ em α, podemos determinar se ocorreu algum erro e, em caso afirmativo, corrigi-lo.

`[Pode provar-se que este código tem distância mínima igual a 3, pelo que corrige erros singulares]`

Exemplo: Para codificar a palavra $(1,1,0,1)$ tomemos o polinómio $p_{\mathcal{C}}(x) = x^6 + x^5 + x^3$ e dividamo-lo por $m(x) = x^3 + x + 1$:

$$x^6 + x^5 + x^3 = (x^3 + x^2 + x + 1)(x^3 + x + 1) + 1.$$

Como o resto $r_{\mathcal{C}}(x)$ é igual a 1, temos $p(x) = x^6 + x^5 + x^3 + 1$. (Note que $p(\alpha) = \alpha^6 + \alpha^5 + \alpha^3 + 1 = (\alpha^2 + 1) + (\alpha^2 + \alpha + 1) + (\alpha + 1) + 1 = 0$.) A palavra do código deverá ser então igual a $(1,1,0,1,0,0,1)$. Se o receptor receber a palavra $(1,1,0,1,\underline{1},0,1)$, considera o polinómio $r(x) = x^6 + x^5 + x^3 + x^2 + 1$ e, usando o quadro da página 138, calcula $r(\alpha)$:

$$\begin{array}{rcccccccc} 1 & = & & & & & & & 1 \\ +\alpha^2 & = & \alpha^2 & & & & & & \\ +\alpha^3 & = & & & & \alpha & + & 1 \\ +\alpha^5 & = & \alpha^2 & + & \alpha & + & 1 \\ +\alpha^6 & = & \alpha^2 & & & & + & 1 \\ \hline r(\alpha) & = & \alpha^2. & & & & & & \end{array}$$

Assim, detecta que ocorreu um erro no coeficiente de x^2 e conclui que a palavra correcta é igual a $(1,1,0,1,\underline{0},0,1)$.

Se o receptor receber a palavra $(1,1,\underline{1},1,0,0,1)$, considera o polinómio $r(x) = x^6 + x^5 + x^4 + x^3 + 1$ e calcula $r(\alpha)$:

$$\begin{array}{rcccccc}
1 & = & & & & & 1 \\
+\alpha^3 & = & & & \alpha & + & 1 \\
+\alpha^4 & = & \alpha^2 & + & \alpha & & \\
+\alpha^5 & = & \alpha^2 & + & \alpha & + & 1 \\
+\alpha^6 & = & \alpha^2 & & & + & 1 \\
\hline
r(\alpha) & = & \alpha^2 & + & \alpha & & \\
 & = & \alpha^4. & & & &
\end{array}$$

Assim, detecta que ocorreu um erro no coeficiente de x^4 e conclui que a palavra correcta é igual a $(1, 1, \underline{0}, 1, 0, 0, 1)$.

Vamos apresentar agora um código deste tipo que detecte erros duplos. Para isso precisamos de um corpo maior (o corpo $\mathbb{F}_{16}$ descrito na página 86). Neste corpo, o elemento g é um elemento primitivo ($g^2 = i$, $g^3 = e$, $g^4 = h$, $g^5 = \alpha$, $g^6 = k$, $g^7 = n$, $g^8 = j$, $g^9 = m$, $g^{10} = \beta$, $g^{11} = c$, $g^{12} = d$, $g^{13} = l$, $g^{14} = f$ e $g^{15} = 1$) que é raiz do polinómio $m(x) = x^4 + x + 1$, irredutível sobre $\mathbb{F}_2$. Portanto $\mathbb{F}_{16}$ pode obter-se como extensão de $\mathbb{F}_2$, através do quociente $\mathbb{F}_2[x]/(m(x))$, e podemos olhar todos os seus elementos não nulos como potências de g (onde $g^{15} = 1$). Uma vez que $m(g) = g^4 + g + 1 = 0$, todo o elemento deste corpo pode exprimir-se como polinómio em g de grau inferior a 4:

0	0	0
1	1	1
g	g	g
i	g^2	g^2
e	g^3	g^3
h	g^4	$g+1$
α	g^5	g^2+g
k	g^6	g^3+g^2
n	g^7	g^3+g+1
j	g^8	g^2+1
m	g^9	g^3+g
β	g^{10}	g^2+g+1
c	g^{11}	g^3+g^2+g
d	g^{12}	g^3+g^2+g+1
l	g^{13}	g^3+g^2+1
f	g^{14}	g^3+1
1	g^{15}	

A ideia para este código é utilizar palavras de comprimento 15 construídas com os coeficientes dos polinómios de grau 14 em $\mathbb{F}_2[x]$ que têm g e g^3 como raízes. Já sabemos que $m(x) = x^4 + x + 1$ é o polinómio mínimo de g sobre $\mathbb{F}_2$. Por outro lado, é fácil provar que $m_3(x) = x^4 + x^3 + x^2 + x + 1$ é o polinómio mínimo de g^3. Então o polinómio $m_{13}(x)$ de menor grau que tem simultaneamente g e g^3 como raízes é o menor múltiplo comum de $m(x)$ e $m_3(x)$; como são ambos irredutíveis,

$$m_{13}(x) = m(x)m_3(x) = x^8 + x^7 + x^6 + x^4 + 1.$$

Como se trata de um polinómio de grau 8, as palavras do código terão comprimento 15, com 7 dígitos de informação e 8 dígitos de controle. Sendo $(a_{14}, a_{13}, \ldots, a_8)$ a palavra com a informação a transmitir, calculamos a respectiva palavra do código do seguinte modo:

Seja $p_{\mathcal{C}}(x) = a_{14}x^{14} + a_{13}x^{13} + \cdots + a_8x^8$. Dividimos $p_{\mathcal{C}}(x)$ por $m_{13}(x)$ (em $\mathbb{F}_2[x]$):

$$p_{\mathcal{C}}(x) = q(x)m_{13}(x) + r_{\mathcal{C}}(x),$$

onde o resto $r_{\mathcal{C}}(x)$ tem grau inferior a 8, isto é, $r_{\mathcal{C}}(x) = a_7x^7 + a_6x^6 + \cdots + a_1x + a_0$. Então

$$\begin{aligned} q(x)m_{13}(x) &= p_{\mathcal{C}}(x) - r_{\mathcal{C}}(x) \\ &= p_{\mathcal{C}}(x) + r_{\mathcal{C}}(x) \\ &= a_{14}x^{14} + a_{13}x^{13} + \cdots + a_1x + a_0. \end{aligned}$$

Este polinómio, que denotaremos por $p(x)$, quando calculado em g e g^3, raízes de $m_{13}(x)$, dá $p(g) = m_{13}(g)q(g) = 0$. Codificaremos a palavra inicial $(a_{14}, a_{13}, \ldots, a_8)$ pelo vector $(a_{14}, a_{13}, \ldots, a_0) \in \mathbb{F}_2^{15}$ definido pelos coeficientes de $p(x)$. Este vector tem 7 dígitos de informação e 8 dígitos de controle e é caracterizado pela seguinte propriedade:

> *Corresponde ao único polinómio de grau inferior a 15 com coeficientes de maior grau $a_{14}, \ldots, a_8$ e tendo g e g^3 como raízes.*

Na descodificação, quando o receptor recebe a palavra $(A_{14}, A_{13}, \ldots, A_0)$, forma o polinómio

$$r(x) = A_{14}x^{14} + A_{13}x^{13} + \cdots + A_1x + A_0.$$

Suponhamos que no canal de comunicação ocorrem, quando muito, erros duplos. Então o vector erro $e(x) = p(x) - r(x)$ é o polinómio nulo, ou consiste num único termo x^e (onde $e \in \{14, 13, \ldots, 1, 0\}$ corresponde ao coeficiente onde ocorreu o erro), ou consiste na soma de dois termos $x^{e_1} + x^{e_2}$ (onde $e_1, e_2 \in \{14, 13, \ldots, 1, 0\}$

correspondem aos coeficientes onde ocorreram os dois erros):

$$e(x) = \begin{cases} 0 & \text{se não ocorreram erros} \\ x^e & \text{se ocorreu um erro na posição } e \\ x^{e_1} + x^{e_2} & \text{se ocorreram erros nas posições } e_1 \text{ e } e_2. \end{cases}$$

Como $m_{13}(x)$ divide $p(x)$, temos:

- $r(g) = e(g)$, porque $m_{13}(g) = 0$;
- $r(g^2) = e(g^2)$, porque $m_{13}(g) = 0$ (logo $m_{13}(g^2) = (m_{13}(g))^2 = 0$);

  [Exercício: Prove, usando o Teorema Binomial e indução sobre o grau, que qualquer polinómio $p(x)$ em $\mathbb{F}_2[x]$ satisfaz a propriedade $(p(x))^2 = p(x^2)$]

- $r(g^3) = e(g^3)$, porque $m_{13}(g^3) = 0$.

Consideremos o polinómio

$$P(x) = r(g)x^2 + r(g^2)x + (r(g^3) + r(g)r(g^2))).$$

- *Caso 1*: Se $e(x) = 0$, então $e(g) = e(g^2) = e(g^3) = 0$; consequentemente, $r(g) = r(g^2) = r(g^3) = 0$ e $P(x) = 0$.
- *Caso 2*: Se $e(x) = x^e$, então

$$P(x) = g^e x^2 + g^{2e} x + (g^{3e} + g^{2e} g) = g^e x(x + g^e).$$

- *Caso 3*: Se $e(x) = x^{e_1} + x^{e_2}$, então

$$\begin{aligned} P(x) &= (g^{e_1} + g^{e_2})x^2 + (g^{2e_1} + g^{2e_2})x + (g^{3e_1} + g^{3e_2}) + (g^{2e_1} + g^{2e_2})(g^{e_1} + g^{e_2}) \\ &= (g^{e_1} + g^{e_2})[x^2 + (g^{e_1} + g^{e_2})x + g^{e_1} g^{e_2}] \\ &= (g^{e_1} + g^{e_2})[(x + g^{e_1})(x + g^{e_2})]. \end{aligned}$$

Isto mostra que, se há raízes de $P(x)$, estas são necessariamente potências de g, cujo expoente indica a posição onde ocorreram os erros. O receptor pode assim descobrir o(s) valor(es) de e (e_1 e e_2) e corrigir automaticamente o(s) erro(s). Só tem que calcular $P(x)$ e determinar as suas raízes.

Exemplo: Suponhamos que pretendemos enviar os dígitos de informação 1101101. Para isso consideramos o polinómio $p_{\mathcal{C}}(x) = x^{14} + x^{13} + x^{11} + x^{10} + x^8$ e dividimo-lo por $m_{13}(x) = x^8 + x^7 + x^6 + x^4 + 1$:

$$p_{\mathcal{C}}(x) = (x^6 + x^4 + x^2 + x)m_{13}(x) + (x^7 + x^5 + x^4 + x^2 + x).$$

Portanto, os dígitos de controle da palavra a enviar são 10110110, ou seja, a palavra codificada a enviar é a palavra

$$(1,1,0,1,1,0,1,1,0,1,1,0,1,1,0).$$

Suponhamos que o receptor recebe

$$(1,1,0,1,1,\underline{1},1,\underline{0},0,1,1,0,1,1,0).$$

Então $r(x) = x^{14} + x^{13} + x^{11} + x^{10} + x^9 + x^8 + x^5 + x^4 + x^2 + x$, donde:

$$\begin{aligned}
r(g) &= g^{14} + g^{13} + g^{11} + g^{10} + g^9 + g^8 + g^5 + g^4 + g^2 + g; \\
r(g^2) &= (r(g))^2 \quad \text{(porque o corpo tem característica 2);} \\
r(g^3) &= g^{42} + g^{39} + g^{33} + g^{30} + g^{27} + g^{24} + g^{15} + g^{12} + g^6 + g^3 \\
&= g^{12} + g^9 + g^3 + 1 + g^{12} + g^9 + 1 + g^{12} + g^6 + g^3 \quad \text{(pois } g^{15} = 1\text{)} \\
&= g^{12} + g^6.
\end{aligned}$$

Usando a tabela da página 141, substituímos todos estes termos por polinómios em g de grau inferior a 4. Por exemplo, em $r(g)$:

Coeficientes de	g^3	g^2	g	1
g^{14}	1			1
g^{13}	1	1		1
g^{11}	1	1	1	
g^{10}		1	1	1
g^9	1		1	
g^8		1		1
g^5		1	1	
g^4			1	1
g^2		1		
g			1	
$r(g)$	0	0	0	1

Assim, $r(g) = 1$. Então $r(g^2) = r(g)^2 = 1$. Por outro lado,

$$r(g^3) = g^{12} + g^6 = (g^3 + g^2 + g + 1) + (g^3 + g^2) = g + 1.$$

Portanto,

$$P(x) = x^2 + x + \left(\frac{g+1}{1} + 1\right) = x^2 + x + g.$$

Para determinar as raízes de $P(x)$ podemos testar todas as hipóteses, usando a tabela da página 141 para exprimir tudo em termos de 1, g, g^2, g^3:

x	x^2	x (pela Tabela p. 141)	x^2	x^2+x+g
0	0	0	0	g
1	1	1	1	g
g	g^2	g	g^2	g^2
g^2	g^4	g	$g+1$	$g+1$
g^3	g^6	g^3	g^3+g^2	g^2+g
g^4	g^8	$g+1$	g^2+1	g^2
g^5	g^{10}	g^2+g	g^2+g+1	$g+1$
g^6	g^{12}	g^3+g^2	g^3+g^2+g+1	1
g^7	g^{14}	g^3+g+1	g^3+1	0

Paramos em g^7 porque se trata de uma raiz. Então $P(x)=(x+g^7)(x+g^{e_1})$ para algum e_1, pelo que $g^7 g^{e_1}=g=g^{16}$, isto é, $e_1=9$. Em conclusão,

$$P(x)=(x+g^9)(x+g^7).$$

Isto significa que os erros ocorreram nas posições de x^9 e x^7.

[São códigos deste tipo que são utilizados na gravação da informação nos discos áudio CD. Mais concretamente, utilizam-se dois códigos sobre o corpo $\mathbb{F}_{256}=\mathbb{F}_{2^8}$, com palavras de comprimento $n=255$. Habitualmente escolhe-se o elemento primitivo α que tem o polinómio mínimo $m(x)=x^8+x^4+x^3+x^2+1$. Estes códigos têm distância mínima igual a 5. Para mais informação, consulte *Error correction and compact discs*, D. Dorninger e H. Kaiser, UMAP Journal 21 (2) (2000) 139-156]

[É possível formalizar estes códigos de modo geral sobre um corpo qualquer $\mathbb{F}_q$ e determinar a sua eficiência na correcção de erros]

Exercícios

4.1. Pode existir um corpo com 6 elementos? E com 12 elementos? Quanto vale $1+1$ num corpo com 64 elementos?

4.2. Seja $K=\{0,1,\alpha,\beta\}$ um corpo. Quanto valem $1+1$, $\alpha+\alpha$, $\beta+\beta$, $\alpha+1$, $\beta+1$, α^2, β^2 e $\alpha\cdot\beta$? Construa as tabelas da adição e da multiplicação em K.

4.3. Seja F a extensão de decomposição de $x^2-2\in\mathbb{Z}_3[x]$.

(a) Descreva o corpo F e indique um gerador de $F^*=F\setminus\{0\}$.

(b) Qual é o subcorpo primo de F?

4.4. Seja F a extensão de decomposição de $f(x) = x^{p^n} - x$ sobre $\mathbb{F}_p$.

(a) Mostre que o conjunto $R = \{a \in F \mid a^{p^n} = a\}$ das raízes de $f(x)$ é um subcorpo de F.

(b) Prove directamente, a partir da definição de raiz dupla, que todas as raízes de $f(x)$ são simples.

(c) Conclua que $R = F$.

4.5. Seja F um corpo com 81 elementos.

(a) Determine a característica de F, indique o seu corpo primo $\mathbb{F}_p$ e determine $[F : \mathbb{F}_p]$.

(b) Justifique a afirmação "*o único subcorpo próprio de F é o seu subcorpo primo*".

4.6. Construa um corpo finito de ordem 16 e determine todos os geradores do seu grupo multiplicativo.

4.7. Construa um corpo com 27 elementos.

4.8. Indique, justificando, o número de corpos não isomorfos de ordem inferior a 100.

4.9. Determine todos os subcorpos de um corpo com 32 e 64 elementos, respectivamente.

4.10. Liste os subcorpos do corpo $\mathbb{F}_{256}$. Qual deles é o subcorpo primo?

4.11. Usando resultados sobre corpos finitos, mostre que se p é um número primo e r divide n, então $p^r - 1$ divide $p^n - 1$.

4.12. Determine o número de elementos do corpo $\mathbb{F}_{11}[x]/\langle x^2 + 1\rangle$.

4.13. Mostre que:

(a) O corpo $\mathbb{F}_{11}[x]/\langle x^2 + x + 4\rangle$ é isomorfo a $\mathbb{F}_{11}[x]/\langle x^2 + 1\rangle$.

(b) A soma de todos os elementos de um corpo finito, com a excepção de $\mathbb{F}_2$, é 0.

4.14. Mostre que num código binário linear, ou todas as palavras têm peso par, ou metade das palavras tem peso par e metade tem peso ímpar.

4.15. Através de um comando à distância de uma televisão podem ser efectuadas 20 operações: escolher entre 18 canais diferentes (0–17), aumentar (A) ou diminuir (D) o volume. A tabela indica três códigos decimais para transmitir essa informação.

	0	1	2	⋯	9	10	11	⋯	17	A	D
C_1	00	01	02	⋯	09	10	11	⋯	17	18	19
C_2	0000	0101	0202	⋯	0909	1010	1111	⋯	1717	1818	1919
C_3	00000	01011	02022	⋯	09099	10109	11118	⋯	17172	18181	19190

(a) Determine a distância mínima de cada um dos três códigos.

(b) Diga quais dos códigos detectam e/ou corrigem erros singulares.

(c) Um receptor de televisão recebe informação do comando utilizando o terceiro código. Sempre que possível diga o efeito gerado pela recepção das seguintes mensagens: 15154, 13144, 19191.

4.16. Seja $\mathcal{C}$ o código $(7,3)$-linear binário definido pela matriz

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

(a) Qual é o número de palavras de $\mathcal{C}$?

(b) Calcule a distância mínima $\delta(\mathcal{C})$. Poderá $\mathcal{C}$ detectar erros singulares? E corrigir?

(c) Corrija, caso tal seja possível, os erros nas seguintes mensagens: 0001000, 1011110.

4.17. Seja $\mathcal{C}$ um código binário de comprimento 7 com matriz

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

(a) Mostre que a distância mínima de $\mathcal{C}$ é 3.

(b) Supondo que, no máximo, um erro singular é introduzido na transmissão, descodifique as mensagens 0010101 e 1000010.

4.18. Seja $\mathcal{C}$ um código binário com matriz

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

(a) Determine uma correspondência bijectiva entre líderes de classes laterais e síndromes.

(b) Descodifique as seguintes mensagens: $r_1 = 10101$, $r_2 = 01111$, $r_3 = 11111$, $r_4 = 11100$.

4.19. As matrizes H_1, H_2 e H_3 seguintes determinam três códigos lineares binários.

$$H_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \quad H_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad H_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Para cada um desses códigos, responda às seguintes questões:

(a) Determine o comprimento do código e o número de dígitos de controlo.

(b) Calcule a distância mínima e descreva o conjunto das mensagens.

(c) Poderão estes códigos ser usados para detectar e/ou corrigir erros singulares?

(d) Supondo que os três últimos dígitos da mensagem são 011, diga se esta mensagem pode pertencer ao código e determine a mensagem completa.

4.20. Para os códigos do Exercício 4.19, determine as síndromes e, se possível, corrija os erros das seguintes mensagens.

(a) Código 1; mensagens: 00000, 11111, 01010.

(b) Código 2; mensagens: 11011, 10011.

(c) Código 3; mensagens: 1000000, 1110101.

4.21. Considere $\mathbb{F}_{16} = \mathbb{F}_2(\alpha)$, com $\alpha^4 = \alpha + 1$, e a matriz do código BCH

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

(a) Faça uma estimativa para a distância mínima deste código.

(b) Codifique a mensagem 1010101 e descodifique 110010110100110 e 100111000000000.

(c) Mostre que se uma mensagem recebida r tem apenas um erro e esse erro é na posição i então $Hr = [\alpha^{(i-1)} \ \ \alpha^{3(i-1)}]^t$.

Bibliografia

[1] E. Artin, *Galois Theory*, Dover, 1998.

[2] R. L. Fernandes e M. Ricou, *Introdução à Álgebra*, IST Press, 2004.

[3] William J. Gilbert, *Modern Algebra with Applications*, Wiley, 1976.

[4] A. Gonçalves, *Introdução à Álgebra*, IMPA, Rio de Janeiro, 1979.

[5] C.H. Hadlock, *Field Theory and Its Classical Problems*, The Carus Mathematical Monographs, 19, The Mathematical Association of America, 2000.

[6] A. Hefez e M. L. Villela, *Códigos Correctores de Erros*, IMPA, Rio de Janeiro, 2002.

[7] T. W. Hungerford, *Algebra*, Springer-Verlag, 1980.

[8] A. Jones, S. Morris e K. Pearson, *Abstract Algebra and Famous Impossibilities*, Universitext, Springer-Verlag, 1994.

[9] R. Lidl e H. Niederreiter, *Introduction to Finite Fields and their Applications*, Cambridge University Press, 2000.

[10] J. Picado e M. Sobral, *Álgebra*, Textos de Apoio, Universidade de Coimbra, 2000.

[11] M. Sobral, *Álgebra*, Universidade Aberta, 1996.

[12] I. Stewart, *Galois Theory*, Chapman & Hall, 1973 (3a ed. 2004).