Introdução à Criptografia com Curvas Elípticas

Sally Andria
Rodrigo Gondim
Rodrigo Salomão

30º Colóquio Brasileiro de Matemática

impa

Introdução à Criptografia com Curvas Elípticas

Sally Andria
Rodrigo Gondim
Rodrigo Salomão

Introdução à Criptografia com Curvas Elípticas

Publicado no Brasil / Published in Brazil.

ISBN 978-85-244-0428-3
MSC (2010) Primary: 94A60, Secondary: 14G50, 11T71, 14H52, 12E20, 14H50

Comissão Editorial
Emanuel Carneiro
S. Collier Coutinho
Lorenzo J. Díaz
Étienne Ghys
Paulo Sad

Produção Books in Bytes **Capa** Sergio Vaz

Realização da Editora do IMPA
IMPA
Estrada Dona Castorina, 110
Jardim Botânico
22460-320 Rio de Janeiro RJ

Telefones: (21) 2529-5005
2529-5276
www.impa.br
ddic@impa.br

Conteúdo

Prefácio

Esse livro é resultado de um trabalho coletivo e foi escrito com a finalidade de servir de base para um curso introdutório, homônimo a ser ministrado no 32^o Colóquio Brasileiro de Matemática, a ser realizado no final de Julho de 2019 no Rio de Janeiro.

A ideia de escrever essas notas conjuntas, por parte dos Rodrigos, nasceu há alguns anos. Enquanto isso ambos rascunhavam notas preliminares. Sally teve exatamente esse como tema de seu trabalho de finalização de curso e foi uma grande motivação para o Rodrigo Gondim.

Finalmente escrevemos o texto final no ano de 2019 e o texto ficou um pouco mais robusto que o que havíamos previsto inicialmente. Nesse sentido oferecemos aos diversos tipos de leitores algumas formas alternativas de leitura, além, é claro, da possibilidade de ler todo o texto da forma como foi escrito.

A primeira forma alternativa de leitura está destinada ao leitor iniciante na Matemática, que tenha feito apenas cursos iniciais, como geometria analítica, aritmética dos inteiros ou álgebra I (ou algo que o equivalha). Nesse caso, uma leitura possível seria o Capítulo 1 somente as partes de Aritmética (Seções 1.1, 1.3 e Grupos (Seção 1.8), o Capítulo 2, o Capítulo 4 (Seção 4.1) e finalmente o Capítulo 5.

A segunda forma alternativa de leitura está destinada a um leitor mais experiente que não queira ”sujar” muito as mãos com contas explícitas sobre curvas. Os pré-requisitos necessários para uma tal leitura é um bom conhecimento de álgebra e um curso introdutório sobre curvas algébricas planas. Nesse caso nossa sugestão seria o Capítulo 1, como uma revisão do curso de álgebra, o Capítulo 2, o Capítulo 3, como uma revisão de alguns aspectos do curso de curvas, o Capítulo 4, excetuando a primeira seção, e o Capítulo 5.

1 Fundamentos Algébricos

A álgebra é a ciência dos símbolos e das operações entre estes. Certa vez, Alexander Dewdney, matemático e cientista da computação canadense, declarou:

"Uma das ideias mais poderosas da matemática é a noção de variável."

Se tal afirmação já era no passado uma verdade valiosa e intuída por muitos, hoje com o advento da computação, tem crescido fortemente. Nessa seção vamos introduzir os ingredientes algébricos para desenvolver a criptografia com curvas elípticas. Tais ingredientes passam por várias estruturas algébricas, desde a aritmética clássica até objetos da chamada álgebra abstrata.

Se por um lado a geometria, durante séculos, ficou refém do sistema axiomático advindo da intuição contida nos primeiros livros dos Elementos de Euclides, a álgebra também tardou em se desvencilhar das amarras da aritmética introduzidas no livro XII dos Elementos de Euclides. Aquilo que chamamos de álgebra abstrata trata de objetos simbólicos sujeitos a novos axiomas de combinação, não necessariamente inspirados na aritmética. Assim se referia Augustus de Morgan ao introduzir tal ponto de vista abstrato:

> *"nenhuma palavra ou sinal de aritmética ou álgebra tem um átomo de significado ao longo deste capítulo, cujo objeto são os símbolos, e suas leis de combinação, dando uma álgebra simbólica a qual pode daqui em diante se tornar a gramática de cem álgebras significativas e distintas."*

O surgimento, por exemplo, de estruturas algébricas não comutativas se deu contemporaneamente ao surgimento das geometrias não euclidianas. Estaremos particularmente

interessados em uma estrutura chamada de grupo. O grupo que queremos analisar com maior profundidade nasce de uma operação geométrica sobre os pontos de uma curva cúbica chamada curva elíptica.

Os principais livros sobre álgebra escritos no Brasil são Gonçalves (2017), Hefez (2016) e Lequain e Garcia (2018).

1.1 Aritmética modular

Um exemplo típico da aritmética modular no nosso cotidiano vem da contagem das horas. Suponhamos que o relógio esteja marcando onze horas da manhã e você tenha combinado de almoçar com uns colegas uma hora da tarde. Então, se alguém perguntar para você quantas horas faltam para o tal almoço, você irá responder duas horas. De certa forma, a conta que você está fazendo é:

$$11 + 2 = 1 \text{ ou } 1 - 11 = 2.$$

Apesar desta conta parecer não ter o menor sentido, veremos que ela é apenas um caso particular do que chamaremos de aritmética modular.

Seja $m > 1$. Dizemos que dois inteiros a e b são *congruentes módulo* m, quando m divide $a - b$. Usamos o seguinte símbolo para denotar isto.

$$a \equiv b \bmod m$$

Note que a conta que fizemos sobre as horas pode ser interpretada como uma conta de congruência módulo 12, já que,

$$11 + 2 = 13 \equiv 1 \bmod 12.$$

Exemplo 1. Temos as seguintes congruências.

- $-4 \equiv 2 \bmod 3$ pois $-4 - 2 = -6 = -2 \cdot 3$;
- $-3 \equiv 0 \bmod 3$ pois $-3 - 0 = -3 = -1 \cdot 3$;
- $-2 \equiv 1 \bmod 3$ pois $-2 - 1 = -3 = -1 \cdot 3$;
- $-1 \equiv 2 \bmod 3$ pois $-1 - 2 = -3 = -1 \cdot 3$;
- $4 \equiv 1 \bmod 3$ pois $4 - 1 = 3 = 1 \cdot 3$;
- $5 \equiv 2 \bmod 3$ pois $5 - 2 = 3 = 1 \cdot 3$;
- $6 \equiv 0 \bmod 3$ pois $6 - 0 = 6 = 2 \cdot 3$;
- $7 \equiv 1 \bmod 3$ pois $7 - 1 = 6 = 2 \cdot 3$;
- $8 \equiv 2 \bmod 3$ pois $8 - 2 = 6 = 2 \cdot 3$;

- $9 \equiv 0 \bmod 3$ pois $9 - 0 = 9 = 3 \cdot 3$;
- $10 \equiv 1 \bmod 3$ pois $10 - 1 = 9 = 3 \cdot 3$;

Ao longo do texto faremos as contas com o software Pari/GP. Para obter o programa e aprender sobre os seus comandos ver Apêndice A Agora, usaremos o comando $Mod(a, m)$. Para -4, -3 e 4 módulo 3, por exemplo:

```
?   Mod(−4, 3)
%1  = Mod(2, 3)
?   Mod(−3, 3)
%2  = Mod(0, 3)
?   Mod(4, 3)
%3  = Mod(1, 3)
```

É natural intuir, pelo exemplo acima, que um número inteiro n sempre será congruente a 0, 1 ou 2 módulo 3. De fato, pelo Teorema da Divisão Euclidiana, que enunciaremos abaixo, temos que $n = 3 \cdot q + r$, com $0 \leqslant r < 3$, isto é, $r = 0$, 1 ou 2. Portanto, $n - 0$, $n - 1$ ou $n - 2$ é um múltiplo de 3.

Teorema 1 (Divisão Euclidiana)**.** *Sejam a e m números inteiros, com $m \neq 0$. Então existem inteiros q e r, unicamente determinados, satisfazendo as seguintes propriedades*

$$a = q \cdot m + r \ \ e \ \ 0 \leqslant r < m.$$

Prova: Ver Hefez (2016), seção 2 do capítulo 3. □

Exemplo 2. Se quisermos realizar a divisão euclidiana de x por y no Pari/GP, utilizamos o comando $divrem(x, y)$ e o programa responde o vetor com duas entradas: o quociente e o resto.

```
?   divrem(4, 2)
%1  = [2, 0] ~
?   divrem(3487510638456, 9345629)
%2  = [373170, 2264526] ~
```

Em geral, segue do Teorema 1.1, que ao fixarmos um número inteiro m maior que 1, então cada número inteiro n será congruente a um dos números inteiros compreendidos entre 0 e $m - 1$.

Desta forma, faz sentido em separar os números inteiros nas seguintes *classes de equivalência módulo m*

$$\begin{aligned}
&\overline{0} := \{n \in \mathbb{Z} \mid n \equiv 0 \bmod m\} = \{n \in \mathbb{Z} \mid n = q \cdot m \text{ para algum } q \in \mathbb{Z}\};\\
&\overline{1} := \{n \in \mathbb{Z} \mid n \equiv 1 \bmod m\} = \{n \in \mathbb{Z} \mid n = q \cdot m + 1 \text{ para algum } q \in \mathbb{Z}\};\\
&\overline{2} := \{n \in \mathbb{Z} \mid n \equiv 2 \bmod m\} = \{n \in \mathbb{Z} \mid n = q \cdot m + 2 \text{ para algum } q \in \mathbb{Z}\};\\
&\overline{3} := \{n \in \mathbb{Z} \mid n \equiv 3 \bmod m\} = \{n \in \mathbb{Z} \mid n = q \cdot m + 3 \text{ para algum } q \in \mathbb{Z}\};\\
&\vdots\\
&\overline{m-1} := \{n \in \mathbb{Z} \mid n \equiv m-1 \bmod m\} = \{n \in \mathbb{Z} \mid n = q \cdot m + (m-1) \text{ para algum } q \in \mathbb{Z}\};
\end{aligned}$$

O conjunto destas classes

$$\mathbb{Z}_m := \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{m-1}\}$$

será chamado de *anel dos inteiros módulo m*.

Observação 1. *Para cada número inteiro k, também podemos definir*

$$\overline{k} := \{n \in \mathbb{Z} \mid n \equiv k \bmod m\} = \{n \in \mathbb{Z} \mid n - k = q \cdot m \text{ para algum } q \in \mathbb{Z}\}$$

Mas, pelo Teorema 1.1, temos que $k = q' \cdot m + r$ *com* $0 \leqslant r < m$. *Portanto, para cada* $n \in \mathbb{Z}$, *temos que*

$$\begin{gathered}
n - k = q \cdot m \text{ para algum } q \in \mathbb{Z}\\
\Longleftrightarrow\\
n = \widetilde{q} \cdot m + r \text{ para algum } \widetilde{q} \in \mathbb{Z}.
\end{gathered}$$

Consequentemente, temos que $\overline{k} = \overline{r}$ *para algum* $r \in \{0, 1, \ldots, m-1\}$.

No que se segue, iremos explicar o motivo da palavra anel. Para isso, vamos ver, primeiramente, que faz sentido somar e multiplicar "dentro" das congruências.

Exemplo 3. Sabemos que $15 \equiv 3 \bmod 4$ e que $2 \equiv 6 \bmod 4$. Agora, $15 + 2 = 17 \equiv 1 \bmod 4$ e $15 \cdot 2 = 30 \equiv 2 \bmod 4$. Por outro lado, $3 + 6 = 9 \equiv 1 \bmod 4$ e $3 \cdot 6 = 18 \equiv 2 \bmod 4$. Portanto,

$$\begin{array}{ccccccc}
15 + 2 & \equiv & 1 & \equiv & 3 + 6 & \bmod 4\\
15 \cdot 2 & \equiv & 2 & \equiv & 3 \cdot 6 & \bmod 4
\end{array}$$

No Pari/GP é possível verificar essas contas

```
? Mod(15 + 2, 4)
%1 = Mod(1, 4)
? Mod(3 + 6, 4)
%2 = Mod(1, 4)
? Mod(15 * 2, 4)
%3 = Mod(2, 4)
? Mod(3 * 6, 4)
%4 = Mod(2, 4)
```

De modo geral, estas propriedades ainda são verificadas, como podemos ver na seguinte proposição.

Proposição 1. *Seja $m > 1$. Se $a_1 \equiv a_2 \bmod m$ e $b_1 \equiv b_2 \bmod m$, então valem as seguintes congruências.*

$$\begin{array}{rcll} a_1 + b_1 & \equiv & a_2 + b_2 & \bmod m \\ a_1 \cdot b_1 & \equiv & a_2 \cdot b_2 & \bmod m \end{array}$$

PROVA: Escrevemos $a_1 - a_2 = m \cdot n_1$ e $b_1 - b_2 = m \cdot n_2$. Então,

$$\begin{array}{rcl} (a_1 + b_1) - (a_2 + b_2) & = & (a_1 - a_2) + (b_1 - b_2) \\ & = & m \cdot n_1 + m \cdot n_2 \\ & = & m \cdot (n_1 + n_2) \end{array}$$

e

$$\begin{array}{rcl} (a_1 \cdot b_1) - (a_2 \cdot b_2) & = & (a_1 \cdot b_1) - (a_2 \cdot b_1) + (a_2 \cdot b_1) - (a_2 \cdot b_2) \\ & = & (a_1 - a_2) \cdot b_1 + a_2 \cdot (b_1 - b_2) \\ & = & m \cdot n_1 \cdot b_1 + a_2 \cdot m \cdot n_2 \\ & = & m \cdot (n_1 b_1 + a_2 n_2) \end{array}$$ □

Com esta proposição, faz sentido definir uma operação no anel dos inteiros módulo m da seguinte forma. Sejam $\overline{a}$ e $\overline{b}$ em $\mathbb{Z}_m$. Definimos, então:

$$\begin{array}{rcll} \overline{a} + \overline{b} & := & \text{classe no qual } a + b \text{ pertence} & = \overline{a + b} \\ \overline{a} \cdot \overline{b} & := & \text{classe no qual } a \cdot b \text{ pertence} & = \overline{a \cdot b} \end{array}$$

Exemplo 4. 1. Vamos fazer a tabela da adição e do produto em $\mathbb{Z}_3$:

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{1}$

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

Podemos obter a tabela da soma com o Pari:

```
? for(k = 0, 2, print(Mod(k + 0, 3), Mod(K + 1, 3), Mod(k + 2, 3)))
Mod(0, 3)Mod(1, 3) * K + Mod(1, 3)Mod(2, 3)
Mod(1, 3)Mod(1, 3) * K + Mod(1, 3)Mod(0, 3)
Mod(2, 3)Mod(1, 3) * K + Mod(1, 3)Mod(1, 3)
```

E a tabela da multiplicação:

```
? for(k = 0, 2, print(Mod(k * 0, 3), Mod(K * 1, 3), Mod(k * 2, 3)))
  Mod(0, 3)Mod(1, 3) * KMod(0, 3)
  Mod(0, 3)Mod(1, 3) * KMod(2, 3)
  Mod(0, 3)Mod(1, 3) * KMod(1, 3)
```

Porém observe que na segunda coluna, as operações com $\overline{1}$ no programa são respondidas em função de k.

2. Vamos fazer a tabela da adição e do produto em $\mathbb{Z}_4$:

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

Não é difícil de verificar que as operações de soma e produto em $\mathbb{Z}_m$ satisfazem as seguintes propriedades.

Propriedades da Soma:

1. $\overline{0} + \overline{a} = \overline{a} + \overline{0} = \overline{a}$ para cada $\overline{a} \in \mathbb{Z}_m$;
2. $\overline{-a} + \overline{a} = \overline{a} + \overline{-a} = \overline{0}$ para cada $\overline{a} \in \mathbb{Z}_m$;
3. $\overline{a} + (\overline{b} + \overline{c}) = (\overline{a} + \overline{b}) + \overline{c}$ para cada $\overline{a}, \overline{b}, \overline{c} \in \mathbb{Z}_m$;
4. $\overline{a} + \overline{b} = \overline{b} + \overline{a}$ para cada $\overline{a}, \overline{b} \in \mathbb{Z}_m$.

Propriedades do Produto:

1. $\overline{1} \cdot \overline{a} = \overline{a} \cdot \overline{1} = \overline{a}$ para cada $\overline{a} \in \mathbb{Z}_m$;
2. $\overline{a} \cdot (\overline{b} \cdot \overline{c}) = (\overline{a} \cdot \overline{b}) \cdot \overline{c}$ para cada $\overline{a}, \overline{b}, \overline{c} \in \mathbb{Z}_m$;
3. $\overline{a} \cdot \overline{b} = \overline{b} \cdot \overline{a}$ para cada $\overline{a}, \overline{b} \in \mathbb{Z}_m$.

Propriedades da Soma e do Produto:

1. $\overline{a} \cdot (\overline{b} + \overline{c}) = (\overline{a} \cdot \overline{b}) + (\overline{a} \cdot \overline{c})$ para cada $\overline{a}, \overline{b}, \overline{c} \in \mathbb{Z}_m$.

Estas propriedades correspondem exatamente aos axiomas que determinam o fato de um conjunto, com duas operações de "soma" e "produto", ser um anel Seção 1.2.

1.2 Anéis e corpos

Seja A um conjunto munido das seguintes operações

$$\begin{array}{rccccc} + & : & A\times A & \to & A \\ & & (a,b) & \mapsto & a+b \end{array} \quad \text{e} \quad \begin{array}{rccccc} \cdot & : & A\times A & \to & A \\ & & (a,b) & \mapsto & a\cdot b \end{array}$$

que chamaremos de soma e produto, respectivamente.

Dizemos que $(A,+,\cdot)$ é um *anel* quando as operações satisfazem os seguintes axiomas.

Propriedades da Soma:

1. (Comutatividade da soma) $a+b=b+a$ para cada $a,b\in A$.
2. (Associatividade da soma) $a+(b+c)=(a+b)+c$ para cada $a,b,c\in A$;
3. (Existência do elemento neutro da soma) Existe um elemento $0\in A$ tal que $0+a=a+0=a$ para cada $a\in A$;
4. (Existência do elemento simétrico da soma) Para cada $a\in A$ existe $-a\in A$ tal que $-a+a=0$;

Propriedades do Produto:

1. (Associatividade do produto) $a\cdot(b\cdot c)=(a\cdot b)\cdot c$ para cada $a,b,c\in A$;
2. (Comutatividade do produto) $a\cdot b=b\cdot a$ para cada $a,b\in A$.
3. (Existência do elemento neutro do produto) Existe um elemento $1\in A$ tal que $1\cdot a=a\cdot 1=a$ para cada $a\in A$;

Propriedades da Soma e do Produto:

1. (Distributividade do produto com relação a soma) $a\cdot(b+c)=(a\cdot b)+(a\cdot c)$ para cada $a,b,c\in A$.

Exemplo 5. Os conjuntos dos números inteiros $\mathbb{Z}$, dos números racionais $\mathbb{Q}$, dos números reais $\mathbb{R}$ e dos números complexos $\mathbb{C}$, com as operações de somas e produtos usuais, são anéis. O conjunto $\mathbb{Z}_m$ definido na Seção 1.1 é um anel.

Dizemos que um anel $(A,+,\cdot)$ é um *domínio de integridade* quando a igualdade $a\cdot b=0$ em A só for satisfeita se a ou b forem iguais a 0.

Exemplo 6. Os anéis $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ são domínios de integridade. Pelas tabelas do Exemplo 4 temos que $\mathbb{Z}_3$ é um domínio de integridade enquanto $\mathbb{Z}_4$ não é.

Dizemos que um elemento a em um anel A é *invertível* quando existe $b \in A$ tal que $a \cdot b = 1$. Um tal elemento é chamado de *inverso* de a e é denotado por a^{-1}. Pelo Exercício 6 temos que o inverso está bem definido. Denotamos A^* pelo conjunto dos elementos inversíveis em A.

Diremos que dois elementos a, b em um anel A são *associados* quando existir um elemento invertível $u \in A$ tal que $a = u \cdot b$.

Dizemos que um domínio de integridade $(K, +, \cdot)$ é um *corpo* quando todo elemento não nulo em K for invertível .

Exemplo 7. Os conjuntos dos números inteiros $\mathbb{Z}$ não é um corpo, mas os domínios de integridade $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ são corpos. Pelas tabelas do Exemplo 4 temos que $\mathbb{Z}_3$ é um corpo. Veremos na Seção 1.3 que $\mathbb{Z}_m$ é um corpo se, e somente se, m é um número primo.

Um subconjunto I de um anel $(A, +, \cdot)$ é dito um *ideal* quando ele satisfizer as seguintes três condições.

1. $0 \in I$;
2. $a + b \in I$, sempre que $a, b \in I$;
3. $a \cdot b \in I$, sempre que $a \in A$ e $b \in I$.

Exemplo 8. Os ideais de $\mathbb{Z}$ são da forma $m\mathbb{Z} = \{mz | z \in \mathbb{Z}\}$ com $m \in \mathbb{Z}$. Os únicos ideais em um corpo K são os ideais triviais, isto é, o ideal nulo (0) e o próprio corpo K.

Sejam A e B anéis e $f : A \to B$ uma função. Dizemos que f é um *homomorfismo de anéis* quando f satisfizer as seguintes condições.

1. $f(a + b) = f(a) + f(b)$ para todo $a, b \in A$;
2. $f(a \cdot b) = f(a) \cdot f(b)$ para todo $a, b \in A$;

O *núcleo* do homomorfismo f é definido por $\text{Nuc}(f) = \{a \in A | f(a) = 0\}$. Não é difícil verificar o seguinte lema, que deixaremos a cargo do leitor.

Lema 1. *Seja $f : A \to B$ um homomorfismo de anéis. Então, valem:*

1. $\text{Nuc}(f)$ *é um ideal de A;*
2. f *é injetiva se, e só se,* $\text{Nuc}(f) = \{0\}$.

Um homomorfismo de anéis $f : A \to B$ é dito um *isomorfismo de anéis* quando f for uma bijeção, isto é, injetiva e sobrejetiva.

Proposição 2. *Seja $f : A \to B$ um homomorfismo de anéis, onde A e B são corpos. Se f não é o homomorfismo nulo, então $f(1) = 1$, f é injetiva e $f(A)$ é um subcorpo de B.*

1.3 O corpo $\mathbb{Z}_p$

Além das propriedades sobre o anel $\mathbb{Z}_m$, evidenciadas na Seção 1.1, ainda vamos observar outras duas propriedades, que distinguem o caso em que m é primo ou não.

A primeira, é muito parecida com o que estamos habituados no conjunto dos números inteiros, que diz que o produto de dois elementos não nulos também é não nulo. Quando isto for satisfeito, iremos dizer que $\mathbb{Z}_m$ é um *domínio de integridade*.

Exemplo 9. Vimos no Exemplo 4 que $\mathbb{Z}_3$ é um domínio de integridade, mas $\mathbb{Z}_4$ não, já que, $\overline{2} \cdot \overline{2} = \overline{0}$.

De forma mais geral, ainda temos o seguinte resultado.

Proposição 3. *Seja $m > 1$ um número inteiro. Então $\mathbb{Z}_m$ é um domínio de integridade se, e só se, m for primo.*

Prova: Suponhamos que $m = p$ seja primo e que $\overline{a} \cdot \overline{b} = \overline{0}$. Por definição, temos que $a \cdot b$ é divisível por p. Sendo p primo, temos, pela unicidade da fatoração em $\mathbb{Z}$, que p deve estar na fatoração de a ou de b, isto é, p divide a ou b. Logo, $\overline{a} = \overline{0}$ ou $\overline{b} = \overline{0}$.

Agora, se m não for primo, digamos $m = a \cdot b$ com a e b inteiros positivos distintos de 1, então $\overline{a} \neq \overline{0}$ ou $\overline{b} \neq \overline{0}$, já que, $1 < a < m$ e $1 < b < m$. Além disso, $a \cdot b = m$ nos diz que $\overline{a} \cdot \overline{b} = \overline{0}$. □

Para destacar a segunda propriedade de $\mathbb{Z}_m$, precisamos estudar as soluções de certas equações envolvendo congruência módulo.

Proposição 4. Sejam $m > 1$ e a números inteiros. Então existe $b \in \mathbb{Z}$ tal que $ab \equiv 1 \bmod m$ se, e só se, $\text{mdc}(a, m) = 1$.

Prova: Se $ab \equiv 1 \bmod m$, então $ab - 1 = qm$ para algum q inteiro, isto é, $ab - qm = 1$. Agora, se d é um divisor comum entre a e m, então pela última igualdade temos que d divide 1. Portanto, temos que $\text{mdc}(a, m) = 1$.

Por outro lado, se $\text{mdc}(a, m) = 1$, então, pelo algoritmo Euclidiano estendido que enunciaremos abaixo, temos que a equação nas variáveis X e Y

$$aX + mY = 1$$

possui soluções inteiras, digamos $X = b$ e $Y = q$, isto é, $ab - 1 = -qm$. Logo, $ab \equiv 1 \bmod m$. □

Teorema 2 (Algoritmo Euclidiano Estendido - Lema de Bézout)**.** *Sejam a e b inteiros positivos. Então a equação nas variáveis X e Y*

$$aX + bY = \text{mdc}(a, b)$$

possui soluções inteiras, digamos $X = u_0$ e $Y = v_0$. Além disso, todas as suas soluções são da forma

$$X = u_0 + \frac{bk}{\text{mdc}(a,b)} \quad e \quad Y = v_0 - \frac{ak}{\text{mdc}(a,b)}$$

com $k \in \mathbb{Z}$.

Ao invés de fazermos a prova do caso geral, iremos descrever o algoritmo em um exemplo. Para consultar a prova o caso geral, ver Hefez (2016) Capítulo 5, Seção 3.

Exemplo 10. Consideremos $a = 12$ e $b = 125$. Como $a = 2^2 \cdot 3$ e $b = 5^3$ temos que $\text{mdc}(a,b) = 1$. Agora procedemos com as divisões Euclidianas abaixo:

125	$=$	$12 \cdot 10 + 5$	
12	$=$	$5 \cdot 2 + 2$	Dividimos o divisor "12" pelo resto "5" da equação acima;
5	$=$	$2 \cdot 2 + 1$	Dividimos o divisor "5" pelo resto "2" da equação acima;
2	$=$	$1 \cdot 2 + 0$	Dividimos o divisor "2" pelo resto "1" da equação acima. Obtemos resto zero. Então paramos!!!

Repare que o resto da penúltima divisão é exatamente o $\text{mdc}(12, 125)$. Isolando os restos não nulos em cada uma das divisões acima obtemos:

$$\begin{aligned} 5 &= 125 - 12 \cdot 10 \\ 2 &= 12 - 5 \cdot 2 \\ 1 &= 5 - 2 \cdot 2 \end{aligned}$$

O que temos que fazer agora é, de baixo para cima, substituir os restos isolados na equação de baixo. Isto é,

1	$=$	$5 - 2 \cdot 2$	
	$=$	$5 - (12 - 5 \cdot 2) \cdot 2$	Substituímos o penúltimo resto na equação acima;
	$=$	$5 \cdot (5) + 12 \cdot (-2)$	Isolamos *a=12* e o antepenúltimo resto 5;
	$=$	$(125 - 12 \cdot 10) \cdot (5) + 12 \cdot (-2)$	Substituímos o antepenúltimo resto 5.
	$=$	$125 \cdot (5) + 12 \cdot (-52)$	Isolamos a=12 e b=125

Portanto, uma solução para a equação $12X + 125Y = 1$ é $X = -52$ e $Y = 5$.

Observação 2. *É importante observar que este processo, descrito no exemplo e chamado de algoritmo Euclidiano estendido, serve inicialmente para encontrar o máximo divisor comum entre dois números inteiros.*

Agora vamos utilizar estas propriedades descritas acima para evidenciar uma estrutura peculiar que $\mathbb{Z}_m$ possui para certos inteiros m.

Um elemento $\overline{a} \neq \overline{0}$ em $\mathbb{Z}_m$ é dito *invertível* quando existe $\overline{b} \in \mathbb{Z}_m$ tal que $\overline{a} \cdot \overline{b} = \overline{1}$.

Neste caso, não é difícil ver que o elemento $\overline{b} \in \mathbb{Z}_m$ é unicamente determinado pela propriedade $\overline{a} \cdot \overline{b} = \overline{1}$. Este elemento é chamado de *inverso multiplicativo* de $\overline{a}$ e é denotado por $\overline{a}^{-1}$.

Exemplo 11. Pelas tabelas da multiplicação em $\mathbb{Z}_3$ e $\mathbb{Z}_4$, que fizemos no Exemplo 4, temos:

1. Em $\mathbb{Z}_3$, temos que $\overline{1}$ e $\overline{2}$ são invertíveis, pois $\overline{1}\cdot\overline{1}=\overline{1}$ e $\overline{2}\cdot\overline{2}=\overline{1}$;
2. Em $\mathbb{Z}_4$, temos que $\overline{1}$ e $\overline{3}$ são invertíveis, pois $\overline{1}\cdot\overline{1}=\overline{1}$ e $\overline{3}\cdot\overline{3}=\overline{1}$. Mas $\overline{2}$ não é invertível pois $\overline{2}\cdot\overline{2}=\overline{0}$, $\overline{2}\cdot\overline{3}=\overline{2}$.

Note que, pela definição, temos que a igualdade $\overline{a}\cdot\overline{b}=\overline{1}$ equivale a dizer que o resto da divisão de $a\cdot b$ por m é 1. Em outras palavras, que $a\cdot b\equiv 1 \bmod m$. Portanto, pela proposição anterior, temos que:

$$\overline{a} \text{ é invertível em } \mathbb{Z}_m \iff \text{mdc}(a,m)=1.$$

Esta equivalência nos fornece a seguinte generalização do primeiro item do exemplo anterior.

Proposição 5. *Seja $p>1$ um número primo. Então, todo elemento não nulo de $\mathbb{Z}_p$ é invertível.*

Prova: É claro que p não aparece na fatoração de cada elemento do conjunto $\{1,2,\ldots,p-1\}$. Portanto, $\text{mdc}(a,p)=1$ para cada $a\in\{1,2,\ldots,p-1\}$. □

Como $\mathbb{Z}_p$, com p primo, é um domínio de integridade e todos os seus elementos são invertíveis, então dizemos que $\mathbb{Z}_p$ é um *corpo*. . Utilizaremos a notação $\mathbb{F}_p$ para designar o corpo finito $\mathbb{Z}_p$.

1.4 Anéis de polinômios

Se D um domínio, então denotaremos por $D[X]$ o conjunto formado por todos os polinômios, na variável X, com coeficientes em D. Cada elemento $f(X)\in D[X]$ se escreve como

$$f(X)=\sum_{i=0}^{n} a_i X^i = a_0+a_1X+\cdots+a_nX^n$$

onde n é um inteiro não negativo e $a_0,\ldots,a_n\in D$. Quando $a_n\neq 0$, dizemos que n é o *grau* de $f(X)$ e denotamos $n=\text{grau } f(X)$. Quando $a_n=1$, dizemos que $f(X)$ é um *polinômio mônico*.

As operações de soma e multiplicação em $D[X]$ são definidas da seguinte forma. Dados $f(X)=\sum_{i=0}^{n}a_iX^i$ e $g(X)=\sum_{j=0}^{m}b_jX^j$, definimos

$$f(X)+g(X)=\sum_{i=0}^{\max\{n,m\}}(a_i+b_i)X^i,$$

$$f(X)\cdot g(X)=\sum_{i=0}^{n+m}c_iX^i, \text{ onde } c_i=\sum_{j=0}^{i}a_jb_{i-j}.$$

Teorema 3. *O conjunto $D[X]$, com as operações de soma e produto definidas acima, é um anel.*

PROVA: Ver Villela e Hefez (2017), Teorema 1, Capítulo 3. □

Em virtude deste teorema, chamamos $D[X]$ de *anel de polinômios* em uma variável com coeficientes em D. Na verdade, $D[X]$ é um domínio.

Os casos em que estaremos mais interessados serão quando $D = K[Y]$ e $D = K$, sendo K é um corpo e Y é uma outra indeterminada sobre K.

A partir de agora consideraremos K sendo um corpo. Também vale o Teorema da Divisão Euclidiana em $K[X]$, assim como em $\mathbb{Z}$.

Teorema 4 (Divisão Euclidiana)**.** *Sejam $f(X)$ e $g(X)$ dois polinômios em $D[X]$, com $f(X)$ tendo coeficiente líder invertível em D. Então existem polinômios $q(X)$ e $r(X)$ em $K[X]$, unicamente determinados, satisfazendo as seguintes propriedades*

$$g(X) = q(X) \cdot f(X) + r(X) \;\; e \;\; 0 \leqslant \text{grau}\, r(X) < \text{grau}\, f(X).$$

PROVA: Ver Lequain e Garcia (2018), Proposição I.3.9, página 24. □

Exemplo 12. O mesmo comando $divrem(x, y)$, que usamos para divisão euclidiana de números inteiros pode ser usado para divisão de polinômios em $\mathbb{Z}[X]$. Podemos ainda escolher qual a variável principal, abaixo realizamos a divisão de $YX^2 + 1$ por X em $\mathbb{Z}[Y][X]$.

```
?   divrem(x^2 − 2, x + 2)
%1  = [x − 2, 2] ~
?   divrem(y ∗ x^2 + 1, x, x)
%2  = [y ∗ x, 1] ~
```

Sejam $f(X), g(X) \in K[X]$. Dizemos que $f(X)$ *divide* $g(X)$, quando $g(X) = f(X)q(X)$, para algum $q(X) \in K[X]$.

Dizemos que $D(X)$ é um *máximo divisor comum* entre $f(X)$ e $g(X)$ quando $D(X)$ divide $f(X)$ e $g(X)$, além de que, todo divisor comum entre $f(X)$ e $g(X)$ tem que dividir $D(X)$.

Exemplo 13. O $mdc(f(X), g(X))$ também pode ser calculado:

```
?   gcd(x^2 − 2, x + 2)
%1  = 1
?   gcd(3 ∗ x^2 + 1, x^3 − 4 ∗ x + 8)
%2  = 1
?   gcd(x^4 + 3 ∗ x^3 − 24 ∗ x^2 − 28 ∗ x + 48, x^3 + 2 ∗ x^2 − 21 ∗ x + 18)
%3  = x^2 + 5 ∗ x − 6
```

Analogamente ao caso dos números inteiros, o algoritmo da divisão Euclidiana também fornece um algoritmo para calcular um máximo divisor comum entre dois polinômios. Resumidamente, temos o seguinte resultado.

Teorema 5 (Algoritmo Euclidiano Estendido - Lema de Bézout)**.** *Seja K um corpo. Dados $f(X), g(X) \in K[X]$, não simultaneamente nulos, e $D(X)$ um máximo divisor comum entre $f(X)$ e $g(X)$, existem $r(X), s(X) \in K[X]$ tais que*

$$D(X) = r(X)f(X) + s(X)g(X).$$

Este resultado garante a existência de um máximo divisor em comum entre dois polinômios não simultaneamente nulos. O único máximo divisor comum e mônico entre $f(X)$ e $g(X)$ será denotado por $\text{mdc}(f(X), g(X))$.

Dizemos que $f(X)$ e $g(X)$ são *primos entre si* quando $\text{mdc}(f(X), g(X)) = 1$.

Corolário 1. *Seja K um corpo. Dois polinômios $f(X), g(X) \in K[X]$, não simultaneamente nulos, são primos entre si se, e só se, existem $r(X), s(X) \in K[X]$ tais que*

$$1 = r(X)f(X) + s(X)g(X).$$

Um polinômio $f(X) \in K[X]$ é dito *irredutível* quando

$$f(X) \text{ divide } g(X)h(X) \Longrightarrow f(X) \text{ divide } g(X) \text{ ou } h(X).$$

Teorema 6. *Seja K um corpo. Todo polinômio em $K[X]$ se escreve, de forma única, como produto de um elemento de K por um produto de polinômios mônicos e irredutíveis.*

Prova: Ver Lequain e Garcia (2018), Teorema II.3.1, página 48. □

Esta escrita do teorema acima é chamada de *fatoração* do polinômio.

Em virtude deste teorema, dizemos que $K[X]$ é um *domínio de fatoração única*.

Seja $f(X) \in K[X]$ e $\alpha \in K$. Dizemos que α é uma *raiz* quando $f(\alpha) = 0$.

Aplicando o Teorema 4 para dividir $f(X)$ por $X - \alpha$, obtemos que α é raiz de $f(X)$ se, e só se, $X - \alpha$ divide $f(X)$.

Note que, $X - \alpha$ é irredutível em $K[X]$, já que, é um polinômio de grau um. Desta forma, se α é uma raiz de $f(X)$, então aparece uma potência de $X - \alpha$ na fatoração de $f(X)$. Esta potência de $X - \alpha$ na fatoração de $f(X)$ é chamada de *multiplicidade* de α como raiz de $f(X)$.

Teorema 7. *Um polinômio de grau n com coeficientes em um corpo K possui, no máximo, n raizes em K contadas com suas respectivas multiplicidades.*

Prova: De fato, a quantidade de raizes de um polinômio, contadas com as suas respectivas multiplicidades, é o número de polinômios de grau um, distintos ou não, que aparece na fatoração do polinômio. Como o grau de um produto de polinômios é a soma dos graus dos polinômios, segue que a quantidade de polinômios de grau um na fatoração de um

polinômio é menor ou igual que o seu grau. □

Faremos agora um estudo das classes residuais em $K[X]$, analogamente ao que originou o anel $\mathbb{Z}_m$ dos inteiros módulo um inteiro positivo m.

Seja $f(X) \in K[X]$ mônico de grau $d \geqslant 1$. Dado $g(X) \in K[X]$, definimos a *classe residual* de $g(X)$ módulo $f(X)$ por

$$\overline{g(X)} = \{h(X) \in K[X] | f(X) \text{ divide } g(X) - h(X)\}.$$

Pelo Teorema 4, existem $q(X), r(X) \in K[X]$ unicamente determinados tais que

$$g(X) = q(X)f(X) + r(X), \text{ com } r(X) = 0 \text{ ou } 0 \leqslant \operatorname{grau} r(X) < d.$$

Desta forma, para cada $g(X) \in K[X]$, existe um único $r(X) \in K[X]$ com $r(X) = 0$ ou com $0 \leqslant \operatorname{grau} r(X) < d$ tal que $\overline{g(X)} = \overline{r(X)}$. De fato, dados dois polinômios $r_1(X)$ e $r_2(X)$ satisfazendo $r_i(X) = 0$ ou $0 \leqslant \operatorname{grau} r_i(X) < d$, então, comparando os graus, vemos que a única possibilidade para $f(X)$ dividir $r_1(X) - r_2(X)$ é com $r_1(X) - r_2(X) = 0$.

Definimos o conjunto das classes residuais em $K[X]$ módulo $f(X)$ por

$$K[X]_{f(X)} = \{\overline{r(X)} | r(X) = 0 \text{ ou } 0 \leqslant \operatorname{grau} r(X) < d\}.$$

As operações em $K[X]$ induzem operações em $K[X]_{f(X)}$. De fato, dados $f(X), g(X) \in K[X]$, definimos

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

Assim como no caso dos anéis de inteiros módulo m, não é difícil verificar que $K[X]_{f(X)}$ com as operações definidas acima é um anel com $\overline{0}$ e $\overline{1}$ sendo os elementos neutros da adição e da multiplicação, respectivamente.

Teorema 8. *O anel $K[X]_{f(X)}$ é um corpo se, e só se, $f(X)$ é um polinômio irredutível.*

Prova: Denotemos por d o grau de $f(X)$.

Suponhamos que $f(X)$ seja irredutível. Então, para cada $r(X) \neq 0$ com $0 \leqslant \operatorname{grau} r(X) < d$ temos que $r(X)$ e $f(X)$ são primos entre si. Pelo Teorema 4 existem $a(X), b(X) \in K[X]$ tais que

$$1 = a(X)r(X) + b(X)f(X).$$

Portanto, concluímos que $\overline{1} = \overline{a(X)} \cdot \overline{r(X)}$, isto é, todo elemento não nulo de $K[X]_{f(X)}$ tem inverso multiplicativo.

Suponhamos agora que $K[X]_{f(X)}$ é um corpo e consideremos $g(X), h(X) \in K[X]$ com $f(X)$ dividindo $g(X)h(X)$ e não dividindo $g(X)$. Desta forma,

$$\overline{0} = \overline{g(X)} \cdot \overline{h(X)}.$$

Note que, temos que provar que $\overline{h(X)} = \overline{0}$.

Como $f(X)$ não divide $g(X)$, então $\overline{g(X)} \neq \overline{0}$ em $K[X]_{f(X)}$. Logo, existe $a(X) \in K[X]$ tal que $\overline{1} = \overline{a(X)} \cdot \overline{g(X)}$. Multiplicando a igualdade destacada acima por $\overline{a(X)}$, concluímos o que queremos provar. □

1.5 Aritmética em domínios

Sejam D um domínio e $a, b \in D$ não nulos. Dizemos que *a divide b*, e escrevemos $a|b$ se existir $c \in D$ tal que $b = ac$. Essa é chamada uma *fatoração* de b. Claramente, se $u \in D$ é um elemento invertível, então $u|b$ para todo $b \in D$, de modo que uma fatoração $b = uc$ é chamada fatoração trivial. Um elemento $a \in D \setminus \{0\}$, não invertível, é chamado *irredutível* se só admitir fatoração trivial. Por outro lado, dizemos que $a \in D \setminus \{0\}$ é um *elemento primo* se a não for invertível e se $a|bc$ implicar que $a|b$ ou $a|c$. Em primeiro lugar vejamos como se relacionam tais conceitos.

Proposição 6. *Sejam D um domínio e $a \in D$. Se a é um elemento primo, então a é irredutível.*

Prova: Seja $a = bc$ uma fatoração de a. Devemos mostrar que tal fatoração é trivial. Como $a|bc$ e a é primo, então $a|b$ ou $a|c$. Sem perda de generalidade, digamos que $a|b$, ou seja, $b = ad$. Daí $a = acd$ e, como $a \neq 0$, obtemos $cd = 1$, implicando que c é invertível. □

Exemplo 14. Considere o domínio $D = \mathbb{Z}[\sqrt{5}] = \{a + b\sqrt{5} | a, b, \in \mathbb{Z}\}$. Considere as seguintes fatorações do elemento $4 = 4 + 0\sqrt{5}$.

$$(2 + 0\sqrt{5})(2 + 0\sqrt{5}) = 4 + 0\sqrt{5} = (1 + \sqrt{5})(-1 + \sqrt{5}).$$

É fácil verificar que $2 \in D$ é irredutível. Por outro lado, $2|(1+\sqrt{5})(-1+\sqrt{5})$, mostrando que 2 não é primo em D.

Em ambientes onde vale uma relação tipo Bézout temos a recíproca. Se o domínio D possuir algum tipo de algoritmo euclidiano, então é possível fazer um algoritmo estendido de Euclides, que é uma relação de Bézout e concluir a equivalência entre os conceitos de primo e irredutível. Como vimos no Exemplo 14 existe uma relação entre a não equivalência desses conceitos e a possibilidade de múltiplas fatorações não equivalentes para um mesmo elemento.

Dizemos que D é um *domínio fatorial* se todo elemento $a \in D$ possuir uma fatoração como produto de elementos irredutíveis, unicamente determinada, a menos de ordem dos elementos e de associados.

Proposição 7. *Seja D um domínio e considere que em D todos os elementos podem ser escritos como um produto finitos de elementos irredutíveis. Então todo elemento irredutível de D é primo se, e somente se, D é um domínio fatorial.*

PROVA: Sejam D um domínio fatorial e $a \in D$ um elemento irredutível. Vamos mostrar que a é primo. Suponha que $a|bc$, isto é, existe $d \in D$ tal que $bc = ad$. Escrevendo a fatoração em irredutíveis de b e c e usando a unicidade da fatoração, é fácil ver que $a|b$ ou $a|c$, uma vez que a é irredutível.

Reciprocamente, suponha que todo elemento irredutível de D seja primo. Como cada elemento de D possui fatoração em irredutíveis, só nos resta mostrar que tal fatoração é única, a menos de ordem e de associados.

Suponha que $a \in D$ possui duas fatorações em irredutíveis

$$p_1 \dots p_n = a = q_1 \dots q_m.$$

Como em D todo irredutível é primo, p_1 é primo. Como $p_1|q_1 \dots q_m$ podemos supor, sem perda de generalidade, que $p_1|q_1$. Por hipótese q_1 é irredutível e assim p_1 e q_1 são associados.

Supondo $n \leqslant m$ e usando o argumento anterior para cada um dos irredutíveis $p_1, \dots, p_n$, obtemos $p_1 = q_1, \dots, p_n = q_n$. Agora, queremos concluir que $m = n$. Se este não for o caso, obtemos que $m > n$ e que

$$q_{n+1} \dots q_m = u.$$

Nessa fatoração u é um invertível. Isso implica que os q_j são invertíveis para $j = n + 1, \dots, m$ que é um absurdo. Logo concluímos que $m = n$ e a unicidade segue. □

Lema 2 (Gauss)**.** *Sejam D um domínio, K seu corpo de frações e $F \in D[X]$ um polinômio irredutível. Então $F \in K[X]$ é irredutível.*

PROVA: Ver Lequain e Garcia (2018, Lemma II.3.6, Pg. 54). □

Teorema 9 (Teorema de Gauss)**.** *Seja D um domínio fatorial. Então $D[X]$ é fatorial. Em particular, se K é um corpo, então $K[X_1, \dots, X_n]$ é um domínio fatorial.*

PROVA: Ver Lequain e Garcia (ibid.) Página 48 Teorema II.3.1. □

1.6 Corpos finitos

Vimos na Seção 1.3 que, para cada número primo positivo p, o conjunto $\mathbb{Z}_p$ é um corpo, que possui p elementos.

Nesta seção descreveremos os possíveis corpos finitos. Para isto, precisaremos estudar a estrutura destes corpos.

Definição 1. Seja K um corpo. A *característica* de K é o menor inteiro não negativo n tal que $0 = n \cdot 1 = 1 + \dots + 1$, n vezes.

Proposição 8. *Se K é um corpo finito, então a característica de K é um número primo.*

PROVA: Suponhamos que a característica n de K não seja um número primo. Então, $n = n_1 n_2$, com $1 < n_i < n$ para $i = 1, 2$. Como $0 = n \cdot 1 = n_1 n_2 \cdot 1 = (n_1 \cdot 1) \cdot (n_2 \cdot 1)$ e K é um domínio de integridade, temos que $n_1 \cdot 1 = 0$ ou $n_2 \cdot 1 = 0$. Entretanto, isto contradiz a minimalidade de n. □

Sejam $F \subseteq K$ dois corpos. Dizemos que F é um *subcorpo* de K quando as operações de F forem as induzidas pelas operações de K.

Lema 3. *Seja K um copro finito de característica p. Então K possui um subcorpo isomorfo a $\mathbb{F}_p$.*

PROVA: Primeiramente, observamos que a aplicação

$$\begin{array}{ccc} \mathbb{Z}_p & \to & K \\ \overline{n} & \mapsto & n \cdot 1 \end{array}$$

independe do representante n da classe $\overline{n}$. De fato, se $\overline{n} = \overline{m}$, então $n - m = qp$ para algum $q \in \mathbb{Z}$ e, portanto, $n \cdot 1 - m \cdot 1 = (n - m) \cdot 1 = qp \cdot 1 = q(p \cdot 1) = 0$.

Verifica-se facilmente que esta aplicação é um homomorfismo. Desta forma, basta verificarmos que ele é injetiva para que a sua imagem seja um corpo contido em K.

Para verificar a injetividade, suponhamos que $n \cdot 1 = 0$ com $0 \leqslant n \leqslant p - 1$. Então, pela minimalidade da característica, obtemos que $n = 0$. □

Teorema 10. *Seja K um corpo finito de característica p. Então K é um espaço vetorial sobre $\mathbb{F}_p$ de dimensão finita, digamos m, e contém p^m elementos.*

PROVA: Pelo Lema 3, temos que K é um espaço vetorial sobre $\mathbb{F}_p$. Sendo K finito, temos que ele possui dimensão finita como $\mathbb{F}_p$-espaço vetorial, digamos m.

Seja $\{c_1, \ldots, c_m\}$ uma base de K. Então, todo elemento de K se escreve de forma única como $a_1 c_1 + \cdots + a_m c_m$, onde $a_1, \ldots, a_m \in \mathbb{F}_p$. Portanto, uma simples contagem nos dá que a quantidade de elementos de K é p^m. □

Antes de lidar com a existência e a unicidade de corpos finitos teremos que observar as suas estruturas de corpos de decomposição de polinômios, isto é, corpos onde certos polinômios possuem todas as raizes.

Lema 4. *Seja K um corpo finito com q elementos. Para todo $\alpha \in K^* = K \setminus \{0\}$, temos que $\alpha^{q-1} = 1$.*

PROVA: Como K é um domínio de integridade, temos que a aplicação

$$\begin{array}{ccc} K^* & \to & K^* \\ a & \mapsto & \alpha a \end{array}$$

é injetiva. Como o domínio e o contradomínio são K^*, que é finito com $q-1$ elementos, segue que esta aplicação é uma bijeção. Desta forma, $K^* = \{a_1, \ldots, a_{q-1}\} = \{\alpha a_1, \ldots, \alpha a_{q-1}\}$. Portanto

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

o que mostra o Lema. □

Corolário 2. *Seja K um corpo finito com q elementos. Então $X^q - X = \prod_{\alpha \in K}(X - \alpha)$ em $K[X]$.*

Proposição 9. *Sejam K um corpo de característica p, $a, b \in K$ e n um inteiro positivo. Então, $(a+b)^{p^n} = a^{p^n} + b^{p^n}$.*

PROVA: Pela fórmula do binômio de Newton, temos que

$$(a+b)^p = a^p + \sum_{i=1}^{p-1} \begin{pmatrix} p \\ i \end{pmatrix} a^{p-i} b^i + b^p.$$

Como $\begin{pmatrix} p \\ i \end{pmatrix}$ é um múltiplo de p, obtemos que $\begin{pmatrix} p \\ i \end{pmatrix} = 0$ em K e, portanto, $(a+b)^p = a^p + b^p$. Agora, o resultado segue por indução em n. □

Exemplo 15. Para criar um elemento num corpo finito primo $\mathbb{F}_p$ basta escolher um número primo p e usar o $Mod(a, p)$:

```
?   isprime(347516038475106384750634 87)
%1  = 0
?   p = nextprime(347516038475106384750634 87)
%2  = 347516038475106384750635 03
?   Mod(4, p)
%3  = Mod(4, 347516038475106384750635 03)
```

Podemos encontrar um polinômio irredutível P de grau n em $\mathbb{F}_p$ usando $ffinit(p, n)$:

```
?   P = ffinit(p, 2)
%4  = Mod(1, 347516038475106384750635 03) * x^2 + Mod(1, 34751603847
      510638475063503) * x + Mod(1, 347516038475106384750635 03)
```

Desta forma obtemos o polinômio $P(X) = X^2 + X + \overline{1}$ em $\mathbb{F}_p$ para $p = 34751603847510638$
Podemos confirmar que P é irredutível:

```
?   polisirreducible(P)
%5  = 1
```

1.6.1 Existência e unicidade de corpos finitos

Primeiramente vamos estudar a existência dos corpos finitos.

Proposição 10. *Seja k um corpo arbitrário e $f(X) \in k[X]$ um polinômio irredutível. Então, existe um corpo K, contendo k, e $\alpha \in K$ talque $f(\alpha) = 0$.*

Prova: Vimos no Teorema 8 que $K = k[X]_{f(X)}$ é um corpo. Consideremos o homomorfismo sobrejetivo

$$\begin{array}{rccc} \phi & : & k[X] & \to & \overline{K} \\ & & g(X) & \mapsto & \overline{g(X)} \end{array} .$$

Note que $\phi(k)$ é um subcorpo de K, isomorfo a k. Para isto, basta ver que ϕ, restrito a k, é injetiva. De fato, dado $c \in k$ tal que $\overline{c} = \overline{0}$, então $c = f(X)g(X)$ para algum $g(X) \in k[X]$. Analisando os graus desta última igualdade, vemos que a única possibilidade para c só pode ser $c = 0$.

Para finalizar, considerando $\alpha = \overline{X} \in K$, obtemos que $\overline{0} = \overline{f(X)} = f(\overline{X}) = f(\alpha)$ em K. □

Denotaremos por $k(\alpha)$ o corpo $k[X]_{f(X)}$ da prova da Proposição anterior, onde $\alpha = \overline{X}$.

Proposição 11. *Seja k um corpo e $f(X)$ um polinômio irredutível em $k[X]$. Se o grau de $f(X)$ é m, então $k(\alpha)$ é um espaço vetorial sobre k com base $\{1, \alpha, \ldots, \alpha^{m-1}\}$.*

Prova: Como vimos na Seção 1.4, cada classe $\overline{g(X)} \in k[X]_{f(X)} = k(\alpha)$ é da forma $\overline{g(X)} = \overline{c_0 + c_1X + \cdots + c_{m-1}X^{m-1}} = c_0 + c_1\overline{X} + \cdots + c_{m-1}\overline{X}^{m-1} = c_0 + c_1\alpha + \cdots + c_{m-1}\alpha^{m-1}$. Desta forma, temos que $\{1, \alpha, \ldots, \alpha^{m-1}\}$ gera $k(\alpha)$ como k-espaço vetorial.

Para mostrar que são linearmente independentes, se $0 = c_0 + c_1\alpha + \cdots + c_{m-1}\alpha^{m-1} = \overline{c_0 + c_1X + \cdots + c_{m-1}X^{m-1}}$, então $c_0 + c_1X + \cdots + c_{m-1}X^{m-1} = f(X)g(X)$, para algum $g(X) \in k[X]$. Comparando os graus nesta última igualdade obtemos que a única possibilidade é de teremos $c_0 + c_1X + \cdots + c_{m-1}X^{m-1} = 0$ □

Corolário 3. *Se $f(X)$ é um polinômio irredutível de grau m em $\mathbb{F}_p[X]$, então $\mathbb{F}_p(\alpha)$ é um corpo finito com p^m elementos.*

Este Corolário, juntamente com o Teorema 10, reduz a busca sobre a existência de corpos finitos na busca sobre a existência de polinômios irredutíveis em $\mathbb{F}_p[X]$ de grau arbitrário.

Para isto, consideremos $F_d(X)$ o produto de todos os polinômios mônicos e irredutíveis em $\mathbb{F}_p[X]$ de grau d.

Teorema 11. $X^{p^m} - X = \prod_{d|m} F_d(X)$ *em* $\mathbb{F}_p[X]$.

Antes de provar este Teorema, faremos dois lemas que ajudarão nesta prova.

Lema 5. *Seja F um corpo. Então $X^l - 1$ divide $X^s - 1$ em $F[X]$ se e só se l divide s.*

PROVA: Se $s = ql$, então $X^s - 1 = (X^l - 1)((X^l)^{q-1} + \cdots + (x^l) + 1)$. Agora, se $s = ql + r$ com $0 \leqslant r < l$, escrevemos

$$X^s - 1 = X^r(X^{ql} - 1) + X^r - 1.$$

Desta forma, se $X^l - 1$ divide $X^s - 1$, então $X^l - 1$ divide $X^r - 1$. Comparando os graus destes dois últimos polinômios, vemos que a única possibilidade para isto acontecer é $r = 0$. □

Lema 6. *Seja a um inteiro positivo. Então $a^l - 1$ divide $a^s - 1$ se e só se l divide s.*

PROVA: Aprova deste Lema é a cópia da prova do Lema acima, trocando X por a. □

PROVA DO TEOREMA 11: Primeiramente, observe que se $f(X)$ é um polinômio não constante que divide $X^{p^m} - X$ em $\mathbb{F}_p[X]$, então $f(X)^2$ não pode dividir $X^{p^m} - X$. De fato, caso contrário, $X^{p^m} - X = f(X)^2 g(X)$. Derivando esta igualdade, obtemos que $-1 = 2f(X)f'(X)g(X) + f(X)^2 g'(X)$, o que implica que $f(X)$ divide 1. Porém, isto contradiz o fato de $f(X)$ ser não constante.

Segue, do que observamos acima, que o polinômio $X^{p^m} - X$ não possui fatores múltiplos na sua fatoração em $\mathbb{F}_p[X]$. Desta forma, basta provarmos que dado um polinômio $f(X)$ mônico, irredutível e de grau $d \leqslant m$ em $\mathbb{F}_p[X]$, então

$$f(X) \text{ divide } X^{p^m} - X \iff d \text{ divide } m.$$

Se $X^{p^m} - X = f(X)g(X)$, então $\alpha^{p^m} - \alpha = f(\alpha)g(\alpha) = 0$ em $\mathbb{F}_p(\alpha)$. Agora, dado um elemento qualquer β de $\mathbb{F}_p(\alpha)$, escrevemos $\beta = a_0 + a_1\alpha + \cdots + a_m\alpha^{m-1}$, com $a_0, \ldots, a_m \in \mathbb{F}_p$. Pela Proposição 9 e pelo Lema 4, aplicado ao corpo $\mathbb{F}_p$, obtemos que

$$\beta^{p^m} = a_0 + a_1\alpha^{p^m} + \cdots + a_m\left(\alpha^{p^m}\right)^{m-1} = a_0 + a_1\alpha + \cdots + a_m\alpha^{m-1} = \beta.$$

Sendo assim, temos que todos os elementos de $\mathbb{F}_p(\alpha)$ são raizes de $X^{p^m} - X$. Como os elementos de $\mathbb{F}_p(\alpha)$ são todas as raizes de $X^{p^d} - X$, obtemos que $X^{p^d} - X$ divide $X^{p^m} - X$. Isto implica que $X^{p^d-1} - 1$ divide $X^{p^m-1} - 1$ e, pelos dois últimos Lemas, obtemos que d divide m.

Suponhamos agora que d divide m. Como $\mathbb{F}_p(\alpha)$ é um corpo com p^d elementos temos, pelo Lema 4, que α é raiz de $X^{p^d} - X$.

Note que, isto implica que $f(X)$ divide $X^{p^d} - X$. De fato, considere

$$n = \min\{\text{grau } g(X) | g(\alpha) = 0\}.$$

Observamos que $n \leqslant d$, já que, $f(\alpha) = 0$ e grau $f(X) = d$. Afirmamos que $n = d$, pois, caso contrário, existiria $g(X)$ anulando α, com grau $n < d$. Pelo Teorema 4, aplicado em

$\mathbb{F}_p[X]$, $f(X) = q(X)g(X) + r(X)$ com $0 \leqslant \operatorname{grau} r(X) < n$. Como $r(\alpha) = 0$ temos, pela minimalidade de n, que $r(X) = 0$, o que contraria o fato de $f(X)$ ser irredutível. Repetindo a argumentação envolvendo o Teorema da Divisão Euclidiana acima, trocando $f(X)$ por $X^{p^d} - X$ e $g(X)$ por $f(X)$, obtemos que $f(X)$ divide $X^{p^d} - X$.

Como d divide m temos, pelos dois Lemas acima, que $X^{p^d-1} - 1$ divide $X^{p^m-1} - 1$, isto é, $X^{p^d} - X$ divide $X^{p^m} - X$. Desta forma, $f(X)$ divide $X^{p^m} - X$. □

Definimos por $I(d)$ o número de polinômios mônicos e irredutíveis de grau d em $\mathbb{F}_p[X]$.

Corolário 4. $p^n = \sum_{d|n} d\,I(d)$

Prova: Basta calcular os graus na igualdade polinomial no último Teorema. □

Note que, pelo Corolário 3, basta provarmos que $I(m) > 0$ para provarmos que existe um corpo finito com p^m elementos.

Teorema 12. *Para cada número inteiro positivo m, existe ao menos um polinômio mônico e irredutível com grau m em* $\mathbb{F}_p[X]$.

Prova: Primeiramente, para $m = 1$ é verdade, já que, X é um tal polinômio.

Consideremos agora $m > 1$ e $1 < d_1 < \cdots < d_s < m$ os divisores de m. Pelo corolário acima, temos que

$$p^m = \sum_{d|m} d\,I(d) = \sum_{i=1}^{s} d_i\,I(d_i) + m\,I(m) \leqslant \sum_{i=1}^{s} \left(\sum_{d|d_i} d\,I(d) \right) + m\,I(m) =$$

$$\sum_{i=1}^{s} p^{d_i} + m\,I(m) < \sum_{i=0}^{d_s} p^i + m\,I(m) = \frac{p^{d_s+1} - 1}{p - 1} + m\,I(m) < p^{d_s+1} + m\,I(m),$$

o que implica que $m\,I(m) > p^m - p^{d_s+1}$. Pela definição de d_s, temos que $m = a d_s$, com $a > 1$. Logo, $d_s = \frac{m}{a} \leqslant \frac{m}{2}$, que implica que,

$$m\,I(m) > p^m - p^{d_s+1} \geqslant p^m - p^{\frac{m}{2}+1}.$$

Como $m \geqslant 2$, temos que $m \geqslant \frac{m}{2} + 1$ e, consequentemente, $I(m) \geqslant 1$. □

Corolário 5. *Seja p um número primo positivo. Então, para cada inteiro positivo m, existe um corpo finito com* p^m *elementos*

Agora vamos investigar a unicidade de tais corpos.

Teorema 13. *Dois corpos finitos com o mesmo número de elementos são isomorfos.*

PROVA: Seja K um corpo finito de característica p. Vimos no Teorema 10 que K é um espaço vetorial sobre $\mathbb{F}_p$ de dimensão finita, digamos m.

Pelo Teorema 12, podemos considerar $f(X) \in \mathbb{F}_p[X]$ mônico, irredutível e de grau m. Pelo Teorema 11, temos que $f(X)$ divide $X^{p^m} - X$.

Pelo Corolário 2 e pelo fato que $X^{p^m} - X$ não possui fatores múltiplos, como vimos no início da prova do Teorema 11, existe $\beta \in K$ raiz de $f(X)$. Pelo mesmo argumento do final da prova do Teorema 2, temos que $f(X)$ é um polinômio irredutível de grau mínimo anulando β. Desta forma, os elementos $1, \beta, \ldots, \beta^{m-1}$ são linearmente independentes sobre $\mathbb{F}_p$ e, logo, formam uma base de K como espaço vetorial sobre $\mathbb{F}_p$.

Por outro lado, a Proposição 11 nos diz que $\mathbb{F}_p[X]_{f(X)} = \mathbb{F}_p(\alpha)$ é um espaço vetorial sobre $\mathbb{F}_p$ com base $\{1, \alpha, \ldots, \alpha^{m-1}\}$.

Guiados pelo isomorfismo natural de espaços vetoriais que envia uma base na outra, definimos a seguinte aplicação.

$$\begin{array}{rccc} \phi & : & \mathbb{F}_p[X]_{f(X)} & \to & K \\ & & \overline{g(X)} & \mapsto & g(\beta) \end{array}$$

Esta aplicação está bem definida, já que, $\overline{g(X)} = \overline{h(X)}$ nos diz que $g(X) - h(X) = f(X)q(X)$. Logo, $g(\beta) - h(\beta) = f(\beta)q(\beta) = 0$.

Concluímos facilmente, pela definição, que ϕ é um homomorfismo de anéis. Para mostrar a sobrejetividade, observamos que todo elemento de K é da forma $a_0 + a_1\beta + \ldots + a_{m-1}\beta^{m-1} = \phi\left(\overline{a_0 + a_1X + \ldots + a_{m-1}X^{m-1}}\right)$.

Para verificar a injetividade, consideramos $\overline{g(X)} \neq \overline{0}$ no corpo $\mathbb{F}_p[X]_{f(X)}$. Como existe $\overline{g(X)}^{-1} \in \mathbb{F}_p[X]_{f(X)}$, obtemos que $1 = \phi(\overline{1}) = \phi\left(\overline{g(X)} \cdot \overline{g(X)}^{-1}\right) = \phi\left(\overline{g(X)}\right) \cdot \phi\left(\overline{g(X)}^{-1}\right)$. Sendo assim, $\phi\left(\overline{g(X)}\right) \neq \overline{0}$. □

Se m é um inteiro positivo, denotaremos por $\mathbb{F}_{p^m}$ o único *corpo finito*, a menos de isomorfismos, com p^m elementos.

1.7 O fecho algébrico de um corpo

Seja K um corpo. Dizemos que K é *algebricamente fechado* quando todo polinômio não constante em $K[X]$ possui raiz em K.

Exemplo 16. O corpo dos números reais $\mathbb{R}$ não é algebricamente fechado, pois $X^2 + 1$ não possui raiz em $\mathbb{R}$.

Teorema 14. *O corpo dos números complexos $\mathbb{C}$ é algebricamente fechado.*

PROVA: Ver Lang 2008, Exemplo 5, página 272. □

Teorema 15. *Todo polinômio em uma variável, com coeficientes em um corpo algebricamente fechado, é escrito unicamente como o produto de um elemento de K e de potências de polinômios mônicos de grau um.*

Prova: Pelo Teorema 6, basta observar que os únicos polinômios mônicos em $K[X]$ são os polinômios de grau um. De fato, se $g(X)$ é um polinômio em $K[X]$ de grau maior que um, então o fato de K ser algebricamente fechado nos diz que existe uma raiz $\alpha \in K$ de $g(X)$. Desta forma, $g(X) = (X - \alpha)q(X)$, com $q(X)$ de grau maior ou igual a um, ou seja, $g(X)$ não é irredutível. □

Dizemos que o corpo L é uma *extensão* do corpo K quando K for um subcorpo de L.

Dizemos que o corpo L é uma *extensão algébrica* do corpo K quando K for um subcorpo de L e quando todo elemento de L é raiz de algum polinômio não constante em $K[X]$.

Teorema 16. *Seja K um corpo. Então existe um corpo $\overline{K}$ que é uma extensão algébrica de K e é algebricamente fechado.*

Prova: Ver Lang 2008, Corolário 2.6, página 232. □

O corpo $\overline{K}$ do teorema acima é chamado de *fecho algébrico* de K.

Exemplo 17. $\overline{\mathbb{R}} = \mathbb{C}$ (ver ibid., Exemplo, página 235.)

Teorema 17. $\overline{\mathbb{F}_p} = \bigcup_{m \in \mathbb{N}} \mathbb{F}_{p^m}$.

Prova: Primeiramente, vamos provar que $K = \bigcup_{m \in \mathbb{N}} \mathbb{F}_{p^m}$ um corpo. De fato, dados $a, b \in K$, temos que $a \in \mathbb{F}_{p^m}$ e $b \in \mathbb{F}_{p^n}$, com m e n inteiros positivos. Como $X^{p^n} - X$ e $X^{p^n} - X$ dividem $X^{p^{mn}} - X$ (ver os dois lemas antes do Teorema 11), temos pelo Corolário 2 que $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^{mn}}$ e $\mathbb{F}_{p^n} \subseteq \mathbb{F}_{p^{mn}}$. Desta forma, $a + b$ e ab pertencem a $\mathbb{F}_{p^{mn}} \subseteq K$ e, caso $a \neq 0$, então $a^{-1} \in \mathbb{F}_{p^{mn}} \subseteq K$.

Note que K é uma extensão algébrica de $\mathbb{F}_p$. De fato, dado $a \in K$, temos que $a \in \mathbb{F}_{p^n}$ para algum n e, portanto, a é raiz do polinômio $X^{p^n} - X$ em $\mathbb{F}_p[X]$.

Por último, temos que provar que K é algebricamente fechado. Para isto, seja $f(X) \in K[X]$ irredutível de grau m. Como $f(X)$ possui finitos coeficientes, obtemos, com o mesmo argumento do primeiro parágrafo desta prova, que $f(X) \in \mathbb{F}_{p^n}[X]$ para algum n. Não é difícil verificar que o Teorema 11 continua valendo se trocarmos p por p^n e $\mathbb{F}_p$ por um corpo finito com p^n elementos. Desta forma, $f(X)$ divide $X^{(p^n)^m} - X = X^{p^{nm}} - X$ em $\mathbb{F}_{p^n}[X]$. Portanto, o Corolário 2, junto com o fato que $X^{p^{nm}} - X$ não possui raizes múltiplas, nos fornece que existe uma raiz de $f(X)$ em $\mathbb{F}_{p^{nm}} \subseteq K$. □

1.8 Grupos

Nesta seção definiremos a estrutura necessária para a definição do problema do logaritmo discreto, que é crucial para definir os sistemas de criptografia que abordaremos.

Seja G um conjunto munido com uma operação

$$\begin{array}{rccc} \cdot & : & G \times G & \to & G \\ & & (a,b) & \mapsto & a \cdot b \end{array}$$

Dizemos que G é um *grupo* quando a operação $\cdot$ satisfizer os seguintes axiomas.

1. (Associatividade) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para todo $a, b, c \in G$;
2. (Existência do Elemento Neutro) Existe $e \in G$ tal que $a \cdot e = e \cdot a = a$ para todo $a \in G$;
3. (Existência do Elemento Inverso) Para cada $a \in G$ existe $a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = e$;

Dizemos ainda que G é um grupo *Abeliano* ou *comutativo* quando $a \cdot b = b \cdot a$ para todo $a, b \in G$.

Exemplo 18. Os seguintes conjuntos são grupos.

1. $\mathbb{Z}$ com a operação de soma;
2. $\mathbb{Z}_m$ com a operação de soma;
3. $K^* = \{a \in K | a \neq 0\}$, com a operação de produto, onde K é um corpo;

Dizemos que um subconjunto H de um grupo G é um *subgrupo* de G, quando a operação de G também faz de H um grupo.

Exemplo 19. Os subgrupos de $\mathbb{Z}$ são da forma $m\mathbb{Z} = \{mz | z \in \mathbb{Z}\}$ com m sendo um inteiro não negativo.

A *ordem* de um grupo G, denotada por $|G|$, é definida pela sua quantidade de elementos.

Seja $a \in G$ e n um número inteiro. Definimos:

$$a^n = \begin{cases} a \cdot \ldots \cdot a & \text{(n vezes)} & \text{se} & n > 0 \\ e & & \text{se} & n = 0 \\ a^{-1} \cdot \ldots \cdot a^{-1} & \text{(-n vezes)} & \text{se} & n < 0 \end{cases} .$$

A *ordem* de $g \in G$ é definida pelo menor inteiro positivo n tal que $g^n = e$, se um tal n existe.

Seja G um grupo e $g \in G$. Note que conjunto $\langle g \rangle = \{g^i | i \in \mathbb{Z}\}$ é um subgrupo de G, chamado de *subgrupo gerado* por g. Um grupo que é gerado por um de seus elementos é

chamado de *grupo cíclico*. Se G for um grupo finito, então $\langle g\rangle = \{g^i | 0 \leqslant i < n\}$, onde n é a ordem de g. Desta forma, a ordem de um elemento é menor ou igual a ordem do grupo.

Note que, se o grupo G for finito e cíclico, com $|G| = n$, então existe $g \in G$ tal que $G = \langle g\rangle$, ou seja, $e, g, g^2, \ldots, g^{n-1}$ são distintos. Portanto a ordem de g é n.

Observamos que os grupos finitos (multiplicativos) $\mathbb{Z}_p^*$ são cíclicos quando p é primo. Isso não é fácil de demonstrar (ver Mullen e Mummert 2007, Teorema 1.2.8). Um gerador de $\mathbb{Z}_p^*$ é chamado uma *raiz primitiva* módulo p.

Na verdade, a relação entre a ordem de um elemento e a ordem do grupo vai além do que a desigualdade observada anteriormente, como podemos ver no seguinte resultado.

Teorema 18 (Lagrange)**.** *Se G é um grupo finito, então a ordem de qualquer um de seus subgrupos divide a sua ordem. Em particular, a ordem de qualquer um de seus elementos divide a sua ordem.*

PROVA: Conferir Lequain e Garcia 2018, Capítulo V, Seção 3. □

1.9 Exercícios

As quetões sinalizadas com ⋆ deverão ser feitas como Pari/GP.

Questão 1. *Verifique a veracidade das seguintes congruências:*

a) $67 \equiv 16\ mod\ 17$

b) $98 \equiv 5\ mod\ 23$

⋆ *c)* $384756023847 \equiv 78952978\ mod\ 374563847$

⋆ *d)* $2375629369 \equiv 10\ mod\ 23$

Questão 2. ⋆ *Estude o anel $\mathbb{Z}_{12}$ determinando os seus elementos invertíveis. Descreva as tablelas da soma e da multiplicação dos elementos de $\mathbb{Z}_{12}$.*

Questão 3. *Mostre que se $a \equiv b\ mod\ m$, então para qualquer c inteiro, tem-se $a \pm c \equiv b \pm c\ mod\ m$ e $ac \equiv bc\ mod\ m$. Quando vale a recíproca em cada caso?*

Questão 4. *Sejam a e b números inteiros que deixam restos 2 e 7, respectivamente, na divisão por 8. Quais os restos da divisão de $a + b$, $a - b$, $b - a$, ab por 8?*

Questão 5. *Encontre as soluções inteiras das equações abaixo:*

a) $327X + 43Y = mdc(327, 43)$.

b) $72634X + 2387Y = mdc(72634, 2387)$.

c) $5394603X + 8347Y = mdc(5394603, 8347)$.

Questão 6. *Mostre que o elemento neutro e o inverso multiplicativo em um anel são únicos.*

Questão 7. *Mostre que as funções a seguir são homormorfismos de anéis e determine o seu núcleo e a sua imagem.*

a) $f : \mathbb{R} \to \mathbb{R}^3$ *dada por* $f(x) = (0, x, x)$.

b) $g : \mathbb{Z} \to M_2(\mathbb{Z})$ *dada por* $g(a) = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$.

Questão 8. *Prove o Lema 1 e a Proposição 2.*

Questão 9. *Seja D um domínio. Mostre que o grau do produto de polinômios em $D[X]$ é a soma dos graus destes polinômios.*

Questão 10. *Mostre que se D é um domínio, então $D[X]$ também é.*

Questão 11. *Mostre que os elementos inversíveis em $D[X]$ correspondem aos polinômios de grau zero dados por elementos inversíveis de D.*

Questão 12. *Determine os os polinômios $q(X)$ e $r(X)$ na divisão euclidiana de $g(X)$ por $f(X)$ nos anéis de polinômios determinados:*

a) $g(X) = x^6 - 5x^3 + 3x^2 - 45$ *e* $f(X) = x^4 + 8x^3 - 3x$ *em* $\mathbb{Z}[X]$.

b) $g(X) = x^4 - \overline{3}x^2 - \overline{5}x - \overline{1}$ *e* $f(X) = \overline{4}x - \overline{2}$ *em* $\mathbb{Z}_7[X]$.

c) $g(X) = 5x^4 - 2x^2 - 3x + 4$ *e* $f(X) = 3x^2 + x - 1$ *em* $\mathbb{Q}[X]$.

Questão 13. *Construa o corpo de frações de um domínio D usando o quociente de $D \times D \setminus 0$ pela relação de equivalência $(a, b) \equiv (c, d)$ se $ad = bc$. O corpo de frações de $D = K[x]$ é $K(x)$.*

Questão 14. *Mostre que todo corpo algebricamente fechado é infinito.*

2 Criptografia

A criptografia trata de esconder informações, pessoais ou corporativas, para garantir privacidade e segurança. Com o passar do tempo a criptografia centrou-se em problemas Matemáticos de difícil solução do ponto de vista computacional. Mesmo a criptografia sendo parte importante de nossas vidas, muitas vezes não nos damos conta do tipo de problema Matemático envolvido naquela operação. Nesse livro trataremos de um caso específico, a criptografia com curvas elípticas. Nessa pequena introdução vamos abordar a questão de um ponto de vista mais amplo e traçar as linhas gerais de uma ideia que será a condutora durante todo o texto, o criptossistema ElGamal.

Sendo parte de nossas vidas, a criptografia, em última instância, está diretamente relacionada à cidadania. A segurança de nossas informações e a nossa privacidade dependem da criptografia. O ativista digital Julian Assange defende que todos devem aprender ao menos o básico sobre criptografia. Ele diz que:

"A criptografia é a forma definitiva de ação direta não violenta."

O melhor livro de introdução à criptografia escrito no Brasil é o Coutinho (2000), que trata prioritariamente de criptografia RSA.

2.1 Introdução

São muitas as situações nas quais se mostra necessário o envio de mensagens, ou outra informação sensível, de forma secreta para certo destinatário. Em períodos de guerra as

mensagens não podem ser decifradas pelo inimigo, mesmo quando encontradas. Questões diplomáticas muitas vezes são sensíveis e portanto, informações trocadas devem ser mantidas confidenciais. Transações financeiras, comerciais e empresariais muitas vezes exigem confidencialidade entre os participantes. Na era da informação, praticamente todos nós, todos os dias compartilhamos informações de forma segura graças aos inúmeros métodos de criptografia desenvolvidos para a internet.

A criptografia (do grego *kryptós*, escondido, e *gráphein*, escrita) consiste no conjunto de técnicas que podem ser utilizadas na codificação de mensagens. Existem registros históricos de utilização da criptografia desde a antiguidade. Em torno de 1900 a.c., no Egito, um escriba usou hieróglifos fora do padrão em uma inscrição. Os espartanos utilizavam um instrumento chamado *cítala* para codificar suas mensagens. A cítala consistia de um bastão de madeira em forma de um prisma sob o qual se enrolava uma tira de couro na qual se escrevia a mensagem. Para decodificá-la era necessária uma cítala igual. Entre 600 a.c. e 500 a.c., a cifra de substituição simples (*Atbash*) era utilizada pelos hebreus. Júlio César (100 a.C.) também utilizava a substituição simples nas trocas de mensagens durante as guerras. Tal cifra ficou conhecida como Cifra de César que se baseava em uma translação nas letras do alfabeto em uma ordem cíclica.

A criptografia de substituição simples pode ser facilmente quebrada utilizando a frequência relativa das letras na língua escrita. Tal sistema de criptoanálise foi desenvolvido pelos árabes. A ideia é simples e consiste em identificar as letras que mais ocorrem na mensagem, substituindo-as pelas letras de maior ocorrência na língua em questão. Decretada a obsolência da cifra de substituição simples, começaram a surgir sistemas criptográficos poli-alfabéticos. O mais conhecido sistema poli-alfabético é chamado cifra de Vigenère, apesar de ter sido descoberto anteriormente por Giovan Battista Bellaso. Ela funciona como uma sucessão de cifras de César e consiste em utilizar uma frase de transição como chave, de forma que cada letra dessa chave indique qual a translação sera utilizada. Inicialmente descrita como indecifrável, mostrou-se mais tarde ser quebrável por um método similar ao de análise de frequência, desenvolvido por Charles Babbage. Durante a segunda guerra mundial o método poli-alfabético mais poderoso a ser utilizado pelos exércitos da Alemanha nazista para encriptação se utilizava de uma máquina chamada *Enigma* que modificava periodicamente a frase de transição, utilizando-se de rotores eletromecânicos. Uma versão anterior desse sistema criptográfico foi quebrada pelos Matemáticos poloneses Marian Rejewski, Jerzy Różycki e Henryk Zygalski. A codificação gerada pela Enigma alemã foi resolvida graças aos esforços de uma equipe de criptoanálise inglesa liderada por Alan Turing, utilizando uma generalização das ideias dos poloneses e um sistema considerado um dos precursores do computador.

Atualmente existem vários algoritmos de criptografia de chave simétrica cuja segurança é garantida. Citamos os mais conhecidos, a saber, DES (Data Encription Standard), 3DES, Blowfish, AES (Advanced Encription Standard), OTP (One Time Pad). Veremos em detalhes o algoritmo OTP, que a pesar de clássico e não mais utilizado, foi a base para a elaboração de outros mais sofisticados ainda em uso. Mostrou-se que todo sistema de criptografia de chave simétrica seguro compartilha as mesmas propriedades do OTP. Por isso ele será o único algoritmo de criptografia de chave simétrica que explicaremos.

Este trabalho trata da criptografia de chave pública, ou criptografia assimétrica. Nela utiliza-se um par de chaves: uma chave pública que é distribuída livremente para todos os correspondentes, via um canal de comunicação não seguro, e uma chave privada, para cada usuário, que deve ser conhecida apenas por ele. A chave pública é utilizada para encriptar a mensagem, enquanto a chave privada é utilizada para decriptá-la. Para esse tipo de sistema criptográfico não é necessária uma combinação prévia de chaves. Os algoritmos mais usados em criptografia de chave pública são RSA (Rivest - Shamir - Adleman), DSA (Digital Signature Algorithm), ECDSA (Eliptic Curve Gigital Signature Algorithm) e GPG (GNU Privacy Guard). A segurança do RSA depende da dificuldade de fatoração de inteiros, enquanto a segurança dos demais algoritmos depende da dificuldade computacional do problema do logaritmo discreto. O RSA, além de ser um sistema criptográfico, também permite desenvolver esquemas de assinatura digital.

Em 1976, Whitfield Diffie e Martin E. Hellman publicaram um artigo, *New Directions in Cryptography* (1976), a fim de definir a criptografia de "chave pública". Eles apresentaram um sistema criptográfico que não necessita de encontro prévio para definições de chaves. Tal sistema criptográfico funciona também para desenvolver um esquema de assinatura digital. O protocolo criptográfico se assegura na dificuldade do problema do logaritmo discreto, mesmo que a comunicação seja estabelecida em um canal inseguro. Eles introduziram a noção *trapdoor one-way function* (traduzindo ao pé da letra "função de um caminho") que significa uma função que é facilmente calculada, porém, o mesmo não ocorre para a sua inversa.

Miller e Koblitz foram os primeiros a propor sistemas criptográficos com curvas elípticas. É importante frisar que eles não criaram uma nova criptografia, apenas implementaram o uso das curvas elípticas em sistemas criptográficos já existentes. Miller (1986) propôs um similar ao Protocolo de Diffie-Hellman, 20% mais rápido, e Koblitz (1987) propôs uma implementação baseada nos criptossistemas de ElGamal e Massey-Omura. A implementação popularizou-se e Koyama, Maurer, Okamoto e Vanstone em (Koyama et al. 1991) apresentaram sistemas análogos ao RSA e três novas "trapdoor one-way functions" baseadas em curvas elípticas.

Assim, um criptossistema com curvas elípticas é um termo referente a um análogo de um sistema criptográfico com a implementação destas curvas.

Segundo a Certicom, consultoria especializada em criptografia com curvas elípticas, três tipos de sistemas criptográficos assimétricos (ou de chave-pública) são considerados seguros e eficientes. Esses sistemas são subdivididos com relação ao problema matemático no qual se baseiam, a saber,

1. Sistemas de fatoração de inteiros (IFS), baseados no problema de fatoração de inteiros (IFP). Por exemplo RSA e Rabin-Williams.

2. Sistemas de logaritmo discreto (DLS), baseados no problema do logaritmo discreto (PLD). Por exemplo ElGamal, DAS – US Gov., Diffie-Hellman/KE e Schnorr.

3. Sistemas de curva elíptica (ECDLS), baseados no problema do logaritmo discreto em curvas elípticas (PLDCE). Por exemplo a implementação do ElGamal com curvas elípticas (CCE).

Atualmente as principais aplicações da criptografia com curvas elípticas são o esquema de assinaturas digitais e o blockchain. Este último é usado principalmente em criptomoedas, como o Bitcoin, e é uma tecnologia que faz registro de transações colocando-as em uma cadeia de dados que não pode ser alterada. A criptografia com curvas elípticas também tem aplicações em desenvolvimento para celulares, internet, dentre outras.

2.2 Princípios básicos da criptografia

A criptografia de dados em Sistemas de informação é o estudo e aplicação de técnicas para transmissão e armazenamento seguros de dados, mesmo na presença de terceiros. A criptografia é um mecanismo de segurança e privacidade que torna determinados dados ininteligíveis para pessoas não autorizadas.

Os seguintes termos são usados normalmente quando tratamos de criptografia.

1. **Texto plano (do inglês plain text), também chamado texto simples ou puro**: Se refere ao conjunto de dados que não foi criptografado ou encriptado, sendo portanto inteligível.

2. **Texto cifrado, encriptado ou codificado**: Se refere ao conjunto de dados após o processo de criptografia, ou seja, não inteligíveis.

3. **Chave (do inglês Key)**: Dados ou algoritmos utilizados para encriptar e decriptar o texto plano.

4. **Invasor, intruso ou adversário**: Qualquer pessoa ou Inteligência Artificial que não tem autorização de acesso à informação e tente roubá-la, ter acesso à mesma fazendo-se passar por outra pessoa ou ainda se utilize de algoritmos para tentar quebrar a criptografia.

Os quatro princípios básicos da segurança da informação são os seguintes:

1. **Confidencialidade**: É a forma de garantir que apenas pessoas autorizadas tenham acesso à informação. Geralmente utiliza-se autenticação a fim de garantir a confidencialidade.

2. **Integridade**: Se refere ao conteúdo dos dados armazenados não poderem ser alterados.

3. **Disponibilidade**: Garantir que as pessoas autorizadas possam ter acesso aos dados armazenados a qualquer momento.

4. **Autenticidade**: Identificação do usuário que está enviando, alterando ou tendo acesso à informação.

Existem dois tipos básicos de cifras, as de fluxo e as de bloco. No caso das cifras de fluxo, mais utilizada em criptossistemas de chave simétrica, combina-se cada bit do texto plano com um fluxo de bits aleatório. Nas cifras de bloco a mensagem é dividida em blocos que são criptografados de forma independente, sendo mais utilizada em criptossistemas de chave pública. Nesse curso trabalharemos quase que exclusivamente com criptossistemas de chave pública. Além disso, limitaremos o comprimento das mensagens de acordo com a chave a ser utilizada, de modo que a própria mensagem seja considerada um único bloco.

2.3 Pré-codificação

A primeira parte da codificação de um texto consiste em transformá-lo em dados numéricos. Os computadores só entendem e processam dados numéricos e, além disso, os algoritmos de criptografia se baseiam em operações matemáticas. Sendo assim, se faz necessário o processo de pré-codificação. Historicamente foram utilizados diversos modelos de pré-codificação e, atualmente, um dos mais utilizados se chama ASCII (American Standard Code for Information Interchange).

A tabela ASCII é um código binário de 7 bits que codifica um conjunto de 128 caracteres incluindo letras, maiúsculas e minúsculas, números, sinais de pontuação, sinais matemáticos e sinais de controle. Como um byte possui 8 bits, o bit excedente dos caracteres da tabela ASCII pode ser utilizado de maneira diferente e não unificada. Por exemplo o padrão $UTF8$, utilizado pelos sistemas operacionais Linux e IOS, difere do padrão utilizado pela Microsoft. Trabalharemos com o sistema de bytes, conforme a tabela ASCII, isto é, trabalharemos com 8 bits, sendo que o primeiro bit sempre será 0.

Observe que a tabela ASCII da Figura 2.3 exibe a codificação dos caracteres em numeração decimal e em binária, sendo que os computadores utilizam a numeração binária. Por outro lado, nesse livro vamos utilizar prioritariamente a decimal. A questão é que para nós é muito mais fácil as contas com números em base 10, além do fato da calculadora algébrica que usamos operar somente em base 10. Na codificação decimal todos os caracteres serão associados a números com 3 algarismos. Neste sentido colocaremos um 0 antes de todos os números decimais de 2 dígitos que são associados a caracteres.

Se Alice deseja enviar uma mensagem para Bob, então ela deverá utilizar a tabela ASCII para transformar os caracteres da mensagem em números decimais de 3 dígitos.

Somente para adiantar o procedimento da encriptação e decriptação, observamos que, após usarmos a nossa calculadora algébrica para criptografar a mensagem obtemos um número decimal. O receptor deve decriptar a mensagem, obtendo um novo decimal. Aí será necessário separar de 3 em 3 dígitos para decriptar a mensagem. Para não haver problemas nesta separação, iremos fazê-la da direita para a esquerda. Caso o último bloco tenha apenas dois dígitos, adicionaremos um dígito 0 a esquerda deste bloco.

Exemplo 20. Vamos usar a tabela ASCII para pré-codificar o texto “Café!”. Pela nossa tabela devemos colocar o número associado ao acento agudo antes do número associado a letra “e”. Desta forma, obtemos:

Binário	Decimal	Hexa	Glifo
0010 0000	32	20	
0010 0001	33	21	!
0010 0010	34	22	"
0010 0011	35	23	#
0010 0100	36	24	$
0010 0101	37	25	%
0010 0110	38	26	&
0010 0111	39	27	'
0010 1000	40	28	(
0010 1001	41	29	)
0010 1010	42	2A	*
0010 1011	43	2B	+
0010 1100	44	2C	,
0010 1101	45	2D	-
0010 1110	46	2E	.
0010 1111	47	2F	/
0011 0000	48	30	0
0011 0001	49	31	1
0011 0010	50	32	2
0011 0011	51	33	3
0011 0100	52	34	4
0011 0101	53	35	5
0011 0110	54	36	6
0011 0111	55	37	7
0011 1000	56	38	8
0011 1001	57	39	9
0011 1010	58	3A	:
0011 1011	59	3B	;
0011 1100	60	3C	<
0011 1101	61	3D	=
0011 1110	62	3E	>
0011 1111	63	3F	?

Binário	Decimal	Hexa	Glifo
0100 0000	64	40	@
0100 0001	65	41	A
0100 0010	66	42	B
0100 0011	67	43	C
0100 0100	68	44	D
0100 0101	69	45	E
0100 0110	70	46	F
0100 0111	71	47	G
0100 1000	72	48	H
0100 1001	73	49	I
0100 1010	74	4A	J
0100 1011	75	4B	K
0100 1100	76	4C	L
0100 1101	77	4D	M
0100 1110	78	4E	N
0100 1111	79	4F	O
0101 0000	80	50	P
0101 0001	81	51	Q
0101 0010	82	52	R
0101 0011	83	53	S
0101 0100	84	54	T
0101 0101	85	55	U
0101 0110	86	56	V
0101 0111	87	57	W
0101 1000	88	58	X
0101 1001	89	59	Y
0101 1010	90	5A	Z
0101 1011	91	5B	[
0101 1100	92	5C	\
0101 1101	93	5D	]
0101 1110	94	5E	^
0101 1111	95	5F	_

Binário	Decimal	Hexa	Glifo
0110 0000	96	60	`
0110 0001	97	61	a
0110 0010	98	62	b
0110 0011	99	63	c
0110 0100	100	64	d
0110 0101	101	65	e
0110 0110	102	66	f
0110 0111	103	67	g
0110 1000	104	68	h
0110 1001	105	69	i
0110 1010	106	6A	j
0110 1011	107	6B	k
0110 1100	108	6C	l
0110 1101	109	6D	m
0110 1110	110	6E	n
0110 1111	111	6F	o
0111 0000	112	70	p
0111 0001	113	71	q
0111 0010	114	72	r
0111 0011	115	73	s
0111 0100	116	74	t
0111 0101	117	75	u
0111 0110	118	76	v
0111 0111	119	77	w
0111 1000	120	78	x
0111 1001	121	79	y
0111 1010	122	7A	z
0111 1011	123	7B	{
0111 1100	124	7C	|
0111 1101	125	7D	}
0111 1110	126	7E	~

Figura 2.1: A tabela ASCII

Texto	*C*	*a*	*f*	*´*	*e*	!
Dec	067	097	102	180	101	033

Mas a verdade é que existem tabelas ASCII estendidas que já tem um código numérico para os "novos caracteres", como por exemplo as letras acentuadas ou caracteres de outros idiomas. Usamos o conversor online

https://www.branah.com/ascii-converter

e obtemos o seguinte:

Texto	*C*	*a*	*f*	*é*	!
Dec	067	097	102	233	033

2.4 Criptossistemas de chave privada

A criptografia de chave privada utiliza uma única chave. O emissor utiliza essa chave para encriptar a mensagem enquanto o receptor utiliza a mesma para decripta-la. Também é conhecida como criptografia simétrica uma vez que tanto o emissor quanto o receptor compartilham a mesma chave.

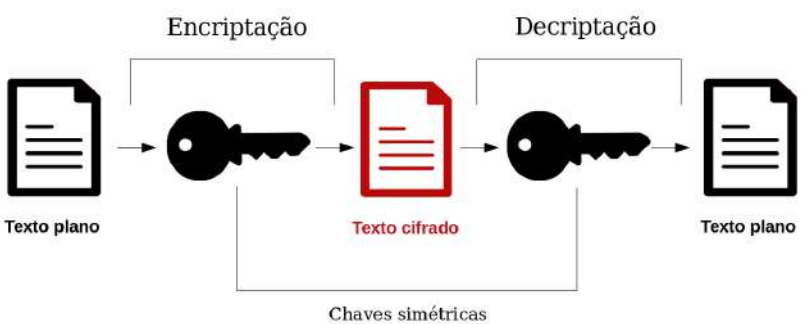


Figura 2.2: Criptosistema de chave privada

A fim de compreender os criptossistemas de chave privada notamos que a cifra depende de um par de algoritmos, sendo E o algoritmo de encriptação e D o algoritmo de desencriptação. Denotamos ainda por k a chave de criptografia, m a mensagem original ou texto plano e por c o texto final encriptado. Naturalmente,

$$c = E(k, m) \text{ e } m = D(k, c).$$

A chamada equação de consistência é dada por:

$$m = D(k, E(k, m)).$$

Vamos inicialmente utilizar a chamada cifra de Vigenère, apenas como ilustração. Como dissemos, esse sistema criptográfico já está obsoleto, mas tem sua importância histórica.

Exemplo 21. Vamos criptografar a frase "CIFRA DE VIGENERE" utilizando a chave BRASIL. Primeiro, repetimos a chave até obtermos o comprimento do texto:

C	*I*	*F*	*R*	*A*		*D*	*E*		*V*	*I*	*G*	*E*	*N*	*E*	*R*	*E*
B	*R*	*A*	*S*	*I*		*L*	*B*		*R*	*A*	*S*	*I*	*L*	*B*	*R*	*A*

Agora usamos a cifra de César para cada letra da seguinte forma: para a letra C usaremos a cifra de César que leva A em B (isto é, $k = 1$); para letra I usaremos a cifra que leva A em R, e assim sucessivamente. Obtemos:

C	*I*	*F*	*R*	*A*		*D*	*E*		*V*	*I*	*G*	*E*	*N*	*E*	*R*	*E*
D	*Z*	*F*	*J*	*I*		*O*	*F*		*M*	*I*	*Y*	*M*	*Y*	*F*	*I*	*E*

Assim a mensagem cifrada é "DZFJI OF MIYMYFIE".

Um dos mais clássicos algoritmos seguros de criptografia de chave simétrica é o OTP (One Time Pad). O OTP foi desenvolvido em 1917 por Gibert Vernan no Bell Labs. Consideramos a mensagem m escrita em base 2 utilizando ASCII e escolhemos para chave k uma sequência de bits aleatória (ou pseudo-aleatória) de mesmo comprimento da mensagem. Os sistemas de criptografia de chave privada atuais foram fortemente inspirados no OTP.

Seja n o comprimento de m e k. Então podemos pensar $m, k \in \mathbb{Z}_2^n$ como uma sequência de elementos de $\mathbb{Z}_2$. Sendo assim, definimos $m \oplus k$ como a soma convencional em $\mathbb{Z}_2^n$ como espaço vetorial sobre $\mathbb{Z}_2$. Esse é o algoritmo de encriptação e em lógica básica, essa operação é chamada XOR, significando o "ou exclusivo" coordenada a coordenada. Assim, definimos:

$$c = E(k, m) = k \oplus m.$$

O algoritmo de desencriptação é igual ao algoritmo de encriptação. Com efeito, fazendo $D(k, c) = k \oplus c$, obtemos:

$$D(k, E(k, m)) = k \oplus E(k, m) = k \oplus (k \oplus M) = (k \oplus k) \oplus m = 2k \oplus m = m.$$

Exemplo 22. Considere que queremos criptografar a palavra "livro". Na tabela ASCII, em base binária ela corresponde a

$$m = 0110110001101001011101100111001001101111$$

de comprimento 40. Seja $k = 0011101100010111000111001111000111100101$ a chave de encriptação. Assim, a mensagem cifrada sera:

$$c = k \oplus m = 0101011101111110011010101000001110001010.$$

Segundo a tabela ASCII este último número binário significa "W∼jâè". É fácil verificar que

$$D(k, E(k, m)) = 01101100011010010111011001110011001101111 = m.$$

2.5 Autenticação e assinatura digital

A autenticação é o processo que busca verificar e validar a identidade digital de um usuário do sistema. Chamamos de reclamante o usuário cuja identidade será verificada e, credenciais, as evidências que o mesmo terá que exibir a fim de comprovar sua identidade. A autenticação pode se referir a uma entidade, como por exemplo, um usuário afirmando ter uma identidade legítima em um sistema, ocorrendo sempre que nos logamos em um sistema. A autenticação pode ainda se referir à origem dos dados, garantindo que determinado conjunto de dados partiu de um certo usuário legítimo. Isso ocorre por exemplo, quando enviamos um e-mail. A autenticação de entidade pode ser unilateral ou multilateral.

Os principais paradigmas de autenticação são: algo que o usuário sabe, como senhas; algo que o usuário possui, como totens; ou ainda algo que o usuário é, geralmente dados biométricos. Tais paradigmas foram apresentados em nível de vulnerabilidade decrescente, sendo as senhas o principal problema de segurança de dados.

Assinatura digital é um esquema matemático para comprovar a autenticidade de um conjunto de dados. As principais características de uma assinatura digital são: a simplicidade de execução por parte do remetente da mensagem; fácil de ser verificada pelo receptor; difícil de ser falsificada por terceiros e, por último, quem a utiliza não pode negar que o fez.

A assinatura digital é utilizada quando, além da necessidade da informação ser mantida em segredo, seja imprescindível identificar exatamente quem foi o remetente da mensagem. Qualquer alteração da mensagem invalida a assinatura digital. A assinatura digital é fortemente utilizada em documentos fiscais eletrônicos. Muitas vezes, dependendo da extensão da mensagem, a assinatura digital pode ser feita em um pequeno número autenticador, que seja em função da mensagem.

Dentre os algoritmos de criptografia mais usados para fins de assinatura digital estão as funções Hash, RSA, DSA, ElGamal e HMAC.

2.6 Criptossistemas de chave pública: as ideias iniciais

A ideia por trás dos criptossistemas de chave pública é que os interessados na comunicação, emissor e receptor não tenham a necessidade de se encontrar antes para combinar a chave de criptografia. Antes de prosseguir com o modelo convencional de criptografia de chave pública, vamos fazer o seguinte exemplo ilustrativo e ingênuo, que não corresponde a um sistema criptográfico real mas que capta a ideia de compartilhar informação sem comunicação prévia.

Exemplo 23. Digamos que Alice e Bob desejam se comunicar secretamente. Para isso, Alice possui um cadeado **A** e uma chave para abri-lo, **a**. Além disso o mesmo ocorre com Bob. Façamos da seguinte maneira: Alice envia para Bob uma caixa com a mensagem, fechada com o cadeado **A** e cuja chave só Alice possui. Bob, ao receber a caixa, coloca também seu cadeado **B** e devolve para Alice a caixa com os dois cadeados **A** e **B**. Desde que a ordem dos cadeados não importe, Alice usa agora a sua chave **a** para abrir seu cadeado **A** e devolve a caixa para Bob apenas com o cadeado **B**. Bob, ao receber a caixa, usa sua chave **b** para abrir o cadeado **B** e, finalmente, poder ler a mensagem que está dentro da caixa.

Uma forma matematicamente eficiente de modelar a situação acima é utilizando a noção de um grupo abeliano agindo em um conjunto.

Sejam $(G, \cdot)$ um grupo abeliano e X um conjunto. Uma *ação* de G em X é uma função

$$*\colon G \times X \to X$$

que a cada par (g, x) associa $g * x \in X$ de modo que

1. $e * x = x$ para todo $x \in X$;
2. Se $a, b \in G$, então $a * (b * x) = (a * b) * x$.

Exemplo 24. No Exemplo 23, com o ponto de vista de ação de grupos, os cadeados e chaves são elementos inversos no grupo. Usando a notação anterior, $A = g, a = g^{-1} \in G$, $B = h$,$b = h^{-1} \in G$ e a mensagem pertence a um conjunto X onde G age. Assim temos:

$Alice$		Bob
x	$\rightarrow$	$g * x$
$h * (g * x)$	$\leftarrow$	$a * x$
$(h * g) * x = (g * h) * x =$		
$= (g * h) * x$	$\rightarrow$	$g^{-1} * [(g * h) * x] =$
		$g^{-1} * = [(g * h) * x] = h * x$
		$h^{-1} * (h * x) = x$

Vamos agora introduzir uma forma um pouco mais realista de modelo de criptossistema assimétrico. A ideia central passa pelo conceito de par de chaves. Assim, cada usuário do sistema terá duas chaves, sendo que uma delas é pública e a outra privada. A chave pública pode transitar em meios inseguros de comunicação, enquanto a chave privada deve ser mantida em segredo por seu proprietário. A chave privada não pode ser derivada da chave pública. Em um criptossistema assimétrico qualquer pessoa pode criptografar uma mensagem utilizando a chave pública do destinatário e, somente este, com sua chave privada, é capaz de decifrar a mensagem.

Exemplo 25 (Troca de Mensagens)**.** Consideremos, novamente, que Alice deseja enviar uma mensagem secreta m para Bob, sem terem se encontrado previamente. Utilizaremos a

ideia do cadeado e da chave, apresentada no Exemplo 23, sob o ponto de vista de ações de grupos. Para isto, consideraremos o cadeado e a sua chave como o par de chaves pública e privada, respectivamente. Desta forma, utilizaremos os pares $(\mathbf{A}, \mathbf{a})$ e $(\mathbf{B}, \mathbf{b})$ para denotar as chaves pública e privada de Alice e Bob, respectivamente. A chave pública serve para encriptar e a secreta para decriptar. Se Alice deseja enviar a mensagem para Bob, ela deve usar a chave pública de Bob para encriptar a mensagem. Nesse caso, Bob receberá $\mathbf{B} * m$ e, com a sua chave secreta, utilizando operações do grupo associado, Bob será capaz de ler a mensagem.

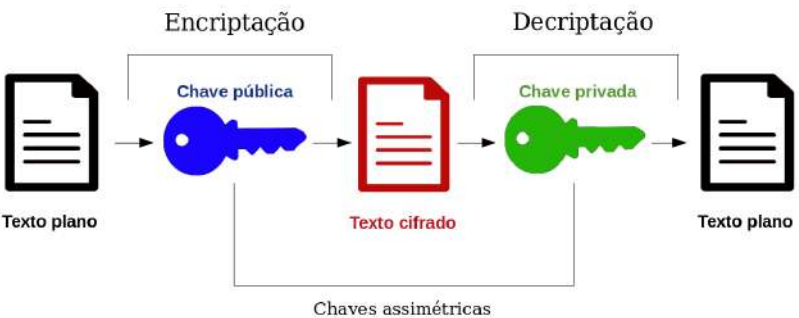


Figura 2.3: Criptosistema de chave pública

A ideia de par de chaves é também muito útil para elaborar esquemas de assinatura digital.

Exemplo 26 (Assinatura Digital)**.** Considere que que Alice envie uma mensagem m, encriptada ou não, para Bob e que ela queira que Bob tenha certeza que ela foi a emissora da mensagem. Para isso ela pode encriptar a mensagem com sua chave privada. Assim Bob receberá $\mathbf{a} * m$ e agora, utilizando a chave pública de Alice ele tem a segurança de que a mensagem foi realmente enviada por ela.

2.7 O problema do logaritmo discreto e o criptossistema ElGamal

Para ser bem realista, os criptossistemas de chave pública utilizados não usam exatamente a ideia inicial apresentada na seção anterior e sim uma variação dela, cuja segurança esteja firmada no problema do logaritmo discreto.

O *problema do logaritmo discreto (PLD)* em um grupo (abeliano) finito G é o seguinte. Dados $a, g \in G$ encontrar, caso exista, um inteiro positivo n tal que

$$a = g^n.$$

Em geral, num grupo finito, é fácil realizar g^n mas é muito difícil realizar o contrário, isto é, resolver o problema do logaritmo discreto.

O criptossistema *ElGamal* é nomeado pelo seu desenvolvedor Taher ElGamal. Ele elaborou este criptossistema em 1985, que é baseado no problema do logaritmo discreto para números inteiros. Este método é comumente aplicado com o grupo multiplicativo $\mathbb{Z}_p^*$ e pode ser generalizado para qualquer grupo cíclico, no qual as operações sejam fáceis e o problema do logaritmo discreto seja intratável. Vejamos o esquema deste criptossistema:

Sejam G um grupo abeliano finito e $g \in G$ um elemento de ordem grande em G. São conhecidos, publicamente, o par (G, g) . Consideremos que G age no espaço X, das mensagens pré-codificadas, via a ação $*$.

- Digamos que Alice deseja enviar uma mensagem para Bob. O primeiro passo é pré-codificar a mensagem, tornando-a um número $m \in X$.

- Alice e Bob escolhem inteiros a e b, como suas respectivas chaves secretas, e divulgam as suas chaves públicas $A = g^a$ e $B = g^b$.

- Alice calcula $C = B^a$ e envia para Bob a mensagem codificada $C * m$.

- Para decifrar a mensagem, Bob calcula $D = A^{-b}$ e $D * [C * m]$. Note que

$$D * (C * m) = A^{-b} * (B^a * m) = (g^{-ab} \cdot g^a b) * m = e * m = m.$$

Note que o problema do logaritmo discreto para o grupo G garante a segurança do sistema. De fato, se um intruso conseguir interceptar os dados trocados pelos usuários Alice e Bob

$$G, g, A, B$$

é preciso ainda resolver o problema do logaritmo discreto

$$g^x = A \text{ ou } g^x = B$$

para poder descobrir a ou b e, consequentemente, decifrar a mensagem m.

Os pré-requisitos num sistema como este, para garantir que isto não ocorra, são: a escolha de um grupo G de ordem muito grande (geralmente $2^{1000} \approx 10^{301}$) e a escolha de g com ordem maior ou igual à metade da ordem do grupo.

Exemplo 27. Vejamos um como utilizar o criptossistema ElGamal para o grupo $G = \mathbb{Z}_p^*$. Utilizaremos o aplicativo livre, para Android, PariDroid 2.9.3.1.4 e começaremos escolhendo um primo de 35 dígitos:

```
(?)   isprime(37462350986754834657874659823456747);
(%1)  = 0
(?)   nextprime(37462350986754834657874659823456747);
(%2)  = 3746235098675483465787465982345684 9
(?)   isprime(37462350986754834657874659823456849);
(%3)  = 1
(?)   p = 37462350986754834657874659823456849;
(%4)  = 37462350986754834657874659823456849
```

Escolhido o primo p, e nomeado para o programa, escolheremos uma mensagem e a pré-codificaremos com ASCII:

C	U	R	V	A	S		E	L	I	P	T	I	C	A	S
67	85	82	86	65	83	32	69	76	73	80	84	73	67	65	83

Assim, a mensagem é $m = 67858286658332697673808473676583$. Precisamos encontrar o elemento g, raiz primitiva de $\mathbb{Z}_p^*$.

```
(?)   g = znprimroot(p);
(%5)  = Mod(17, 37462350986754834657874659823456849)
```

Para completarmos a chave pública de cada indivíduo, precisamos adicionar ao par $(\mathbb{Z}_p^*, g)$ (com $g = 17 (\text{mod } p)$) as chaves g^x, com x a chave privada de cada indivíduo.

- Alice escolhe $a = 3759874672$ e Bob escolhe $b = 5466093453756547$, como chaves secretas, calculam $A = g^a$ e $B = g^b$:

```
(?)   a = 3759874672;
(%6)  = 3759874672
(?)   A = g ^ a;
(%7)  = Mod(76158292201512696920319099622479 86,
      37462350986754834657874659823456849)
(?)   b = 5466093453756547;
(%8)  = 5466093453756547
(?)   B = g ^ b;
(%9)  = Mod(19705776624261971174311130304749901,
      37462350986754834657874659823456849)
```

  e publicam A e B por um canal não seguro de comunicação.

- Alice deseja enviar a mensagem $m \in G$ para Bob. Então ela calcula $C = B^a$, e para codificar a mensagem m, faz $C * m$:

```
(?)     C = B ∧ a;
(%10)   = Mod(169046451959319452704497040917566 35,
        37462350986754834657874659823456849)
(?)     m = 67858286658332697673808473676583;
(%11)   = 67858286658332697673808473676583
(?)     M = C ∗ m;
(%12)   = Mod(285945760366402946167209829376451 78,
        37462350986754834657874659823456849)
```

A mensagem codificada e enviada para Bob é $M = C * m$:

$$M = \overline{285945760366402946167209829376451 78}$$

- Para decifrar a mensagem, Bob calcula $D = A^{-b}$:

```
(?)     D = A ∧ (−b);
(%12)   = Mod(11481422382737038502191092737089223,
        37462350986754834657874659823456849)
```

e então $D * M$:

```
(?)     D ∗ M;
(%14)   = Mod(67858286658332697673808473676583,
        37462350986754834657874659823456849)
```

E desta forma temos a mensagem recuperada

$$m = 67858286658332697673808473676583$$

que decodificada com a tabela ASCII é CURVAS ELIPTCAS.

2.8 Exercícios

Questão 15. *Decifre*

$$ELHQDOGDPDWHPDWLFD$$

utilizando a cifra de César.

Questão 16. *Decifre "Siis mofiw ps mfezlwyfgveg rge olenedgf qoelgi vsffgqstlgi mwqzolgmewtgei gnsq pg qglsqglemg strwnrepg cos isfg gfelqslemg pwi etlsefwi s mofrgi zngtgi" que está cifrada por substituição simples. Se necessário use um analisador de frequência on-line. http://www.numaboa.com.br. para decifrar o texto*

Questão 17. *Decifre*

$$WOOGOFJBDVPYAOSGQMJBQKUGVNEBFWNMLJOS$$

sabendo que utilizamos a cifra de Vigènere com palavra chave de 3 letras.

Questão 18. *Encripte e posteriormente decripte a mensagem*

$$RIO\ DE\ JANEIRO$$

utilizando OTP com a chave

$$k = 11100101001010110110010101011010111110100101011101011100$$

$$11100101010100010001101000010101000101110001011100010101.$$

Questão 19. *Encripte a mensagem "Coloquio" com o sistema ElGamal, escolhendo o primo p, calculando a raiz primitiva em $\mathbb{Z}_p^*$, escolhendo a sua chave privada a e a chave pública do destinatário B.*

3 Curvas algébricas planas

Uma curva algébrica plana, intuitivamente, é um subconjunto do plano cartesiano que pode ser definido a partir de uma equação polinomial do tipo $f(x, y) = 0$. Essa ideia originalmente se deve a René Descartes, quando da introdução de sua Geometria Analítica em *La Geometrie* (Descartes 1954). No clássico filosófico *Discurso do Método (para bem conduzir a própria razão e procurar a verdade nas ciências)*, (Descartes 2002) – um prefácio de suas obras científicas - Descartes comenta o que o levou a introduzir a Geometria Analítica:

> *"com respeito à Análise dos Antigos (Geometria) e à Álgebra dos modernos, ..., a primeira permanece sempre tão adstrita à consideração das figuras que não pode exercitar o entendimento sem fatigar muito a imaginação; e esteve-se de tal forma sujeito, na segunda, a certas regras e certas cifras, que se fez dela uma arte confusa e obscura que embaraça o espírito... por esse meio, tomaria de empréstimo o melhor da Análise geométrica e da Álgebra, e corrigiria todos os defeitos de uma pela outra." - René Descartes*

Deve-se salientar que foi nesse período que o mesmo formalizou grande parte da notação algébrica moderna. Isso explica seus comentários sobre a confusão das regras e cifras na álgebra, que à altura não estavam completamente estabelecidas.

Exemplo 28. Retas, circunferências e cônicas em $\mathbb{R}^2$ são exemplos de curvas algébricas planas.

1. Reta: $\ell = \{(x, y) \in \mathbb{R}^2 | ax + by + c = 0\}$.

2. Circunferência: $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = r^2\}$.

3. Cônicas: $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 | ax^2 + by^2 + cxy + dx + ey + f = 0\}$.

Por outro lado, a Geometria Analítica tem uma limitação que gostaríamos de suprimir, a saber, o plano em questão sempre foi considerado o $\mathbb{R}^2$. Para nossos interesses será adequado generalizar a ideia de plano e de curva plana, permitindo assim, que ao invés de nos restringirmos aos reais, possamos escolher qualquer corpo para definir o plano (afim). Além disso, em algumas situações gostaríamos ainda de incluir um conjunto de pontos especiais, "impróprios " ou "no infinito", construindo assim o denominado Plano Projetivo. O Plano projetivo sobre um corpo é o ambiente unificador para a teoria que pretendemos abordar.

Independente do corpo sobre o qual estivermos trabalhando, quando fizermos uma figura, pensaremos na curva definida sobre $\mathbb{R}$. Isso ajuda nossa intuição.

3.1 Os planos afim e projetivo

Nessa seção vamos tratar do ambiente no qual iremos trabalhar uma versão do problema do logaritmo discreto. A modelagem do sistema de criptografia de curvas elípticas tem como base esse problema aplicado ao grupo de pontos de curvas sobre corpos finitos. Nesse sentido vamos introduzir, de forma mais geral, os conceitos de plano e curva.

Seja K um corpo, a *reta afim* é o conjunto $\mathbb{A}^1_K = K$ e o *plano afim* sobre K é o conjunto

$$\mathbb{A}^2_K = \{(x, y) \mid x, y \in K\}.$$

Quando o corpo no qual estamos trabalhando estiver claro escrevemos apenas $\mathbb{A}^1$ e $\mathbb{A}^2$. Podemos ainda definir, de maneira análoga, o *espaço afim* de dimensão n por $\mathbb{A}^n = \{(x_1, \ldots, x_n) | x_i \in K\}$.

Uma *reta* no plano afim é definida por uma equação linear do tipo

$$\ell = \{(x, y) \in \mathbb{A}^2 | ax + by + c = 0\}$$

com $a, b, c \in K$ e $(a, b) \neq (0, 0)$. Note que tais retas são sempre imagem da reta afim $\mathbb{A}^1$ via uma parametrização (afim). Com efeito, seja $(x_0, y_0) \in \ell$, então todo ponto de ℓ pode ser escrito da forma $x = x_0 + bt$ e $y = y_0 - at$ com $t \in K$ (verifique!).

A fim de definir o plano projetivo, iniciamos com a definição de reta projetiva. Fixado o corpo K considere em $\mathbb{A}^2 \setminus (0, 0)$ a relação de equivalência:

$$v \equiv w \Leftrightarrow v = \lambda w, \text{ com } \lambda \in K^*.$$

A reta projetiva

$$\mathbb{P}^1_K = (\mathbb{A}^2 \setminus (0, 0))/\equiv$$

é o espaço quociente de $\mathbb{A}^2 \setminus (0,0)$ pela relação de equivalência $\equiv$. Quando o corpo no qual estamos trabalhando estiver claro escrevemos apenas $\mathbb{P}^1$. Podemos pensá-la como o conjunto das direções no plano afim, ou equivalentemente, o conjunto das retas passando pela origem.

Em coordenadas, os pontos de $\mathbb{P}^1$, que são classes de equivalência, são denotados por $(x : y)$ para deixar clara a proporcionalidade. Podemos naturalmente identificar uma cópia de $\mathbb{A}^1 \subset \mathbb{P}^1$, fazendo $y \neq 0$, neste caso, $(x : y) = (x/y : 1)$. O único ponto tal que $y = 0$ é o ponto $(x : 0) = (1 : 0)$, chamado ponto no infinto.

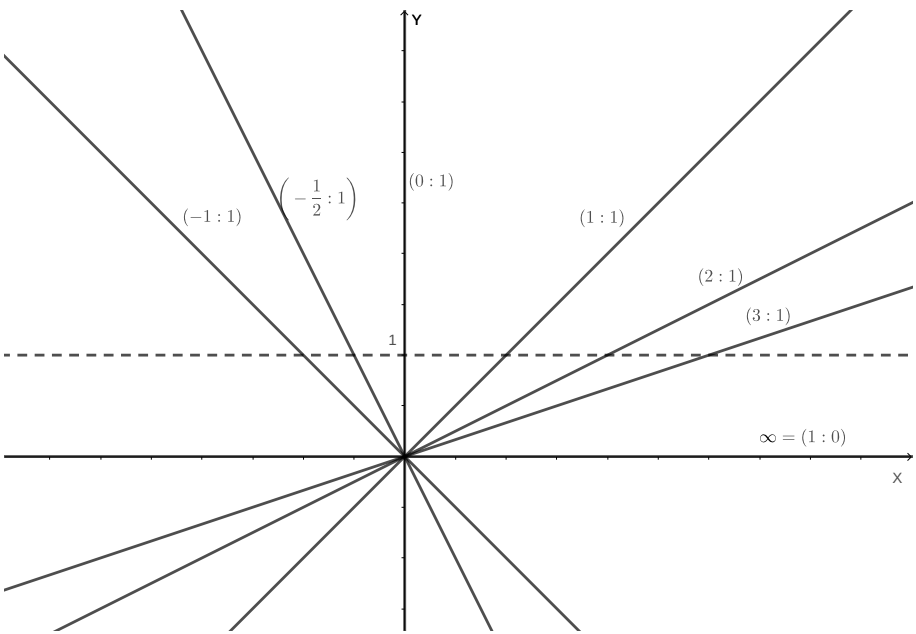


Figura 3.1: A reta projetiva

Exemplo 29. A primeira aparição, para a maior parte dos estudantes, da reta projetiva foi, provavelmente, a esfera de Riemman, ou plano complexo estendido, $\hat{\mathbb{C}}$, isto é, a compactificação do plano complexo, obtida pela adjunção de um ponto "no infinito". Muitas vezes tal processo é feito utilizando a projeção estereográfica da esfera de $\mathbb{R}^3$.
Em coordenadas projetivas um ponto na esfera de Riemman se escreve como $(z : w)$, e, claramente, se $w \neq 0$, podemos escrever $(z : w) = (z/w : 1)$ identificando assim, tais pontos, com os pontos do plano complexo. Por outro lado, se $w = 0$, temos $\infty = (1 : 0)$. Aqui, nesse texto, $\mathbb{P}^1_{\mathbb{C}}$ será pensado como uma reta (de dimensão complexa 1) e não como um plano estendido.

O plano projetivo pode ser visto como o conjunto de todas as direções do espaço afim

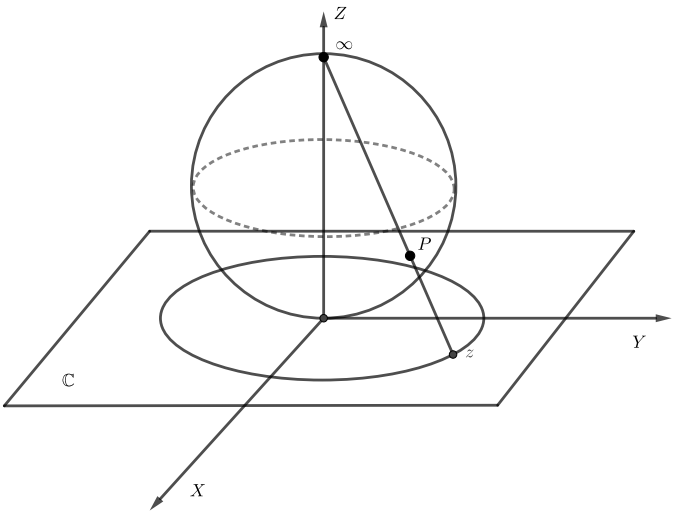


Figura 3.2: A esfera de Riemman

$\mathbb{A}^3$. De forma análoga, considerando em $\mathbb{A}^3 \setminus (0,0,0)$ a seguinte relação de equivalência:

$$v \equiv w \Leftrightarrow v = \lambda w \; \lambda \in K^*$$

definimos o *plano projetivo*

$$\mathbb{P}^2_K = (\mathbb{A}^3 \setminus (0,0,0))/\equiv$$

como o espaço quociente de $\mathbb{A}^3 \setminus (0,0,0)$ pela relação de equivalência $\equiv$. Quando não houver confusão sobre o corpo em questão, escreveremos simplesmente $\mathbb{P}^2$.

Novamente, em coordenadas, $\mathbb{P}^2 = \{(x : y : z) | (x, y, z) \in \mathbb{A}^3 \setminus (0,0,0)\}$. Além disso, se $z \neq 0$, então $(x : y : z) = (x/z : y/z : 1)$, que podemos pensar como coordenadas afins x/z e y/z; estamos assim identificando uma cópia de $\mathbb{A}^2$ em $\mathbb{P}^2$. Reciprocamente, cada ponto do espaço afim com coordenadas (x, y) induz o ponto projetivo $(x : y : 1)$. Note que se $z = 0$, podemos considerar o conjunto $(x : y : 0)$ como uma cópia de $\mathbb{P}^1$, representando a reta "no infinito"; assim:

$$\mathbb{P}^2 = \mathbb{A}^2 \cup \mathbb{P}^1.$$

Na Figura 3.3 destacamos o plano afim α, mergulhado no espaço tridimensional $\mathbb{A}^3$.

$$\alpha = \mathbb{A}^2 = \{(x, y, 1) \in \mathbb{A}^3\} \subset \mathbb{A}^3$$

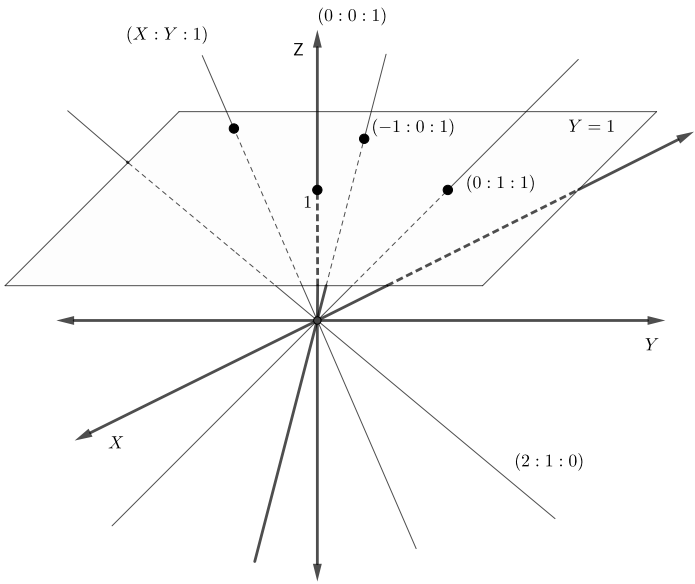


Figura 3.3: O plano projetivo

Cada ponto do plano α determina uma reta passando pela origem e pelo ponto.

Reciprocamente, cada reta pela origem em $\mathbb{A}^3$ que não seja paralela a α, determina um único ponto em α.

Uma *reta* no plano projetivo é definida por uma equação linear homogênea, do tipo;

$$\ell = \{(x : y : z) \in \mathbb{P}^2 | ax + by + cz = 0\},$$

com $a, b, c \in K$ não todos nulos. Neste caso, dizemos que a ℓ é uma reta de equação $aX + bY + cZ = 0$.

Note que, assim sendo, toda reta projetiva é a imagem, via relação de equivalência de um plano pela origem em $\mathbb{A}^3$. Note também que a interseção de uma reta projetiva com o plano afim, definido por $Z = 1$ em $\mathbb{A}^3$, nos fornece uma reta afim de equação $aX + bY + c = 0$, desde que tenhamos $(a, b) \neq (0, 0)$. O caso da reta projetiva de equação $Z = 0$ é especial, representando exatamente o $\mathbb{P}^1$ dos pontos no infinito.

Reciprocamente, toda reta afim no plano, digamos, com equação $aX + bY + c = 0$ induz uma reta projetiva $aX + bY + cZ = 0$. O processo algébrico de passar da equação de uma reta afim à equação correspondente no plano projetivo é chamado homogeneização e o processo contrário deshomogeneização.

3.2 Curvas algébricas planas

Primeiramente vamos observar atentamente as curvas que já conhecemos advindas da geometria analítica, dos cursos de cálculo e da geometria diferencial. Aquelas que chamamos curvas algébricas consistem em um conjunto do tipo

$$C = \{(x, y) \in \mathbb{R}^2 | f(x, y) = 0\}.$$

Nesse caso $f \in \mathbb{R}[X, Y]$ é um polinômio em duas indeterminadas.

Por um lado, muitas das curvas algébricas com as quais nos deparamos eram dadas por uma parametrização e, posteriormente encontramos sua forma implícita. Para essa classe de curvas essa ideia intuitiva não apresenta grandes problemas como veremos a seguir. Claramente, nem todas as curvas dadas por uma parametrização possuem uma equação algébrica implícita, tomemos como exemplo a curva exponencial dada por $x(t) = t$ e $y(t) = e^t$. Por outro lado algumas curvas que possuem parametrizações transcendentais clássicas muitas vezes são algébricas, como o próprio círculo cuja parametrização clássica é $x(t) = \cos(t)$ e $y(t) = \text{sen}(t)$, mas tem equação $X^2 + Y^2 = 1$. O fato relevante é que se tal parametrização for dada por funções racionais, então certamente existe uma equação implícita algébrica. Gostaríamos de salientar ainda que nem toda curva algébrica possui parametrização racional.

Observemos alguns inconvenientes com nossa definição. O primeiro deles é que existem polinômios em duas indeterminadas sobre $\mathbb{R}$ cuja curva associada é o conjunto vazio ou um conjunto finito de pontos, que são considerados casos degenerados.

Exemplo 30.

$$\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = -1\} = \emptyset.$$

$$\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 0\} = \{(0, 0)\}.$$

$$\{(x, y) \in \mathbb{R}^2 | 2x^2 + 3y^2 = 0\} = \{(0, 0)\}.$$

Claramente, polinômios associados em $R = \mathbb{R}[X, Y]$ definem o mesmo conjunto em $\mathbb{R}^2$. Entretanto, existem polinômios que não são associados em R e que ainda assim definem o mesmo conjunto, como mostra o exemplo anterior.

Dado que estamos interessados em generalizar essa ideia para um corpo arbitrário, poderíamos tomar, como primeira, e certamente a mais intuitiva, definição de curva plana (conjuntista) a que segue.

Seja K um corpo. Uma *curva algébrica plana* (conjuntista) definida sobre K consistirá, provisoriamente, num conjunto do tipo

$$C = V(f) := \{(x, y) \in K^2 | f(x, y) = 0\}$$

com $f \in K[X, Y]$.

Como discutimos anteriormente, apesar dessa definição nos trazer uma ideia daquilo que gostaríamos de chamar de curva, ela tem alguns problemas técnicos. Vamos abordar

tais problemas e encontrar uma definição alternativa, um pouco mais técnica, mas que nos será útil no que segue. A primeira solução encontrada é a mesma que foi dada a fim de resolver equações em uma variável, isto é, considerar pontos no fecho algébrico do corpo em questão.

Como vimos na seção 1.7 todo corpo K possui um fecho algébrico $L = \overline{K}$. Dado um polinômio $f \in K[X, Y]$, gostaríamos de pensar em curvas $V(f) \subset \mathbb{A}^2_L$, que dizemos serem definidas sobre K. A proposição a seguir nos garante que, nesse contexto, não temos degenerações.

Proposição 12. *Sejam L um corpo algebricamente fechado e $f \in L[X, Y]$ um polinômio não constante então a curva conjuntista $C = V(f) \subset \mathbb{A}^2_L$ é infinita.*

PROVA: Suponhamos, sem perda de generalidade, que f depende da variável Y. Segue que existe uma infinidade de $a \in L$ tais que o polinômio $f(a, Y)$ é não nulo. Como L é algebricamente fechado, a $f(a, Y) = 0$ possui solução. Isso nos fornece uma infinidade de pontos em C (ver Exercício 14). □

Mais que isso, quando o polinômio for irredutível e o corpo algebricamente fechado, veremos que a classe do polinômio (módulo associados) determina a curva e *vice versa*.

Sejam K um corpo e $f \in R = K[X, Y]$ um polinômio não constante. Sendo R um domínio de fatoração única, podemos escrever

$$f = \lambda f_1^{e_1} f_2^{e_2} \dots f_r^{e_r}. \tag{3.1}$$

Nesta fatoração, os polinômios $f_i \in R$ são irredutíveis e não associados (unicamente determinados a menos de associados). Dizemos que f é *reduzido* se não possui fatores múltiplos, isto é, na fatoração 3.1 tivermos $e_i = 1$ para cada $i = 1, \dots, r$. Claramente

$$V(f) = V(f_1) \cup \dots \cup V(f_r) \subset \mathbb{A}^2_K. \tag{3.2}$$

Com efeito, se $(x, y) \in \mathbb{A}^2_K = K^2$, então:

$$f(x, y) = 0 \Leftrightarrow f_i(x, y) = 0 \text{ para algum } i \in \{1, \dots, r\}.$$

Cada um dos conjuntos $V(f_i)$ representa o que chamamos uma *componente irredutível* da curva (conjuntista). Se, por um lado, classicamente, nos primórdios da geometria algébrica se trabalhava (prioritariamente) com as curvas sobre corpos algebricamente fechados e o foco estava em sua estrutura conjuntista, por outro lado, modernamente, notou-se a importância de trabalhar também com as multiplicidades dos fatores irredutíveis na Equação 3.1 associando assim multiplicidades às componentes irredutíveis da curva. Esse será nosso ponto de vista. Nossa escolha se justifica por diversas razões sendo a definição de grau da curva aquela mais evidente. Além disso temos a seguinte:

Proposição 13. *Sejam L um corpo algebricamente fechado e $f \in L[X, Y]$ um polinômio irredutível. Então a curva conjuntista $C = V(f) \subset \mathbb{A}^2_L$ determina f a menos de associados.*

PROVA: Seja g um outro polinômio irredutível que define C. Sem perda de generalidade podemos supor que f depende de Y. Seja $D = L[X]$ e seja $K = L(X)$ seu corpo de frações. Pelo Lema de Gauss (ver Teorema 2), f e g são irredutíveis em $K[Y]$ (uma vez que são irredutíveis em $D[Y]$). Vamos mostrar que $f|g$ em $K[Y]$ e analogamente, $g|f$, pela simetria do problema.
Suponhamos, por absurdo, que $f \not| g$. Como g é irredutível, segue que $\text{mdc}(f,g) = 1$. Pelo Lema de Bézout, Teorema 2 existem polinômios $f', g' \in K[Y]$ tais que

$$ff' + gg' = 1.$$

Após limpar denominadores temos uma expressão do tipo

$$fp + gq = d \in D[Y]$$

com $p, q \in D[Y]$ e $d \in D$, isto é $d = d(X) \in L[X]$. De nossa hipótese que f depende de Y, segue que existe uma infinidade de $a \in L$ tais que a equação $f(a, Y) = 0$ possui solução. Como por hipótese f e g representam a mesma curva conjuntistica, existe uma infinidade de $a \in K$ tais que $d(a) = 0$, portanto $d = 0 \in L[X]$ e isto é um absurdo. E o resultado segue. □

Note que a hipótese do corpo ser algebricamente fechado foi usada duas vezes na demonstração. Primeiramente para evitar o caso da curva conjuntista ser vazia. Em seguida pelo fato de todo corpo algebricamente fechado ser infinito.

Seja L um corpo algebricamente fechado. Uma *curva algébrica plana afim* definida sobre L corresponde a uma classe de equivalência de um polinômio não constante $f \in L[X, Y]$, módulo associados, isto é, f e g definem a mesma curva se, e somente se, $f = \lambda \cdot g$ com $\lambda \in L^*$.

Se K é um corpo arbitrário e $\overline{K}$ seu fecho algébrico, dizemos que a curva algébrica plana $C \subset \mathbb{A}^2_{\overline{K}}$ é *definida sobre* K se possui alguma equação $f \in K[X, Y]$. O conjunto

$$C(K) = V(f) \subset \mathbb{A}^2_K$$

é chamado conjunto dos *pontos K-racionais* da curva C.

O *grau* de um polinômio $f \in K[X, Y]$,

$$f = \sum a_{ij} X^i Y^j$$

é o máximo dos $i + j$ para os quais $a_{ij} \neq 0$. O *grau* de uma curva é o grau do polinômio que a define.

Retas, cônicas e *cúbicas* são curvas de grau 1, 2 e 3, respectivamente.

No caso em que $f \in K[X, Y]$ é irredutível, diremos que a curva f é *irredutível* sobre K.

Note que pode ocorrer de f ser irredutível sobre K mas não irredutível sobre $\overline{K}$, como $X^2 + Y^2 \in \mathbb{R}[X, Y]$ que é irredutível sobre $\mathbb{R}$ mas se fatora sobre $\mathbb{C}$ em $X^2 + Y^2 =$

$(X + iY)(X - iY)$. No caso em que f é irredutível sobre $\overline{K}$ dizemos que a curva é *geometricamente irredutível*. Em geral, toda curva tem uma decomposição em componentes irredutíveis, unicamente determinadas pela curva via 3.2.

$$C = C_1 \cup C_2 \cup \ldots \cup C_k.$$

Dada uma curva algébrica plana afim f definida sobre K e um subcorpo $k \subset K$, um ponto $(a, b) \in V(f) \cap \mathbb{A}^2_k$ será chamado de ponto k-racional.

Exemplo 31. Seja $K = \mathbb{R}$. No plano afim $\mathbb{A}^2 = \mathbb{R}^2$, a parábola, a elipse, a hipérbole são exemplos de curvas algébricas irredutíveis. Reciprocamente, uma cônica é uma curva definida por um polinômio de grau 2. Se o polinômio não for reduzido, então ele é o quadrado de um polinômio linear, nesse caso representa uma reta dupla. Há ainda casos degenerados em que o conjunto de pontos $\mathbb{R}$ racionais da curva é vazio ou um número finito de pontos, que dizemos serem curvas degeneradas. Verifique que uma cônica redutível mas reduzida consiste, necessariamente, em um par de retas. Verifique que toda cônica irredutível não degenerada representa uma circunferência, uma elipse, uma parábola ou uma hipérbole.

Seja f uma curva plana definida sobre K. Dizemos que f é uma curva racional se existir uma bijeção $\phi : \mathbb{A}^1 \setminus \{p_1, \ldots, p_k\} \to V(f) \setminus \{q_1, \ldots, q_s\}$ tal que ϕ e ϕ^{-1} sejam dadas por funções racionais com coeficientes em K.

As curvas racionais são, para a geometria algébrica, aquelas parametrizáveis. Com efeito, para a geometria algébrica não faz sentido trabalhar com parametrizações transcendentais. Notamos que muitas vezes curvas que classicamente conhecemos uma parametrização transcendental são efetivamente racionais, possuindo, portanto uma parametrização por funções racionais. O primeiro célebre exemplo é o círculo, de equação $X^2 + Y^2 = 1$. Sua parametrização mais popular é aquela dada por cossenos e senos, por outro lado ele possui uma parametrização racional descrita por Fermat , a saber, $x(t) = \frac{2t}{t^2+1}$ e $y(t) = \frac{t^2-1}{t^2+1}$. A técnica desenvolvida por Fermat é chamada método das tangentes e das secantes de Fermat e possui diversas aplicações em Teoria dos Números. Voltaremos a esse exemplo após a definição de reta tangente e a apresentação do método na proposição que segue.

Sejam $C \in \mathbb{A}^2$ uma curva definida sobre K por um polinômio $f \in K[x, y]$ e $P = (x_0, y_0) \in C(K)$ um ponto K-racional. Dizemos que P é *regular* se $f_X(x_0, y_0) \neq 0$ ou $f_Y(x_0, y_0) \neq 0$. Neste caso, definimos a *reta tangente* a C em P como sendo a reta de equação

$$T_P C = f_X(x_0, y_0)(X - x_o) + f_Y(x_0, y_0)(Y - y_0).$$

O ponto P é dito *singular* se

$$f_X(x_0, y_0) = f_Y(x_0, y_0) = 0.$$

A curva C é dita *lisa* se ela não possui ponto singular. A curva C é dita *singular* se, pelo menos, um dos seus pontos for singular.

A *multiplicidade de interseção* de uma reta ℓ com uma curva $C \subset \mathbb{A}^2$ num ponto K-racional $P = (x_0, y_0) \in \ell \cap C$, denotada por $\text{Mult}_P\, \ell \cap C$, é definida como a multiplicidade da raiz $t = 0$ na equação obtida pela substituição de uma parametrização $X = x_0 + aT$ e $Y = y_0 + bT$ de ℓ, na equação de C.

Note que, se P é um ponto simples de f, então o único caso em que $\text{Mult}_P\, \ell \cap C \geqslant 2$ é quando $\ell = T_P C$ é a tangente à C em P. Para ver isto basta ver que f se escreve como

$$f = f_X(x_0, y_0)(X - x_o) + f_Y(x_0, y_0)(Y - y_0) + g, \text{ com grau } g \geqslant 2$$

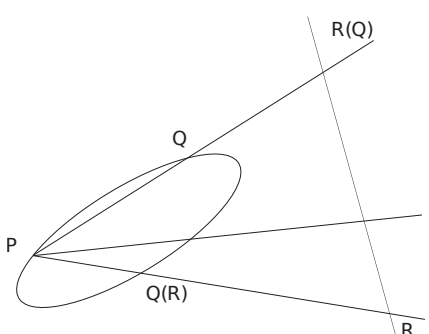
e substituir a parametrização $X = x_0 + f_Y(x_0, y_0)T$ e $Y = y_0 - f_X(x_0, y_0)T$ de $T_P C$. Se $\text{Mult}_P\, \ell \cap C \geqslant 3$, dizemos que a reta tangente é *inflexionária*. Neste caso, dizemos que o ponto P é um ponto de *inflexão* de f.

Proposição 14. *Seja $\mathcal{C} \subset \mathbb{A}^2_K$ uma cônica irredutível definida sobre K que possui algum ponto K-racional e liso. Então $\mathcal{C}$ é racional.*

Demonstração. Sejam $\mathcal{C}$ a cônica, $P \in \mathcal{C}(K)$ e ℓ uma reta que não passa por $P = P_1$. Seja $\widetilde{\ell}$ a reta paralela a ℓ passando por P e seja $\mathcal{C} \cap \widetilde{\ell} = \{P_1, P_2\}$ (não necessariamente distintos). Defina, ainda, $\widetilde{P} = T_P C \cap \ell$. Então a função:

$$\begin{array}{rccc} \phi: & \mathcal{C} \setminus \{P_1, P_2\} & \longrightarrow & \ell \setminus \{\widetilde{P}\} \\ & Q & \mapsto & \overline{PQ} \cap \ell = R(Q) \end{array}$$

está bem definida é racional e inversível.



De fato, para cada $Q \in \mathcal{C} \setminus \{P_1, P_2\}$, a reta $\overline{PQ}$ não é paralela a ℓ (a única paralela a ℓ passando por P é $\overline{P_1 P_2} = \ell$). Assim, a intersecção $\overline{PQ} \cap \ell = R(Q)$ define um ponto, $R(Q)$.
A inversa de ϕ é a seguinte função:

$$\begin{array}{rccc} \phi^{-1}: & \ell \setminus \{\widetilde{P}\} & \longrightarrow & \mathcal{C} \setminus \{P_1, P_2\} \\ & R & \mapsto & \overline{PR} \cap \mathcal{C} = Q(R) \end{array}$$

de fato, para que ϕ^{-1} pudesse ser definida, tivemos que excluir $\widetilde{P}$, caso contrário a reta $\overline{\widetilde{P}P}$ seria tangente a C.

Só nos falta mostrar que ambas são funções racionais, iniciemos com a inversa, ϕ^{-1}. Utilizando uma parametrização da reta ℓ, o ponto $R = R_t = (at + b, ct + d)$ com $a, b, c, d \in K$, fica parametrizado. Assim, a reta PR tem equação paramétrica $x = x_0(t) + v(t)s$ e $y = y_0(t) + u(t)s$ com $s \in K$, para cada $t \in K$, em que as funções $x_0(t), v(t), y_0(t), u(t)$ são lineares. A interseção $PQ \cap C$ nos fornece uma equação quadrática em s com coeficientes lineares em t, tendo uma das raizes dadas por P a outra pode ser obtida pelo coeficiente livre que é o produto das raizes. Assim encontramos uma expressão racional para s que nos fornece $Q(R)$ como função racional. Um raciocínio similar funciona para a ϕ. □

Note que se K for um corpo infinito, isso implica que uma cônica com um ponto K-racional possui uma infinidade de pontos K-racionais. Esse argumento é a primeira encarnação do chamado método das tangentes e das secantes de Fermat, utilizado por este, a fim de encontrar soluções de equações diofantinas não lineares.

Exemplo 32. Um primeiro exemplo de aplicação do método das tangentes e das secantes de Fermat é aquele que já havíamos comentado, da parametrização do círculo. Considere o círculo C de equação $X^2 + Y^2 = 1$ definido sobre $\mathbb{Q}$, e o ponto $\mathbb{Q}$-racional $P = (-1, 0)$. Considere ainda a reta ℓ de equação X com parametrização $(0, T)$. Cada ponto $R = (0, t) \in \ell$, determina uma reta PR de equação paramétrica $X = 0 + S$ e $Y = -1 + tS$. A interseção $PQ \cap C$ determina a equação quadrática dada pela substituição:

$$S^2 + (-1 + St)^2 = 1.$$

Ou ainda

$$(1 + t^2)S^2 + (-2t)S = 0$$

Como a solução $S = 0$ nos fornece P, temos $S = \frac{2t}{1+t^2}$, que nos fornece $x_t = \frac{2t}{1+t^2}$ e $y_t = \frac{t^2-1}{t^2+1}$.

Exemplo 33. Considere $K = \mathbb{R}$. Sejam $\mathcal{C}_1$ e $\mathcal{C}_2$ curvas de equações $Y^2 - X^3$ e $Y^2 - X^2(X - 1)$, respectivamente. Verifique que essas curvas são singulares na origem. A singularidade de $\mathcal{C}_1$ é chamada uma cúspide (a curva é chamada cuspidal) enquanto a de $\mathcal{C}_2$, um nó (a curva é chamada nodal). Essas curvas são racionais, verifique!

Para definir curva projetiva, inicialmente pensemos em seu conjunto de pontos K-racionais. Um ponto em $\mathbb{P}^2$ é uma classe de equivalência de um ponto não nulo no espaço tridimensional K^3. Assim sendo, não faz sentido definir uma avaliação de $P \in \mathbb{P}^2$ por um polinômio $F \in K[X, Y, Z]$. Por outro lado podemos definir o lugar de zeros em $\mathbb{P}^2$ de um polinômio *homogêneo*, isto é, um polinômio $F = \sum a_{ijk} X^i Y^j Z^k$ cujos monômios

$X^iY^jZ^k$ sejam todos de mesmo grau, digamos $d = i + j + k$. Com efeito, nesse caso, se F é homogêneo de grau d, então $F(\lambda a, \lambda b, \lambda c) = \lambda^d F(a, b, c)$, portanto,

$$F(\lambda a, \lambda b, \lambda c) = o \Leftrightarrow F(a, b, c) = 0.$$

Seja L um corpo algebricamente fechado. Uma *curva algébrica plana projetiva* C definida sobre L, corresponde a uma classe de equivalência de um polinômio homogêneo não constante $F \in L[X, Y, Z]$ módulo associados. Se K é um corpo arbitrário e $\overline{K}$ é um fecho algébrico para K, temos a curva algébrica plana $C \subset \mathbb{P}^2_{\overline{K}}$, e definimos o conjunto $C(K) = V(F) \subset \mathbb{P}^2_K$ como sendo o conjunto dos pontos projetivos K-*racionais* da curva C.

Sejam $C \in \mathbb{P}^2$ uma curva definida sobre K por um polinômio homogêneo $F \in K[X, Y, Z]$ e $P = (a : b : c) \in C(K)$ um ponto K-racional. O ponto P é dito *regular* quando $f_X(a, b, c) \neq 0$, $f_Y(a, b, c) \neq 0$ ou $f_Z(a, b, c) \neq 0$. Neste caso, a *reta tangente* a C em P, denotada por T_PC, é a reta definida pela equação

$$f_X(a, b, c)X + f_Y(a, b, c)Y + f_Z(a, b, c)Z.$$

O ponto P é dito *singular* se

$$F_X(a, b, c) = F_Y(a, b, c) = F_Z(a, b, c) = 0.$$

A curva C é dita *lisa* se não possui ponto singular. Uma curva C é dita *singular* se pelo menos um dos seus pontos é singular.

Dado um polinômio $f \in K[X, Y]$ de grau d então podemos escrever

$$f = \sum_0^d f_i$$

com $f_i \in K[X, Y]$, homogêneo de grau i e $f_d \neq 0$. A *homogeneização* de f é o polinômio homogêneo, também de grau d,

$$f^* = \sum Z^{d-i} f_i(X, Y) \in K[X, Y, Z].$$

Via a identificação $\mathbb{A}^2 \subset \mathbb{P}^2$, podemos ver que toda curva algébrica afim $\mathcal{C} \subset \mathbb{A}^2$ possui uma curva projetiva $\overline{\mathcal{C}}$ associada, com $\mathcal{C} \subset \overline{\mathcal{C}} \subset \mathbb{P}^2$, chamada *projetivização* de $\mathcal{C}$. Para obter a projetivização de uma curva afim basta homogeneizar o polinômio que define $\mathcal{C}$. Reciprocamente, dada uma curva projetiva F encontramos uma curva afim associada deshomogeneizando F, isto é, considerando a *deshomogeneização* $F_* = F(X, Y, 1)$ de F. As curvas algébricas planas afins f, com $f \in K[X, Y]$, serão consideradas implicitamente como $f^*(X, Y, 1)$, onde f^* é a homogeinização de f. Os conceitos de curva irredutível e componente irredutível de uma curva também valem para curvas projetivas e são compatíveis com a projetivização. Isto é, se $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \ldots \cup \mathcal{C}_k$ são as componentes

irredutíveis da curva afim $\mathcal{C} \subset \mathbb{A}^2_K$, então $\overline{\mathcal{C}} = \overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2} \cup \ldots \cup \overline{\mathcal{C}_k}$ são as componentes irredutíveis da curva projetiva $\overline{\mathcal{C}} \subset \mathbb{P}^2_K$.

Após essa breve introdução às curvas algébricas, podemos enunciar o Teorema de Bézout. Antes disso vamos necessitar um lema.

Lema 7. *Sejam K um corpo algebricamente fechado e $F \in K[U, V]$ um polinômio homogêneo de grau d. Então f possui d raizes em $\mathbb{P}^1$ contadas com multiplicidade.*

Prova: Escrevendo $F = V^a G(U, V)$, com V não dividindo G, obtemos que a multiplicidade de $(1 : 0)$ como raiz de F é a. Como $G(U, 1)$ é um polinômio com coeficientes em um corpo algebricamente fechado, então podemos escrever

$$G(U, 1) = c \prod_i (U - a_i)^{m_i}$$

com $c, a_i \in K$ e m_i inteiro positivo. Como V não divide G, temos que

$$G(U, V) = c \prod_i (U - a_i V)^{m_i}.$$

Portanto, $(a_i : 1)$ é raiz de F com multiplicidade m_i. □

Teorema 19 (Teorema de Bézout para Retas). *Seja K um corpo algebricamente fechado. Sejam ℓ uma reta projetiva e C uma curva plana projetiva de grau d, ambas definidas sobre K e suponha que ℓ não é uma componente de C. Então, o número de pontos na interseção entre ℓ e C, contados com multiplicidades, é d.*

Demonstração. Escrevendo a equação da reta ℓ por $aX + bY + cZ$ e supondo, sem perda de generalidade, que $c \neq 0$ obtemos que esta reta pode ser parametrizada por $X = U$, $Y = V$ e $Z = -(aU + bV)/c$. Desta forma, se F é uma equação para C, então $F(U, V, -(aU + bV)/c)$ é um polinômio homogêneo de grau d. Pelo lema anterior, este polinômio possui d raizes contadas com multiplicidades. Por fim, temos uma bijeção entre as raizes $(x : y)$ de $F(U, V, -(aU + bV)/c)$ e os pontos $(x : y : -(ax + by)/c)$ da interseção entre ℓ e C. □

Na verdade, o Teorema de Bézout é mais geral do que o enunciamos acima. Enunciaremos ele da forma mais geral possível abaixo, mas só provaremos esta versão para retas, além da versão para cônicas no fim da seção seguinte.

Teorema 20 (Teorema de Bézout). *Seja K um corpo algebricamente fechado. Sejam C e D curvas planas projetivas sobre K de graus m e n, respectivamente, sem componentes em comum. Então, o número de pontos na interseção entre C e D, contados com multiplicidades, é mn.*

Não entraremos nos detalhes sobre o significado das multiplicidades citadas no enunciado acima. Para mais detalhes sobre este significado e sobre a prova deste teorema o leitor pode consultar Vainsencher (2017), Capítulo 5.

Uma consequência interessante deste fato é que as componentes de uma curva plana sempre se intersectam. A partir disto, é possível ver que toda curva plana redutível é singular (ver Exercício 26).

3.3 Afinidades e projetividades

No plano afim as mudanças de coordenadas que vamos considerar serão as chamadas afinidades. Uma afinidade é uma plicação de K^2 em si mesmo definida por uma composição de uma translação com um isomorfismo K linear.

Mais precisamente, seja K um corpo. Uma *afinidade* ou *transformação afim* em $\mathbb{A}^2$ é uma aplicação do tipo:

$$T : \mathbb{A}^2 \to \mathbb{A}^2$$

definida por

$$T(x_1, x_2) = (ax_1 + bx_2 + s, cx_1 + dx_2 + t).$$

com $a, b, c, d, s, t \in K$ satisfazendo $ad - bc \neq 0$.

No plano projetivo a mudança de coordenadas que iremos utilizar é chamada projetividade ou transformação projetiva.

Mais precisamente, seja K um corpo. Uma *projetividade* ou *transformação projetiva* em $\mathbb{P}^n$ consiste em uma transformação

$$\tau : \mathbb{P}^n \to \mathbb{P}^n$$

induzida por uma transformação linear bijetiva $T : K^{n+1} \to K^{n+1}$.

Note que para todo $\lambda \in K^*$ a transformação linear λT induz a mesma projetividade τ.

Exemplo 34. Na esfera de Riemman, ou equivalentemente, na reta projetiva complexa, toda transformação projetiva é do tipo:

$$\tau(Z : W) = (aZ + bW : cZ + dW)$$

com a matriz correspondente $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ sendo inversível, isto é, com $ad - bc \neq 0$. Sua restrição ao plano complexo (reta afim) nos fornece, para $W \neq 0$:

$$\tau(Z : 1) = (aZ + b : cZ + d)$$

que no caso em que $cZ + d \neq 0$, é precisamente, uma transformação de Möbius $\tau(Z) = \frac{aZ+b}{cZ+d}$.

Dizemos que duas curvas $C, C' \in \mathbb{P}^2$ são *projetivamente equivalentes* se existir uma transformação projetiva $\tau : \mathbb{P}^2 \to \mathbb{P}^2$ tal que $\tau(C) = C'$.

Note que, se $\tau : \mathbb{P}^2 \to \mathbb{P}^2$ é uma transformação projetiva definida por

$$\tau(x : y : z) = (T_1(x, y, z) : T_2(x, y, z) : T_2(x, y, z))$$

com $T_1, T_2, T_3 \in K[X, Y, Z]$ polinômios de grau um, então a imagem inversa por τ da curva de equação F é a curva de equação $F^\tau = F(T_1, T_2, T_3)$. Desta forma, as curvas de equação F e F^τ são projetivamente equivalentes.

Proposição 15. *Sejam $P_1, P_2, P_3, P_4 \in \mathbb{P}^2$ quatro pontos em posição geral, isto é, dados quaisquer três deles, seus representantes em K^3 são linearmente independentes. Então existe uma única transformação projetiva $\tau : \mathbb{P}^2 \to \mathbb{P}^2$ tal que $\tau(P_1) = (1 : 0 : 0)$, $\tau(P_2) = (0 : 1 : 0)$, $\tau(P_3) = (0 : 0 : 1)$ e $\tau(P_4) = (1 : 1 : 1)$.*

Demonstração. Sejam $w_i \in K^3$ representantes de P_i e suponha que $w_4 = a_1 w_1 + a_2 w_2 + a_3 w_3$ com $a_i \in K$. Defina $v_i = a_i w_i$ para $i = 1, 2, 3$ e $v_4 = w_4$, assim, v_i são ainda representantes de P_i. Como v_1, v_2, v_3 formam uma base do espaço linear K^3, existe uma única transformação linear $T : K^3 \to K^3$ tal que $T(v_1) = (1, 0, 0)$, $T(v_2) = (0, 1, 0)$ e $T(v_3) = (0, 0, 1)$, além disso, $T(v_4) = T(v_1) + T(v_2) + T(v_3) = (1, 1, 1)$. O resultado segue. □

Proposição 16. *Seja K um corpo e seja C uma cônica projetiva redutível sobre K. Então, C é projetivamente equivalente a X^2 ou a XY.*

Prova: Seja $F \in K[X, Y, Z]$ um polinômio homogêneo de grau dois que define a curva C.

Suponhamos que $F = aL^2$ é associado ao quadrado de um polinômio linear. Escolhendo quatro pontos em $\mathbb{P}^2$ de modo que dois deles estejam na reta L e que todos estejam em posição geral, podemos aplicar a Proposição 15 para encontrar uma transformação projetiva que leve a reta definida por L na reta definida por X.

Suponhamos que $F = LM$ com L e M polinômios lineares não associados. Escolhendo quatro pontos em $\mathbb{P}^2$ de modo que um deles esteja na interseção de L com M, um deles esteja em $L \setminus M$, um deles esteja em $M \setminus L$ e o outro esteja em $\mathbb{P}^2 \setminus (L \cup M)$, podemos aplicar a Proposição 15 para encontrar uma transformação projetiva que leve a reta definida por L na reta definida por X e que leve a reta definida por M na reta definida por Y. □

Proposição 17. *Seja K um corpo de característica diferente de 2 e seja $C \subset \mathbb{P}^2$ uma cônica irredutível definida sobre K. Então C é projetivamente equivalente a $aX^2 + bY^2 + cZ^2$, com $abc \neq 0$. Se K ainda for algebricamente fechado, então C é projetivamente equivalente a $X^2 + Y^2 + Z^2$.*

PROVA: Considere $F = aX^2 + bXY + cXZ + dY^2 + eYZ + fZ^2$ uma equação que define a cônica C. Se $b = c = e = 0$, não temos nada a fazer. Suponhamos agora que este não seja o caso. Sem perda de generalidade, suponhamos que $b \neq 0$.

Se $a = d = 0$, então teremos que $b \neq 0$ para termos F irredutível. Neste caso, aplicamos a projetividade definida por $\tau(x : y : z) = (x - y : x + y : z)$ para obtermos F^τ com $abd \neq 0$. Desta forma, vamos supor que F possui estes três coeficientes não nulos.

Completando o quadrado em $ax^2 + bxy = a(x^2 + ba^{-1}xy)$ obtemos $ax^2 + bxy = a[(x + b(2a)^{-1}y)^2 - b^2(2a)^{-2}y^2]$. Aplicando a projetividade definida por $\tau(x : y : z) = (x + b(2a)^{-1}y : y : z)$ podemos supor que $a \neq 0$ e $b = 0$. Desta forma,

$$F = aX^2 + cXZ + dY^2 + eYZ + fZ^2 \text{ com } a \neq 0.$$

Se c for não nulo, procedemos de forma análoga para obter

$$F = aX^2 + dY^2 + eYZ + fZ^2 \text{ com } a \neq 0.$$

Se $d = f = 0$, então, pela irredutibilidade de de F, obtemos $e \neq 0$. Aplicando a projetividade definida por $\tau(x : y : z) = (x : y - z : y + z)$ podemos supor $def \neq 0$. Completando os quadrados em $dY^2 + eYZ$, assim como fizemos acima, obteremos a equação desejada.

Por fim, se K for algebricamente fechado, então a projetividade $\tau(x : y : z) = (x/\sqrt{a} : y/\sqrt{b} : z/\sqrt{c})$ nos fornece o desejado. □

Proposição 18. *Seja K um corpo algebricamente fechado e seja $C \subset \mathbb{P}^2$ uma cônica lisa definida sobre K. Então C é projetivamente equivalente a $Z^2 - XY$.*

PROVA: Seja P um ponto da cônica lisa. Note que o Teorema 19 nos diz que a reta $T_P C$ intersecta C somente em P. Se $Q \neq P$ é outro ponto de C, então temos $T_P C \cap T_Q C = \{R\}$.

Seja π uma projetividade tal que $\pi(1 : 0 : 0) = Q$, $\pi(0 : 1 : 0) = P$ e $\pi(0 : 0 : 1) = R$. A cônica $\pi^{-1}C$ passa por $(0 : 1 : 0)$ com reta tangente X e por $(1 : 0 : 0)$ com reta tangente Z. Seja F a equação da cônica $\pi^{-1}C$.

$$F = aX^2 + bXY + cXZ + dY^2 + eYZ + fZ^2.$$

Como $F(1 : 0 : 0) = a$ e $F(0 : 1 : 0) = d$ segue que $a = d = 0$. Portanto, $F = bXY + cXZ + eYZ + fZ^2$. A condição de tangência nos dá $0 = F_Z(0 : 1 : 0) = e$ e $0 = F_Z(1 : 0 : 0) = c$, isto é,

$$F = bXY + fZ^2.$$

Note que $bf \neq 0$, já que, F é irredutível. Sendo K algebricamente fechado, podemos aplicar a transformação projetiva $\tau(x : y : z) = (-x/\sqrt{b} : y/\sqrt{b} : z/\sqrt{f})$ que nos fornece a equação desejada. □

Teorema 21 (Teorema de Bézout para Cônicas Lisas). *Seja K um corpo algebricamente fechado. Sejam C uma cônica lisa projetiva e D uma curva plana projetiva de grau d, ambas definidas sobre K. Se C não é uma componente de D, então, o número de pontos na interseção entre C e D, contados com multiplicidades, é $2d$.*

Demonstração. Se C é uma cônica lisa, então a proposição anterior nos diz que existe uma projetividade $\tau : \mathbb{P}^2 \to \mathbb{P}^2$ tal que $F^\tau = Z^2 - XY$. Escrevemos $\tau(x : y : z) = (T_1(x, y, z), T_2(x, y, z), T_3(x, y, z))$ com $T_1, T_2, T_3 \in K[X, Y, Z]$ polinômios de grau um.

Note que temos uma parametrização da forma $X = U^2$, $Y = V^2$ e $Z = UV$ para a curva de equação $F^\tau = Z^2 - XY$, já que, $F^\tau(U^2, V^2, UV) = 0$. Desta forma,

$$F(T_1(U^2, V^2, UV), T_2(U^2, V^2, UV), T_3(U^2, V^2, UV)) = F^\tau(U^2, V^2, UV) = 0$$

nos diz que temos uma parametrização $X = T_1(U^2, V^2, UV)$, $Y = T_2(U^2, V^2, UV)$ e $Z = T_3(U^2, V^2, UV)$ para a curva de equação F com estes polinômios sendo homogêneos de grau dois.

Substituindo esta parametrização na equação G de D, obtemos

$$G(T_1(U^2, V^2, UV), T_2(U^2, V^2, UV), T_3(U^2, V^2, UV))$$

que é um polinômio homogêneo de grau $2d$. Pelo Lema 7, este polinômio possui $2d$ raizes contadas com multiplicidades. Por fim, temos uma bijeção entre as sua raizes e os pontos da interseção entre C e D. □

3.4 Curvas elípticas

As curvas elípticas são o principal objeto geométrico nesse curso. Nosso objetivo é fazer criptografia com curvas elípticas sobre corpos finitos.

Seja K um corpo. Uma *curva elíptica* definida sobre K é uma curva cúbica lisa $C \subset \mathbb{P}^2$ possuindo algum ponto K-racional.

Note que, no caso em que K é algebricamente fechado, temos pelas considerações feitas após o Teorema 20 que uma curva elíptica é irredutível, já que, ela é lisa.

Uma outra consequência importante do Teorema 20, no sentido de explorar uma forma normal para as curvas elípticas, é o fato de sempre existir um ponto de inflexão em uma curva elíptica (ver Exercício 27).

Observação 3. Vimos que as cúbicas irredutíveis singulares podem ser cuspidais ou nodais e ambas são racionais. Por outro lado, as cúbicas irredutíveis não singulares não são racionais.

Proposição 19. *Seja K um corpo algebricamente fechado, de característica diferente de 2 e de 3. Toda curva elíptica E sobre K é projetivamente equivalente a uma cúbica da forma*

$$ZY^2 - X(X - Z)(X - \lambda Z)$$

com $\lambda \in K$ *e* $\lambda \neq 0, 1$.

PROVA: Seja P um ponto de inflexão de F. Após uma projetividade podemos supor que $P = (0 : 1 : 0)$ e $T_P F = Z$. Desta forma,

$$F = ZY^2 + a_0X^2Y + a_1XYZ + a_2Z^2Y + a_3X^3 + a_4X^2Z + a_5XZ^2 + a_6Z^3.$$

Como $\text{Mult}_P\, Z \cap F \geqslant 3$, obtemos que $a_0 = 0$. Como F é irredutível temos que $a_3 \neq 0$.

Fazendo a mudança $\tau(x : y : z) = (x : y - \frac{a_1}{2}x - \frac{a_2}{2}z : z)$ podemos supor que $a_1 = a_2 = 0$. Assim

$$F = ZY^2 + a_3X^3 + a_4X^2Z + a_5XZ^2 + a_6Z^3 = ZY^2 + a_3(X - \lambda_1 Z)(X - \lambda_2 Z)(X - \lambda_2 Z)$$

com $\lambda_1, \lambda_2, \lambda_3$ distintos dois a dois (ver Exercício 28).

Fazendo a mudança $\tau(x : y : z) = (x - \lambda_1 z : y : z)$ podemos supor que $\lambda_1 = 0$, $\lambda_2\lambda_3 \neq 0$ e que $\lambda_2 \neq \lambda_3$.

Fazendo a mudança $\tau(x : y : z) = (x/b : y : z)$ com $b^3 = -a_3$ podemos supor que $a_3 = -1$.

Fazendo a mudança $\tau(x : y : z) = (x : y : z/\lambda_2)$ podemos supor que

$$F = \frac{1}{\lambda_2}ZY^2 - X(X - Z)(X - \lambda Z)$$

com $\lambda \neq 0, 1$.

Fazendo a mudança $\tau(x : y : z) = (x : by : z)$ com $b^2 = \lambda_2$ podemos supor que

$$F = ZY^2 - X(X - Z)(X - \lambda Z)$$

com $\lambda \neq 0, 1$. □

Observação 4. Pode ser mostrado ainda que toda curva elíptica E, sobre um corpo K de característica maior que 3, com um ponto K-racional é equivalente, em um sentido que não mencionaremos aqui, a uma curva plana afim de equação

$$Y^2 - X^3 - aX - b$$

com $a, b \in K$ e com $\Delta = 4a^3 + 27b^2 \neq 0$. Esta nova equação é chamada de *forma normal de Weierstrass* da curva elíptica.

Observemos que o fato do discriminante $\Delta = 4a^3 + 27b^2 \neq 0$ assegura a não-singularidade desta última curva em todos os pontos. Relembremos, da definição que:

$P = (x_0, y_0) \in E(K)$ é singular se, e somente se, $f_X(x_0, y_0) = f_Y(x_0, y_0) = 0$

No nosso caso, a equação da curva é $f = Y^2 - (X^3 + aX + b)$. Desta forma, se (x_0, y_0) é ponto singular de f, então (x_0, y_0) satisfaz o seguinte sistema.

$$\begin{cases} f_X(x_0, y_0) = -(3x_0^2 + a) = 0; \\ f_Y(x_0, y_0) = 2y_0 = 0; \\ f(x_0, y_0) = y_0^2 - (x_0^3 + ax_0 + b) = 0. \end{cases}$$

Pelo sistema acima, temos que $y_0 = 0$ e que x_0 é solução de $X^3 + aX + b$ e de sua derivada, *i.e.*, x_0 é raiz dupla de $X^3 + aX + b$. Mas isto acontece se, e somente se, $\Delta = -(4a^3 + 27b^2) = 0$.

Notemos ainda que a homogenização $ZY^2 - X^3 - aXZ^2 - bZ^3$ da forma normal de Weierstrass nos fornece um ponto K-racional distinguível de E, a saber, o ponto no infinito $(0 : 1 : 0)$, que é um ponto de inflexão.

3.5 Exercícios

Questão 20. *Verifique que por dois pontos no plano projetivo passa uma única reta projetiva. Verifique também que quaisquer duas retas projetivas em $\mathbb{P}^2$ se encontram em um único ponto. Explique o que ocorre no caso da projetivização de duas retas no plano afim que sejam paralelas, isto é, que não se intersectam em $\mathbb{A}^2$.*

Questão 21. *Mostre que todas as cônicas reais geometricamente irredutíveis não degeneradas são projetivamente equivalentes. Mais precisamente, mostre que as seguintes curvas são projetivamente equivalentes:*

a) $x^2 + y^2 = 1$.

b) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

c) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

d) $y = x^2$.

Questão 22. *Estude os pontos singulares de cada das seguintes curvas em $\mathbb{P}^2_{\mathbb{C}}$, com coordenadas homogêneas $(X : Y : Z)$.*

a) Folha de equação $x^4 + 2x^2y^2 + y^4 + 4x^3 = 0$. (Figura 3.4)

b) Bifólio de equação $x^4 + 2x^2y^2 + y^4 - 4xy^2 = 0$. (Figura 3.5)

c) Trevo alongado de equação $x^4 + 2x^2y^2 + y^4 + 2x^3 - 2xy^2 = 0$. (Figura 3.6)

d) Caracol de Etienne Pascal de equação $(x^2 + y^2 - 4x)^2 - 4(x^2 + y^2) = 0$. (Figura 3.7)

e) Nefroide de equação $(x^2 + y^2 - 1)^3 - y^2 = 0$. (Figura 3.8)

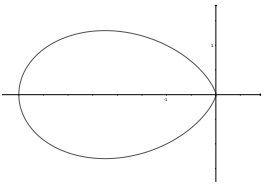


Figura 3.4: Folha (Questão 22)

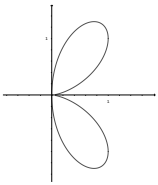


Figura 3.5: Bifólio (Questão 22)

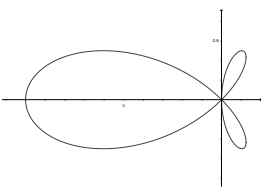


Figura 3.6: Trevo alongado (Questão 22)

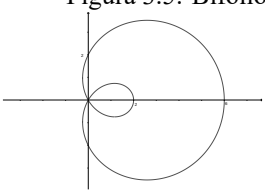


Figura 3.7: Caracol de Etienne Pascal (Questão 22)

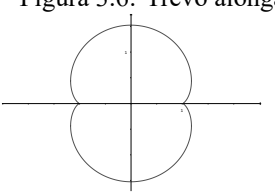


Figura 3.8: Nefroide (Questão 22)

Questão 23. *Para cada curva, encontre os pontos singulares com as respectivas multiplicidades e retas tangentes. Faça um esboço.*

a) $x^5 - 5yx^2 + 2x^5 + 2y^5 = 0$.

b) $x^3 - y^2 + 2yx^2 + 2xy^2 - xy = 0$.

Questão 24. *Considere a curva algébrica plana* C *em* $\mathbb{P}^2$, *com coordenadas homogêneas (X:Y:Z), com equação* $X^2Y^3 - X^2YZ^2 + X^2Y^2Z - X^2Z^3 + YZ^4 = 0$.

a) Encontre os pontos singulares de C *com suas respectivas multiplicidades.*

b) Existem linhas retas que sejam tangentes a C *em dois pontos distintos, que passem pelo ponto* $(1:0:0)$?

Questão 25. *Mostre que a curva projetiva associada ao polinômio* $f(X,Y) = X^3Y + Y^3 + X + X^2Y^2 + Y^2 + X^2 + X^2Y + XY^2 \in \mathbb{F}_2[X,Y]$ *é não-singular.*

Questão 26. *Seja* $P \in \mathbb{P}^2$ *um ponto de interseção das curvas planas projetivas de equação* F *e* G. *Mostre que* P *é um ponto singular de* FG.

Questão 27. *Seja F uma curva plana projetiva de grau d sobre um corpo algebricamente fechado. Definimos a curva Hessiana associada a F pela curva plana projetiva dada por*

$$H(F) = \det \begin{bmatrix} F_{XX} & F_{XY} & F_{XZ} \\ F_{YX} & F_{YY} & F_{YZ} \\ F_{ZX} & F_{ZY} & F_{ZZ} \end{bmatrix}.$$

1. *Mostre que $H(F)$ é uma curva de grau $3(d-2)$;*
2. *Mostre que os pontos de interseção entre F e $H(F)$ consistem dos pontos singulares de F e dos pontos de inflexão de F.*
3. *Mostre que, se F for uma cúbica lisa, então existe um ponto de inflexão em F. (Dica: Utilize o Teorema 20)*

Questão 28. *Seja $F = ZY^2 + a(X-\lambda_1 Z)(X-\lambda_2 Z)(X-\lambda_2 Z)$ uma curva plana projetiva sobre um corpo K. Mostre que F é lisa se, e somente se, $\lambda_1, \lambda_2, \lambda_3$ são dois a dois distintos.*

4 Curvas Elípticas

Neste capítulo introduziremos a estrutura de grupo em uma curva elíptica, de duas formas distintas. A primeira utilizará a forma normal de Weierstrass de uma curva elíptica afim. Já na segunda, abordaremos estas curvas no plano projetivo sem a necessidade de tal forma normal. Em ambos os casos veremos que a dificuldade será a prova da associatividade da soma de pontos. No primeiro caso isto é feito somente utilizando contas e no segundo, utilizaremos o auxílio dos sistemas lineares de curvas para provar que vale a associatividade no caso genérico.

4.1 Lei de grupo

Nesta seção introduziremos a estrutura de grupo em uma curva elíptica da forma mais ingênua possível, retirando qualquer necessidade de pré-requisitos sobre a teoria de curvas algébricas planas. Para isto, consideraremos somente as curvas elípticas afins que estão definidas sob a forma normal de Weierstrass.

4.1.1 Curvas elípticas sobre corpos algebricamente fechados

Seja E uma *curva elíptica* sobre um corpo algebricamente fechado K de característica diferente de 2 e 3, isto é, o conjunto de soluções de uma equação da forma

$$Y^2 = X^3 + AX + B,$$

em $\mathbb{A}^2_K$ satisfazendo a condição do discriminante $\Delta = 4A^3 + 27B^2$ ser não nulo. Esta equação é chamada de forma normal de Weierstrass para E.

Exemplo 35. As figuras Fig. 4.1 e Fig. 4.2 representam os traços reais das curvas elípticas definidas sobre o corpo dos números complexos. Já as Fig. 4.3 e Fig. 4.4 representam os traços reais das curvas, também definidas sobre o corpo dos números complexos, com equações semelhantes às curvas elípticas mas com discriminante nulo.

Figura 4.1: Curva de equação $Y^2 = X^3 - X + 1$

Figura 4.2: Curva de equação $Y^2 = X^3 - 3X$

Figura 4.3: Cúbica nodal $Y^2 = X^3 + X^2$

Figura 4.4: Cúbica cuspidal $Y^2 = X^3$

Agora vamos definir a estrutura de grupo de uma curva elíptica. Para isso, vamos definir a soma de dois pontos.

Sejam P e Q pontos na curva elíptica E e L uma reta passando por P e Q, como na Figura 4.5.

Para pontos P e Q genéricos, a reta L ainda vai cortar a curva em um terceiro ponto, digamos R. Por este ponto passa uma reta perpendicular ao eixo horizontal, que vai cortar a curva em mais um ponto, digamos R'.

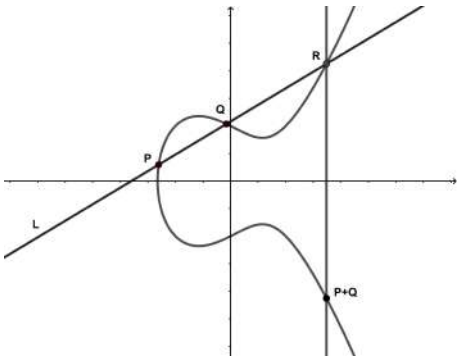


Figura 4.5: Soma de pontos distintos em um Curva Elíptica

O ponto R' será chamado de *soma* entre P e Q, e será denotado por

$$P + Q.$$

Antes de especificar a genericidade mencionada acima vamos fazer um exemplo.

Exemplo 36. Consideremos a curva elíptica

$$E: \quad Y^2 = X^3 - 15X + 18.$$

Sejam $P = (7, 16)$ e $Q = (1, 2)$ dois pontos de E. A reta passando por P e Q tem equação

$$Y = \frac{7}{3}X - \frac{1}{3}.$$

Para encontrar o ponto em comum entre E e L resolvemos a equação abaixo:

$$\begin{aligned} \left(\tfrac{7}{3}x - \tfrac{1}{3}\right)^2 &= x^3 - 15x + 18 \\ \tfrac{49}{9}x^2 - \tfrac{14}{9}x - \tfrac{1}{9} &= x^3 - 15x + 18 \\ 0 &= x^3 - \tfrac{49}{9}x^2 - \tfrac{121}{9}x - \tfrac{161}{9} \end{aligned}$$

Note que, já sabemos que $x = 7$ e $x = 1$ são duas raizes desta equação polinomial. Desta forma, é mais fácil encontrar a terceira, já que, podemos dividir o polinômio $x^3 - \frac{49}{9}x^2 - \frac{121}{9}x - \frac{161}{9}$ por $(x - 7)(x - 1)$.

Fazendo esta conta temos:

$$x^3 - \frac{49}{9}x^2 - \frac{121}{9}x - \frac{161}{9} = (x-7)(x-1)\left(x + \frac{23}{9}\right).$$

Assim, $R = (-\frac{23}{9}, ?)$. Para encontrar a coordenada y basta substituir $x = -\frac{23}{9}$ na equação da reta e obter $y = -\frac{170}{27}$. Refletindo ao longo do eixo horizontal, obtemos

$$P + Q = \left(-\frac{23}{9}, \frac{170}{27}\right).$$

Para fazermos estas contas no Pari, primeiro nomeamos a curva e usando o comando $ellinit([x])$, depois verificamos se os pontos P e Q pertencem a curva e por fim somamos:

```
?   e = ellinit([0, 0, 0, −15, 18])
%1  = [0, 0, 0, −15, 18, 0, −30, 72, −225, 720, −15552, 76032, 54000/11,
    Vecsmall([1]), [Vecsmall([128, 1])], [0, 0, 0, 0, 0, 0, 0, 0]]
?   ellisoncurve(e, [7, 16])
%2  = 1
?   ellisoncurve(e, [1, 2])
%3  = 1
?   elladd(e, [7, 16], [1, 2])
%4  = [−23/9, 170/27]
```

Após este exemplo podemos repetir a argumentação e provar os seguinte resultado.

Proposição 20. *Seja E uma curva dada pela equação $Y^2 = X^3 + AX + B$ em $\mathbb{A}^2_K$ com $\Delta = 4A^3 + 27B^2 \neq 0$ e sejam $P = (x_1, y_1)$ e $Q = (x_2, y_2)$ em E. Se $x_1 \neq x_2$, então a reta L, passando por P e Q, corta E no ponto $R = (x_3, -y_3)$, onde*

$$x_3 = a^2 - (x_1 + x_2),\ y_3 = a(x_1 - x_3) - y_1 \text{ e } a = \frac{y_2 - y_1}{x_2 - x_1}.$$

Portanto, $P + Q = (x_3, y_3)$.

Prova: Note que a reta L, que passa por P e Q, tem equação

$$Y = \frac{y_2 - y_1}{x_2 - x_1}X + \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1}.$$

Desta forma, o terceiro ponto em comum entre E e L vem como solução da equação $\left(\frac{y_2-y_1}{x_2-x_1}X + \frac{x_2y_1-x_1y_2}{x_2-x_1}\right)^2 = X^3 + AX + B$, isto é,

$$0 = X^3 - a^2X^2 + (A - 2ab)X + (B - b^2),$$

onde $a = \frac{y_2-y_1}{x_2-x_1}$ e $b = \frac{x_2y_1-x_1y_2}{x_2-x_1}$. Por outro lado, $X = x_1$ e $X = x_2$ são duas soluções desta equação polinomial. Desta forma, temos a seguinte fatoração.

$$X^3 - a^2X^2 + (A - 2ab)X + (B - b^2) = (X - x_1)(X - x_2)(X - x_3)$$

Comparando os coeficientes do termo de grau dois nesta igualdade, obtemos que

$$x_3 = a^2 - (x_1 + x_2).$$

Para encontrar o terceiro ponto $R = (x_3, -y_3)$, em comum entre E e L, e consequentemente $P + Q = (x_3, y_3)$, calculamos y_3 substituindo $X = x_3$ na equação da reta. Fazendo isso, obtemos $-y_3 = \frac{y_2-y_1}{x_2-x_1}x_3 + \frac{x_2y_1-x_1y_2}{x_2-x_1} = a(x_3 - x_1) + y_1$. □

Por outro lado, existem algumas sutilezas quando variamos a escolha dos pontos à somar e variamos as possíveis curvas elípticas. Além disso, se quisermos que esta soma forneça uma estrutura de grupo, então ainda temos que definir o elemento neutro e o inverso.

Primeiro vamos definir a adição de um ponto com ele mesmo.

Suponhamos que queremos somar o ponto P da curva E com ele mesmo. Olhando a Figura 4.5 e fazendo o ponto Q deslizar na direção de P, intuímos que a reta L, a ser considerada, é a reta tangente a E em P.

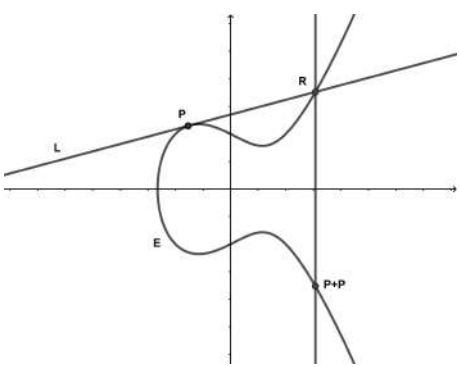


Figura 4.6: Soma entre pontos iguais em um Curva Elíptica

Consideremos a reta tangente L a E em P. Esta reta ainda corta E em um outro ponto, digamos R. Considerando a reta vertical que passa por R, então a soma de P com ele mesmo será definido pelo ponto de encontro desta reta com E (ver Figura 4.6).

Antes de fazer um exemplo, vamos fazer uma observação. Note que, nas curvas representadas nas Figuras 4.4 e 4.3, temos problemas para falar da tangente em todos os pontos de E. Este é o exato motivo para pedirmos na definição a condição $\Delta \neq 0$.

Esta desigualdade é uma forma de impor, nos coeficientes da curva, que ela não tenha "singularidades". Isto, por sua vez, vai implicar na possibilidade de encontrar a reta tangente em todos os pontos.

Vamos fazer um exemplo utilizando, novamente a curva do exemplo anterior.

Exemplo 37. Consideremos a curva elíptica $E: \quad Y^2 = X^3 - 15X + 18$ e o ponto $P = (7, 16)$. Utilizamos as técnicas de derivação implícita do curso de cálculo para determinar a reta tangente a E em P. Para isso, veremos Y como função de X na igualdade que define E e derivamos esta igualdade em X para obter:

$$\frac{dY}{dX} = \frac{3X^2 - 15}{2Y}.$$

Avaliando isto no ponto P, podemos deduzir que a inclinação da reta tangente em P é $\frac{33}{8}$. Portanto a equação da reta tangente é dada por $Y - 16 = \frac{33}{8}(X - 7)$, isto é,

$$Y = \frac{33}{8}X - \frac{103}{8}.$$

Para descobrir o outro ponto de encontro entre a reta tangente e E calculamos:

$$\begin{aligned} \left(\tfrac{33}{8}X - \tfrac{103}{8}\right)^2 &= x^3 - 15x + 18 \\ 0 &= x^3 - \tfrac{1089}{64}x^2 - \tfrac{2919}{32}x - \tfrac{9457}{64} \\ 0 &= (x-7)^2\left(x - \tfrac{193}{64}\right). \end{aligned}$$

Note que $x = 7$ é uma raiz deste polinômio com multiplicidade 2. Logo a solução $x = \frac{193}{64}$ nos dá exatamente a abscissa do ponto que queremos determinar.

Substituindo $x = \frac{193}{64}$ na equação da reta, obtemos $y = -\frac{223}{512}$. refletindo em torno do eixo $y = 0$ obtemos:

$$P + P = \left(\frac{193}{64}, \frac{223}{512}\right).$$

No Pari, já nomeada a curva e, basta fazer a soma $P + P$:

```
? elladd(e, [7, 16], [7, 16])
%5 = [193/64, 223/512]
```

Após este exemplo podemos repetir a argumentação e provar os seguinte resultado.

Proposição 21. *Seja E uma curva dada pela equação $Y^2 = X^3 + AX + B$ em $\mathbb{A}^2_K$, com $4A^3 + 27B^2 \neq 0$, e seja $P = (x_1, y_1)$ em E com $y_1 \neq 0$. Se L é a reta tangente a E em P, então L corta E no ponto $R = (x, -y)$, onde*

$$x = a^2 - 2x_1,\ y = a(x_1 - x) - y_1 \text{ e } a = \frac{3x_1^2 + A}{2y_1}.$$

Portanto, $P + P = (x, y)$.

PROVA: Olhando Y como função de X na equação que define E e derivando implicitamente esta equação em X obtemos:

$$\frac{dY}{dX} = \frac{3X^2 + A}{2Y}.$$

Avaliando isto no ponto P, podemos deduzir que a inclinação da reta tangente a E em P é $a = \frac{3x_1^2+A}{2y_1}$. Ainda utilizando as técnicas do curso de cálculo, obtemos que a equação da reta tangente a E em P é dada por

$$Y - y_1 = a(X - x_1).$$

Para encontrar o outro ponto em comum entre E e L resolvemos a equação $(a(X - x_1) + y_1)^2$ $X^3 + AX + B$, isto é,

$$X^3 - a^2X^2 + (A + 2a^2x_1 - 2ay_1)X + (B - a^2x_1^2 + 2ax_1y_1 - y_1^2) = 0$$

Neste caso, como L é tangente a E em P, teremos que x_1 é raiz deste último polinômio com multiplicidade 2, isto é,

$$X^3 - a^2X^2 + (A + 2a^2x_1 - 2ay_1)X + (B - a^2x_1^2 + 2ax_1y_1 - y_1^2) = (X - x_1)^2(X - x).$$

Analisando os coeficientes dos termos de grau dois nesta igualdade, obtemos que

$$x = a^2 - 2x_1.$$

Para encontrar o outro ponto em comum $R = (x, -y)$ entre E e L e, consequentemente $P + P = (x, y)$, basta substituir as coordenadas de R na equação da reta L e obter $y = a(x_1 - x) - y_1$. □

O próximo problema que iremos abordar é quando queremos somar um ponto $P = (a, b)$ com a sua reflexão, em torno do eixo $y = 0$, digamos $Q = (a, -b)$.

A solução para este impasse é criar um ponto $\mathcal{O}$ que "mora" no infinito! Isto é, ele não está no plano cartesiano XY e , além disso, queremos que ele esteja na interseção entre toda reta vertical e E. Esta propriedade desejada se torna natural quando pensamos que estamos em uma estrada que vai até o "infinito", já que, as margens são retas paralelas e se encontram no horizonte. Neste caso, definimos

$$P + Q = \mathcal{O}.$$

Pela Figura 4.8, ainda podemos deduzir pelas construções geométricas anteriores que, para todo $P \in E$, vale:

$$P + \mathcal{O} = P.$$

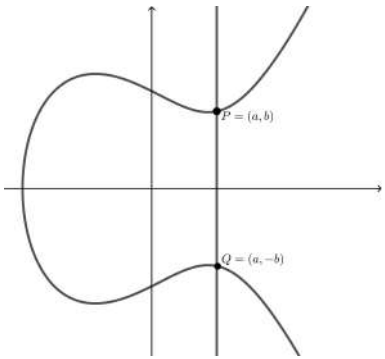


Figura 4.7: Soma entre pontos simétricos em um Curva Elíptica

Portanto, $\mathcal{O}$ é o *elemento neutro* que estamos procurando para esta soma que estamos definindo.

Observa ainda que se $P = (a, b)$, então o *inverso aditivo* de P é o ponto $Q = (a, -b)$. Usamos a seguinte notação

$$-P := Q.$$

Resumiremos todos os resultados e definições acima da seguinte forma.

Teorema 22. *Seja* $E : Y^2 = X^3 + AX + B$ *uma curva elíptica e sejam* P_1 *e* P_2 *em* E. *Então:*

- *Se* $P_1 = \mathcal{O}$, *então* $P_1 + P_2 = P_2$;
- *Se* $P_2 = \mathcal{O}$, *então* $P_1 + P_2 = P_1$;

Se nenhum destes casos ocorrem, escrevemos $P_1 = (x_1, y_1)$ *e* $P_2 = (x_2, y_2)$ *e então:*

- *Se* $x_1 = x_2$ *e* $y_1 = -y_2$, *então* $P_1 + P_2 = \mathcal{O}$;
- *Caso contrário,* $P_1 + P_2 = (x_3, y_3)$ *onde*

$$x_3 = a^2 - x_1 - x_2 \;\; e \;\; y_3 = a(x_1 - x_3) - y_1$$

e

$$a = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & se\ P_1 \neq P_2 \\ \frac{3x_1^2 + A}{2y_1} & se\ P_1 = P_2. \end{cases}$$

Corolário 6 (Unicidade do Inverso Aditivo). *Se* $P + Q = \mathcal{O}$, *então* $Q = -P$.

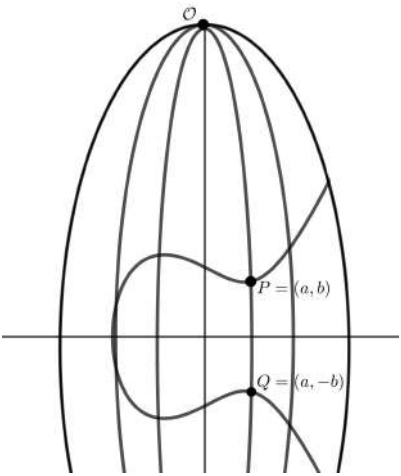


Figura 4.8: O ponto no infinito

Prova: De fato,basta considerar o caso em que $P \neq \mathcal{O}$. Pelas fórmulas da soma, se $Q \neq -P$, então $P + Q$ é um ponto do plano afim, isto é, é diferente de $\mathcal{O}$. □

Agora vamos mostrar que esta operação define uma estrutura de grupo na curva elíptica.

Teorema 23. *A operação de soma no Teorema 22 faz de uma curva elíptica E um grupo abeliano, com elemento neutro $\mathcal{O}$.*

Prova: Pela construção acima, temos que a existência do elemento neutro e do inverso aditivo. Pela construção geométrica, não é difícil de ver que a soma que definimos satisfaz:

$$P + Q = Q + P$$

para cada P e Q em E. A propriedade que necessita ser verificada, e que não é nada simples, é a associatividade, isto é,

$$(P + Q) + R = P + (Q + R)$$

para cada P, Q e R em E. Como esta prova é bem cumprida e técnica, deixaremos ela em uma subseção separada. □

4.1.2 A associatividade da soma

Nesta seção, utilizaremos as fórmulas do Teorema 22 para dar uma prova da associatividade, da soma definida em uma curva elíptica, que utiliza somente cálculos explícitos.

Considere $P = (x_P, y_P)$, $Q = (x_Q, y_Q)$ e $R = (x_R, y_R)$, três pontos na curva elíptica E.

É claro que a associatividade vale, se um dos pontos for igual a $\mathcal{O}$. Vamos, então, supor que $P, Q, R \in E \setminus \{\mathcal{O}\}$.

Por conveniência, vamos recordar a definição e reestabelecer a notação da soma $P+Q$ com estas notações.

Se $(x_P, y_P) = (x_Q, -y_Q)$, então $P+Q = \mathcal{O}$. Caso contrário, $P+Q = (x_{PQ}, y_{PQ})$, onde

$$\begin{array}{rcl} x_{PQ} & = & \alpha(P,Q)^2 - x_P - x_Q \\ y_{PQ} & = & -y_P + \alpha(P,Q)(x_P - x_{PQ}) \end{array} \quad \text{e } \alpha(P,Q) = \begin{cases} \frac{y_Q - y_P}{x_Q - x_P} & \text{se } x_P \neq x_Q \\ \frac{3x_P^2 + A}{2y_P} & \text{se } x_P = x_Q \end{cases}. \tag{4.1}$$

Dividiremos a verificação da associatividade em vários casos. O primeiro caso a ser considerado será o caso em que a fórmula 4.1 é usada para o cálculo de $P+Q, Q+R, (P+Q)+R$ e $P+(Q+R)$.

Lema 8. *Sejam $P, Q, R \in E \setminus \{\mathcal{O}\}$. Se $P \neq \pm Q$, $Q \neq \pm R$, $P+Q \neq \pm R$ e $Q+R \neq P$, então*

$$(P+Q)+R = P+(Q+R).$$

Prova: Escrevendo $(P+Q)+R = (x_1, y_1)$, $P+(Q+R) = (x_2, y_2)$, $\alpha = \alpha(P,Q)$, $\beta = \alpha(P+Q, R)$, $\gamma = \alpha(R,Q)$ e $\tau = \alpha(Q+R, P)$, obtemos que

$$\begin{array}{rclcrcl} x_1 & = & \beta^2 + x_P + x_Q - x_R - \alpha^2, & & y_1 & = & -y_R + \beta(2x_R - x_P - x_Q - \beta^2 + \alpha^2), \\ x_2 & = & \tau^2 + x_Q + x_R - x_P - \gamma^2, & & y_2 & = & -y_P + \tau(2x_P - x_Q - x_R - \tau^2 + \gamma^2). \end{array}$$

Escrevemos ainda

$$\begin{array}{rclcrcl} \widetilde{\alpha} & = & y_Q - x_P, & & \widetilde{\beta} & = & (y_P + y_R)(x_Q - x_P)^3 - \widetilde{\alpha}((2x_P + x_Q)(x_Q - x_P)^2 - \widetilde{\alpha}^2), \\ \widetilde{\gamma} & = & y_Q - y_R, & & \widetilde{\tau} & = & (y_P + y_Q)(x_Q - x_R)^3 - \widetilde{\gamma}((2x_Q + x_R)(x_Q - x_R)^2 - \widetilde{\gamma}^2), \\ \widetilde{\eta} & = & x_Q - x_P, & & \widetilde{\mu} & = & x_Q - x_R. \end{array}$$

Expandindo $x_1 - x_2$, obtemos que $x_1 = x_2$ se, e somente se,

$$\begin{array}{l} (\widetilde{\beta}^2(x_Q - x_R)^2 + (((2x_P - 2x_R)(x_Q - x_R)^2 + \widetilde{\gamma}^2)(x_Q - x_P)^2 - \widetilde{\alpha}^2(x_Q - x_R)^2) \\ ((x_P + x_Q + x_R)(x_Q - x_P)^2 - \widetilde{\alpha}^2)^2)((x_P + x_Q + x_R)(x_Q - x_P)^2 - \widetilde{\gamma}^2)^2 \\ -\widetilde{\tau}^2((x_P + x_Q + x_R)(x_Q - x_P)^2 - \widetilde{\alpha}^2)^2(x_Q - x_P)^2 = 0. \end{array}$$

Expandindo $y_1 - y_2$, obtemos que $y_1 = y_2$ se, e somente se,

$$\begin{array}{l} (y_P - y_R)((x_P + x_Q + x_R)\widetilde{\eta}^2 - \widetilde{\alpha}^2)^3((x_P + x_Q + x_R)\widetilde{\mu}^2 - \widetilde{\gamma}^2)^3\widetilde{\eta}^3\widetilde{\mu}^3 \\ +\widetilde{\beta}(((2x_R - x_P - x_Q)\widetilde{\eta}^2 + \widetilde{\alpha}^2)((x_P + x_Q + x_R)\widetilde{\eta}^2 - \widetilde{\eta}^2)^2 - \widetilde{\beta}^2) \\ ((x_P + x_Q + x_R)\widetilde{\mu}^2 - \widetilde{\gamma}^2)^3\widetilde{\mu}^3 \\ -\widetilde{\tau}(((2x_P - x_Q - x_R)\widetilde{\mu}^2 + \widetilde{\gamma}^2)((x_P + x_Q + x_R)\widetilde{\eta}^2 - \widetilde{\gamma}^2)^2 - \widetilde{\tau}^2) \\ ((x_P + x_Q + x_R)\widetilde{\eta}^2 - \widetilde{\alpha}^2)^3\widetilde{\eta}^3 = 0. \end{array}$$

Por fim, estas duas igualdades destacadas podem ser verificadas, substituindo as expressões de $\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\gamma}, \widetilde{\tau}, \widetilde{\eta}$ e $\widetilde{\mu}$ e expandindo tudo como polinômio nas variáveis x_P, x_Q, x_R, y_P, y_Q e y_R. □

Os próximos dois casos particulares da associatividade também podem ser provados de forma similar ao caso acima e deixaremos a prova a cargo do leitor.

Lema 9. *Se $P, Q \neq \mathcal{O}$, $P \neq -P$, $P \neq \pm Q$, $P + P \neq \pm Q$ e $P + Q \neq \pm P$, então*

$$(P + P) + Q = P + (P + Q).$$

Lema 10. *Se $P \neq \mathcal{O}$, $P \neq -P$, $P + P \neq -(P + P)$, $(P + P) + P \neq \pm P$ e $P + P \neq \pm P$, então*

$$(P + P) + (P + P) = P + (P + (P + P)).$$

A partir de agora, faremos oito resultados sobre propriedades básicas da soma em uma curva elíptica, para conseguir provar os casos que restam da associatividade, sem abrir a prova em contas explícitas nas coordenadas.

Lema 11.

$$-(P + Q) = -P - Q.$$

Prova: Se $P = \mathcal{O}$, então $P + Q = Q$ e $-P = \mathcal{O}$. Logo vale a igualdade do enunciado. Além disso, o caso $Q = \mathcal{O}$ é análogo a este último caso.

Se $P = -Q$, então $-(P + Q) = \mathcal{O} = -P - Q$. Sendo assim, nos resta considerar o caso onde $P \neq \mathcal{O}$, $Q \neq \mathcal{O}$ e $P \neq -Q$. Neste caso, devemos utilizar a fórmula 4.1 para calcular $P + Q = (x_{PQ}, y_{PQ})$ e $-P - Q = (x'_{PQ}, y'_{PQ})$.

Note que

$$\begin{array}{rcl} x_{PQ} & = & \alpha(P,Q)^2 - x_P - x_Q \\ y_{PQ} & = & -y_P + \alpha(P,Q)(x_P - x_{PQ}) \end{array} \quad \text{e} \quad \begin{array}{rcl} x'_{PQ} & = & \alpha(-P,-Q)^2 - x_P - x_Q \\ y'_{PQ} & = & y_P + \alpha(-P,-Q)(x_P - x_{PQ}) \end{array}$$

Como $-P = (x_P, -y_P)$ e $-Q = (x_Q, -y_Q)$, temos que $\alpha(-P,-Q) = -\alpha(P,Q)$ e, portanto, $x_{PQ} = x'_{PQ}$ e $y_{PQ} = -y'_{PQ}$. □

Lema 12. *Se $P + Q = P - Q$ e $P \neq -P$, então $Q = -Q$.*

Prova: Os casos $P = \mathcal{O}$ e $Q = \mathcal{O}$ são imediatos. Se $P = Q$, então a hipótese nos diz que $Q + Q = \mathcal{O}$ e, consequentemente, $Q = -Q$. Analogamente, concluímos o mesmo se $P = -Q$.

Agora, vamos considerar os casos complementares a estes, isto é, $P, Q \neq \mathcal{O}$ e $P \neq \pm Q$. Neste caso, podemos utilizar a fórmula 4.1. Igualando as primeiras coordenadas dos pontos $P + Q$ e $P - Q$, obtemos que

$$\left(\frac{y_Q - y_P}{x_Q - x_P}\right)^2 - x_P - x_Q = \left(\frac{-y_Q - y_P}{x_Q - x_P}\right)^2 - x_P - x_Q.$$

Expandindo as potências e simplificando esta igualdade, obtemos que $-2y_P y_Q = 2y_P y_Q$. Como $P \neq -P$, obtemos que $y_P \neq 0$. Desta forma, utilizando o fato que o corpo que estamos trabalhando não possui característica 2, obtemos que $y_Q = -y_Q$, isto é, $y_Q = 0$, que finaliza a prova do lema. □

Lema 13 (Unicidade do Elemento Neutro). *Se $P + Q = P$, então $Q = \mathcal{O}$.*

PROVA: Os casos $P = \mathcal{O}$ e $P = -Q$ são imediatos. Desta forma, podemos assumir que $P \neq \mathcal{O}$ e $P \neq -Q$, ou seja, devemos usar a fórmula 4.1 para calcular $P + Q$.

Suponhamos que $Q \neq \mathcal{O}$. Escrevendo $(x_{PQ}, y_{PQ}) = P + Q = P = (x_P, y_P)$, obtemos que

$$y_P = y_{PQ} = -y_P + \alpha(P, Q)(x_P - x_{PQ}) = -y_{PQ} = -y_P,$$

já que, $x_P = x_{PQ}$ e $y_P = y_{PQ}$. Desta forma, obtemos que $y_P = 0$.

Dividiremos agora em dois casos. Primeiramente assumiremos que $x_P = x_Q$. Neste caso, $y_Q^2 = x_Q^3 + Ax_Q + B = x_P^3 + Ax_P + B = y_P^2 = 0$, ou seja, $y_Q = 0$. Assim, $Q = P = P + Q = P + P$. Portanto, $Q = \mathcal{O}$, já que, $y_P = 0$ nos diz que $P + P = \mathcal{O}$.

Por fim, temos que considerar o caso em que $x_P \neq x_Q$. Note que $x_P^3 + Ax_P + B = 0$ implica que

$$x_P^3 = -(Ax_P + B).$$

Por outro lado, $x_P = x_{PQ}$ nos diz que

$$\alpha(P, Q)^2 - x_Q - 2x_P = 0.$$

Substituindo a expressão de $\alpha(P, Q)$ e as igualdades $x_P^3 = -(Ax_P + B)$ e $y_P = 0$ na última igualdade destacada, obtemos, após simplificações que $(x_P - x_Q)(2Ax_P + 3B) = 0$. Como $x_P \neq x_Q$, temos que

$$2Ax_P + 3B = 0.$$

Esta igualdade, junto com a hipótese sobre a curva elíptica $4A^3 + 27B^2 \neq 0$, nos diz que $A \neq 0$. Sendo assim, $x_P = \frac{-3B}{2A}$. Substituindo este valor de x_P na igualdade $x_P^3 + Ax_P + B = 0$ e simplificando a expressão, obtemos que $4A^3 + 27B^2 = 0$, o que é um absurdo. □

Lema 14. *Se $P \neq -P$ e $P + P \neq -P$, então $(P + P) - P = P$.*

PROVA: Note que $P \neq \mathcal{O}$, já que, caso contrário $P = -P$. Este fato, em conjunto com as hipóteses, nos diz que as somas envolvidas em $(P + P) - P$ são todas feitas via a fórmula 4.1. Sendo assim,

$$x_{PP} = \alpha(P, P)^2 - 2x_P \text{ e } y_{PP} = -y_P + \alpha(P, P)(x_P - x_{PP}).$$

Escrevendo $(P + P) - P = (x, y)$, obtemos que

$$x = \alpha(P + P, -P)^2 - 2x_{PP} \text{ e } y = -y_{PP} + \alpha(P + P, -P)(x_{PP} - x).$$

Substituindo as expressões de x_{PP} e y_{PP} em,

$$\alpha(P + P, -P) = \frac{-y_P - y_{PP}}{x_P - x_{PP}}$$

obtemos que $\alpha(P + P, -P) = -\alpha(P, P)$. Substituindo esta igualdade e a expressão de x_{PP} em x, obtemos que $x = x_P$. Temos, então, uma das seguintes opções.

$$(P + P) - P = P \text{ ou } (P + P) - P = -P.$$

No entanto, a unicidade do elemento neutro nos diz que $P + P = \mathcal{O}$, caso tivéssemos $(P + P) - P = -P$. Porém, isto não pode ocorrer pelo fato que $P \neq -P$. □

Lema 15. *Se* $P + Q = -P$, *então* $Q = -P - P$.

Prova: Os casos $P = \mathcal{O}$, $Q = \mathcal{O}$, $P = Q$ e $P = -Q$ são simples de serem verificados. Se $P = -P$, então $-P + Q = P + Q = -P$. Segue do Lema 13 que $Q = \mathcal{O}$ e, logo, $Q = \mathcal{O} = P - P = -P - P$.

Agora iremos assumir que $P \neq \mathcal{O}$, $Q \neq \mathcal{O}$, $P \neq \pm Q$ e $P \neq -P$. Desta forma, podemos utilizar as fórmulas 4.1 para calcular todas as somas envolvidas no enunciado.

Como $P + Q = -P$, temos que

$$x_P = x_{PQ} = \left(\frac{y_Q - y_P}{x_Q - x_P}\right)^2 - x_P - x_Q,$$

já que, $P + Q \neq \mathcal{O}$ e $P \neq Q$. Após substituir nesta igualdade as expressões de y_P^2 e y_Q^2 em função de x_P e x_Q, respectivamente, obtemos que

$$2y_P y_Q = 3x_Q x_P^2 - x_P^3 + Ax_P + Ax_Q + 2B.$$

Elevando esta igualdade ao quadrado e substituindo as mesmas expressões de y_P^2 e y_Q^2, como fizemos acima, obtemos que

$$\begin{aligned} -x_P^6 + 6x_P^5 x_Q + 2Ax_P^4 - 9x_P^4 x_Q^2 + 8Bx_P^3 + 4x_P^3 x_Q^3 - A^2 x_P^2 - 6Ax_P^2 x_Q^2 - 12Bx_P^2 x_Q \\ + 2A^2 x_P x_Q + 4Ax_P x_Q^3 - A^2 x_Q^2 + 4Bx_Q^3 = 0 \end{aligned}$$

Esta igualdade, por sua vez, é equivalente a igualdade

$$\left(x_Q - \left(\left(\frac{3x_P^2 + A}{2y_P}\right)^2 - 2x_P\right)\right)(x_Q - x_P)^2 = 0.$$

Como $x_Q - x_P \neq 0$, obtemos que

$$x_Q = \left(\frac{3x_P^2 + A}{2y_P}\right)^2 - 2x_P = x_{PP},$$

isto é, $Q = P + P$ ou $Q = -(P + P) = -P - P$. Vamos verificar que a possibilidade $Q = P + P$ também implica em $Q = -P - P$, o que finaliza a prova do lema.

Note que $P + P = Q \neq -P$, pela suposição que fizemos acima. Sendo assim, $-P = P + Q = P + (P + P)$. Pelo lema anterior e pelo Lema 11, temos que $-P = P - (P + P) = P - Q$. Desta forma, $P + Q = P - Q$ e, pelo Lema 12, obtemos que $Q = -Q = -(P + P) = -P - P$. □

Lema 16 (Lei do Cancelamento). *Se $P + Q = P + R$, então $Q = R$.*

Prova: Claramente podemos considerar somente o caso em que $P \neq \mathcal{O}$. Note que, os casos $Q = \mathcal{O}$ e $P + Q = \mathcal{O}$ seguem imediatamente dos Lemas 13 e 6, respectivamente. No caso em que $P + Q = P + R = -P$, temos, pelo Lema 15, que $Q = -P - P = R$.

Portanto, podemos assumir que $P, Q, R \neq \mathcal{O}$ e que $P + Q = P + R \neq -P, \mathcal{O}$. Escrevendo $P + Q = P + R = (x, y)$, obtemos, pelas fórmulas 4.1, as seguintes igualdades.

$$\begin{array}{rcccl} x & = & \alpha(P,Q)^2 - x_P - x_Q & = & \alpha(P,R)^2 - x_P - x_R \\ y & = & -y_P + \alpha(P,Q)(x_P - x) & = & -y_P + \alpha(P,R)(x_P - x) \end{array}.$$

Como $Q \neq \mathcal{O}$, temos que $P + Q \neq \pm P$. Logo $x \neq x_P$ e, pela segunda equação evidenciada acima, obtemos que $\alpha(P, Q) = \alpha(P, R)$. Esta fato, juntamente com a primeira equação acima, nos diz que $x_Q = x_R$, isto é, $Q = R$ ou $Q = -R$.

Agora vamos considerar dois casos a serem tratados. O primeiro é quando $P = -P$. Neste caso, $Q, R \neq -P = P$, já que, temos que $P + Q = P + R \neq \mathcal{O}$. Sendo assim,

$$\frac{y_Q - y_P}{x_Q - x_P} = \alpha(P, Q) = \alpha(P, R) = \frac{y_R - y_P}{x_R - x_P},$$

junto com o fato que $x_Q = x_R$, nos diz que $y_Q = y_R$.

Por último, temos que considerar o caso $P \neq -P$. Como vimos anteriormente, $Q = R$ ou $Q = -R$. Caso ocorresse $Q = -R$, teríamos que $P + R = P + Q = P - R$. Como $P \neq -P$, então o Lema 12 nos garante que $R = -R = Q$. □

Lema 17.

$$(P + Q) - Q = P.$$

Prova: Consideremos primeiramente o caso $P + Q = -Q$. Pelo Lema 15 obtemos que $P = -Q - Q$ e portanto $(P + Q) - Q = -Q - Q = P$. Desta forma, podemos supor que $P + Q \neq -Q$.

Se $P + Q = Q$, então $(P + Q) - Q = \mathcal{O}$ e, pelo Lema 13, também obtemos que $P = \mathcal{O}$. Desta forma, podemos supor que $P + Q \neq \pm Q$.

Os casos $P = \mathcal{O}$, $Q = \mathcal{O}$ e $P = -Q$ são imediatos. Suponhamos então que $P \neq \mathcal{O}$, $Q \neq \mathcal{O}$ e $P \neq -Q$. Neste caso, se $P = Q$, então $P \neq -P$ e $P + P \neq -P$. Logo, podemos usar o Lema 14 para obter $(P + P) - P = P$. Desta forma, podemos supor que $P \neq \mathcal{O}$, $Q \neq \mathcal{O}$, $P \neq \pm Q$ e $P + Q \neq \pm Q$.

Escrevendo $(P + Q) - Q = (x, y)$ obtemos que

$$x = \alpha(P + Q, -Q)^2 - x_{PQ} - x_Q = \left(\frac{y_{PQ}+y_Q}{x_{PQ}-x_Q}\right)^2 - (x_{PQ} + x_Q) =$$
$$\left(\frac{-y_Q+\alpha(Q,P)(x_Q-x_{QP})+y_Q}{x_{PQ}-x_Q}\right)^2 - (x_{QP} + x_Q) =$$
$$\alpha(Q, P)^2 - (\alpha(Q, P)^2 - x_Q - x_P + x_Q) = x_P.$$

Analogamente, podemos verificar que $y = y_P$. □

Corolário 7. *Se $P + Q = R$, então $P = R - Q$.*

Prova: De fato, pelo lema acima, temos que $R - Q = (P + Q) - Q = P$. □

Por último, faremos um lema para abordar alguns casos que não foram contemplados ainda na prova da associatividade.

Lema 18. *Suponhamos que vale um dos seguintes itens.*

1. $P + Q \neq R$ *e* $P \neq Q + R$;
2. $P = Q$ *ou* $Q = R$ *ou* $R = P$;
3. $\mathcal{O} \in \{P, Q, R, P + Q, Q + R, (P + Q) + R, P + (Q + R)\}$.

Então,

$$(P + Q) + R = P + (Q + R).$$

Prova: Primeiramente, vamos verificar que vale a associatividade em todos os casos do terceiro item. Os casos $P = \mathcal{O}$, $Q = \mathcal{O}$ e $R = \mathcal{O}$ são imediatos. O caso $P = R$ segue da comutatividade da soma. Os casos $P = -Q$ e $R = -Q$ seguem diretamente do Lema 17. Se $P + Q = -R$, então, pelo Lema 17 e pelo Lema 11, temos que

$$(P+Q)+R = \mathcal{O} = P-P = P+(Q-(Q+P)) = P+(Q-(-R)) = P+(Q+R).$$

O caso $Q + R = -P$ é análogo a este último caso, o que termina a verificação deste item.

Agora podemos assumir que nenhum dos casos estudados acima ocorrem, isto é, $P, Q, R \neq \mathcal{O}$, $P \neq R$, $Q \neq -P$, $Q \neq -R$, $P + Q \neq -R$ e $Q + R \neq -P$.

Vamos verificar agora que a associatividade vale nos casos do segundo item.

Suponhamos primeiramente que $P = Q$. Note que, a associatividade neste caso é exatamente a igualdade mostrada no Lema 9. Pelas suposições acima, temos que as

seguintes hipóteses do Lema 9 são satisfeitas: $P, R \neq \mathcal{O}$, $P = Q \neq -P$, $P = Q \neq \pm R$, $P + P = P + Q \neq -R$ e $P + R = Q + R \neq -P$. Note que, a hipótese $P + R \neq P$ também vale, já que, caso contrário teríamos pelo Lema 13 que $R = \mathcal{O}$. Portanto, temos dois casos a considerar: $P + P \neq R$ e $P + P = R$. O primeiro caso é exatamente a hipótese que falta para podermos aplicar o Lema 9 e, portanto, vale a associatividade. Já no segundo caso, as hipóteses verificadas acima nos fornecem exatamente as hipóteses do Lema 10. Portanto, $(P + Q) + R = (P + P) + (P + P) = P + (P + (P + P)) = P + (Q + R)$.

Como já verificamos a associatividade no caso $P = R$ no início, então o único caso que nos resta, para terminar a verificação da associatividade no segundo item, é quando $Q = R$. Este caso é análogo ao de cima e deixaremos como exercício para o leitor.

Agora podemos considerar que todos os casos feitos acima não são verificados, isto é, $P, Q, R \neq \mathcal{O}$, $P \neq R$, $P \neq \pm Q$, $Q \neq \pm R$, $P + Q \neq -R$ e $Q + R \neq -P$. Desta forma, adicionando as hipóteses do primeiro item recaímos exatamente nas hipóteses do Lema 8, onde já sabemos que vale a associatividade. □

Agora já temos todos os ingredientes para provar a associatividade em todos os casos.

Teorema 24. *Sejam $P, Q, R \in E$. Então*

$$(P + Q) + R = P + (Q + R).$$

Prova: Temos que verificar a associatividade nos casos que não foram contemplados no Lema 18. Pelo item 1 deste lema temos que considerar $P, Q, R \in E$ satisfazendo $P + Q = R$ ou $P = Q + R$. Nestes casos, a associatividade se resume a provar as igualdades

$$\begin{aligned} (P + Q) + (P + Q) &= P + (Q + (P + Q)) \\ ((Q + R) + Q) + R &= (Q + R) + (Q + R), \end{aligned}$$

respectivamente. Note que eles são similares, já que, temos a comutatividade da soma. Desta forma, iremos considerar apenas o caso $P + Q = R$.

Pelos itens 2 e 3 do Lema 18 temos que considerar $P, Q, R \in E$ satisfazendo

$$P, Q, P + Q \neq \mathcal{O}, P \neq \pm Q, Q \neq \pm P + Q, P + Q \neq P, (P + Q) + (P + Q) \neq \mathcal{O}, \\ P + (Q + (P + Q)) \neq \mathcal{O}.$$

Pelo segundo item do Lema 18 temos que $((P + Q) + (P + Q)) - P = (P + Q) + ((P + Q) - P)$. Aplicando o Lema 17 duas vezes, obtemos que $(P + Q) + ((P + Q) - P) = (P + Q) + Q = (P + (Q + (P + Q))) - P$, onde a segunda igualdade foi obtida com $Q + (P + Q)$ no lugar de P e P no lugar de Q. Resumindo, obtemos que $((P + Q) + (P + Q)) - P = (P + (Q + (P + Q))) - P$. Por fim, pelo Lema 16, obtemos que $(P + Q) + (P + Q) = P + (Q + (P + Q))$. □

4.1.3 Curvas elípticas sobre corpos finitos

Seja $q = p^n$, com $p > 3$ primo e n um inteiro positivo. Uma *curva elíptica* sobre $\mathbb{F}_q$ é uma curva sobre $\overline{\mathbb{F}_p}$ cuja equação é da forma

$$Y^2 = X^3 + AX + B \text{ com } A, B \in \mathbb{F}_q \text{ e } 4A^3 + 27B^2 \neq 0.$$

Esta equação que define a curva elíptica é chamada de *forma de Weierstrass* de E.

Já o conjunto dos pontos que satisfazem a equação de E e que tem coordenadas em $\mathbb{F}_q$ será denotado por

$$E(\mathbb{F}_q) = \{(x, y) \mid x, y \in \mathbb{F}_q \text{ e } y^2 = x^3 + Ax + B\} \cup \{\mathcal{O}\}$$

e será denominado por *pontos racionais* de E sobre $\mathbb{F}_q$.

Considere a curva elíptica $E : Y^2 = X^3 + 3X + 8$ sobre $\mathbb{F}_{13}$. Podemos encontrar os pontos de $E(\mathbb{F}_{13})$ substituindo cada $x \in \mathbb{F}_{13}$ no polinômio $X^3 + 3X + 8$ e decidindo se este valor é um quadrado ou não.

y	y^2
0	0
1	1
2	4
3	9
4	3
5	12
6	10
7	10
8	12
9	3
10	9
11	4
12	1

x	$x^3 + 3x + 8$
0	8
1	12
2	9
3	5
4	6
5	5
6	8
7	8
8	11
9	10
10	11
11	7
12	4

Pela tabela que montamos, temos que

$$E(\mathbb{F}_{13}) = \{\mathcal{O}, (\overline{1}, \overline{5}), (\overline{1}, \overline{8}), (\overline{2}, \overline{3}), (\overline{2}, \overline{10}), (\overline{9}, \overline{6}), (\overline{9}, \overline{7}), (\overline{12}, \overline{2}), (\overline{12}, \overline{11})\}.$$

Podemos construir uma estrutura de grupo para o conjunto $E(\mathbb{Z}_p)$ de duas formas.

1. Definir geometria sobre corpos quaisquer, o que é o propósito da Geometria Algébrica;
2. Utilizar as formulas que observamos no Teorema 22. Como as formulas deste teorema envolvem apenas as operações de soma, subtração, produto, produto por $2 \neq 0$ em $\mathbb{F}_q$ ($p > 3$) e divisão nas coordenadas, então temos que a soma de dois pontos de $E(\mathbb{F}_q)$ é um ponto em $E(\mathbb{F}_q)$, já que, $\mathbb{F}_q$ é um corpo.

Portanto, $E(\mathbb{F}_q)$ com as operações definidas no Teorema 22 é um grupo finito. Com relação a ordem deste grupo, podemos enunciar o seguinte resultado.

Teorema 25 (Hasse). *Seja E uma curva elíptica sobre $\mathbb{F}_q$. Então $|E(\mathbb{F}_q)| = q + 1 - t_q$, com $|t_q| \leqslant 2\sqrt{q}$.*

A prova deste teorema pode ser vista no Capítulo V de Silverman (2009).

Este valor t_q é chamado de *traço de Frobenius* de E e é obtido como o traço de uma matriz.

É possível generalizar a construção do Teorema 22 para curvas elípticas definidas sobre corpos de característica $p = 2$ ou $p = 3$. Primeiro o que é uma tal curva.

Uma *curva elíptica* sobre $\mathbb{F}_q$ é uma curva sobre $\overline{\mathbb{F}_p}$ cuja equação é da forma

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6 \text{ com } a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}_q$$

satisfazendo

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6 \neq 0,$$

onde $b_2 = a_1^2 + 4a_2$, $b_4 = 2a_4 + a_1a_3$, $b_6 = a_3^2 + 4a_6$, $b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2$.

Esta equação que define a curva elíptica é chamada de *forma de Weierstrass generalizada* de E.

No Exemplo 36 que utilizamos o software Pari, o comando $ellinit([x])$ fornece os dados da curva elíptica com equação na forma generalizada, onde o vetor $[x]$ nada mais é que o vetor de coeficientes $[a_1, a_2, a_3, a_4, a_6]$, ou apenas $[a_4, a_6]$, no caso em que $a_1 = a_2 = a_3 = 0$.

A estrutura de grupo sobre as curvas elípticas dadas por estas novas equações são construídas de forma análoga a estrutura de grupo que construímos anteriormente, com as seguintes alterações.

Ao invés de considerar o simétrico $(x, -y)$ de um ponto $(x, y) \in E$, consideramos o ponto $(x, -y - a_1x - a_3)$. Este ponto entra na definição da estrutura de grupo, assim como fornece o inverso aditivo de (x, y). Com esta modificação, a estrutura de grupo em tais curvas são resumidas da seguinte forma.

Teorema 26. *Seja* $E : Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$ *uma curva elíptica e sejam* P_1 *e* P_2 *em* E. *Então:*

- *Se* $P_1 = \mathcal{O}$, *então* $P_1 + P_2 = P_2$;
- *Se* $P_2 = \mathcal{O}$, *então* $P_1 + P_2 = P_1$;

Se nenhum destes casos ocorrem, escrevemos $P_1 = (x_1, y_1)$ *e* $P_2 = (x_2, y_2)$ *e então:*

- *Se* $x_1 = x_2$ *e* $y_1 = -y_2 - a_1x_2 - a_3$, *então* $P_1 + P_2 = \mathcal{O}$;
- *Caso contrário,* $P_1 + P_2 = (x_3, y_3)$ *onde*

$$x_3 = \lambda^2 + \lambda a_1 - a_2 - x_1 - x_2 \;\; e \;\; y_3 = -(\lambda + a_1)x_3 - \nu - a_3,$$

com

$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & se\ P_1 \neq P_2 \\ \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} & se\ P_1 = P_2. \end{cases}$$

e

$$\nu = \begin{cases} \frac{y_1x_2 - y_2x_1}{x_2 - x_1} & se\ P_1 \neq P_2 \\ \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3} & se\ P_1 = P_2. \end{cases}$$

As provas deste teorema e do fato que esta soma fornece uma estrutura de grupo em E não serão feitas.

4.2 A estrutura de grupo em uma cúbica lisa

Nesta seção abordaremos, de forma mais geral do que foi feito anteriormente, a soma de pontos em uma curva elíptica, sem utilizar a forma de Weierstrass. Neste sentido, precisamos fazer a definição mais geral de tais curvas.

Uma *curva elíptica* E é uma cúbica plana projetiva e lisa sobre um corpo algebricamente fechado K.

Dados $P, Q \in E$, consideremos a reta L_{PQ} em $\mathbb{P}^2(K)$ definida da seguinte forma.

1. Se $P \neq Q$, então L_{PQ} é a única reta contendo P e Q;
2. Se $P = Q$, então L_{PQ} é a reta tangente a E em P.

Pelo Teorema de Bézout temos que, em ambos os casos, existe um terceiro ponto na interseção entre L_{PQ} e E, digamos R. Desta forma, temos uma aplicação

$$\begin{array}{rccc} \varphi & : & E \times E & \to & E \\ & & (P, Q) & \mapsto & R \end{array}$$

Fixado um ponto $\mathcal{O} \in E$ definimos a adição $\oplus$ em E por $P \oplus Q = \varphi(\mathcal{O}, \varphi(P, Q))$.

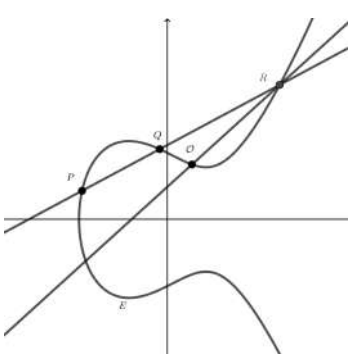


Figura 4.9: Soma entre pontos em um Curva Elíptica

Teorema 27. *A curva elíptica E, com a estrutura de soma $\oplus$, é um grupo abeliano cujo elemento neutro é $\mathcal{O}$.*

Antes de provar o teorema acima, vamos enunciar um resultado que nos auxiliará.

Proposição 22 (dos 9 pontos). *Considere C uma cúbica plana irredutível e C' e C'' cúbicas. Se $C \cap C' = \{P_1, \ldots, P_9\}$, com $P_1, \ldots, P_9$ pontos simples de C, e $C \cap C'' = \{P_1, \ldots, P_8, Q\}$, então $Q = P_9$.*

A prova desta proposição é bem trabalhosa e utiliza resultados mais avançados de um curso de curvas algébricas. No entanto, com a utilização de argumentos geométricos mais simples, poderemos fazer uma prova de um caso particular desta proposição, que fornece a associatividade para três pontos genéricos na curva elíptica. Tal demonstração segue as ideias apresentadas em Reid 2013.

Prova do Teorema 27: Primeiramente vamos provar que $\mathcal{O}$ é o elemento neutro desta soma. Para isto, seja $R = \varphi(\mathcal{O}, P)$. Então, $P \oplus \mathcal{O} = \varphi(\mathcal{O}, \varphi(\mathcal{O}, P)) = \varphi(\mathcal{O}, R) = P$.

A comutatividade da soma $\oplus$ segue do fato que $\varphi(P, Q) = \varphi(Q, P)$. Já construção do inverso de um ponto P se faz da seguinte forma. Consideremos $R = \varphi(\mathcal{O}, \mathcal{O})$ e $Q = \varphi(P, R)$. Desta forma,

$$P \oplus Q = \varphi(\mathcal{O}, \varphi(P, Q)) = \varphi(\mathcal{O}, R) = \mathcal{O},$$

isto é, $Q = -P$.

Para a prova da associatividade, consideremos $P, Q, R \in E$. Escrevemos $S' = \varphi(P, Q)$, $S = \varphi(\mathcal{O}, S')$, $T' = \varphi(S, R)$, $U' = \varphi(Q, R)$, $U = \varphi(\mathcal{O}, U')$ e $T'' = \varphi(P, U)$. Note quebrada

$$(P \oplus Q) \oplus R = \varphi(\mathcal{O}, S') \oplus R = S \oplus R = \varphi(\mathcal{O}, T')$$

e

$$P \oplus (Q \oplus R) = P \oplus \varphi(\mathcal{O}, U') = P \oplus U = \varphi(\mathcal{O}, T'').$$

Desta forma, basta provar que $T' = T''$. Para isto, considere as cúbicas $C' = L_{PQ} L_{SR} L_{\mathcal{O}U'}$ e $C'' = L_{\mathcal{O}S'} L_{QR} L_{PU}$. Note que

$$E \cap C' = \{P, Q, S', S, R, T', \mathcal{O}, U', U\} \text{ e } E \cap C'' = \{\mathcal{O}, S', S, Q, R, U', P, U, T''\}.$$

Comparando estas duas interseções e usando a Proposição 22, obtemos que $T' = T''$. □

4.3 Sistemas lineares de cônicas

Antes de entrar no tema da seção, vamos provar o seguinte resultado.

Proposição 23. *Sejam $P_1, \ldots, P_5 \in \mathbb{P}^2(K)$ pontos distintos tais que não tenhamos quatro deles colineares. Então existe uma única cônica plana projetiva contendo $P_1, \ldots, P_5$.*

PROVA: Primeiramente vamos verificar que temos uma tal cônica. Escrevemos $P_i = (x_i : y_i : z_i)$ e a equação geral de uma cônica em $\mathbb{P}^2(K)$ da forma

$$F = a_{20}X^2 + a_{11}XY + a_{02}Y^2 + a_{10}XZ + a_{01}YZ + a_{00}Z^2$$

com $a_{ij} \in K$. Note que, uma cônica F satisfazendo $F(x_i, y_i, z_i) = 0$, para cada $i = 1, \ldots, 5$, existe se, e somente se, existe um vetor $(a_{20}, a_{11}, a_{02}, a_{10}, a_{01}, a_{00}) \in K^6$ no núcleo da transformação linear de K^6 para K^5 induzida pela matriz

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1z_1 & y_1z_1 & z_1^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_5^2 & x_5y_5 & y_5^2 & x_5z_5 & y_5z_5 & z_5^2 \end{bmatrix}.$$

Pelo Teorema do Núcleo e da Imagem, esta transformação linear possui pelo menos um vetor no núcleo e, portanto, pelo menos uma tal cônica existe.

Agora vamos verificar a unicidade. Suponhamos que existam duas cônicas $C_1 \neq C_2$ passando por $P_1, \ldots, P_5$.

Suponhamos primeiramente que C_1 é uma cônica lisa. Então, após uma mudança de coordenadas projetivas, temos que C_1 é equivalente a curva parametrizada

$$C_1 = \{(u^2 : v^2 : uv) | (u : v) \in \mathbb{P}^1(K)\}.$$

Como $\{P_1, \ldots, P_5\} \subset C_1 \cap C_2$ temos, pelo Teorema de Bézout que $C_1 \subset C_2$.

Como consequência desta inclusão temos que se $F_1 = Z^2 - XY$ e F_2 são as equações que determinam C_1 e C_2, respectivamente, então

$$F_2 = \lambda F_1 \text{ com } \lambda \in K,$$

contrariando o fato que $C_1 \neq C_2$. De fato, para provar esta igualdade evidenciada acima, escrevemos $Z^2 = F_1 - XY$ e substituímos na equação $F_2 = F = a_{20}X^2 + a_{11}XY + a_{02}Y^2 + a_{10}XZ + a_{01}YZ + a_{00}Z^2$ para obter

$$F_2 = \lambda F_1 + A(X, Y) + ZB(X, Y)$$

com $\lambda = a_{00}$, $A(X, Y) = a_{20}X^2 + (a_{11} - a_{00})XY + a_{02}Y^2$ e $B(X, Y) = a_{10}X + a_{01}Y$. Sendo assim,

$$\begin{aligned} C_1 \subset C_2 &\iff A(u^2, v^2) + uvB(u^2, v^2) = 0 \text{ para cada } (u : v) \in \mathbb{P}^1(K) \\ &\iff A(U^2, V^2) + UVB(U^2, V^2) = 0 \text{ em } K[U, V]. \end{aligned}$$

Comparando os termos de grau par e grau ímpar nesta última igualdade obtemos que

$$C_1 \subset C_2 \iff A(X, Y) = B(X, Y) = 0.$$

Agora suponhamos que C_1 não é uma cônica lisa. Então, após uma mudança de coordenadas projetivas, temos que $C_1 = L_0 \cup L_1$ com L_0 e L_1 podendo ser retas iguais ou não. Note que, três dos pontos $P_1, \ldots, P_5$ têm que pertencer a uma destas retas. Sem perda de generalidade, suponhamos que eles pertençam a reta L_0. Aplicando o Teorema de Bézout para a reta L_0 e a cônica C_2 obtemos que $C_2 = L_0 \cup L_2$. Como $C_1 \neq C_2$, temos que $L_1 \neq L_2$. Desta forma, como

$$\{P_1, \ldots, P_5\} \subset C_1 \cap C_2 = L_0 \cup (L_1 \cap L_2)$$

e $L_1 \cap L_2$ consiste de um único ponto concluímos que quatro dos cinco pontos $P_1, \ldots, P_5$ estão em L_0, que é um absurdo. □

Consideremos o conjunto S_2 dos polinômios homogêneos de grau dois em $K[X, Y, Z]$. Este conjunto é chamado de *sistema linear de cônicas* no plano projetivo. Como todo polinômios de grau dois é dado por uma equação da forma

$$a_{20}X^2 + a_{11}XY + a_{02}Y^2 + a_{10}XZ + a_{01}YZ + a_{00}Z^2$$

com $a_{ij} \in K$, temos que S_2 pode ser identificado com o espaço vetorial K^6.

Fixemos agora $P_0 = (x_0 : y_0 : z_0) \in \mathbb{P}^2(K)$. Então, a identificação $S_2 \simeq K^6$ identifica o conjunto

$$S_2(P_0) = \{F \in S_2 | F(x_0, y_0, z_0) = 0\}$$

com o núcleo da transformação linear de K^6 em K induzida pela matriz

$$\begin{bmatrix} x_0^2 & x_0y_0 & y_0^2 & x_0z_0 & y_0z_0 & z_0^2 \end{bmatrix},$$

que é claramente uma transformação linear não nula. Logo,

$$S_2(P_0) \simeq K^5 \subset K^6 = S_2.$$

Similarmente, fixados $P_1, \ldots, P_n \in \mathbb{P}^2(K)$, definimos o *sistema linear de cônicas passando pelos pontos* $P_1, \ldots, P_n$, pelo conjunto

$$S_2(P_1, \ldots, P_n) = \{F \in S_2 | F(P_i) = 0 \text{ para todo } i = 1, \ldots, n\},$$

que é identificado com o núcleo da transformação linear de K^6 em K^n induzida pela matriz

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1z_1 & y_1z_1 & z_1^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_ny_n & y_n^2 & x_nz_n & y_nz_n & z_n^2 \end{bmatrix}.$$

Pelo Teorema do Núcleo e da Imagem obtemos o seguinte resultado.

Proposição 24. $\dim_K S_2(P_1, \ldots, P_n) \geqslant 6 - n$.

Podemos dizer mais no caso em que os pontos $P_1, \ldots, P_n$ são suficientemente gerais.

Corolário 8. *Se $n \leqslant 5$ e não temos quatro dos pontos $P_1, \ldots, P_n$ que sejam colineares, então*

$$\dim_K S_2(P_1, \ldots, P_n) = 6 - n$$

Prova: Pela Proposição 23, temos que $\dim_K S_2(P_1, \ldots, P_5) = 1$. Suponhamos agora que $n \leqslant 4$.

Note que, podemos adicionar pontos $P_{n+1}, \ldots, P_5$ de modo que não tenhamos quatro dos pontos $P_1, \ldots, P_5$ que sejam colineares. De fato, admitindo o lema que provaremos logo após esta prova, podemos proceder por indução da seguinte forma. Consideramos P_{n+1} que esteja fora de todas as retas que ligam cada par de pontos em $P_1, \ldots, P_n$, depois consideremos P_{n+2} que esteja fora de todas as retas que ligam cada par de pontos em $P_1, \ldots, P_{n+1}$ e assim sucessivamente.

Aplicando sucessivamente o Teorema do Núcleo e da Imagem obtemos

$$\begin{aligned} 1 &= \dim_K S_2(P_1, \ldots, P_5) \geqslant \dim_K S_2(P_1, \ldots, P_4) - 1 \geqslant \dim_K S_2(P_1, \ldots, P_3) - 2 \\ &\geqslant \cdots \geqslant \dim_K S_2(P_1, \ldots, P_n) - (5 - n). \square \end{aligned}$$

Lema 19. *Suponhamos que K seja infinito e sejam $L_1, \ldots, L_n$ retas em $\mathbb{P}^2(K)$. Então existe $P \in \mathbb{P}^2(K) \setminus \bigcup_{i=1}^n L_i$.*

Prova: Considere $F \in K[X, Y, Z]$ o produto das equações que definem as retas $L_1, \ldots, L_n$. Note que

$$\bigcup_{i=1}^{n} L_i = \{P \in \mathbb{P}^2(K) | F(P) = 0\}.$$

Desta forma, $\mathbb{P}^2(K) = \bigcup_{i=1}^n L_i$ nos diria que $F(x, y, z) = 0$ para todo $x, y, z \in K$.

Afirmamos que este fato implica que $F = 0$ em $K[X, Y, Z]$, o que é um absurdo. De fato, escrevendo $F = \sum A_i(X, Y)Z^i$ com $A_i(X, Y) \in K[X, Y]$ obtemos que, para cada $x, y \in K$ fixado, então o polinômio

$$F(x, y, Z) = \sum A_i(x, y)Z^i \in K[Z]$$

possui infinitas raizes. Logo $F(x, y, Z) = 0$, isto é, $A_i(x, y) = 0$ para cada i. Procedendo de forma indutiva obtemos que cada $A_i(X, Y) = 0$, isto é, $F = 0$. $\square$

4.4 Sistemas lineares

Consideremos d um inteiro positivo e S_d o conjunto de todos os polinômios homogêneos de grau d. Este conjunto é chamado de *sistema linear* de curvas de grau d no plano

projetivo. Cada elemento $F \in S_d$ pode ser escrito de forma única como

$$F = \sum_{i+j+k=d} a_{ijk} X^i Y^j Z^k$$

com $a_{ijk} \in K$. Desta forma, S_d é um espaço vetorial sobre K tendo como base o conjunto formado pelos seguintes elementos.

$$Z^d \longleftarrow \text{1 elemento}$$

$$Z^{d-1}X \quad Z^{d-1}Y \longleftarrow \text{2 elementos}$$

$$Z^{d-2}X^2 \quad Z^{d-2}XY \quad Z^{d-2}Y^2 \longleftarrow \text{3 elementos}$$

$$\vdots$$

$$ZX^{d-1} \quad ZX^{d-2}Y \quad \cdots \quad ZY^{d-1} \longleftarrow d \text{ elementos}$$

$$X^d \quad X^{d-1}Y \quad X^{d-2}Y^2 \quad \cdots \quad Y^d \longleftarrow d+1 \text{ elementos}$$

Desta forma, $\dim_K S_d = 1 + 2 + 3 + \cdots + d + 1 = \begin{pmatrix} d+1 \\ 2 \end{pmatrix}$.

Fixados $P_1, \ldots, P_n \in \mathbb{P}^2(K)$, definimos o *sistema linear de curvas de grau d passando pelos pontos* $P_1, \ldots, P_n$, pelo conjunto

$$S_d(P_1, \ldots, P_n) = \{F \in S_d | F(P_i) = 0 \text{ para todo } i = 1, \ldots, n\}.$$

Note que, a identificação de S_d com $\begin{pmatrix} d+1 \\ 2 \end{pmatrix}$, identifica $S_d(P_1, \ldots, P_n)$ com o núcleo da transformação linear de $\begin{pmatrix} d+1 \\ 2 \end{pmatrix}$ em K^n induzida pela seguinte matriz.

$$\begin{bmatrix} z_1^d & z_1^{d-1}x_1 & z_1^{d-1}y_1 & \cdots & y_1^d \\ \vdots & \vdots & \vdots & & \vdots \\ z_n^d & z_n^{d-1}x_n & z_n^{d-1}y_n & \cdots & y_n^d \end{bmatrix}.$$

Pelo Teorema do Núcleo e da Imagem obtemos o seguinte resultado.

Proposição 25. $\dim_K S_d(P_1, \ldots, P_n) \geqslant \begin{pmatrix} d+1 \\ 2 \end{pmatrix} - n.$

Lema 20. *Suponha que K seja um corpo infinito e que $F \in S_d$.*

1. *Seja $L \subset \mathbb{P}^2(K)$ uma reta. Se F se anula em todos os pontos de L, então $F = HF'$, onde $H, F' \in K[X, Y, Z]$, H é a equação que define a reta L e $F' \in S_{d-1}$.*

2. *Seja $C \subset \mathbb{P}^2(K)$ uma cônica lisa. Se F se anula em todos os pontos de C, então $F = QF'$, onde $Q, F' \in K[X,Y,Z]$, Q é a equação que define a cônica C e $F' \in S_{d-2}$.*

PROVA: (1) Após uma mudança de coordenadas, podemos assumir que $H = X$. Agrupando todos os monômios de F que possuem alguma potência não nula de X, podemos escrever $F = XF' + G(Y,Z)$ com $F' \in S_{d-1}$ e com $G(Y,Z)$ sendo um polinômio só nas variáveis Y e Z.

Desta forma, $L = \{(0:y:z) | (y:z) \in \mathbb{P}^1(K)\}$ e portanto

$$F(P) = 0 \text{ para todo } P \in L \iff G(y,z) = 0 \text{ para todo } (y:z) \in \mathbb{P}^1(K).$$

Como K é um corpo infinito, isto só pode ocorrer se $G(Y,Z) = 0$.

(2) Após uma mudança de coordenadas, podemos assumir que $Q = Z^2 - XY$, isto é, $Z^2 = Q + XY$. Observe que esta escrita de Z^2 nos permite reescrever todas as possíveis potências de Z. De fato,

$$\begin{aligned} Z^{2n} &= \left(Z^2\right)^n = (Q+XY)^n = QG + X^nY^n \\ Z^{2n+1} &= \left(Z^2\right)^n Z = (Q+XY)^n Z = QG' + X^nY^nZ \end{aligned}$$

onde $G, G' \in K[X,Y,Z]$ são polinômios homogêneos. Reescrevendo os monômios de F com as novas escritas das potências de Z, podemos escrever $F = QF' + A(X,Y) + ZB(X,Y)$ com $F' \in S_{d-2}$ e com $A(X,Y), B(X,Y)$ polinômios homogêneos só nas variáveis X e Y.

Como $Q = \{(u^2 : v^2 : uv) | (u:v) \in \mathbb{P}^1(K)\}$ temos que $F(P) = 0$ para todo $P \in Q$ se, e somente se,

$$A(u^2, v^2) + uvB(u^2, v^2) = 0 \text{ para todo } (u:v) \in \mathbb{P}^1(K).$$

Como K é um corpo infinito, isto só pode ocorrer quando

$$A(U^2, V^2) + UVB(U^2, V^2) = 0 \text{ em } K[U,V].$$

Escrevendo $A(X,Y) = \sum A_i(X)Y^i$ e $B(X,Y) = \sum B_j(X)Y^j$, com $A_i(X), B_i(X) \in K[X]$, obtemos que

$$0 = A(U^2, V^2) + UVB(U^2, V^2) = \sum A_i(U^2)V^{2i} + \sum UB_j(U^2)V^{2j+1}$$

que, por sua vez, implica que $A_i(U^2) = UB_j(U^2) = 0$ para cada i e cada j, isto é, $A_i(X) = B_j(X) = 0$ para cada i e cada j, isto é, $A(X,Y) = B(X,Y) = 0$. □

Corolário 9. *Sejam $L \subset \mathbb{P}^2(K)$ uma reta dada por uma equação H, $C \subset \mathbb{P}^2(K)$ uma cônica lisa dada por uma equação Q e $P_1, \ldots, P_n \in \mathbb{P}^2(K)$. Então*

1. *Se* $P_1, \ldots, P_a \in L$, $P_{a+1}, \ldots, P_n \notin L$ *e* $a > d$, *então*

$$S_d(P_1, \ldots, P_n) = H \cdot S_{d-1}(P_{a+1}, \ldots, P_n)$$

2. *Se* $P_1, \ldots, P_a \in C$, $P_{a+1}, \ldots, P_n \notin C$ *e* $a > 2d$, *então*

$$S_d(P_1, \ldots, P_n) = Q \cdot S_{d-2}(P_{a+1}, \ldots, P_n)$$

Prova: Faremos somente a prova do primeiro item, já que, a prova do segundo é igual. Seja $F \in S_d(P_1, \ldots, P_n)$ e D a curva de equação F. Como $\{P_1, \ldots, P_a\} \subset L \cap D$ e $a > d$ temos, pelo Teorema de Bézout, que $L \subset D$. Pelo lema anterior, $F = HF'$ com $F' \in S_{d-1}$. Como $P_{a+1}, \ldots, P_n$ anulam F e não anulam H, temos que eles deverão anular F', isto é, $F' \in S_{d-1}(P_{a+1}, \ldots, P_n)$. □

Agora, faremos o resultado chave para conseguirmos provar a Proposição 22, que ficou restando para terminar a prova da associatividade da soma em uma curva elíptica.

Proposição 26. *Seja* K *um corpo infinito e sejam* $P_1, \ldots, P_8 \in \mathbb{P}^2(K)$ *pontos distintos sem que quatro deles sejam colineares e sem que sete deles sejam concônicos, isto é, estejam em uma cônica lisa. Então,*

$$\dim_K S_3(P_1, \ldots, P_8) = 2.$$

Prova: Dividiremos a prova em três casos. Primeiramente note que, em geral, a Proposição 25, temos que

$$\dim_K S_3(P_1, \ldots, P_8) \geqslant 2.$$

Caso genérico: Quaisquer três dos oito pontos não são colineares e quaisquer seis dos oito pontos não são concônicos.

Suponhamos que $\dim_K S_3(P_1, \ldots, P_8) \geqslant 3$ e sejam $P_9 = (x_9 : y_9 : z_9)$ e $P_{10} = (x_{10} : y_{10} : z_{10})$ pontos distintos na reta L que passa por P_1 e P_2. Note que,

$$S_3(P_1, \ldots, P_{10}) \subset S_3(P_1, \ldots, P_8)$$

e que a identificação de S_3 com o K^{10} identifica $S_3(P_1, \ldots, P_{10})$ com o núcleo da transformação linear de $S_3(P_1, \ldots, P_8)$ em K^2 cuja matriz é dada por

$$\begin{bmatrix} z_9^3 & z_9^2 x_9 & z_9^2 y_9 & \cdots & y_9^3 \\ z_{10}^3 & z_{10}^2 x_{10} & z_{10}^2 y_{10} & \cdots & y_{10}^3 \end{bmatrix}.$$

Pelo Teorema do Núcleo e da Imagem, temos que

$$\dim_K S_3(P_1, \ldots, P_{10}) \geqslant S_3(P_1, \ldots, P_8) - 2 \geqslant 1.$$

Logo, existe $F \in S_3(P_1, \ldots, P_{10})$ não nulo. Observe que $P_1, P_2, P_9, P_{10} \in L$ e que, o fato de cada três dos oito pontos $P_1, \ldots, P_8$ não serem colineares nos diz que, $P_3, \ldots, P_8 \notin$

L. Desta forma, o primeiro item do corolário anterior nos fornece que $F = HQ$, onde H é a equação da reta L e $Q \in S_2(P_3, \ldots, P_8)$.

Se a curva dada por Q fosse uma cônica lisa teríamos que os pontos $P_3, \ldots, P_8$ seriam concônicos. Como a hipótese do caso em questão nos diz que isto não é possível, temos que a curva dada por Q é uma união de retas. Porém, isto também não é possível, já que, teríamos três dos pontos $P_3, \ldots, P_8$ sendo colineares. Portanto, $\dim_K S_3(P_1, \ldots, P_8) = 2$.

Primeiro caso especial: Suponhamos que três dos oito pontos são colineares.

Sem perda de generalidade, vamos supor que P_1, P_2 e P_3 pertencem a uma reta L de equação H. Note que, a hipótese da proposição nos diz que $P_4, \ldots, P_8 \notin L$. Considere P_9 um ponto na reta L distinto dos pontos P_1, P_2 e P_3. Pelo corolário anterior temos que

$$S_3(P_1, \ldots, P_9) = H \cdot S_2(P_4, \ldots, P_8).$$

Novamente pela hipótese desta proposição, não podemos ter quatro dos pontos $P_4, \ldots, P_8$ que sejam colineares. Portanto, o Corolário 8 nos diz que $\dim_K S_2(P_4, \ldots, P_8) = 1$, isto é, $\dim_K S_3(P_1, \ldots, P_9) = 1$. Por outro lado, a identificação entre S_2 e K^6 também identifica $S_3(P_1, \ldots, P_9)$ com o núcleo da transformação linear de $S_3(P_1, \ldots, P_8)$ em K definida pela matriz

$$\begin{bmatrix} z_9^3 & z_9^2 x_9 & z_9^2 y_9 & \cdots & y_9^3 \end{bmatrix},$$

onde $P_9 = (x_9 : y_9 : z_9)$. Portanto, o Teorema do Núcleo e da Imagem implica que $\dim_K S_3(P_1, \ldots, P_8) \leqslant 2$.

Segundo caso especial: Suponhamos que seis dos oito pontos são concônicos.

Sem perda de generalidade, vamos supor que $P_1, \ldots, P_6$ pertencem a uma cônica lisa C de equação Q. Note que a hipótese da proposição nos diz que $P_7, P_8 \notin C$. Considere P_9 um ponto em C distinto dos pontos $P_1, \ldots, P_6$. Pelo corolário anterior temos que

$$S_3(P_1, \ldots, P_9) = Q \cdot S_1(P_7, P_8).$$

Como $P_7 \neq P_8$ temos que só existe uma única reta L que passa por P_7 e P_8. Desta forma $\dim_K S_3(P_1, \ldots, P_9) = \dim_K S_1(P_7, P_8) = 1$. Por outro lado, a identificação entre S_2 e K^6 também identifica $S_3(P_1, \ldots, P_9)$ com o núcleo da transformação linear de $S_3(P_1, \ldots, P_8)$ em K definida pela matriz

$$\begin{bmatrix} z_9^3 & z_9^2 x_9 & z_9^2 y_9 & \cdots & y_9^3 \end{bmatrix},$$

onde $P_9 = (x_9 : y_9 : z_9)$. Portanto, o Teorema do Núcleo e da Imagem implica que $\dim_K S_3(P_1, \ldots, P_8) \leqslant 2$. □

Como consequência desta proposição vamos provar um caso particular da Proposição 22.

Corolário 10. *Considere C_1, C_2 duas cúbicas com interseção $C_1 \cap C_2 = \{P_1, \ldots, P_9\}$, com $P_1, \ldots, P_9$ pontos distintos. Então, uma cúbica C passando por $P_1, \ldots, P_8$ também passa por P_9.*

PROVA: Se quatro dos oito pontos $P_1, \ldots, P_8$ estivessem em uma reta L, então o Teorema de Bézout nos diria que $L \subset C_1 \cap C_2$, contrariando o fato que $C_1 \cap C_2 = \{P_1, \ldots, P_9\}$. Com o mesmo argumento, podemos concluir que não podemos ter sete destes oito pontos pertencendo a uma cônica lisa. Desta forma, as hipóteses da proposição anterior são satisfeitas e, portanto, $\dim_K S_3(P_1, \ldots, P_8) = 2$.

Considere F_1 e F_2 equações para as curvas C_1 e C_2, respectivamente. Note que, $F_1, F_2 \in S_3(P_1, \ldots, P_8)$ são linearmente independentes, já que, caso contrário teríamos $F_1 = \lambda F_2$, contrariando a igualdade $C_1 \cap C_2 = \{P_1, \ldots, P_9\}$. Sendo assim, podemos escrever a equação G de C por $G = \alpha F_1 + \beta F_2$. Como F_1 e F_2 se anulam em P_9, então o mesmo ocorre com G. □

Por fim, observamos que este corolário fornece a prova da associatividade no caso em que todos os pontos $P, Q, S', S, R, T', \mathcal{O}, U', U$, da prova do Teorema 27 , são distintos. Neste sentido, temos uma prova mais simples da associatividade para uma escolha genérica de três pontos na curva elíptica.

4.5 Exercícios

Questão 29. ⋆ *Vamos estudar alguns grupos sobre curvas elípticas com a ajuda do Pari.*

Dada a curva $E : Y^2 = X^3 + 2X + 1$ *e o corpo finito* $\mathbb{F}_{13}$, *o grupo* $E(\mathbb{F}_{13})$ *tem 8 elementos.*

Vamos começar nomeando a curva e informando ao programa:

```
?   E = ellinit([0, 0, 0, 2, 1], 13)
%1  = [Mod(0, 13), Mod(0, 13), Mod(0, 13), Mod(2, 13), Mod(1, 13), Mod(0, 13),
    Mod(4, 13), Mod(4, 13), Mod(9, 13), Mod(8, 13), Mod(7, 13), Mod(5, 13), Mod(1, 13),
    Vecsmall([3]), [13, [5, 12, [6, 0, 0, 0]]], [0, 0, 0, 0]]
```

Existe uma função que nos diz se existe uma ordenada para a abscissa que fornecermos. Desta forma podemos testar todos os elementos de $\mathbb{F}_{13}$ *e verificar se existem pontos com essas abscissas:*

```
?    ellordinate(E, Mod(0, 13))
%2   = [Mod(1, 13), Mod(12, 13)]
?    ellordinate(E, Mod(1, 13))
%3   = [Mod(2, 13), Mod(11, 13)]
?    ellordinate(E, Mod(2, 13))
%4   = [Mod(0, 13)]
?    ellordinate(E, Mod(3, 13))
%5   = []
?    ellordinate(E, Mod(4, 13))
%6   = []
?    ellordinate(E, Mod(5, 13))
%7   = []
?    ellordinate(E, Mod(6, 13))
%8   = []
?    ellordinate(E, Mod(7, 13))
%9   = []
?    ellordinate(E, Mod(8, 13))
%10  = [Mod(3, 13), Mod(10, 13)]
?    ellordinate(E, Mod(9, 13))
%11  = []
?    ellordinate(E, Mod(10, 13))
%12  = []
?    ellordinate(E, Mod(11, 13))
%13  = []
?    ellordinate(E, Mod(12, 13))
%14  = []
```

Desta forma obtivemos 7 pontos, mas o ponto no infinito não entrou nesta contagem. Assim obtemos

$$E(\mathbb{F}_{13}) = \{\mathcal{O}, (\overline{0}, \overline{12}), (\overline{0}, \overline{1}), (\overline{1}, \overline{2}), (\overline{1}, \overline{11}), (\overline{2}, \overline{0}), (\overline{8}, \overline{3}), (\overline{8}, \overline{10})\}.$$

Podemos gerar a figura da curva E e marcar os seus pontos com o Geogebra (29).

Poderíamos também usar o Teorema de Hasse 25 para estimar a quantidade de elementos do grupo. Primeiro calculamos o t_{13} e depois utilizamos na equação $13+1-t_{13}$:

```
?    ellap(E)
%15  = 6
?    13 + 1 - 6
%16  = 8
```

Vamos verificar essa ordem, procurando os geradores do grupo

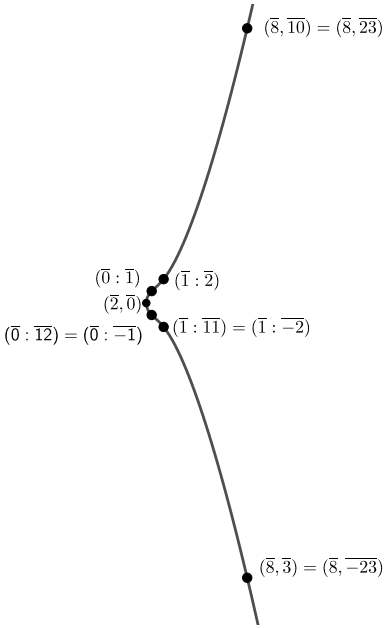


Figura 4.10: $E : \ y^2 = x^3 + 2x + 1$

```
?    ellgenerators(E)
%17  = [[Mod(0, 13), Mod(12, 13)]]
?    P = [Mod(0, 13), Mod(12, 13)]
%18  = [Mod(0, 13), Mod(12, 13)]
```

O programa nos forneceu o elemento $P = (\overline{0}, \overline{12})$, *podemos verificar que ele pertence a curva, encontrar a sua ordem e ainda calcular as suas potências:*

```
?  ellisoncurve(E, P)
%19 = 1
?  ellorder(E, P)
%20 = 8
?  for(k = 1, 8, print(ellpow(E, P, k)))
%21 = [Mod(0, 13), Mod(12, 13)]
     [Mod(1, 13), Mod(2, 13)]
     [Mod(8, 13), Mod(3, 13)]
     [Mod(2, 13), Mod(0, 13)]
     [Mod(8, 13), Mod(10, 13)]
     [Mod(1, 13), Mod(11, 13)]
     [Mod(0, 13), Mod(1, 13)]
     [0]
```

Note que temos os 7 pontos encontrados antes mais o ponto no infinito representado por [0]. *Com a ajuda do Pari ainda podemos fazer operações com estes pontos*

```
?  elladd(E, [Mod(1, 13), Mod(2, 13)], [Mod(8, 13), Mod(10, 13)])
%22 = [Mod(0, 13), Mod(1, 13)]
?  ellsub(E, [Mod(1, 13), Mod(2, 13)], [Mod(1, 13), Mod(11, 13)])
%23 = [Mod(2, 13), Mod(0, 13)]
```

E assim fazer mos uma tabela da soma dos pontos

$\oplus$	$\mathcal{O}$	$(\overline{0},\overline{1})$	$(\overline{0},\overline{12})$	$(\overline{1},\overline{2})$	$(\overline{1},\overline{11})$	$(\overline{2},\overline{0})$	$(\overline{8},\overline{3})$	$(\overline{8},\overline{10})$
$\mathcal{O}$	$\mathcal{O}$	$(\overline{0},\overline{1})$	$(\overline{0},\overline{12})$	$(\overline{1},\overline{2})$	$(\overline{1},\overline{11})$	$(\overline{2},\overline{0})$	$(\overline{8},\overline{3})$	$(\overline{8},\overline{10})$
$(\overline{0},\overline{1})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{11})$	$\mathcal{O}$	$(\overline{0},\overline{12})$	$(\overline{8},\overline{10})$	$(\overline{8},\overline{3})$	$(\overline{1},\overline{2})$	$(\overline{2},\overline{0})$
$(\overline{0},\overline{12})$	$(\overline{0},\overline{12})$	$\mathcal{O}$	$(\overline{1},\overline{2})$	$(\overline{8},\overline{3})$	$(\overline{0},\overline{1})$	$(\overline{8},\overline{10})$	$(\overline{2},\overline{0})$	$(\overline{1},\overline{11})$
$(\overline{1},\overline{2})$	$(\overline{1},\overline{2})$	$(\overline{0},\overline{12})$	$(\overline{8},\overline{3})$	$(\overline{2},\overline{0})$	$\mathcal{O}$	$(\overline{1},\overline{11})$	$(\overline{8},\overline{10})$	$(\overline{0},\overline{1})$
$(\overline{1},\overline{11})$	$(\overline{1},\overline{11})$	$(\overline{8},\overline{10})$	$(\overline{0},\overline{1})$	$\mathcal{O}$	$(\overline{2},\overline{0})$	$(\overline{1},\overline{2})$	$(\overline{0},\overline{12})$	$(\overline{8},\overline{3})$
$(\overline{2},\overline{0})$	$(\overline{2},\overline{0})$	$(\overline{8},\overline{3})$	$(\overline{8},\overline{10})$	$(\overline{1},\overline{11})$	$(\overline{1},\overline{2})$	$\mathcal{O}$	$(\overline{0},\overline{1})$	$(\overline{0},\overline{12})$
$(\overline{8},\overline{3})$	$(\overline{8},\overline{3})$	$(\overline{1},\overline{2})$	$(\overline{2},\overline{0})$	$(\overline{8},\overline{10})$	$(\overline{0},\overline{12})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{11})$	$\mathcal{O}$
$(\overline{8},\overline{10})$	$(\overline{8},\overline{10})$	$(\overline{2},\overline{0})$	$(\overline{1},\overline{11})$	$(\overline{0},\overline{1})$	$(\overline{8},\overline{3})$	$(\overline{0},\overline{12})$	$\mathcal{O}$	$(\overline{1},\overline{2})$

Siga o exemplo e verifique:

a) A curva $Y^2 + Y = X^3 + X + 1$ *sobre o corpo finito* $\mathbb{F}_2$ *só tem o ponto no infinito. (Esse nem precisa do software Pari.)*

b) A curva $Y^2 = X^3 + 3X + 8$ *sobre o corpo* $\mathbb{F}_{13}$ *tem 9 pontos. (É possível fazer esse exemplo sem o Pari).*

⋆*c) A curva* $Y^2 = X^3 + X + 1$ *sobre o corpo* $\mathbb{F}_{61}$ *tem 50 pontos.*

⋆*d) A curva* $Y^2 = X^3 + X + 1$ *sobre o corpo* $\mathbb{F}_{23}$ *tem 24 pontos.*

Questão 30. *Suponha que a cúbica* $X^3 + AX + B$ *se fatore* $X^3 + AX + B = (X - x_1)(X - x_2)(X - x_3)$. *Prove que* $4A^3 + 27B^2 = 0$ *se e somente se duas (ou mais) raizes dentre* x_1, x_2 *e* x_3 *são iguais. Desenvolva a fatoração e compare os coeficientes para relacionas* A *e* B *com* x_1, x_2 *e* x_3.

Questão 31. *Considere* E *uma curva elíptica sobre um corpo algebricamente fechado* K. *O objetivo deste exercício é considerar alguns casos especiais e simples que não foram contemplados na prova da associatividade, já que, para isto utilizamos o Corolário 10 e não demos uma prova da Proposição 22. Para isto, consideremos a seguinte configuração de pontos, retas e cúbicas no plano projetivo.*

$$P, Q, R \in E, S' = \varphi(P, Q), S = \varphi(\mathcal{O}, S'), T' = \varphi(S, R),$$

$$U' = \varphi(Q, R), U = \varphi(\mathcal{O}, U') \text{ e } T'' = \varphi(P, U).$$

e

$$C' = L_{PQ}L_{SR}L_{\mathcal{O}U'} \text{ e } C'' = L_{\mathcal{O}S'}L_{QR}L_{PU}$$

Vimos na prova do Teorema 27 que a prova da associatividade se resumia a provar que $T' = T''$. *Vimos ainda que*

$$E \cap C' = \{P, Q, S', S, R, T', \mathcal{O}, U', U\} \text{ e } E \cap C'' = \{\mathcal{O}, S', S, Q, R, U', P, U, T''\}.$$

1. *Mostre que se* $L_{SR} = L_{UP}$, *então* $T' = T''$;
2. *Mostre que se* $L_{PQ} = L_{SR} = L_{\mathcal{O}U'}$, *então* $T' = T''$;
3. *Mostre que se* $L_{PQ} = L_{SR}$ *ou* $L_{PQ} = L_{\mathcal{O}U'}$, *então* $L_{PQ} = L_{SR} = L_{\mathcal{O}U'}$;

5 Criptografia com curvas elípticas

Neste capítulo juntaremos as ideias desenvolvidas nos Capítulos 2 e 4 para implementar a estrutura de grupo das curvas elípticas no criptossistema ElGamal. Após isto, observaremos a comparação deste criptossistema com o RSA, que é um dos mais utilizados, apesar de não termos definido este sistema neste livro.

É fácil construir um análogo ao criptossistema ElGamal(2.7), que vimos no Capítulo 2, para o grupo das curvas elípticas sobre corpos finitos. Devemos apenas ter o cuidado na pré-codificação da mensagem que Alice deseja enviar, tornando-a um ponto M na curva elíptica $E(\mathbb{F}_p)$. O primo p e a curva elíptica E, sobre $\mathbb{F}_p$, são previamente fixados por um canal não seguro de comunicação. Além disso eles ainda compartilham um ponto $P \in E(\mathbb{F}_p)$.

É claro que a pergunta natural é: como podemos associar mensagens de texto a pontos da curva elíptica?

Uma resposta para isso, pode ser de associar aleatoriamente caracteres a pontos. No entanto, existe outra forma de fazer um análogo do ElGamal sem que surja este problema. Isto foi sugerido por Menezes e Vanstone (1993) em um trabalho que reduzia o problema do logaritmo discreto em curvas elípticas para o problema em corpos finitos.

5.1 Problema do logaritmo discreto para curvas elípticas

Vamos entender o problema do logaritmo discreto para curvas elípticas.

Sejam E a curva elíptica e P um ponto do grupo $E(\mathbb{F}_p)$. Dado um outro ponto $Q \in E$, o inteiro n tal que $nP = Q$, se existir, é dito o logaritmo discreto de Q na base P.

O problema do logaritmo discreto em $\mathbb{F}_p$ (PLD) é complicado de se resolver e o problema para curvas elípticas (PLDCE) é muito mais intratável. Para se ter uma ideia, muitos dos fortes algoritmos utilizados para resolver o PLD não são adaptáveis para PLDCE.

Assim o PLDCE consiste em garantir que seja difícil encontrar um inteiro n tal que $Q = nP$. A forma mais ingênua de resolver o problema $Q = nP$ seria determinar $2P, 3P, 4P, \ldots$, até obter o ponto Q. Mas este caminho é inviável, já que, em média são necessárias $n/2$ somas para obter o ponto Q. Isto é muito difícil quando o n é muito grande, que é o caso da vida real, em que um criptossistema baseado em curvas elípticas utiliza $n \geqslant 2^{160}$.

Exemplo 38. Sejam $E : Y^2 = X^3 - X + 2$ sobre $\mathbb{F}_{47}$. Vamos encontrar o gerador de $E(\mathbb{F}_{47})$

```
?   e = ellinit([0, 0, 0, −1, 2], 47)
%1  [Mod(0, 47), Mod(0, 47), Mod(0, 47), Mod(46, 47), Mod(2, 47), Mod(0, 47),
    Mod(45, 47), Mod(8, 47), Mod(46, 47), Mod(1, 47), Mod(11, 47), Mod(28, 47),
    Mod(42, 47), Vecsmall([3]), [47, [20, 17, [6, 0, 0, 0]]], [0, 0, 0, 0]]

?   ellgenerators(e)
%2  [[Mod(6, 47), Mod(27, 47)]]

?   G = [Mod(6, 47), Mod(27, 47)]
%3  [Mod(6, 47), Mod(27, 47)]
```

Sabendo que o gerador do grupo é $G = (\overline{6}, \overline{27})$, vamos calcular a sua ordem e depois calcular alguma potência

```
?   ellorder(e, G)
%4  53

?   Q = ellpow(e, G, 48)
%5  [Mod(28, 47), Mod(27, 47)]
```

Escolhemos a potência $48G = (\overline{28}, \overline{27})$ que chamamos de Q. Para que o pari faça o logaritmo discreto de Q na base G usamos o comando $elllog()$.

```
?   elllog(e, Q, G, 53)
%6  48

?   ##
%7  *** last result computed in 0 ms.
```

O último comando nos fornece o tempo que foi preciso para resolver o PLDCE. Neste caso foi rápido, pois escolhemos um elemento de ordem pequena. Mas você verá na Ques-

tão 32 um exemplo em que essa conta já não é resolvida tão rapidamente pelo software.

Os algoritmos que são utilizados na resolução destes problemas podem ser divididos em duas classes:

1. Os algoritmos de caráter específico. O tempo de execução e a própria aplicação dependem de certos tipos de parâmetros da curva elíptica. Exemplo: o ataque de MOV.

2. Os algoritmos de caráter geral. O tempo de execução depende apenas do tamanho de cada parâmetro da curva elíptica. Exemplos: métodos de Pollard, a simplificação de Pohlig e Hellman.

5.2 O sistema criptográfico CCE

Para a realização de uma comunicação através de um criptossistema baseado em uma curva elíptica (CCE), serão necessários alguns parâmetros.
É preciso escolher:

1. Um corpo finito $\mathbb{F}_p$ no qual se irá trabalhar;

2. Uma curva elíptica E e seu grupo aditivo $E(\mathbb{F}_p)$ sobre $\mathbb{F}_p$;

3. Um ponto P de ordem suficientemente grande no grupo $E(\mathbb{F}_p)$;

4. As chaves secretas a, b escolhidas pelo emissor e o receptor;

5. A mensagem devidamente pré-codificada, $m \in \mathcal{M}$, no qual $\mathcal{M}$ é o conjunto das mensagens.

Já sabemos que é preciso transformar o texto plano em números. Num criptossistema baseado em curvas elípticas existe mais um passo na pré-codificação para que possamos aplicar o sistema criptográfico. Podemos encaixar/mergulhar o texto simples em uma curva elíptica E (*texto plano mergulhado*, do inglês "embedding plaintext"), ou utilizar uma curva elíptica para "mascarar" o texto original (*texto plano mascarado*, do inglês "masking plaintext"). Em detalhes, isto que dizer:

- **Mergulho:** Os utilizadores desejam um sistema simples para fazer isto, de maneira que a relação entre o texto simples e o seu ponto correspondente na curva seja clara. Deve ser fácil para qualquer usuário autorizado converter o texto simples (inteiros) em coordenadas dos pontos em E, e vice-versa. Segundo, esta conversão deve ser rápida e sistemática. Por último, não existe um algoritmo de tempo polinomial que "mergulhe" um grande número de pontos em uma curva elíptica arbitrária $E(\mathbb{F}_q)$.

- **Mascaramento:** Mascarar um par ordenado (m_1, m_2) com uma curva elíptica significa alterar o par multiplicando as coordenadas deste ponto com as respectivas coordenadas de algum ponto na curva. Consequentemente, textos planos e mensagens cifradas não são obrigados a ser embutidos como pontos em uma curva elíptica. Eles podem ser qualquer par ordenado de elementos não nulos do corpo. Mascarar em vez de mergulhar mantém o sistema de criptografia simples e também reduz o tempo computacional.

Observe que o mascaramento não parece mais ou menos seguro que o mergulho, uma vez que ambos dependem da segurança do problema do logaritmo discreto.

Existem muitas maneiras de realizar esta pré-codificação, no sentido de que a mesma dependa do conjunto de mensagens $\mathcal{M}$ que iremos utilizar. O conjunto $\mathcal{M}$ pode ser de palavras, símbolos ou apenas letras.

Agora vamos implementar as curvas elípticas no Criptossistema ElGamal, nos dois casos de pré-codificação da mensagem.

5.2.1 ElGamal com curvas elípticas

Começaremos com o análogo do ELGamal com curvas elípticas no caso em que a mensagem será um ponto mergulhado na curva.

Seja $E(\mathbb{F}_p)$ uma curva elíptica sobre o corpo finito de característica maior que 3. Escolhe-se um ponto base $P \in E(\mathbb{F}_p)$ que seja preferencialmente um gerador de $E(\mathbb{F}_p)$. Os dados que irão compor a chave pública de cada usuário do sistema serão o par $\big(E(\mathbb{F}_p), P\big)$ juntamente com $C = cP$, onde c é a chave privada de cada usuário. Vejamos como funciona o sistema:

- Alice e Bob escolhem um inteiro a e b, respectivamente, como chave secreta, calculam o ponto $A = aP$ e $B = bP$, respectivamente, e publicam como chave pública;
- Digamos que Alice deseja enviar a mensagem $M \in E(\mathbb{F}_p)$ para Bob;
- Alice calcula $S = M + aB$ e, em seguida, envia para Bob;
- Para decifrar a mensagem, Bob calcula

$$S - bA = M + aB - bA = M + (abP) - (baP) = M.$$

No exemplo a seguir, vamos detalhar a pré-codificação realizando o mergulho da mensagem na curva. E isto é feito encontrando um ponto na curva cuja abcissa seja a mensagem.

Exemplo

Vamos cifrar a palavra "Criptografia". Pela tabela ASCII decimal (lembrando de colocar um zero à esquerda de todo número de 2 dígitos) temos o seguinte

C	r	i	p	t	o	g	r	a	f	i	a
067	114	105	112	116	111	103	114	097	102	105	097

Assim a mensagem pré-codificada é $m = 067114105112116111103114097102105097$ e tem 36 dígitos. Precisamos de uma curva elíptica e um número primo com mais de 36 dígitos. Vamos escolher um primo com 40 dígitos:

```
?   p = nextprime(6846869858332693264879382366866797734568)
%1  684686985833269326487938236686679773456 9
?   isprime(p)
%2  1
```

Vamos escolher inicialmente a curva $E : y^2 - (x^3 + x + 1) = 0$.

```
?   e = ellinit([0, 0, 0, 1, 1], p)
%3  = [Mod(0, 6846869858332693264879382366866797734569),
    Mod(0, 6846869858332693264879382366866797734569),
    Mod(0, 6846869858332693264879382366866797734569),
    Mod(1, 6846869858332693264879382366866797734569),
    Mod(1, 6846869858332693264879382366866797734569),
    Mod(0, 6846869858332693264879382366866797734569),
    Mod(2, 6846869858332693264879382366866797734569),
    Mod(4, 6846869858332693264879382366866797734569),
    Mod(6846869858332693264879382366866797734568,
    6846869858332693264879382366866797734569),
    Mod(6846869858332693264879382366866797734521,
    6846869858332693264879382366866797734569),
    Mod(6846869858332693264879382366866797733705,
    6846869858332693264879382366866797734569),
    Mod(6846869858332693264879382366866797734073,
    6846869858332693264879382366866797734569),
    Mod(5300802470967246398616296025961391794728,
    6846869858332693264879382366866797734569),
    Vecsmall([3]), [6846869858332693264879382366866797734569,
    [1296, 46656, [6, 0, 0, 0]]], [0, 0, 0, 0]]
```

Vamos verificar que E é realmente uma curva elíptica sobre $\mathbb{Z}_p$. Para isto, calculemos o seu discriminante

```
?   e.disc
%4  = Mod(6846869858332693264879382366866797734073,
    6846869858332693264879382366866797734569)
```

De fato, $\Delta = 6846869858332693264879382366866797734073$ é não nulo e E é curva elíptica. Agora calculemos a ordem de $E(\mathbb{Z}_p)$ usando o Teorema de Hasse. O traço t da curva e a ordem

```
?    t = ellap(e, p)
%5   = -19129140746580970256
?    p + 1 - t
%6   = 6846869858332693264898511507613378704826
```

Assim calculamos $\sharp E(\mathbb{Z}_p) = 6846869858332693264898511507613378704826$. Agora escolhemos um ponto. É fácil ver que $P = (0, 1)$ pertence a curva, mas vamos confirmar que realmente $P \in E(\mathbb{F}_0)$ e verificar que sua ordem é ao menos a metade da ordem do grupo

```
?    P = [Mod(0, p), Mod(1, p)]
%7   = [Mod(0, 6846869858332693264879382366866797734569),
     Mod(1, 6846869858332693264879382366866797734569)]
?    ellisoncurve(e, P)
%8   1
?    ellorder(e, P)
%9   = 342343492916634663244925575380668935241 3
?    (%6/2) - %9
%10  = 0
```

A ordem é grande o suficiente $\left(\dfrac{\sharp E(\mathbb{Z}_p)}{2}\right)$. Assim definimos $(E(\mathbb{Z}_p), (0, 1))$.

Vamos mergulhar a mensagem $m = 0671141051121161111031140971021050 97$ na curva pelo **método de Koblitz**:

Temos que escolher um inteiro λ e encontrar um ponto $(x, y) \in E(\mathbb{Z}_p)$ no qual para $x = m\lambda + n$, com $1 \leqslant n < \lambda - 1$, e exista y.

Vamos escolher $\lambda = 28734$ testar $x = m\lambda + 1$. Usaremos o comando que calcula a ordenada dado o x:

```
?    m = 0671141051121161111031140971021050 97
%11  = 671141051121161111031140971021050 97
?    l = 28734
%12  = 28734
?    ellordinate(e, Mod(m * l + 1, p))
%13  = [Mod(8844965836556166826638394798975511 01806,
     6846869858332693264879382366866797734569),
     Mod(5962373274677076582215542886969246632763,
     6846869858332693264879382366866797734569)]
```

Conseguimos duas ordenadas para o ponto na primeira tentativa. Vamos escolher a primeira. Caso não desse certo para $x = m\lambda + 1$, tentaríamos $x = m\lambda + 2$ e assim por

diante até encontrar. Vamos denotar a mensagem cifrada de $M \in E(\mathbb{F}_p)$

```
?    M = [Mod(m * l + 1, p), Mod(8844965836556166826638
     39479897551101806, p)]
%14  = [Mod(192845669629154433643688046613188785719 9,
     68468698583326932648793823668667977345 69),
     Mod(8844965836556166826638394798975511018 06,
     68468698583326932648793823668667977345 69)]
```

Podemos verificar que M está na curva:

```
?    ellisoncurve(e, M)
%15  = 1
```

Assim a menssagem mergulhada na curva é o ponto

$$M = (19284566962915443364368804661318878571 99, 88449658365561668266383947989755110180 6)$$

- Suponha que Alice deseja mandar a mensagem M para Bob. Alice tem a chave privada $a = 98675436754635$ e chave pública $A = aP$. Calculemos:

```
?    a = 98675436754635
%16  98675436754635
?    A = ellpow(e, P, a)
%17  = [Mod(48876656292631249763001461537302827880 86,
     68468698583326932648793823668667977345 69),
     Mod(58323550624841079484377831513648317068,
     68468698583326932648793823668667977345 69)]
?    ellisoncurve(e, A)
%18  1
```

Deste modo,

$$A = aP = (48876656292631249763001461537302827880 86, 58323550624841079484377831513648317068).$$

- Bob tem como chave privada $b = 7684945762574$, e sua chave pública é $B = bP$. Calculemos:

```
?    b = 7684945762574
%19  7684945762574
?    B = ellpow(e, P, b)
%20  = [Mod(67724043750481033508818200399523785914 23,
     68468698583326932648793823668667977345 69),
     Mod(15538270837022498720443619019482512738 41,
     68468698583326932648793823668667977345 69)]
?    ellisoncurve(e, B)
%21  1
```

Deste modo,
$$B = bP = (677240437504810335088182003995237859142 3,\ 155382708370224987204436190194825127384 1).$$

- Para codificar a mensagem, Alice calcula $S = M + aB$

```
?    S = elladd(e, M, ellpow(e, B, a))
%22  = [Mod(484758068570898041095004946263622835406 5,
     684686985833269326487938236686679773456 9),
     Mod(509565822332816578366860753718921231450,
     684686985833269326487938236686679773456 9)]
```

Assim a mensagem codificada e enviada é
$$S = (484758068570898041095004946263622835406 5,\ 509565822332816578366860753718921231450).$$

- Para decifrar a mensagem, Bob calcula calcula $S - bA$, que é

```
?    ellsub(e, S, ellpow(e, A, b))
%23  = [Mod(192845669629154433643688046613188785719 9,
     684686985833269326487938236686679773456 9),
     Mod(884496583655616682663839479897551101806,
     684686985833269326487938236686679773456 9)]
```

Assim Bob obtém $M = (192845669629154433643688046613188785719 9, 884496583655$
Para decriptar ele apenas precisa da primeira entrada do ponto e calcular $\frac{x-1}{\lambda}$:

```
?    (19284566962915443364368804661318878571 99 - 1)/1
%24  = 671141051121161111031140971021050 97
```

Usando a tabela ASCII temos a mensagem m:

C	r	i	p	t	o	g	r	a	f	i	a
67	114	105	112	116	111	103	114	097	102	105	097

5.2.2 Variante de Menezes e Vanstone para o ElGamal

Menezes e Vanstone (1993) elaboraram uma variação em que não é preciso pré-codificar a mensagem como um ponto da curva. A mensagem é mascarada para a aplicação do criptossistema.

Seja $E(\mathbb{F}_p)$ uma curva elíptica sobre o corpo finito de característica prima bem grande. Escolhe-se um ponto base $P \in E(\mathbb{F}_p)$, de ordem grande em $E(\mathbb{F}_p)$. Agora os dados que irão compor a chave pública de cada usuário do sistema são a tripla $\left(\mathbb{F}_p, E(\mathbb{F}_p), P\right)$ junto com $C = cP$, onde c é a chave privada. Vejamos como funciona o sistema:

- Alice e Bob escolhem um inteiro a e b, respectivamente, como chave secreta, calculam o ponto $A = aP$ e $B = bP$, respectivamente, e publicam.

- Digamos que Alice deseja enviar a mensagem m para Bob. A mensagem é pré-codificada usando a tabela ASCII e é quebrada em dois blocos de números, m_1 e m_2 módulo p, e assim a mensagem pronta para criptografar é o par $m = (m_1, m_2)$ (note que m_1 e m_2 não podem iniciar com zeros e devem ser menores que p).

- Alice opera $C = aB = (x_c, y_c)$ e depois calcula

$$s_1 \equiv x_c m_1 \bmod p \quad s_2 \equiv y_c m_2 \bmod p$$

  e envia para Bob o par $[s_1, s_2]$.

- Para decifrar a mensagem, Bob calcula $D = bA = (x_d, y_d)$. Como $C = D$ a mensagem é decifrada calculando-se

$$x_d^{-1} s_1 \equiv x_d^{-1} x_c m_1 \bmod p = m_1 \bmod p$$

$$y_d^{-1} s_2 \equiv y_d^{-1} y_c m_2 \bmod p = m_2 \bmod p$$

Este método de pré-codificação não altera a segurança do criptossistema. Vejamos um exemplo concreto:

Exemplo

Vamos utilizar a mesma curva $E : y^2 - (x^3 + x + 1) = 0$, o mesmo primo $p = 68468698583326932648793823668667977345 69$ o mesmo ponto $P = (0, 1)$. Mas trocaremos a mensagem e as chaves secretas.

Assim a chave pública é $\big(\mathbb{Z}_p, E(\mathbb{Z}_p), P = (0, 1)\big)$

Antes vamos pré-codificar a mensagem mascarando-a: Assim como antes, reescrevemos a mensagem usando o código decimal ASCII. Depois quebramos a mensagem em dois blocos m_1 e m_2, de modo que $m_1, m_2 \in \mathbb{Z}_p^*$ (não podem iniciar com zeros e devem ser menores que p). Vamos utilizar a seguinte mensagem:

M	a	t	e	m	á	t
077	097	116	101	109	225	116
				m	a	s
			032	109	097	115

E assim temos a mensagem com 57 caracteres

$$m = 077097116101109225116105099097032233032109097115115097033$$

que quebraremos da seguinte forma
$(m_1, m_2) = (77097116101109225116105099 09, 703223303210909711511509 7033)$

Uma vez nomeados o primo p, a curva e e o ponto P no Pari, vamos criptografar.

- Suponha que Alice deseja enviar a mensagem m para Bob. Alice escolhe sua chave secreta $a = 394756376$ e tem como a chave pública $A = aP$. Calculemos:

```
?   a = 394756376
%4  394756376
?   A = ellpow(e, P, a)
%5  = [Mod(132155833514596227411159749086721101325 5,
    6846869858332693264879382366866797734569),
    Mod(465112924000968106419957886949913791803 3,
    6846869858332693264879382366866797734569)]
?   ellisoncurve(e, A)
%6  1
```

Deste modo,

$$A = aP = \begin{array}{l}(132155833514596227411159749086721101325 5,\\ 465112924000968106419957886949913791803 3).\end{array}$$

- Bob escolhe $b = 4857628576$ como chave secreta e publica $B = bP$ como chave pública. Calculemos:

```
?   b = 4857628576
%7  4857628576
?   B = ellpow(e, P, b)
%8  = [Mod(422061900257492441516394928629041653952 3,
    6846869858332693264879382366866797734569),
    Mod(179076010304836427257727577914388125458 0,
    6846869858332693264879382366866797734569)]
?   ellisoncurve(e, B)
%9  1
```

$$B = bP = \begin{array}{l}(422061900257492441516394928629041653952 3,\\ 179076010304836427257727577914388125458 0).\end{array}$$

- Alice calcula $C = aB$

```
?   C = ellpow(e, B, a)
%6  = [Mod(338859256239172459526871882992404398021 3,
    6846869858332693264879382366866797734569),
    Mod(129495594649667554566611905764330384728,
    6846869858332693264879382366866797734569)]
```

obtendo $x_c = 338859256239172459526871882992404398021 3$ e $y_c = 129495594649667$
Depois calcula

$$s_1 \equiv x_c m_1 \bmod p \quad s_2 \equiv y_c m_2 \bmod p$$

sabendo que $m_1 = 770971161011092251161050990 9$ e $m_2 = 703223303210909711511 50$
Ou seja

```
?    m1 = 770971161011092251161050990 9
%10  = 770971161011092251161050990 9
?    m2 = 703223303210907 9711511509703 3
%10  = 703223303210907 9711511509703 3
?    xc = 338859256239172459526871882992404398021 3
%10  = 338859256239172459526871882992404398021 3
?    yc = 129495594649667554566611905764330384728
%10  = 129495594649667554566611905764330384728
?    s1 = Mod(xc * m1, p)
%10  = Mod(713927432385338296065501390169095967192,
     68468698583326932648793823668667977345 69)
?    s2 = Mod(yc * m2, p)
%10  = Mod(302271421864626112142805122850983071870 4,
     68468698583326932648793823668667977345 69)
```

e envia para Bob o par $[s_1, s_2]$:
[713927432385338296065501390169095967192 mod p,
302271421864626112142805122850983071870 4 mod p]

- Para decifrar a mensagem, Bob primeiro calcula $D = bA = (x_d, y_d)$:

```
?    ellpow(e, A, b)
%14  = [Mod(175566359845744532348191584517287070238 3,
     68468698583326932648793823668667977345 69),
     Mod(244841254775284957623968677978060786656 5,
     68468698583326932648793823668667977345 69)]
```

Como $C = D$, agora basta efetuar $x_d^{-1}s_1$ e $y_d^{-1}s_2$:

```
?    xd = 175566359845744532348191584517287070238 3
%10  = 175566359845744532348191584517287070238 3
?    yd = 244841254775284957623968677978060786656 5
%10  = 244841254775284957623968677978060786656 5
?    Mod(xc ∧ (−1) * s1, p)
%15  = Mod(770971161011092251161050990 9,
     68468698583326932648793823668667977345 69)
?    Mod(yc ∧ (−1) * s2, p)
%15  = Mod(703223303210907 9711511509703 3,
     68468698583326932648793823668667977345 69)
```

E assim recupera a mensagem $m = (770971161011092251161050990 9, 703223303210909$

M	a	t	e	m	á	t	i	c	a		é
077	097	116	101	109	225	116	105	099	097	032	233
				m	a	s	s	a	!		
			032	109	097	115	115	097	033		

5.3 Comparações

Quando se avalia um sistema criptográfico para decidir o quão eficiente este é, dentre muitas considerações, as mais importantes são o nível de segurança (que consiste na dificuldade de resolução do problema matemático em que se baseia o sistema) e o tempo de resposta (o quão rápido é o algoritmo).

Quanto à segurança de um criptossistema de chave pública, poderíamos nos perguntar se para quebrar um sistema é necessário resolver o problema matemático em que ele se baseia. A questão é que matematicamente já foi provado a impossibilidade da violação de criptossistemas que seguem certos pré-requisitos. A única maneira de quebrar esses criptossistemas seria através de algoritmos que tentam resolver o problema com a maior eficiência possível. Na criptoanálise de sistemas de chave pública, esses algoritmos podem ser divididos em duas classes:

1. Algoritmos específicos, que são criados para casos isolados, *i.e.*, para criptossistemas com determinados parâmetros. Estes são baseados em alguns aspectos de "fraqueza" apresentados pelo parâmetro.

2. Algoritmos genéricos, que não restringem qualquer parametrização.

Os parâmetros que oferecem fraqueza devem ser evitados ao se projetar um criptossistema. Por isso existem "regras" para estes sistemas, como tamanho da chave, por exemplo.

Os algoritmos específicos atingem as fraquezas, sabemos que no caso da fatoração de inteiros e no problema do logaritmo discreto de números, quando se utilizam fatores primos pequenos o problema se torna mais fácil. No caso de curvas elípticas, foram descobertas duas pequenas classes de curvas que apresentam aspectos de vulnerabilidade e, consequentemente, também podem ser tratadas por algoritmos específicos, são elas: *supersingulares* ("supersingular elliptic curves"- quando a característica de $\mathbb{F}q$ divide o traço de Frobenius) e *anômalas* ("anomalous elliptic curves").

Já os algoritmos genéricos são elaborados para resolver qualquer configuração encontrada em cada um desses problemas, e independente dos parâmetros utilizados devem chegar à resposta. Mas, mesmo sendo capazes de quebrar o criptossistema, há uma questão importante: o tempo que estes algoritmos levam para encontrar a solução. É justamente a ordem de grandeza deste tempo que permite avaliar a segurança do sistema criptográfico, baseados nos parâmetros de entrada.

O RSA é considerado como o primeiro sistema criptográfico de chave assimétrica realizável e utilizado na vida real. Portanto, torna-se um padrão para a comparação de criptossistemas de chave pública. Sua segurança está no problema da fatoração de números inteiros, e o processo para decriptar mensagens no RSA não é tão eficiente como seu processo de criptografia. Já se sabe que para aumentar a segurança na troca de dados, são necessários tamanhos de chave maiores. Entretanto, aumentar o tamanho da chave significa mais sobrecarga nos sistemas de computação.

Atualmente, na era dos smartphones, tablets, smartcards e smartwatches, pequenos dispositivos estão desempenhando um papel importante no mundo digital. Mesmo tendo

uma capacidade de memória reduzida, necessitam de segurança para lidar com a demanda do mercado. É neste cenário que o RSA vem se tornando uma segunda opção.

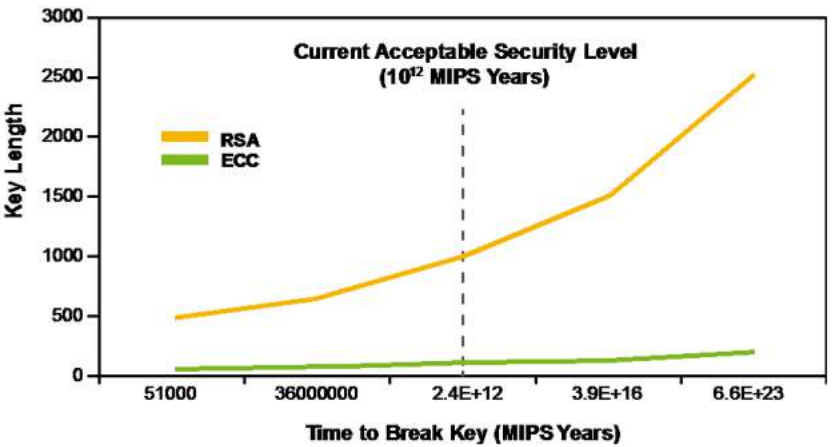


Figura 5.1: Performances do RSA e CCE na tentativa de quebra do sistema, Certicom (1997).

Para se ter uma idéia, um tempo razoável de segurança é 10^{12} MIPS (no qual MIPS é o número de anos que uma máquina, capaz de executar um milhão de instruções por segundo, leva para resolver o problema). Neste nível de segurança, enquanto o RSA e o DSA necessitam de chaves com 1024 bits, o CCE precisa de somente 160 bits. Além disto, quando exigimos o aumento do nível de segurança (aumentamos o MIPS), o tamanho das chaves do RSA e DSA necessitam de um aumento bem mais expressivo do que as da CCE.

Security Bit Level	RSA	ECC
80	1024	160
112	2048	224
128	3072	256
192	7680	384
256	15360	512

Figura 5.2: Níveis de segurança do RSA e CCE de acordo com a recomendação do National Institute of Standards and Technology(NIST) Mahto e Yadav (2017)

Pela tabela acima é fácil perceber a vantagem do uso de curvas elípticas quanto ao tamanho da chave utilizada. A diminuição no 'tamanho' da chave para as curvas elípticas

deve-se ao fato da estrutura diferenciada do grupo $E(\mathbb{F}_p)$. A maioria dos algoritmos criados para atacar os sistemas criptográficos são para grupos mais comuns. O melhor que se conseguiu até os dias de hoje é um algoritmo com tempo de execução exponencial para resolver o problema do logaritmo discreto para curvas elípticas.

Ainda analisando a eficiência de uma criptossistema, também precisamos considerar o tempo gasto para encriptar e o tempo gasto para decriptar uma mensagem. Os gráficos abaixo confirmam que a implementação com curvas elípticas ainda vence o RSA neste sentido.

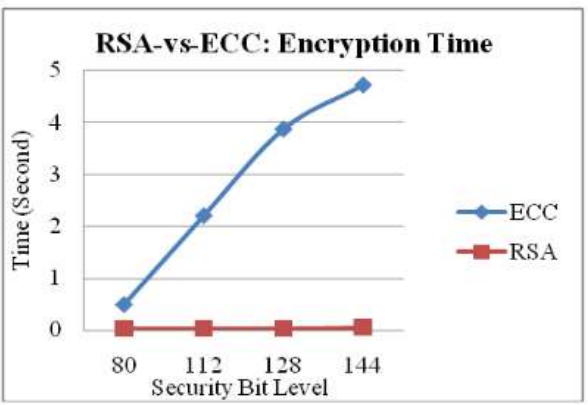


Figura 5.3: Tempo de encriptação (em segundos)Mahto e Yadav (2017)

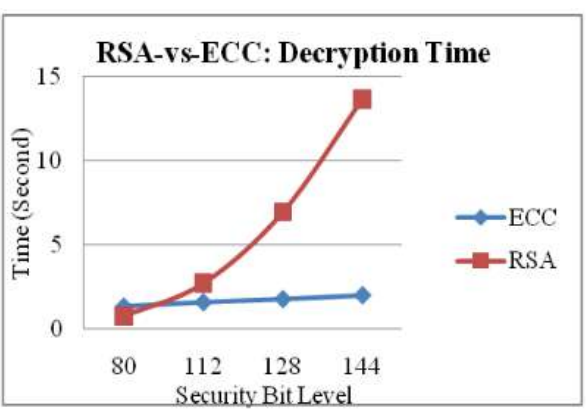


Figura 5.4: Tempo de decriptação (em segundos)Mahto e Yadav (2017)

No geral, se formos resumir as vantagens e desvantagens da CCE, temos:

- Vantagens

  A chave pública tem um tamanho menor, e segurança;

  Requer menores requisitos de velocidade de processador, de memória e de largura de banda;

  O processo de implementação, do ponto de vista de software e hardware, é facilitado.

- Desvantagens

  Os algoritmos, pela sua natureza matemática, são computacionalmente intensos;

  Requer uma autoridade certificadora para garantir a identidade e confiabilidade das chaves públicas.

De fato a implementação das curvas elípticas nos criptossistemas é uma ótima opção para aumentar a segurança do sistema. E a matemática envolvida é muito bonita e complexa o suficiente para dificultar um pouco mais a criação de algoritmos que tentem quebrar o sistema.

5.4 Exercícios

Questão 32. *Nos exemplos a seguir resolva o problema do logaritmo discreto para o grupo das curvas elípticas.*

a) Considere a curva $E : Y^2 = X^3 + X + 1$ *sobre o corpo* $\mathbb{F}_5$. *Sejam* $P = (\overline{4}, \overline{2})$ *e* $Q = (\overline{0}, \overline{1})$, *encontre o inteiro positivo* n *tal que* $Q = nP$.

⋆*b) Seja* E *a mesma curva elíptica, sendo agora* $\mathbb{F}_p$ *com o primo*

$$p = 65473634859875634563728473 4703.$$

Para os pontos $P = (x_1, y_1)$ *e* $Q = (x_2, y_2)$ *com*

$$x_1 = \overline{140875671916033324085065454905}$$

$$y_1 = \overline{568839630795772774828669387959}$$

$$x_2 = \overline{332790426556264071423402397765}$$

e

$$y_2 = \overline{537775934447539930519735581041}$$

use a função $elllog(E, Q, P, ord(P))$ *para resolver* $Q = nP$, *com* n *inteiro positivo. Use o comando* ♯♯ *para ver o tempo exato que o programa levou para resolver o problema do logaritmo discreto.*

Questão 33. ⋆ *Use os dois modelos de criptografia com curva elíptica para encriptar a mensagem com 22 caracteres(contando letras, espaços e acentos) a seguir*

"A matemática não mente".

Use a curva $Y^2 = X^3 + X + 2$ *que não seja singular, e primo adequado. Lembre-se de considerar o comprimento da mensagem após prepará-la usando a tabela ASCII.*

A

Para por em prática tudo que aprendemos durante o curso, vamos conhecer e entender como utilizar a calculadora algébrica Pari/GP. Ela pode ser obtida através do site http://pari.math.u-bordeaux.fr na versão para computadores ou para Android, chamada PariDroid.

A.1 Apresentando a Pari/GP

PARI/GP é um sistema de álgebra computacional que visa facilitar alguns cálculos de teoria dos números, como fatorações, problemas de teoria dos números algébricos, curvas elípticas, formas modulares, contendo também outras funções úteis para calcular matrizes, polinômios, séries de potência, números algébricos, muitas funções transcendentais e outras coisas mais.

Originalmente desenvolvido por Henri Cohen e seus colegas de trabalho (Université Bordeaux I, França), o PARI está agora sob a GPL e mantido por Karim Belabas com a ajuda de muitos colaboradores voluntários.

O sistema PARI/GP consiste nos seguintes componentes padrão:

- O PARI é uma biblioteca C, que permite cálculos rápidos e pode ser chamada a partir de um aplicativo de linguagem de alto nível (por exemplo, escrito em C, C ++, Pascal, Fortran, Perl ou Python).

- O gp é uma interface de linha de comando interativa fácil de usar que dá acesso às

funções PARI. Funciona como uma sofisticada calculadora programável que contém a maioria das instruções de controle de uma linguagem padrão como C.

- GP é o nome da linguagem de script do gp que pode ser usada para programar gp.
- O gp2c, o compilador GP-para-C, combina o melhor dos dois mundos, compilando scripts GP para a linguagem C e carregando transparentemente as funções resultantes no gp. (Normalmente, os scripts compilados por gp2c são executados 3 ou 4 vezes mais rápido). Atualmente, o gp2c só entende um subconjunto do idioma do GP.

As versões 2.1.0 e superiores são distribuídas sob a Licença Pública Geral GNU. Ele é executado nos sistemas operacionais mais comuns.

Já o PariDroid é um app para Android que permite o uso do Pari/GP, iniciado por Charles Boyd e atualmente mantido por Andreas Enge. Ele é distribuído sob a Licença Pública Geral GNU, seja versão 3 da licença, ou (a seu critério) qualquer versão posterior (GPLv3 +).

A.2 Aprendendo a usar a Pari/GP

Quando abrimos o sistema encontramos o seguinte:

```
GP/PARI CALCULATOR Version 2.9.4 (released)
amd64 running linux (x86-64/GMP-6.1.2 kernel) 64-bit version
compiled: Dec 19 2017, gcc version 7.3.0 (Ubuntu 7.3.0-1ubuntu1)
threading engine: pthread
(readline v7.0 enabled, extended help enabled)

Copyright (C) 2000-2017 The PARI Group

PARI/GP is free software, covered by the GNU General Public License, and
comes WITHOUT ANY WARRANTY WHATSOEVER.

Type ? for help,  q to quit.
Type ?15 for how to get moral (and possibly technical) support.

parisize = 8000000, primelimit = 500000, nbthreads = 4
?
```

As linhas de entrada são iniciadas por ?. Se você observou a instrução na antepenúltima linha, viu que a função ? sozinha nos fornece ajuda do programa. Quando utilizamos ? antes de outra função, o resultado é a descrição da mesma.

Ao colocar apenas ? o programa fornece o seguinte guia de funções por tópico:

```
? ?
Help topics: for a list of relevant subtopics, type ?n for n in
   0: user-defined functions (aliases, installed and user functions)
   1: Standard monadic or dyadic OPERATORS
   2: CONVERSIONS and similar elementary functions
   3: TRANSCENDENTAL functions
   4: NUMBER THEORETICAL functions
   5: ELLIPTIC CURVES
   6: L-FUNCTIONS
   7: MODULAR SYMBOLS
   8: General NUMBER FIELDS
   9: Associative and central simple ALGEBRAS
  10: POLYNOMIALS and power series
  11: Vectors, matrices, LINEAR ALGEBRA and sets
  12: SUMS, products, integrals and similar functions
  13: GRAPHIC functions
  14: PROGRAMMING under GP
  15: The PARI community
Also:
  ? functionname (short on-line help)
  ?              (keyboard shortcuts)
  ?.             (member functions)
Extended help (if available):
  ??             (opens the full user's manual in a dvi previewer)
  ?? tutorial / refcard / libpari (tutorial/reference card/libpari manual)
  ?? Keyword     (long help text about "keyword" from the user's manual)
  ??? Keyword    (a propos: list of related functions).
```

É possível saber todas as funções referentes a cada tópico, por exemplo, nós utilizaremos a calculadora para fazer contas com curvas elípticas, então basta colocar ?5 para obter a lista de funções relacionadas às ELLIPTIC CURVES:

```
ellL1           elladd          ellak           ellan
ellanalyticrank ellap           ellbil          ellchangecurve
ellchangepoint  ellconvertname  elldivpol       elleisnum
elleta          ellgenerators   ellglobalred    ellgroup
ellheight       ellheightmatrix ellidentify     ellinit
ellisoncurve    ellj            elllocalred     elllog
elllseries      ellminimalmodel ellmodulareqn   ellorder
ellordinate     ellpointtoz     ellpow          ellrootno
ellsearch       ellsigma        ellsub          elltaniyama
elltatepairing  elltors         ellweilpairing  ellwp
ellzeta         ellztopoint
```

Para saber o que cada função fornece, por exemplo, ellinit, basta utilizar o comando

$$?ellinit$$

que o PARI apresenta a descrição completa da função:

```
?  ?ellinit
ellinit(x,{D = 1}) : let x be a vector [a1,a2,a3,a4,a6], or [a4,a6]
ifa1 = a2 = a3 = 0, defining the curve Y^2 + a1.XY + a3.Y = X^3 + a2.X^2
+a4.X + a6; x can also be a string, in which case the curve with
matching name is retrieved from the elldata database, if
available. This function initializes an elliptic curve over the
domain D (inferred from coefficients if omitted).
```

Os resultados que não são a descrição de funções são antecedidos pelo simbolo %n, onde n é um número natural. Nós podemos utilizá-los para chamar o resultado quanto estivermos mais adiante nas contas.

Nós utilizamos o programa no Linux, nele a calculadora é acessada pelo terminal: basta abrir o terminal e colocar gp.

Antes que comecemos a criptografar, vejamos exemplos com as funções que vamos utilizar nas nossas contas.

Como estamos trabalhando com primos, a primeira coisa que precisamos saber é como escolher primos grandes com a ajuda da calculadora. Nós podemos chutar um número com a quantidade de dígitos que desejamos e verificar se este número é primo através da função isprime(). Começaremos escolhendo um primo de 35 dígitos, o programa responde a esta função com 1, se o número for primo, e 0 no caso contrário:

```
?    isprime(37462350986754834657874659823456747);
%1   = 0
```

Se não acertarmos de primeira, podemos pedir que o PARI diga qual o próximo número primo após o número que chutamos, basta utilizar a função nextprime() e até verificarmos a sua resposta. Depois podemos nomear de p o primo escolhido:

```
?    nextprime(37462350986754834657874659823456747);
%2   37462350986754834657874659823456849

?    isprime(%2);
%3   = 1

?    p = %2;
%4   = 37462350986754834657874659823456849
```

Outra dado que utilizamos no criptossistema ElGamal é a raiz primitiva de $\mathbb{Z}_p$. Com a função znprimroot() obtemos esta raiz. Observamos que a notação do PARI para um número ‘‘x módulo n” é `Mod(x,n)`. Vejamos a raiz primitiva para o primo escolhido anteriormente:

```
?  znprimroot(p);
%5 Mod(17, 37462350986754834657874659823456849)
```

Agora vamos tratar das curvas elípticas. A primeira função, para definir a curva que escolhemos para o programa, é:

$$ellinit([a_1, a_2, a_3, a_4, a_6]),$$

onde a_1, a_2, a_3, a_4, a_6 são coeficientes dos termos que compõem a equação

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

duma curva elíptica $E(K)$. A resposta é o vetor:

$$[a_1; a_2; a_3; a_4; a_6; b_2; b_4; b_6; b_8; c_4; c_6; \Delta; j(E); (...)]$$

onde b_2, b_4, b_6, b_8, c_4 e c_6 são as constantes apresentadas na definição no início do parágrafo III.1 de Silverman 2009, Δ é o discriminante da curva elíptica, $j(E)$ é o $j-invariante$ e o símbolo (...) representa os outros elementos apresentados no vetor que não tem relevância para nosso trabalho.

Exemplo 39. *A partir de agora vamos estudar a curva elíptica $E(\mathbb{F}_{11})$ definida pela equação $y^2 = x^3 + x + 3$, como também o grupo $E(0;1;3)_{\mathbb{F}_{11}}$ sobre curva $E(F_{11})$. Vamos definir as funções e exemplificar com esta curva.*

Se queremos as informações sobre a curva elíptica $E(\mathbb{F}_{11})$ definida pela equação $y^2 = x^3 + x + 3$, devemos colocar a seguinte expressão:

$$e = ellinit([Mod(0,11), Mod(0,11), Mod(0,11), Mod(1,11), Mod(3,11)]) \quad \text{(A.1)}$$

onde a função Mod(x, y) é o inteiro x módulo y, e nomeamos esta curva de e. O programa responde:

```
?  e = ellinit([Mod(0, 11), Mod(0, 11), Mod(0, 11), Mod(1, 11), Mod(3, 11)])
%6 = [Mod(0, 11), Mod(0, 11), Mod(0, 11), Mod(1, 11), Mod(3, 11), Mod(0, 11),
   Mod(2, 11), Mod(1, 11), Mod(10, 11), Mod(7, 11), Mod(4, 11), Mod(8, 11),
   Mod(3, 11), 0, 0, 0, 0, 0, 0]
```

O valor de discriminante da curva elíptica e é obtido adicionando o nome que demos a curva “e” seguido de “.disc”:

```
?   e.disc
%7  = Mod(8, 11)
```

Note que o valor acima é igual ao 12º o valor do vetor gerado pela função da equação A.1. Temos a função ellisoncurve(e, p), para verificar se o ponto representado por p pertence a uma curva elíptica representada por e, o programa responde 1 em caso afirmativo e 0 em caso negativo.

Vamos verificar se os pontos $q = (4; -4)$ e $r = (4; 4)$ pertencem ao grupo $E(0; 1; 3)_{\mathbb{F}_{11}}$ da curva elíptica $E(F_{11})$:

```
?    q = [Mod(4, 11), Mod(−4, 11)]
%8   = [Mod(4, 11), Mod(7, 11)]
?    ellisoncurve(e, q)
%9   = 1
?    r = [Mod(4, 11), Mod(4, 11)]
%10  = [Mod(4, 11), Mod(4, 11)]
?    ellisoncurve(e, r)
%11  = 1
```

Quando desejamos realizar operações com dois pontos p e q sobre a curva elíptica e, utilizamos as funções elladd(e, p, q) e ellsub(e, p, q) para somar e subtrair estes pontos respectivamente. Se desejamos multiplicar um ponto p por um número natural n utilizamos ellpow(e, p, n).

Vamos somar e subtrair os pontos $q = (4; -4)$ e $r = (4; 4)$:

```
?    elladd(e, q, r)
%12  = [0]
?    ellsub(e, q, r)
%13  = [Mod(7, 11), Mod(10, 11)]
```

Quando somamos os pontos acima, a resposta foi o vetor [0], que é o elemento neutro do grupo $E(0; 1; 3)_{\mathbb{F}_{11}}$. Observe que podemos obter o simétrico de um ponto utilizando a função ellsub e o elemento neutro. No nosso exemplo, já vimos que o simétrico de q é o r :

```
?    ellsub(e, [0], q)
%14  = [Mod(4, 11), Mod(4, 11)]
```

Vamos calcular agora o dobro do ponto $q = (4; -4)$:

```
?    ellpow(e, q, 2)
%15  = [Mod(7, 11), Mod(10, 11)]
```

O dobro de q também pode ser calculado somando q consigo, isto é:

```
?   elladd(e, q, q)
%15 = [Mod(7, 11), Mod(10, 11)]
```

Também é possível determinar a ordem do grupo sobre os pontos de uma curva elíptica $y^2 = x^3 + ax^2 + bx + c$. Através da equação dada pelo Teorema de Helmut Hasse

$$\#E(a;b;c)_{\mathbb{F}_p} = p + 1 - t$$

onde t é o traço da curva. No nosso exemplo, queremos achar o traço da curva elíptica $E(\mathbb{F}_{11})$, e basta utilizar a função ellap :

```
?   ellap(e, 11)
%16 = −6
```

Aplicando o Teorema de Helmut Hasse:

$$\#E(0;1;3)_{\mathbb{F}_{11}} = 11 + 1 - t = 11 + 1 - (-6) = 18.$$

Para saber a ordem de um determinado ponto p pertencente a um grupo $E(K)$ de uma curva elíptica, pode-se utilizar a função ellorder. Por exemplo, a ordem do ponto $q = (4;-4)$, pertencente ao grupo $E(0;1;3)_{\mathbb{F}_{11}}$

```
?   ellpow(e, q)
%17 = 9
```

Porém, devemos ressaltar que o programa PARI determina a ordem de um determinado ponto operando sobre o corpo $\mathbb{Q}$ — corpo dos números racionais. Entretanto, pode-se verificar se um número natural m é a ordem de um determinado ponto p pertencente a um determinado grupo $E(\mathbb{F}_q)$. De volta ao exemplo,a ordem do ponto $q = (4;-4)$, é de fato 9, pois ao usarmos ellpow pra calcular todas as potências de q^k para $k = 1, ..9$ obtemos

```
?   for(k = 1, 9, print(ellpow(e, q, k)))
%18 = [Mod(4, 11), Mod(7, 11)],
      [Mod(7, 11), Mod(10, 11)],
      [Mod(1, 11), Mod(7, 11)],
      [Mod(6, 11), Mod(4, 11)],
      [Mod(6, 11), Mod(7, 11)],
      [Mod(1, 11), Mod(4, 11)],
      [Mod(7, 11), Mod(1, 11)],
      [Mod(4, 11), Mod(4, 11)],
      [0]
```

A.3 Colocando a mão na massa: criptografando com a Pari/GP

Agora sim, vamos fazer criptografar com as curvas elípticas. Aqui fizemos um exemplo do ElGamal, um pouco diferente do apresentado no texto, com mais uma chave (efêmera).

Exemplo 40. *Vamos começar criptografando a palavra IMPA. Primeiro é preciso reescrevê-la em números, e para isso utilizamos a Tabela ASCII A.4:*

I	M	P	A
73	77	80	65

Agora vamos escolher um número primo e uma curva. Devemos recordar que o primo escolhido deve ter mais dígitos do que a mensagem. Vamos escolher com o Pari um primo com 10 dígitos:

```
?   isprime(8746593847)
%1  0
?   nextprime(8746593847)
%2  8746593881
?   isprime(%2)
%3  1
```

Nomeamos o primo p escolhido(para facilitar as nossas ações):

```
?   p = %2
%4  = 8746593881
```

Escolhemos $E : y^2 = x^3 + 2x + 1$ como a curva elíptica. Vamos indicar ao programa a curva escolhida e nomeá-la de e.

```
?   e = ellinit([Mod(0, p), Mod(0, p), Mod(0, p), Mod(2, p), Mod(1, p)])
%5  = [Mod(0, 8746593881), Mod(0, 8746593881), Mod(0, 8746593881),
    Mod(2, 8746593881), Mod(1, 8746593881), Mod(0, 8746593881),
    Mod(4, 8746593881), Mod(4, 8746593881), Mod(8746593877,
    8746593881), Mod(8746593785, 8746593881), Mod(8746593017,
    8746593881), Mod(8746592937, 8746593881), Mod(5633400386,
    8746593881), Vecsmall([3]), [8746593881, [2592, 46656, [6, 0, 0, 0]]],
    [0, 0, 0, 0]]
```

Vamos verificar que E é realmente uma curva elíptica sobre $\mathbb{Z}_{8746593881}$. Para isto calculemos o seu discriminante

```
?   e.disc
%6  = Mod(8746592937, 8746593881)
```

A curva de fato é elíptica pois $\Delta = 8746592937$ *é não nulo. Agora calculemos a ordem de* $E(\mathbb{Z}_{8746593881})$ *usando o Teorema de Hasse. O traço* t *da curva é dado por*

```
?    ellap(e,p)
%7   = 307000
```

Assim calculamos $\sharp E(\mathbb{Z}_{8746593881}) = 8746593881 + 1 - 307000 = 8746286882$. *Agora vamos escolher um ponto da curva. Testemos* $P = (0,1)$*, que realmente está na curva*

```
?    P = [Mod(0,p),Mod(1,p)]
%8   = [Mod(0,8746593881),Mod(1,8746593881)]
?    ellisoncurve(e,P)
%9   1
```

Calculemos a ordem de $P = (0,1)$

```
?    ellorder(e,P)
%10  = 8746563182
```

A ordem é grande o suficiente $\left(\frac{\sharp E(\mathbb{Z}_p)}{2}\right)$. *Assim definimos a chave pública* $(E(\mathbb{Z}_{8746593881}),$

Sendo $m = 73778065$ *a mensagem em ASCII decimal. Vamos mergulhar a mensagem na curva pelo método de Koblitz:*
Temos que escolher um inteiro λ *e encontrar um ponto* $(x,y) \in E(\mathbb{Z}_{8746593881})$ *no qual para* $x = m\lambda + n$*, com* $1 \leqslant n < \lambda - 1$*, e exista* y.

Vamos escolher $\lambda = 34751$ *testar* $x = m\lambda + 1$. *Usaremos o comando que calcula a ordenada dado o* x*:*

```
?    m = 73778065
%11  = 73778065
?    l = 34751
%12  = 34751
?    Mod(m*l+1,p)
%13  = Mod(1109529683,8746593881)
?    ellordinate(e,Mod(1109529683,8746593881))
%14  = [Mod(144072207,8746593881),Mod(8602521674,8746593881)]
```

Para $x = m\lambda + 1$ *existem dois valores para* y*, escolhemos o segundo. Podemos verificar que este ponto está na curva:*

```
?    M = [Mod(1109529683,8746593881),
     Mod(8602521674,8746593881)]
%17  [Mod(1973985093316083281651120,p),
     Mod(1725504466570196217458724150922 4339,p)]
?    ellisoncurve(e,M)
%18  = 1
```

Assim a mensagem mergulhada na curva é o ponto

$$M = (\overline{1109529683}, \overline{8602521674})$$

A *deseja mandar a mensagem* M *para* **B**. *Então* **B** *tem* $a_B = 357364$ *e publica* $A_B = a_B P$. *Calculemos:*

```
? A = ellpow(e, P, 357364)
%19 = [Mod(6848433471, 8746593881), Mod(7995053469, 8746593881)]
? ellisoncurve(e, ellpow(e, P, 357364))
%20 1
```

Deste modo,

$$A_B = A = a_B P = (\overline{6848433471}, \overline{7995053469}).$$

A *escolhe um inteiro* $k = 92843$ *para codificar a mensagem, e calcula* $C_1 = C = kP$ *e* $C_2 = S = M + kA_B$, *que são dados por*

```
? ellpow(e, P, 92843)
%21 = [Mod(2680262005, 8746593881), Mod(2920461035, 8746593881)]
? C = ellpow(e, P, 92843)
%22 = [Mod(2680262005, 8746593881), Mod(2920461035, 8746593881)]
? D = ellpow(e, A, 92843)
%23 [Mod(1428360812, 8746593881), Mod(2196344602, 8746593881)]
? elladd(e, M, D)
%25 = [Mod(1120832971263709802193778439414 2988, p),
      Mod(1624463261274460021357396700486 2786, p)]
? S = elladd(e, M, D)
%26 = [Mod(3204022292, 8746593881), Mod(5962494320, 8746593881)]
```

Daí **A** *envia para* **B** *o par* $[C_1, C_2] = [C, S]$

$$[2680262005, 2920461035, 3204022292, 5962494320]$$

Para decifrar a mensagem, **B** *calcula* $C_2 - a_B C_1$, *que é*

```
? u = ellpow(e, C, 357364)
%27 = [Mod(1428360812, 8746593881), Mod(2196344602, 8746593881)]]
? ellsub(e, S, u)
%29 = [[Mod(1109529683, 8746593881), Mod(8602521674, 8746593881)]
```

Assim **B** *obtém* $M = (1109529683, 8602521674)$. *Para decriptar* **B** *apenas precisa da primeira entrada do ponto e calcular* $\dfrac{x-1}{\lambda}$:

```
? (1109529683 - 1)/1
%30 = 73778065
```

Usando a tabela ASCII temos a mensagem m:

I	*M*	*P*	*A*
73	77	80	65

A.4 Tabela ASCII

Binário	Decimal	Hexa	Glifo
0010 0000	32	20	
0010 0001	33	21	!
0010 0010	34	22	"
0010 0011	35	23	#
0010 0100	36	24	$
0010 0101	37	25	%
0010 0110	38	26	&
0010 0111	39	27	'
0010 1000	40	28	(
0010 1001	41	29	)
0010 1010	42	2A	*
0010 1011	43	2B	+
0010 1100	44	2C	,
0010 1101	45	2D	-
0010 1110	46	2E	.
0010 1111	47	2F	/
0011 0000	48	30	0
0011 0001	49	31	1
0011 0010	50	32	2
0011 0011	51	33	3
0011 0100	52	34	4
0011 0101	53	35	5
0011 0110	54	36	6
0011 0111	55	37	7
0011 1000	56	38	8
0011 1001	57	39	9
0011 1010	58	3A	:
0011 1011	59	3B	;
0011 1100	60	3C	<
0011 1101	61	3D	=
0011 1110	62	3E	>
0011 1111	63	3F	?

Binário	Decimal	Hexa	Glifo
0100 0000	64	40	@
0100 0001	65	41	A
0100 0010	66	42	B
0100 0011	67	43	C
0100 0100	68	44	D
0100 0101	69	45	E
0100 0110	70	46	F
0100 0111	71	47	G
0100 1000	72	48	H
0100 1001	73	49	I
0100 1010	74	4A	J
0100 1011	75	4B	K
0100 1100	76	4C	L
0100 1101	77	4D	M
0100 1110	78	4E	N
0100 1111	79	4F	O
0101 0000	80	50	P
0101 0001	81	51	Q
0101 0010	82	52	R
0101 0011	83	53	S
0101 0100	84	54	T
0101 0101	85	55	U
0101 0110	86	56	V
0101 0111	87	57	W
0101 1000	88	58	X
0101 1001	89	59	Y
0101 1010	90	5A	Z
0101 1011	91	5B	[
0101 1100	92	5C	\
0101 1101	93	5D	]
0101 1110	94	5E	^
0101 1111	95	5F	_

Binário	Decimal	Hexa	Glifo
0110 0000	96	60	`
0110 0001	97	61	a
0110 0010	98	62	b
0110 0011	99	63	c
0110 0100	100	64	d
0110 0101	101	65	e
0110 0110	102	66	f
0110 0111	103	67	g
0110 1000	104	68	h
0110 1001	105	69	i
0110 1010	106	6A	j
0110 1011	107	6B	k
0110 1100	108	6C	l
0110 1101	109	6D	m
0110 1110	110	6E	n
0110 1111	111	6F	o
0111 0000	112	70	p
0111 0001	113	71	q
0111 0010	114	72	r
0111 0011	115	73	s
0111 0100	116	74	t
0111 0101	117	75	u
0111 0110	118	76	v
0111 0111	119	77	w
0111 1000	120	78	x
0111 1001	121	79	y
0111 1010	122	7A	z
0111 1011	123	7B	{
0111 1100	124	7C	|
0111 1101	125	7D	}
0111 1110	126	7E	~

B por Ulisses de Sá Alencar e Moraes

Aqui faremos uma breve leitura a respeito das diferenças entre aplicação e modelagem de criptografia com a chegada do período moderno.

B.1 Contextualização, jitter e aplicações de tempo real

Existem três elementos base que diferenciam a criptografia clássica e a moderna: identidade, tempo e meio.

Para definir que tipo de criptografia será utilizada é necessário entender por quanto tempo uma informação deve ser protegida e em quanto tempo ela deve ser transmitida. Quanto mais rápida a transmissão, maior é o custo para mantê-la protegida e maior será o custo para cifrá-la e decifrá-la.

Até a criação de criptografia por pares de chave, não havia nenhuma maneira de dois entes confirmarem a identidade um do outro. Hoje possuímos entidades que indiretamente garantem a identidade de um terceiro, sem mesmo participar da transação.

Um outro ponto relevante é que anteriormente, nos preocupávamos em criptografar o objeto à ser transferido, porém o próprio meio onde o objeto será transferido pode ser o melhor alvo.

Neste próximo exemplo vamos tratar do problema de alto custo computacional na comunicação de tempo real.

Exemplo 41. Uma conexão de video-chamada não pode atrasar ou sofrer alto jitter. No qual Jitter é uma variação estatística do atraso na entrega de dados em uma rede, ou seja,

pode ser definida como a medida de variação do atraso entre os pacotes sucessivos de dados. Uma aplicação real disso é o skype. Utiliza-se a criptografia RSA de 1536 ou 2048 bits para negociar um meio com cifragem de bloco AES, uma vez que o custo de usar RSA seria inviável em uma aplicação de tempo real.

B.2 Requisitos e restrições

De acordo com a CISCO uma ligação em tempo real deve ter delay máximo de 150ms para o tempo criptografia-transmissão-descriptografia. O ato de encriptar e decriptar para uma aplicação de tempo real sempre vai demandar o maior índice de segurança que atenda esse tempo. Neste contexto, a implementação com curvas elípticas se torna naturalmente mais eficiente que RSA, que por sua lentidão, é menos usado para criptografar diretamente os dados, e sim para compartilhar uma chave simétrica que, por sua vez, pode executar operações de criptografia-descriptografia em massa a uma velocidade muito maior.

No próximo exemplo, trataremos da melhor solução para os requisitos de um determinado cenário.

Exemplo 42. Os dados são medidos por seu nível de sensibilidade. Se um país declara que após 50 anos todos os dados sejam públicos, não haveria lógica em usar uma chave inquebrável por 200 anos, pois seu custo de uso superaria o requisito. Um caso real foi a ordem executiva do governo Clinton, para reavaliar a cada 25 anos se um dado ainda precisa ser mantido não público https://fas.org/sgp/clinton/eo12958.html.

B.3 Exemplos

As perguntas que temos de responder são: qual o tempo de sigilo do dado em questão e qual o tempo de resposta aceitável na comunicação?

Exemplo 43. Uma transação bancária pode ser feita completamente em sistema de par de chaves (garantia de chave complexa em toda transação).

Exemplo 44. TSL usa chave simétrica negociada por chaves assimétricas que são baseadas em chaves pré-compartilhadas (problema do túnel com chave mais simples).

Exemplo 45. Quando seremos capazes de usar uma chave assimétrica para uma conversa em tempo real?

Bibliografia

Certicom (1997). "Current Public-Key Cryptographic Systems". *A Certicom Whitepaper* April (ver p. 110).

S. C. Coutinho (2000). *Números inteiros e criptografia RSA*. IMPA/SBM: Série de Computação Matemática. MR: 1928229 (ver p. 30).

R. Descartes (1954). *The Geometry of René Descartes: With a facsimile of the first edition*. Traduzido do Francês e do Latin por David E. Smith e Marcia L. Latham. New York: Dover Publications, Inc, p. 272 (ver p. 45).

— (2002). São Paulo: Paulus (ver p. 45).

W. Diffie e M. E. Hellman (1976). "New directions in cryptography". *IEEE Trans. Information Theory* IT-22(6), pp. 644–654. MR: 0437208 (ver p. 32).

A. Gonçalves (2017). *Introdução à Álgebra, Volume 1*. IMPA: Projeto Euclides. MR: 0651519 (ver p. 5).

A. Hefez (2016). *Curso de Álgebra, Volume 1*. IMPA: Coleção Matemática Universitária (ver pp. 5, 6, 13).

J. Hoffstein, J. Pipher e J. H. Silverman (2008). *An introduction to mathematical Cryptography*. Springer: Undergraduate Texts in Mathematics, pp. xvi, 524. MR: 2433856.

N. Koblitz (1987). "Elliptic Curve Cryptosystems". *Mathematics of Computation* 48 No. 177, pp. 203–209. MR: 0866109 (ver p. 32).

K. Koyama, U. M. Maurer, T. Okamoto e S. A. Vanstone (1991). "New Public-Key Schemes Based on Elliptic Curves over the Ring $\mathbb{Z}_n$". *Advances in Cryptology - CRYPTO* 91, pp. 252–265. MR: 1243652 (ver p. 32).

S. Lang (2008). *Algebra*. Springer: Graduate Texts in Mathematics (ver pp. 25, 26).

Y. Lequain e A. Garcia (2018). *Elementos de Álgebra*. IMPA: Projeto Euclides. MR: 3838336 (ver pp. 5, 15, 16, 19, 28).

D. Mahto e D. K. Yadav (2017). "RSA and ECC: A Comparative Analysis". *International Journal of Applied Engineering Research* 12, pp. 9053–9061 (ver pp. 110, 111).

A. Menezes e S. Vanstone (1993). "Elliptic Curve Cryptosystems and Their Implementation". *Journal of Cryptology*, pp. 209–224. MR: 1248082 (ver pp. 98, 105).

V. S. Miller (1986). "Use of Elliptic Curves in Cryptography". *Advances in Cryptology-CRYPTO* 85, pp. 417–426. MR: 0851432 (ver p. 32).

G. L. Mullen e C. Mummert (2007). *Finite Fields and Applications*. American Mathematical Society: Student Mathematical Library, Mathematics Advanced Study Semesters. MR: 2358760 (ver p. 28).

M. Reid (2013). "Undergraduate Algebraic Geometry" (ver p. 85).

J. H. Silverman (2009). *The Arithmetic of Elliptic Curves*. Springer: Graduate Texts in Mathematics (ver pp. 83, 118).

J. H. Silverman e J. Tate (1992). *Rational Points of Elliptic Curves*. Springer: Undergraduate Texts in Mathematics.

I. Vainsencher (2017). *Introdução às Curvas Algébricas Planas*. IMPA: Coleção Matemática Universitária (ver p. 58).

M. L. T. Villela e A. Hefez (2017). *Códigos Corretores de Erros*. IMPA: Séries de Computação e Matemática (ver p. 15).

Índice de Notações

Índice de Autores

Índice Remissivo

Títulos Publicados — 32º Colóquio Brasileiro de Matemática

Emergence of Chaotic Dynamics from Singularities – *Pablo G. Barrientos, Santiago Ibáñez, Alexandre A. Rodrigues e J. Ángel Rodríguez*

Nonlinear Dispersive Equations on Star Graphs – *Jaime Angulo Pava e Márcio Cavalcante de Melo*

Scale Calculus and M-Polyfolds An Introduction – *Joa Weber*

Real and Complex Gaussian Multiplicative Chaos – *Hubert Lacoin*

Rigidez em Grafos de Proteínas – *Carlile Lavor*

Gauge Theory in Higher Dimensions – *Daniel G. Fadel e Henrique N. Sá Earp*

Elementos da Teoria de Aprendizagem de Máquina Supervisionada – *Vladimir Pestov*

Función Gamma: Propriedades Clásicas e Introducción a su Dinâmica – *Pablo Diaz e Rafael Labarca*

Introdução à Criptografia com Curvas Elípticas – *Sally Andria, Rodrigo Gondim e Rodrigo Salomão*

O Teorema dos Quatro Vértices e sua Recíproca – *Mário Jorge Dias Carneiro e Ronaldo Alves Garcia*

Uma Introdução Matemática a Blockchains – *Augusto Teixeira*

ISBN 978-85-244-0428-3

Instituto de
Matemática
Pura e Aplicada